THE AUTOIMMUNE DISEASES

Edited by

NOEL R. ROSE
Department of Immunology and Infectious Diseases
and Department of Medicine
The Johns Hopkins Medical Institutions
Baltimore, Maryland

IAN R. MACKAY
Clinical Research Unit
of The Walter and Eliza Hall Institute of Medical Research
and the Royal Melbourne Hospital
Melbourne, Victoria, Australia

1985

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
Orlando San Diego New York Austin
London Montreal Sydney Tokyo Toronto

ACADEMIC PRESS, INC.
Orlando, Florida 32887

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24-28 Oval Road, London NW1 7DX

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Main entry under title:

The Autoimmune diseases.

Includes index.
1. Autoimmune diseases. I. Rose, Noel R.
II. Mackay, Ian R. [DNLM: 1. Autoimmune Diseases.
WD 305 A9375]
RC600.A83 1985 616.97 85-1406
ISBN 0-12-596920-1 (alk. paper)
ISBN 0-12-596921-X (paperback)

PRINTED IN THE UNITED STATES OF AMERICA

85 86 87 88 9 8 7 6 5 4 3 2 1

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Donato Alarcón-Segovia (119), Department of Immunology and Rheumatology, Instituto Nacional de la Nutricion "Salvador Zubiran," México City, Distrito Federal 14000, México

Giuseppe A. Andres (339), Department of Microbiology, Department of Pathology, and Department of Medicine, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Grant J. Anhalt (443), Department of Dermatology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Barry G. W. Arnason (399), Department of Neurology, The University of Chicago, Chicago, Illinois 60637

Richard H. Aster (493), The Blood Center of Southeastern Wisconsin, Inc., and Medical College of Wisconsin, Milwaukee, Wisconsin 53233

Pierluigi E. Bigazzi (161), Department of Pathology, University of Connecticut Health Center School of Medicine, Farmington, Connecticut 06032

Robert M. Blizzard (201), Department of Pediatrics, University of Virginia Medical Center, Charlottesville, Virginia 22908

Gian Franco Bottazzo (227), Department of Immunology, The Middlesex Hospital Medical School, London W1P 9PG, England

Jan R. Brentjens (339), Department of Pathology, Department of Microbiology, and Department of Medicine, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Roger L. Dawkins (591, 669), Departments of Clinical Immunology, Royal Perth Hospital, and The Queen Elizabeth II Medical Centre, Perth, 6001, Australia

Luis A. Diaz (443), Department of Dermatology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Deborah Doniach (227), Department of Immunology, The Middlesex Hospital Medical School, London W1P 9PG, England

P. Brent Ferrell[1] (29), George Washington University School of Medicine, Washington, D.C. 20037

Michael J. Garlepp (591), Departments of Clinical Immunology, Royal Perth Hospital, and The Queen Elizabeth II Medical Centre, Perth, Western Australia 6001, Australia

Stephen B. Hanauer (267), Department of Medicine, The University of Chicago, Chicago, Illinois 60637

Leonard C. Harrison (617), Department of Diabetes and Endocrinology, The Royal Melbourne Hospital, Melbourne, Victoria 3050, Australia

Sumner C. Kraft (267), Department of Medicine, The University of Chicago, Chicago, Illinois 60637

Parviz Lalezari (523), Division of Immunohematology, Department of Medicine, Montefiore Hospital, and Albert Einstein College of Medicine, The Bronx, New York 10467

Ian R. Mackay (1, 243, 291), Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital, Melbourne, Victoria 3050, Australia

Noel K. Maclaren (201), Department of Pathology, and Department of Pediatrics, College of Medicine, University of Florida, Gainesville, Florida 32610

Alan C. Menge (537), Department of Obstetrics and Gynecology, The University of Michigan Medical Center, Ann Arbor, Michigan 48104

Bernice Noble (339), Department of Microbiology and Department of Pathology, School of Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Sir Gustav Nossal (695), The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria 3050, Australia

Robert B. Nussenblatt (371), Clinical Ophthalmic Immunology Section, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20205

Harish P. Patel (443), Department of Dermatology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Bernard Pirofsky (469), Division of Immunology, Allergy, and Rheumatology, Department of Medicine, The Oregon Health Sciences University, Portland, Oregon 97201

Thomas T. Provost (443), Department of Dermatology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

[1] Present address: Shelby Medical Associates, Shelby, North Carolina 28150.

Noel R. Rose (1, 161), Department of Immunology and Infectious Diseases, and Department of Medicine, The Johns Hopkins Medical Institutions, Baltimore, Maryland 21205

Gordon C. Sharp (81), Division of Immunology and Rheumatology, Department of Medicine, University of Missouri–Columbia School of Medicine, Columbia, Missouri 65212

Arthur M. Silverstein (371), The Wilmer Ophthalmological Institute, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Bernhard H. Singsen (81), Departments of Child Health, Medicine, and Pathology, University of Missouri Health Sciences Center, Columbia, Missouri 65212

Kate M. Spencer (227), Department of Diabetes and Immunology, St. Batholomew's Hospital, London, England

Norman Talal (145), Division of Clinical Immunology, Department of Medicine, The University of Texas Health Science Center at San Antonio, and Clinical Immunology Section, Audie L. Murphy Memorial Veterans Administration Hospital, San Antonio, Texas 78284

Eng M. Tan (29), W. M. Keck Foundation Autoimmune Disease Center, Research Institute of Scripps Clinic, Scripps Clinic and Research Foundation, La Jolla, California 92037

Kenneth S. K. Tung (537), Department of Pathology, University of New Mexico School of Medicine, Albuquerque, New Mexico 87131

Mathew A. Vadas[2] (429), Institute for Medical and Veterinary Science, Adelaide 5000, Australia

Senga Whittingham (243), Clinical Research Unit, The Walter and Eliza Hall Institute of Medical Research, The Royal Melbourne Hospital, Melbourne 3050, Australia

Mathew A. Vadas[2] (429), Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital, Melbourne, Victoria 3050, Australia

Senga Whittingham (243), Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital, Melbourne, Victoria 3050, Australia

Morris Ziff (59), Department of Internal Medicine, The University of Texas Health Science Center at Dallas, Dallas, Texas 75235

[2] Present address: Institute for Medical and Veterinary Science, Adelaide, South Australia 5000, Australia.

Foreword

Like many other medical scientists I first became aware of autoimmune disease as a clinical entity from the work of Dameshek and Schwartz in 1937, and I have been interested in it ever since although always mainly from its bearing on immunological theory. During 1962 I collaborated on the theoretical side with Dr. Mackay in writing one of the earlier texts in English, which was published in 1963. In the ensuing 20 years experimental and clinical research has enriched the field immensely and widened the range of diseases attributable wholly or in part to autoimmunity. At least we were thinking along the right lines, and I am happy to contribute a foreword to the present volume.

To one who has been away from active research in immunology since 1965, the dominant impression of what has happened since is the daunting complexity of the genetic and biochemical processes that are concerned with the regulation of the immune system. Almost all recent experimental work designed to elucidate basic immunology has centered on the use of mouse strains of known genetic composition. As a result, a much more precise understanding of the structural chemistry and genetic control of the diversity of specific patterns in antibodies is now available. This holds also for the nature and developmental sequence of the several species of immunoglobulins and their respective functions. Many other relevant substances, particularly cell surface antigens and receptors, lymphokines, and other cell secretions, have also been effectively studied. All of these body components concerned in immune responses are proteins directly or indirectly coded for by germ line structural genes. In the case of antibodies at least, diversity is achieved by modification at some stage of differentiation, within the lines of somatic stem cells, of one or more germ line genes by transposition or recombination, or by somatic mutation.

It must be remembered, however, that the mammalian genome contains a great deal of DNA not demonstrably coding for specific protein and presumably concerned with regulatory processes, including the timing

and correlation of activity of each structural gene during the processes of growth, differentiation, and repair. Far less is known about the detailed structure and function of regulatory DNA than about structural DNA and the peptide sequences that it codes for. Biologically, however, it is probably just as important as the protein-synthetic system, particularly in determining the details of morphological and functional differentiation during development. It is well known that the chemical and serological qualities of the proteins of humans and chimpanzees are so extremely similar that these can be regarded as sibling species. The differences to be found in almost every detail of morphology, therefore, must be derived from mutational or other types of modification of the regulatory segments of the genome of the common ancestral species.

Most experimental and theoretical immunologists now recognize that the immune system is an immensely complex web of interactions that evolved to deal with a wide variety of exceptional situations. Its regulation obviously requires a sophisticated system of communication, with active responses and feedbacks involving either a proliferative response, a step in differentiation, or an activation of secretory function, and, on the negative side, inhibitory or destructive responses of corresponding type. The concept of immune surveillance was introduced to account for a variety of oncological phenomena but nowadays there is a growing opinion that it is principally concerned with, or is a by-product of, the internal specific control of all types of lymphocytes and their derivatives. Katz has summarized the position by saying that all populations of functioning lymphocytes must be under positive and negative control coordinated to optimize the survival of the individual exposed to infection or to some internally induced change of "self" pattern. Jerne's concept of immune regulation by antibody, or suppressor cells with specificity for idiotypes, is one example of how this may be achieved.

In the absence of any accepted interpretation of immune regulatory processes, I still prefer to think of them broadly as complex homeostatic and self-monitoring functions. Very early in this century Ehrlich clearly expressed the logical necessity that the vertebrate body must in some way be inhibited from developing antibodies reactive against its own components, and this constraint has been more or less consciously recognized ever since. It can be expressed broadly in the form that the function of the immune system is to recognize the presence of substances or chemical patterns foreign to the body and to eliminate them or nullify their harmful actions.

At the risk of being too easily satisfied, one can still look at the problem in general biological terms rather than at a molecular level and use terms more appropriate perhaps to the 1950s than to the 1980s. The immune

system in my view exists primarily to maintain the structural and chemical integrity of the body. The facts of infectious disease, of allergic complaints, and of the troubles that may follow injection or transplantation of cells or tissues from other individuals all suggest that the entry of things that are "not-self" parenterally is harmful and, at least in relation to microbial infection, has needed the evolution of mechanisms to avoid or deal with any such entry into the tissues. In broad terms what is needed is for cells to recognize the foreign quality of the material, to counteract any damaging quality, and in some way to segregate and to destroy or eliminate the offending particles or molecules.

From what we now know of the production and functioning of antibody it is clear that what is needed is a repertoire of epitopes (immune specificities) on immunoglobulins and immunocytes, sufficient to cover as many types of foreign molecular groups as is physically possible but without such action on all molecular groupings of body ("self") components that are accessible to circulating cells or proteins in the blood or lymph. Production of such a repertoire demands a set of mechanisms by which a very wide—virtually infinite—diversity of molecular patterns can be conferred on immunoglobulins or the related cell receptors, and also that any of those patterns reactive with self-components are in one way or another effectively inhibited or eliminated from the repertoire or rendered incapable of being stimulated to proliferation or other significant activity. If this is to be accomplished without the general metabolism being disturbed, an elaborate communications system is clearly needed. Much of the control of the antibody system is at the genetic level. The random transposition of alternative DNA segments within a well-ordered framework plus accelerated point mutation in the *V* gene segments indicates some of the subtlety of the processes.

When we come to cellular aspects of immune responses, however, experimental immunology is much less revealing. It is evident that many classes of lymphocytes carry receptors of similar specificity to those forming the variable regions of antibodies. The B-cell surface receptor responsible for specific cellular activation after contact with antigen appears to be monomeric IgM synthesized by the cell itself. Must less is known of the specific T-cell receptor. There is considerable evidence, which I find impressive, that points to the passive transfer of the whole or part of B-cell receptors to T cells, which then take on the characteristic specificity in their functional responses. Others consider that T-cell specificity is autonomous to the cell concerned.

The activities of higher cells, suppressor cells, cytotoxic and nonspecific killer cells among the T cells have not yet been fully defined. The implication is that within the immune system there must be means by

which the numbers and distribution of every recognizably different type of circulating cell can be monitored and, if necessary, be stimulated to proliferate, functionally inhibited, or destroyed. In its own way communication and control of the immune system may be as complex as its social analogs, the control of crime and delinquency or the economics of an industrial society. Autoimmune disease has often been likened to the revolutionary violence of a rebellious section of the national police force, and there may be just as many possible reasons for the conflict to become clinically overt.

Immune tolerance to self-components is mediated through more than one mechanism, but when antibody production is being considered I believe that clonal deletion is the dominant process. If a precursor B cell at its first surface expression of the specific antibody receptor makes contact with the corresponding antigenic determinant in adequate concentration, the precursor cell is lethally damaged. As the clone matures, its cells become less susceptible to damage and are more likely to show a proliferative response to the same specific stimulus.

Autoimmune disease is not easy to define or to interpret. For fairly obvious reasons the original hallmark of autoimmune disease was the presence of autoantibodies, that is, antibodies reactive with normal constituents of the body. The prototype was "warm-type" autoimmune hemolytic anemia as described by Dameshek and Schwartz in 1950. Virtually all the generally recognized autoimmune diseases have shown one or more types of autoantibody not always directly related to the symptoms of the condition. In healthy subjects the appearance of any of the common autoantibodies tested for in clinical laboratories is quite unusual, but there is a distinct increase in their incidence in old age.

In some ways the simplest interpretation of the appearance of an active clone of B cells producing autoantibody is that a mutation (germ line or somatic) has significantly increased its resistance to tolerization by the corresponding autoantigen. In view of the complexities, known in detail or recognized in principle, that are needed for a normally functioning immune system, it is only to be expected that a wide range of somatic genetic errors may be needed for the appearance of an autoimmune clone. The mutations or other type of change in DNA pattern may involve either germ line cells or somatic cells at various stages of differentiation; and, for overt manifestations of the clone, appropriate functional states of the internal environment will be necessary.

Autoimmune disease has many resemblances to a conditioned malignancy and modern thinking on the initiation of malignancy may well be relevant. There is a general agreement that the final step in the development of at least one important group of cancers is the undue activity of the

product of an oncogene, itself derived from a normal gene which is presumably concerned with cell proliferation in development and repair and which is referred to as a proto-oncogene. According to work in Weinberg's laboratory, the change in structure associated with the proto-oncogene to oncogene change is located in a single nucleotide, the codon GGC being changed by a point mutation to GTC, corresponding to a glycine–valine change in the gene product. Scolnick's group, however, has published evidence that a regulatory change, by which abnormally large amounts of the gene product are synthesized, can produce the same cell-transforming effect.

The eventual interpretation of the appearance of autoimmune clones of B or T cells will almost certainly be expressed in molecular terms that may involve a sequence of point mutations, anomalous transpositions, or some other intragenomic change, functionally equivalent to somatic mutation. It remains to be seen whether, when such investigations have been successfully completed in mouse model systems of autoimmune disease, the results can be usefully applied to autoimmune conditions in humans.

Perhaps the main lesson to be learned from the modern elucidation of the mammalian immune system is that of the extraordinary complexity of its regulatory mechanisms at both genetic and physiological levels. Clinical experience suggests that every case of serious autoimmune disease has its own individual peculiarities, and it will probably always be impracticable to provide a complete picture of the primary and secondary processes responsible for the patient's condition. In a work which must be firmly directed toward the clinical realities of diagnosis and treatment in the various manifestations of autoimmune disease, theoretical studies in immunology and immunopathology have only a very limited place. The handling of individual patients must conform to the current consensus of clinical opinion. Basic research in immunopathology will continue, and if past experience is any guide such research will provide at intervals ideas that will modify the clinical approach. Equally important is the potential of clinical and epidemiological studies in human disease to influence general thinking on the nature of autoimmune disease. The distribution of HLA types in relation to clinical conditions, particularly of immunopathological character, is one of the most important findings of this sort.

Every chapter of this book represents a fusion of acts and ideas from both the clinical and experimental fields. It is both inevitable and essential that this cooperation go on indefinitely.

Sir Macfarlane Burnet
Canterbury, Victoria
Australia

Editors' Tribute to Sir Macfarlane Burnet

We record with great sadness that Frank Macfarlane Burnet, Nobel Laureate, died on August 31, 1985, a few days before his 86th birthday and a few months before the publication of this book. His remarkable research career spanned six decades, sequentially moving from microbiology to virology to immunology. In his prophetic monograph with F. Fenner, "The Production of Antibodies" (1949), he introduced the concept of immunological tolerance, in discussing how mice infected in vitro *with lymphocytic choriomeningitis virus remain free of illness, resistant to virus challenge, and lack evidence of any immune response to the virus: "These phenomena are obviously complex but there is the development of a tolerance to the foreign microorganism during embryonic life . . . ."*

Burnet's conceptual framework for immunology was constructed from two dominant themes: the principle of self/non-self discrimination as a mandatory condition for survival, and the application of principles of micro-Darwinian evolution to explain how lymphoid cells proliferate clonally to provide for those populations of lymphocytes best equipped to serve the needs of the organism—clonal selection. *His prescient application of these themes to immunopathology led to his enduring interest in autoimmune disease—an interest in which the Editors were privileged to participate—and is epitomized in his most valued Introduction to this book.*

Preface

The impetus for the book arose when one of us (N.R.R.) was fortunate enough to spend a sabbatical year working in the laboratory of the other (I.R.M.). During our many spirited discussions, we realized how much our thinking about autoimmunity and autoimmune disease had expanded since Mackay and Burnet published their landmark book on the subject in 1963. Conceptually, study of self-recognition and autoimmunity has produced fresh insights into normal, as well as abnormal, immunoregulation; future clinical possibilities for specific intervention have increased the need for accurate and early diagnosis of autoimmune diseases.

This book, then, is designed to provide a deeper understanding of this increasingly important field of medical science for physicians and investigators involved in the diagnosis, treatment, or research of autoimmune disease. Each chapter is written by an experienced clinician/investigator, but to obtain some measure of uniformity, each author was encouraged to follow a common outline. We are pleased that Sir Macfarlane Burnet has written a foreword that furnishes an historical perspective of our current concepts of self/non-self discrimination. A final chapter by Sir Gustav Nossal offers a glimpse into the future, relating the newer understanding of molecular immunology to innovative approaches to the treatment of autoimmune disorders. We have also included an introduction, giving accounts of recent investigations into cellular interactions, and a chapter discussing genetic regulation, as these subjects pertain to the induction of autoimmune disease and the identification of populations or individuals at greatest risk. A separate chapter by Roger Dawkins examines the practical problems of the evaluation and interpretation of immunological tests for autoantibodies.

All of us who participated in the preparation of this book hope that the readers will find it both useful and stimulating.

Noel R. Rose
Ian R. Mackay

Introduction

By the Editors

AUTOIMMUNITY VERSUS AUTOIMMUNE DISEASES

The term *autoimmunity* carries the general connotation of an immune response to constituents of the body (self-constituents) to which the immune system should be inert. This idea goes back to the expression *horror autotoxicus* of Ehrlich in 1901, strengthened in succeeding years by the general failure of experimentalists to evoke immune responses in animals by injection of their own tissue extracts. Indeed, the inviolate character of immune tolerance to autologous tissues led Burnet to suggest that self/non-self discrimination was the primary function of the immune system. It is now known that nonreactivity to self is not absolute, since sensitive assay systems usually reveal low levels of autoantibody in healthy members of most species studied. In population surveys, there is a proportion of subjects in whom autoantibodies are readily evident but in whom there is no demonstrable disease (see Chapter 24). Some of the subjects will be in what is now recognized as the long (3–5 years) preclinical phase of autoimmune disease, whereas others may have a transient expression of autoantibodies in their blood. The term *autoimmunity* can be applied to these states.

The definition of autoimmune disease is more difficult. The experimentalist would demand not only that the state of autoimmunity be demonstrable, whether experimentally induced or naturally occurring, but also that the autoimmune response can unequivocally be shown to cause disease, usually by transfer of cells or serum to another animal. In the clinical setting, the rigorous demonstration of an autoimmune effector mechanism is seldom possible. There are some instances of natural placental transfer of autoimmune disease from mother to fetus, exemplified by neonatal myasthenia gravis, thyrotoxicosis, thrombocytopenic purpura, and systemic lupus erythematosus (SLE), and there are a few exam-

ples of deliberate production of autoimmune disease in humans by transfer of plasma, for example, thrombocytopenic purpura (see Chapter 19). It is frequently necessary, therefore, to resort to experimental models to establish the autoimmune etiology of a human disease. In general, clinicians apply the term *autoimmune disease* to conditions of uncertain etiology in which there is a ready demonstrability of autoantibody of a relevant specificity. In most of these diseases, the circumstantial evidence linking the state of autoimmunity and the disease is so strong that a causal association can reasonably be assumed; for example, chronic thyroiditis, SLE, pemphigus vulgaris, and myasthenia gravis. For other diseases the evidence is more tenuous; among these we include ulcerative colitis, multiple sclerosis, alopecia totalis, and vitiligo.

NATURAL SELF-TOLERANCE AND AUTOIMMUNITY

The state of natural self-tolerance is defined by the absence of immune reactivity to autologous constituents. Explanations for self-tolerance are difficult because what can be demonstrated *in vitro* may not necessarily have particular relevance *in vivo*. The possible mechanisms, cited in historical order of their development, are three:

1. *Antigen "overload."* This concept has developed in various directions, including blockade of antigen receptors, immunological paralysis of the Felton type, and elimination of free antigen by conjugation with antibody or immunocytes in the circulation. Of these ideas, the concept of antigen blockade of receptors on lymphocytes is still current (Nossal, 1983),

2. *Deletion of self-reactive cells.* This notion was one of the corollaries of the initial clonal selection theory. It was assumed that lymphocytes with a capacity to react with self-constituents were deleted by exposure to antigen in excess at a critical stage of their maturation. This process was held to occur predominantly in fetal life but operated throughout life to maintain tolerance in the event of self-reactive cells arising by somatic mutation. There has been some degree of adherence to deletional theories of self-tolerance, at least in modified forms, to explain B-cell tolerance, although we recognize at the same time that low levels of autoantibodies in blood are the rule rather than the exception. For the T cell the likelihood of deletional tolerance is more plausible. The simplest formulation of the theory is that there is diversification during ontogeny in the thymus, perhaps under the influence of products of the major histocompatibility complex (MHC), and cells which express self-reactivity (in association with self-MHC) are eliminated, at least functionally. Consistent with this

notion, but not serving as direct proof, is the high level of death *in situ* of fetal thymocytes. We express the reservation that, experimentally, exposure of lymphoid cells to antigen in excess has not yet been shown to carry lethal consequences. Along these lines, there are strains of mice that have an innate lack of T-cell responsiveness to particular antigens; this is explained when developing thymocytes are exposed to a particular antigenic specificity, which results in deletion of that reactivity. The resulting failure to respond creates what are described as "holes" in the T-cell repertoire (Klein, 1982).

3. *Active regulation.* This "third generation" concept to explain self-tolerance stems from the two mainstream ideas: the "infectious tolerance" of Gershon and Kondo as the forerunner of immune suppression (Gershon, 1974) and the anti-idiotype network of Jerne (1974). In regard to the first mainstream idea, the phenomenology of T-cell suppression can no longer by doubted, and there are unequivocal markers for cells that transfer unresponsiveness in experimental animals. In fact, a complex infrastructure of cellular suppressor systems has been built up, including antigen-specific suppressor cells, idiotype-specific suppressor cells, and nonspecific suppressor cells; added to these, but with opposing effects, is the putative contrasuppressor cell that can preempt the activities of suppressor cells. We acknowledge a minority viewpoint that activities attributed specifically to suppression can be explained without invoking the existence of a distinct subset of T cells. The second mainstream idea is that of anti-idiotype regulation. "Idiotype" refers to the antigenic specificity of antibody produced in response to a given antigenic determinant or epitope (see Chapter 23). Such antibody itself generates an antibody response: anti-idiotypic antibody. The potential regulatory capacity of such antibody, through its action on receptors on B cells or even T cells, was drawn upon by Jerne (1974) as the basis for a network concept of idiotype/anti-idiotype interactions, providing an essential regulatory process in immune function. The existence of endogenously arising anti-idiotypic antibody has been established but the importance of its role in maintaining natural tolerance is still *sub judice*.

It is worth noting that maintenance of self-tolerance is clearly of such importance that, on a fail-safe basis, several coexisting processes could well be expected to operate.

CIRCUMVENTION OF SELF-TOLERANCE

Circumvention of self-tolerance is synonymous with autoimmunity. Historically, the earliest idea was that some self-constituents were "sequestered." Such constituents were shielded from the immune system in

the sense of having no opportunity to establish tolerance. Examples were thought to include lens protein, certain brain proteins, thyroglobulin (an erroneous assumption), and sperm. When such constituents were released in the course of tissue damage, they would be processed as "foreign" antigens and would generate an (auto)immune response. This mechanism is still not excluded and may be the explanation for immunization to sperm antigens after vasectomy (see Chapter 21).

On clonal selection-*cum*-deletional theories of self-tolerance, circumvention of self-tolerance was explained by abnormality on the part of the lymphoid system, independent of any role of antigen. It was assumed that self-reactivity occurred *sui generis*, based on somatic mutations that influenced the structure of the antigen-binding receptor of lymphocytes. Thus, there emerged cells capable of responding to self-antigens and having a resistance to normal homeostatic controls. In other words, an autonomous clone of lymphocytes arose that could inflict damage to autologous tissues and, in the process, cause release of further antigen, so sustaining the process.

The concept that natural tolerance is based on deletion or inactivation of self-reactive T cells gave rise to the T-cell bypass theory, developed concurrently by Allison *et al.* (1971) and Weigle (1980). This theory, utilizing the carrier–hapten model of B-cell activation, specified that the presence of self-reactive B cells is permissible so long as self-reactive T cells are "silenced" and so are unable to provide help to B cells for autoantibody formation. However, the presentation of a foreign carrier molecule derived, for example, from a virus or drug in association with a self-hapten would allow for a T-cell response to the carrier, thereby providing help for B cells reactive with the hapten. Of the several experimental models illustrating T-cell bypass, the most convincing is that based on the liver-specific F antigen (see Chapter 12). In the mouse, F antigen exists as two allotypes. Immunization of responder mice with allotypic F antigen, which provides the foreign carrier, induces antibody that is both allo- and auto-reactive; that is, stimulation by the T-cell carrier component results in break of self-tolerance to the hapten component of F antigen. However, the degree to which T-cell bypass can account for the wide range of self-sustaining autoimmune diseases is uncertain.

Another of the earlier concepts on induction of autoimmune responses was molecular mimicry, with rheumatic carditis as the most familiar example (see Chapter 16). It was shown that a particular streptococcal cell-wall antigen has a molecular configuration similar to that of a sarcolemmal antigen of heart muscle and so, in the course of streptococcal infection, antibody produced against the streptococcal antigen is incidentally reactive with heart. Other examples based on this idea have been cited from

time to time, including the demonstration by immunoblotting of epitopes common to bacteria and the acetylcholine receptor (Stefansson *et al.*, 1985). To support the concept of molecular mimicry for induction of a given autoimmune disease, one should be able to discern both a demonstrable clinical association between that disease and the presence in the patient of the offending microorganism, as well as demonstrable cross-reactivity between antigens of the microorganism and host tissues.

If the reality of regulatory influences on immune responses is accepted, then failure of suppressor mechanisms must contribute to states of autoimmunity. Unfortunately, in humans, the role of antigen-specific suppressor circuits is still poorly understood because reliable assays are yet to be developed. Functional assays for nonspecific suppressor effects show evident depression corresponding, in general, with disease exacerbations, indicative at least of a contributory role for T-suppressor dysfunction. Enumeration in disease of T cells bearing markers for helper and suppressor subsets has added little information, because markers that precisely define the relevant human suppressor populations are lacking.

In the recent past, attention has turned to the question of antigen presentation in the induction and expression of autoimmunity. There is, in fact, relatively little information on the processing of self-antigens by antigen-presenting cells, and it could even be argued that a strong element in the maintenance of self-tolerance is exercised at the level of the handling of self-antigens by macrophages or other antigen-presenting cells. Of equal interest is the question of the expression of MHC antigens on target tissues involved in autoimmune reactions. It would be expected that class II MHC antigens (I-A in mice, DR in humans) would be involved in the presentation of autoantigen, and there is clear evidence that tissues showing destructive reactions express class II MHC antigens (see Chapter 23). Whether expression of class II MHC antigens is a primary event or is secondary to the release of interferons or other cytokines in the course of cytodestructive processes is uncertain.

GENETIC DETERMINANTS OF AUTOIMMUNE DISEASES

We recognize a degree of genetic predisposition to autoimmune disease, with the genetic loading differing for each disease and varying according to the contribution from either parent (see Chapter 1). There must be an interplay between genetic predisposition and environmental determinants, and in some instances a sufficiently strong environmental stimulus—infectious, physical, or chemical—will induce autoimmunity in the

absence of apparent genetic predisposition. There are also somatic genetic influences which could account for the chance component in autoimmune disease, if the likelihood is accepted that mutational events in replicating lymphoid cells will generate receptors with particularly high affinities for autologous components. Evidence supporting this concept is provided by the well-recognized existence of autoimmune reactivity in some B-cell lymphomas and the occurrence of monoclonal proteins of the myeloma or macroglobulinemia types with high affinity for autoantigens (see Chapter 6). Reactivities described include those against IgG Fc determinants (rheumatoid factors, Chapter 3), erythrocyte antigens in the form of cold-reactive autoantibodies (Chapter 18), thyroglobulin, and the cytoskeletal filaments actin and vimentin. These considerations warrant reference to the possible involvement of an oncogene to establish the autonomous potential of an autoimmune lymphoid clone; this possibility seems worthy of exploration by application of the now-available techniques of molecular biology.

CAUSE OF TISSUE DAMAGE

Despite all that has been learned about immunopathology, we still know relatively little about the mechanisms of tissue damage in human autoimmune disease. The initial idea that damage was mediated by circulating autoantibody tended to be discarded in favor of cellular immune processes because titers of autoantibody did not seem to correspond with tissue damage. However, it seems clear that humoral antibody can be incriminated in several types of autoimmune disease, including hemocytolytic diseases (see Chapters 18–20), antireceptor diseases (see Chapter 22), and the bullous skin diseases (see Chapter 17). The evidence is based on natural transfer of these diseases transplacentally or by inoculation of serum into mice with reproduction of the characteristic clinical and histological features. Whether or not these antibody-dependent diseases require the participation of complement for effective induction of lesions is uncertain, but this is highly likely, for some at least. However, when pemphigus vulgaris is transferred to a neonatal mouse by injection of serum from affected patients, the lesions on the mouse's skin appear so readily and quickly that a role for complement seems unlikely; moreover, induction of lesions *in vitro* or by transfer to mice by IgG occurs despite absence or depletion of complement (see Chapter 17).

Another antibody-dependent mechanism of tissue damage is that mediated by K cells, which bind by Fc receptors to cells sensitized by attached autoantibody (see Chapters 7 and 21). The effect is more readily demon-

strable *in vitro* and has been invoked to explain damage to thyroid epithelial cells in autoimmune thyroiditis. Although it is true that the thyroid-specific antigen can be expressed on the thyroid cell surface, and so represent a target for antibody-dependent cytotoxicity, the evidence for a K-cell effector is perhaps better for virus-infected cells or tumor cells than that for cells expressing autoantigens. Moreover, histochemical studies using monoclonal antibodies show few cells with K-cell markers at sites of tissue destruction.

The participation of macrophages in autoimmune tissue lesions must be mentioned since such cells can be activated by lymphokines or become "armed" with cytophilic antibodies. Such cells are frequently demonstrable among the cellular aggregates in autoimmune lesions but their actual contribution to tissue damage is not known (see Chapters 3 and 11).

Immune complexes are a major mediator of tissue damage in some autoimmune diseases, for example, SLE and related disorders and rheumatoid arthritis (see Chapters 2 and 3). It can be readily shown with immunofluorescence that antibody and complement are deposited in tissue lesions and, in the case of SLE, that depletion of serum complement accompanies active disease. In SLE the usual autoantigen, DNA, has been eluted from immune complexes in the kidney. It is generally assumed that damaging immune complexes are formed at sites remote from those at which damage is occurring, and perhaps certain tissues such as kidney or brain have an affinity for circulating immune complexes. Perhaps the character of the antibody component of the complex, such as the isoelectric charge, is an important factor (Gavalchin *et al.*, 1985). Another possibility is that the immune complex forms *in situ*. This could occur if an antigen, for example, DNA, were itself deposited in, say, the renal glomerulus, or if antibody were present to high titer in the circulation or produced locally, and antigen was being released locally as a result of inflammatory activity in the target tissue. The latter appears to be one of several mechanisms involved in the lesions of autoimmune thyroiditis (see Chapter 7).

The T lymphocytes are likely effector agents in some autoimmune diseases, at least as judged by cell counts, using monoclonal antibodies, in affected tissues. Although such studies generally show a relative but slight excess of the suppressor/cytotoxic population, this immunohistological evidence is not decisive as to the likely T-cell subset involved. It can definitely be stated that T lymphocytes are the effector agents in certain experimentally induced autoimmune diseases and, in mice, that disease is transferable to syngeneic animals by an inoculum of T lymphocytes, either from diseased animals or from cultured T-cell lines (see Chapter 15). Of interest, the phenotype of the cells that transfer disease is that of the

helper subset, for example, in experimental autoimmune encephalomyelitis (EAE) in the mouse. This implies the necessity for an induction of cytotoxic cells in the recipient, or that cells of T-helper phenotype are themselves capable of causing damage, presumably by release of cytokines with effector or facilitatory properties. Such cytokines might include interferon, which could cause MHC class II antigen expression on target cells, or cytotoxic lymphotoxins (see Chapter 7). However, the limited cell specificity of cytotoxic lymphotoxins calls for a guarded assessment of their role in autoimmune tissue damage.

A general reservation which can be expressed regarding the involvement of the T-helper subset in autoimmune disease is the difficulty in detecting the "footprints" of this subset by assays *in vivo* or *in vitro* in humans. Although there have been many descriptions of assays based on lymphokine release, with apparently positive results in many instances, clinical immunologists have not found it easy to establish robust lymphokine release assays that regularly and reliably demonstrate cell-mediated immunity to autoantigens in human autoimmune disease. Moreover, skin tests for delayed-type hypersensitivity to autoantigens, in the limited studies in which these have been done, have been usually transient or negative.

The activity of cytolytic MHC-restricted T lymphocytes is well illustrated under defined experimental conditions *in vitro*. For example, cytotoxic T lymphocytes of Lyt-2 phenotype, induced in the course of experimental autoimmune thyroiditis in the mouse, can readily be shown to destroy thyroid cell monolayers but are subject to H-2 restriction (see Chapter 7). Parallel demonstrations of this effect in human autoimmune disease have been sought assiduously, but workers have been confronted with serious technical problems. First, the requirements of MHC compatibility demand that autologous cultured target cells be used, and these are seldom readily available. Second, the cells must be maintained during culture in a differentiated state and continue to express their relevant tissue-specific and MHC antigens. Third, the cells must be sufficiently stable to permit reliable isotope release assays to quantitate the cytoxicity of effector T lymphocytes. Since these conditions can seldom be met, the question of the activity of cytotoxic T cells in human autoimmune disease is unanswered.

The last effector candidate is the natural killer (NK) cell, which has long been studied in the setting of tumor immunology, but the scope of activities of NK cells *in vivo*, and the cell types on which these cells operate, are still unresolved issues. As for K cells, there is a scarcity of cells with the phenotypic markers of NK cells among the lymphoid infiltrates in autoimmune diseases (see Chapter 7). There is current interest in

the function of activated NK cells and the possible origin of such cells either from NK cells or from a differentiation step from cytotoxic T cells. These questions so far have been examined only by studies *in vitro*

WHAT ARE THE "AUTOIMMUNE DISEASES"?

Any selection of diseases to be included in a text or review as being "autoimmune" must of necessity be arbitrary. For most of the diseases discussed in this volume the evidence for an autoimmune contribution to pathogenesis, direct or indirect, is substantial. However, for some of the diseases included here, the evidence for an autoimmune contribution is still controversial, for example, scleroderma, multiple sclerosis, inflammatory bowel diseases, and, especially, regional ileitis and cardiovascular diseases. On the other hand, there are certain diseases, some perhaps of lesser clinical importance, for which an autoimmune component has been suggested, yet these have not been described in any detail in this volume. While the evidence for autoimmunity in such diseases is still only circumstantial, it is toward these diseases that particular attention could be directed in the future. Examples might include the following:

Bone marrow hypoplasia can be attributed to immunopathy, with a particular role for T lymphocytes with suppressor activity, shown in both granulopoietic failure by Bagby *et al.* (1983) and in aplastic anemia by Zoumbos *et al.* (1985).

Alopecia totalis and *alopecia universalis,* which tend to be associated with markers of thyroid and gastric autoimmunity, are implicated on the basis of "guilt by association"; as additional evidence there may be a marked regrowth of hair in these diseases in response to prednisolone.

Vitiligo has likewise been regarded as autoimmune on the basis of clinical associations with thyrogastric and adrenal autoimmunity, but puzzling, and inexplicable immunologically, is the remarkable bilateral symmetry of the lesions; there have been reports from time to time of serum antibody to melanocytes in vitiligo.

Idiopathic deafness and *Ménière's disease* have been implicated as autoimmune disorders, based on clinical associations and on the presence of serum antibodies to type II collagen and the induction in rats of cochlear and vestibular damage by immunization with collagen (Yoo *et al.*, 1982).

Hypopituitarism is a relatively rare disease and the examples studied are usually reported as sporadic cases, numbering 16 so far, and attention can be directed to the presence of serum antibodies, demonstrable by

immunofluorescence, to specialized cells in the anterior lobe of the pituitary gland, including cells secreting prolactin or growth hormone; the subject has been reviewed by Josse (1985).

There are probably various other examples of minor autoimmunopathies of which the present writers are unaware, and there will undoubtedly be new examples as diseases of obscure pathogenesis are scrutinized in clinical immunology laboratories.

FORECASTS FOR AUTOIMMUNITY

As a preamble to this closing section, we can distinguish four eras (or decades) in the history of autoimmunity: 1955–1965, the decade marked by the question, Does autoimmunity exist?; 1965–1975, the decade marked by the upsurge of diagnostic activity associated particularly with the use of diagnostic assays based on immunofluorescence, the establishment of clinical immunology services, and the acceptance of various diseases as being autoimmune in character; 1975–1985, the decade marked by the consolidation of the concept of autoimmune disease and the introduction of new ideas, including that of antireceptor diseases and the wider range of possible autoimmune diseases thereby opened up. We suggest that 1985–1995 will be the decade in which attention is directed to targeted intervention, and specific immunological approaches for prevention or treatment will be used in place of the nonspecific anti-inflammatory and cytotoxic therapy of the past. In this forthcoming decade, in which several promising leads may come to clinical application, we see as prominent the definition of the major autoepitopes relevant to autoimmune diseases, the development of marker-specific monoclonal antibodies, and the utilization of anti-idiotypic antibodies and specific immunotherapy based on the following considerations:

Autoepitopes. The use of immunofluorescent diagnostic serology has led clinical immunologists to refer to macromolecular structures such as nuclei, microsomes, mitochondria, and so on, as "antigens." However, advances in analytical biochemistry and molecular biology should allow precise dissection of these macromolecular "antigens" and the eventual resolution of the epitopes that are relevant to autoimmune reactions. If so, these could be produced *en masse* by gene cloning techniques and applied, using appropriate immunization schedules, to realign the immune system in the direction of tolerance rather than reactivity. This approach is under investigation at present, using a nucleic acid antigen attached to an autologous carrier molecule, γ-globulin.

Monoclonal antibodies. The potential application of monoclonal antibodies to immunotherapy is underscored by success in several experimental model situations in which both prevention and reversal of established disease have been reported. Reference can be made to the use of monoclonal antibody to murine I-A determinants in the prevention and arrest of EAE in mice. The premise is that the encephalitogenic autoantigen must be presented in association with I-A but, with specific blockade of I-A on cells which present neural antigen, induction of disease is abrogated. Experimental autoimmune encephalomyelitis has been best studied, but other experimental examples include myasthenia gravis, thyroiditis, and murine lupus. The use of monoclonal antibodies to the helper/inducer subset of T lymphocytes also appears promising, since the monoclonal reagent could bind specifically to that particular population of T cells subserving induction for help. Further to the above, there will be developments based on the anti-idiotypic approach that is being exploited following the development of cell lines derived from autoreactive T cells. So far these lines are of helper phenotype and, when injected into animals, will reproduce autoimmune disease corresponding to their antigenic specificity, EAE, peripheral neuritis, thyroiditis, and so on. However, of particular interest, cells from these cell lines can be attenuated by various physical or chemical means so that they lose their capacity to induce disease but retain the ability to establish a state of protection against subsequent challenge. Presumably, the inoculation of attenuated cells stimulates the host to produce cells with the ability to suppress cells of the effector class, perhaps via expression of an anti-idiotypic specificity.

Anti-idiotypic antisera. The capability of such antisera to inhibit immunological reactions has been shown in mice by the prevention of responses to phosphoryl choline. Closer to autoimmunity, there is the example of anti-idiotypic (anti-id) antisera raised to the major idiotype of antinuclear antibodies in lupus mice (NZB/NZW). Such anti-id antibodies proved to be capable of markedly reducing levels of autoantibody in such mice, and although there was a reduction in amount of immune complexes in the kidney, the course of disease did not appear to be correspondingly influenced. The production of anti-id antisera, as yet another immunomanipulative approach to human autoimmune disease, is associated with two challenging problems. One is the difficulty in identifying the disease-relevant idiotype among the large spectrum of antibodies against a macromolecular autoantigen in human disease. The other is the likelihood of idiotype escape, meaning that antibody of other idiotypes will increase in response to suppression of the major idiotype. To summarize, the future therapy of human autoimmune disease, based on immunoregulatory intervention, will need to be of a "microsurgical" character, so that particular

pathogenic clones or their products will be specifically and precisely excised from the total immunological repertoire of the patient. For this to become a reality, considerably more work must be done with animal models, and broader application must be made of the outstanding technological advances that have characterized immunology in the 1980s.

Noel R. Rose
Ian R. Mackay

REFERENCES

Allison, A. C., Denman, A. M., and Barnes, R. D. (1971). *Lancet* **2,** 135–142.

Bagby, G. C., Jr., Lawrence, H. J., and Neerhout, R. C. (1983). *N. Engl. J. Med.* **309,** 1073–1078.

Gavalchin, J., Nicklas, J. A., Eastcott, J. W., Madaio, M. P., Stollar, B. D., Schwartz, R. S., and Datta, S. K. (1984). *J. Immunol.* **134,** 885–894.

Gershon, R. K. (1974). *Contemp. Top. Immunobiol.* **3** 1–40.

Jerne, N. K. (1974). *Ann. Immunol. (Paris)* **125c,** 375–389.

Josse, R. G. (1985). *In* "Autoimmunity and Endocrine Disease" (R. Volpe, ed.), pp. 405–426. Dekker, New York.

Klein, J. (1982). *Adv. Cancer Res.* **37,** 234–317.

Nossal, G. J. V. (1983). *Annu. Rev. Immunol.* **1,** 33–62.

Stefansson, K., Dieperink, M., Richman, D. P., Gomex, C. M., and Marton, L. S. (1985). *N. Engl. J. Med.* **312,** 221–225.

Weigle, W. O. (1980). *Adv. Immunol.* **30,** 159–273.

Yoo, T. J., Stuart, J., M., Kang, A. H., Townes, A. S., Tomoda, K., and Dixit, S. (1982). *Science* **217,** 1153–1154.

Zoumbos, N. C., Gascon, P., Djeu, J. Y., Trost, S. R., and Young, N. S. (1985). *N. Engl. J. Med.* **312,** 257–265.

CHAPTER 1

Genetic Predisposition to Autoimmune Diseases

NOEL R. ROSE*
Department of Immunology and Infectious Diseases, and Department of Medicine
The Johns Hopkins Medical Institutions
Baltimore, Maryland

IAN R. MACKAY†
Clinical Research Unit
of The Walter and Eliza Hall Institute of Medical Research
and The Royal Melbourne Hospital
Melbourne, Victoria, Australia

I. INTRODUCTION

Many clinicians have noticed the appearance of multiple cases of autoimmune disease in members of the same family. Sometimes the same

* Aided by PHS research grants AM31632 and AGO4342 from the National Institutes of Health.
† Aided by a grant from the National Health and Medical Research Council of Australia.

disease appears in related individuals; more often, several different autoimmune diseases are found. Generally, the families fall into two major groups. Some families are subject to the occurrence of the organ-specific autoimmune diseases, such as chronic thyroiditis, Graves' disease, autoimmune gastritis, insulin-dependent diabetes mellitus, and related autoimmune endocrinopathies. In other families, one sees a heightened incidence of the systemic autoimmune diseases, such as lupus erythematosus and other multisystem immunopathies; hemolytic anemias, thrombocytopenias and other blood dyscrasias are relatively common in these families. The predisposition to such diseases has been designated an autoimmune diathesis.

In this chapter, we shall concentrate on a limited number of autoimmune diseases in an effort to delineate how the development of autoimmunity is influenced by genetics and to distinguish the types of genetic determinants that may be involved in disease production. Many other chapters contain a section particularly devoted to the genetic aspects of the disease under consideration. We will consider here the genetic aspects of diseases that are exemplars for the two major groups, the organ-specific thyroiditis-associated group and the non-organ-specific multisystem or lupus-associated group, recognizing that there is a spectrum from organ-specific to non-organ-specific autoimmunity (Roitt, 1984).

There are special advantages in considering autoimmune thyroiditis in that this condition is available for genetic studies in three different models. (1) The human disease, chronic thyroiditis, is a classic example of the organ-specific autoimmunities, and numerous family studies are available (Burek *et al.*, 1984); (2) many of the essential features of the disease can be reproduced in the experimental laboratory by immunization with the principal thyroid antigen, thyroglobulin (Witebsky and Rose, 1956; Witebsky *et al.*, 1957); and (3) spontaneous forms of thyroiditis occur in certain genetically predisposed animals, such as obese strain (OS) chickens, Buffalo (BUF) rats, and the Argonne colony of beagle dogs (Bigazzi and Rose, 1975). These advantages do not hold up as well for the multisystem, lupus-associated autoimmune diseases in that these diseases are less precisely defined, do not cluster quite so unequivocally, and cannot be reproduced in animals by immunizing procedures, although natural models do exist.

Fascinating puzzles for immunologists and geneticists have been posed by the association of immune-mediated diseases with genes of the major histocompatibility complex (MHC) that determine acceptance or rejection of tissue grafts. The story began with the discovery in animals that the capacity to respond to simple antigens has a genetic basis, and in mice the genes that determine immune responsiveness were found to be linked

to the *H-2* genes of the MHC (McDevitt and Benacerraf, 1969). Such genes were referred to as *Ir* genes, and these coded for products—antigens—on the surface of cells called Ia, detectable serologically with specific antisera.

Concurrently with research on the role of *H-2* genes in mice, scientists were engaged in the mapping of genes in humans that specified histocompatibility and that were important in organ transplantation. Products of two major human leukocyte antigen (*HLA*) loci could be detected serologically, and these loci, which came to be known as *HLA-A* and *HLA-B*, corresponded to *H-2K* and *H-2D* in mice; a third serologically detected minor locus was called *HLA-C*. Another major locus was later detected in man by using stimulation in mixed lymphocyte cultures, and this became known as *HLA-D;* a gene locus that was closely related (*D*-related, *DR*) was later identified serologically and found to be the analog of the murine Ir-Ia system. The *D* region in man has been found to contain three loci, *DP, DQ,* and *DR*. The MHC in man occupies a locus on the short arm of chromosome 6 and in mice occupies a locus on chromosome 17.

It is now conventional to refer to gene products of loci within the MHC as class I, class II and class III, since their tissue expression and function are quite different. The class I products, HLA A, B are expressed on most tissues in the body and have two effects, both of which are determined by their capacity to attract T lymphocytes: (1) they determine rejection of organ grafts, and (2) they associate with foreign, usually virally encoded, antigens on the cell surface to form a compound antigen that becomes a target for T-lymphocyte attack. It is necessary that the class I antigens of the attacking T lymphocyte match those of the cellular target, an effect known as MHC restriction (Doherty and Zinkernagel, 1975). The class II (D, Dr, Ia) products are expressed predominantly on B cells, macrophages, dendritic cells, endothelial cells, and activated T lymphocytes; these associate with antigens in such a way as to attract helper T lymphocytes, thus promoting amplification of immune responses. The class III products are coded for by genes of the complement system.

II. THYROID DISEASE AS ILLUSTRATING ORGAN-SPECIFIC AUTOIMMUNITY

A. HUMAN THYROIDITIS

Chronic thyroiditis is primarily a disease of the fourth, fifth, and sixth decades of life and shows a strong female:male predominance. Studies conducted over 20 years ago (Hall, *et al.*, 1960) clearly established that

there was a familial predisposition to autoimmune thyroid diseases, Hashimoto's disease, and Graves' disease. Subsequent studies showed that unaffected relatives had a higher frequency of thyroid autoantibodies than did population controls and that among these thyroiditis-prone families there was also an increased predisposition to develop thyroiditis-associated diseases and/or the corresponding autoantibodies (Irvine, 1965; Hall and Stanburg, 1967; Roitt and Doniach, 1967).

We would draw attention here to the fact that autoimmune thyroid disease sometimes occurs in a juvenile form (Burek *et al.*, 1982). In cases starting before puberty, the sex ratio is nearly equal because the differential effects of sex hormones are not yet apparent. It seems likely that the genetic determinants in autoimmune thyroiditis can be more readily sorted out in the juvenile than in the adult form of thyroiditis. First, the premature occurrence of the disease may signify a particularly strong genetic predisposition, making this form of the human disease especially suitable for investigations and, secondly, the chance for confounding environmental factors to interfere with investigations is probably less in children than in individuals with longer life spans.

A study of juvenile thyroid disease raises two primary questions. How strong are heritable as opposed to other factors in the initiation of disease? What types of genetic factors are responsible for the development of thyroiditis? In order to determine the importance of genetic versus environmental factors in the development of disease, Burek *et al.* (1982) analyzed a group of children aged between 8 and 16 years with chronic thyroiditis or thyrotoxicosis. Immunological and clinical data were assembled on the probands as well as on their parents and siblings. Immunological studies included autoantibodies to thyroid antigens, thyroglobulin, and thyroid microsomes, as well as to other autoantibodies. Clinical studies included evaluation of thyroid function with a standard battery of biochemical tests. In each family, only the proband had clinical evidence of thyroid disease. However, many other children were found to have thyroid autoantibodies. Some of the parents were also found to have thyroid autoantibodies. The proband was then eliminated and all of the clinically normal siblings were divided into three groups: those in which neither parent had thyroid autoantibodies; those in which a single parent had thyroid autoantibodies; and a third group in which both parents had thyroid autoantibodies. The incidence of thyroid autoimmunity as attested by thyroid autoantibodies was then determined in each group. As indicated in Table I, there are significant differences in the incidence of thyroid autoantibodies in these groups of clinically normal children. Since all three groups were exposed equally to environmental influences, the

TABLE I

Prevalence of Thyroid Autoantibodies in Unaffected Siblings of Children and Adolescents with Autoimmune Thyroid Disease[a]

Thyroid autoantibodies found in	Number positive/number tested	Percentage positive
Both parents	17/24	71
One parent	14/26	54
Neither parent	4/14	29

[a] From Burek *et al.* (1982).

conclusion must be drawn that genetic factors undoubtedly do play a major role in the initiation of thyroid autoimmunity.

In order to carry out investigations of the kinds of genes involved in the production of disease in a human population, Burek *et al.* (1984) performed family studies, using a series of serological and biochemical markers. These markers are listed in Table II and comprise mainly blood group antigens, blood enzymes, and *HLA* types. In each case, a statistical evaluation was made with the association of thyroid autoimmunity, as measured by the presence of autoantibodies, and the genetic determinant. Four markers were found to be associated with the production of thyroid autoantibodies, namely, *HLA*, Gm, Rh, and ABO, and, although each of these associations individually was rather weak, a composite influence of these could provide a substantial component of the genetic risk (Rose and Burek, 1985). At the practical level, knowledge of these markers could be informative in identifying children in those families most likely to develop thyroid autoimmunity.

The association of *HLA* with autoimmune disease has been described frequently (Spielman *et al.*, 1981) and holds for both the organ-specific and multisystem autoimmune diseases, discussed later in this chapter. The closest association is the relationship of *HLA-B27* with ankylosing spondylitis, although this is not conventionally included among the autoimmune disorders, and various autoimmune diseases have been associated with the *HLA* haplotype-*A1, B8, DR3*. It is perhaps rather surprising that human chronic thyroiditis shows only a weak relationship to particular *HLA* types, although the association between *HLA-B8* and Graves' disease is quite strong (see Chapter 4). On the other hand, the *HLA* type is the most important single determinant in the development of thyroid immunity. That observation is attested to in Table III, in which the occurrence of thyroid autoantibodies is related to the sharing of haplo-

TABLE II
Association of Thyroid Autoimmunity with Genetic Determinants[a]

System	Allele	Relative association with thyroglobulin antibody	Relative association with microsomal antibody
HLA	*B8*	13.37	7.56
Gm	*g*	8.38	14.37
Rh	*R'(CDe)*	7.07	—
ABO	*2*	—	7.89

[a] From Burek *et al.* (1984); "relative association" defined in source article.

types with the proband having chronic thyroiditis (Burek *et al.*, 1982). If a clinically normal sibling shares both haplotypes with his or her affected brother or sister, that sibling has an approximately 90% chance of having autoantibodies. A normal sibling who shares but one haplotype has a 70% chance, and a sibling who shares neither haplotype has only a 56% chance of developing autoantibodies. Although 56% is still high, one must realize that these figures are based on families in which there is a case of juvenile thyroiditis. Even more striking are the figures of subclinical disease measured by biochemical abnormalities or goiter, as observed on follow-up clinical studies. An initially normal sibling who shares both haplotypes with an affected brother or sister has a 32% chance of having subclinical thyroid disease. These observations indicate that *HLA* provides powerful prognostic information on the development of significant thyroid autoimmunity.

Recent studies by Farid and colleagues (Farid *et al.*, 1979, 1981) have suggested that there is some heterogeneity of thyroid diseases based on their association with HLA. The more common goiterous form of thyroiditis, Hashimoto's thyroiditis, is associated with *HLA-DR5*, whereas the atrophic form is associated with *HLA-DR3*.

B. Spontaneous Autoimmune Thyroiditis in Chickens

The second model for the study of the genetic control of thyroid autoimmunity is spontaneous thyroiditis in the OS chicken (see Chapter 4). This closed flock, originally developed by R. F. Cole of Cornell University, now develops autoimmune thyroiditis with an incidence of over 90% (Cole, 1966). These animals develop autoantibodies to thyroglobulin and to thyroid epithelial cells, simulating those found in human patients.

TABLE III
Occurrence of Thyroid Autoantibodies in Relation to Shared Haplotypes[a]

HLA-A and *HLA-B* shared with proband	Siblings		
	Total number	with antibody Number (%)	with subclinical disease Number (%)
Both haplotypes	19	17 (90)	6 (32)
One haplotype	13	9 (70)	1 (8)
Neither haplotype	9	5 (56)	0 (0)

[a] From Burek *et al.* (1984).

The lesions in the thyroid gland closely resemble those in human Hashimoto's disease. The chickens themselves show signs of profound hypothyroidism. The disease itself is clearly genetic in origin (Bacon *et al.*, 1977). Efforts to discover an infectious agent responsible for this disease have been uniformly negative. Other environmental factors, such as food or environmental toxicants, have also been eliminated as causative agents of the disease.

In order to trace the genetic basis of this autoimmune disease, the noninbred OS strain was studied for a variety of genetic markers (Bacon *et al.*, 1981). Of the 12 blood groups segregating in the OS strain, only one, the B blood group, was closely associated with the development of thyroiditis. This association could be demonstrated in terms of the severity of disease, the age of onset of disease, and the levels of antibody produced (Fig. 1). The most susceptible birds were those with the *B* genotype $B^{13}B^{13}$; birds with the *B* genotype B^5B^5 developed only mild disease, whereas the heterozygote B^5B^{13} chickens were intermediate in susceptibility. Breeding experiments were carried out to demonstrate formally that these alleles are linked to the susceptibility to disease.

The basis of this B blood group association with disease was investigated by means of cell transfer studies (Livezey *et al.*, 1981). It was found that the greater susceptibility to disease in the B^{13} birds is transferable with thymocytes but not with bursal cells. This result shows that the major defect is not in the thyroid gland or in the B cells, even though both B cells and the thyroid gland are essential for the development of disease, but in the thymus-derived lymphocytes. The B blood group of the chicken is the major histocompatibility complex, analogous to human *HLA* or murine *H-2;* therefore, it is logical to propose that the B-associated susceptibility to thyroiditis is based on an immune-response gene.

A second genetic abnormality was discovered in the most susceptible

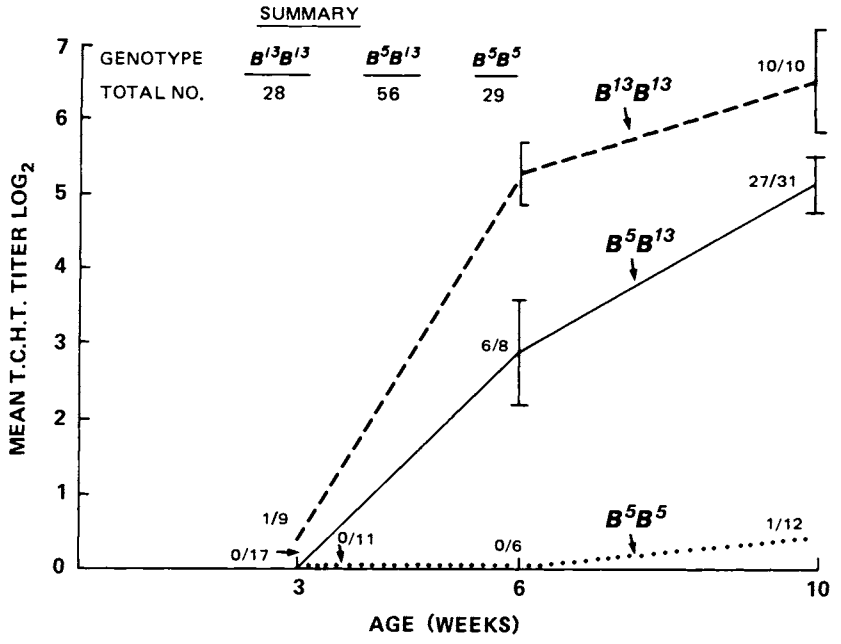


FIG. 1. Relationship of the titer of thyroglobulin autoantibodies from $B^5B^{13} \times B^5B^{13}$ genotype of OS matings. Ratios at each time point represent the number of chickens with histological evidence of thyroiditis over the total number examined. T.C.H.T., Tanned cell hemagglutination test.

OS chickens (Jakobisiak *et al.*, 1976). It is based on an abnormality of thymus function. The first evidence for this abnormality arose from studies of skin graft rejection based on mismatches at minor histocompatibility loci. Normally, neonatally thymectomized chickens accept grafts mismatched at these minor loci. The OS failed to retain such grafts; however, after irradiation, the neonatally thymectomized OS chickens did accept the mismatched skin grafts. Therefore, the conclusion was drawn that an effector cell population residing in peripheral locations was prematurely released by the thymus of the OS compared with that of normal chickens. Several other pieces of information suggesting a thymic abnormality in the OS have arisen. First, the T lymphocytes of OS chickens are more reactive to mitogens than are T cells of normal chickens (Kite *et al.*, 1979). Second, the number of thymic nurse cells is greater in OS chickens; while the nature of these thymic cells is unknown, they undoubtedly are important in timing the maturation of the thymus-derived lymphocytes. Third, OS chickens have a significantly higher incidence of natural antibodies to sheep and rabbit red blood cells as well as autoantibodies to

organ-specific antigens of liver, kidney, stomach, and pancreas (Wick, 1970; Khoury *et al.*, 1982).

The most likely explanation for the thymic abnormality in the OS chicken is that it matures in an abnormal manner. The normal sequence of events in chickens seems to be that a preponderance of the lymphocytes exiting the thymus during embryonic life are of the suppressor T-cell type. Inducer/helper T cells probably leave the thymus somewhat later and are functionally in the minority in the periphery. The situation may be reversed in OS birds; that is, a preponderance of helper/inducer T cells may be the first to leave the thymus. This early advantage of helper/inducer cells may be maintained by contact with low levels of thyroglobulin during later life.

A third abnormality has been found in the thyroid gland of the OS. The first evidence for this abnormality was discovered by Sundick and Wick (1974) as an increased uptake of radioactive iodide. This uptake occurred even if the OS thyroid were placed on the allantoic membrane of a chick egg, indicating that it was not due to a humoral factor (Sundick and Wick, 1976). Other experiments using transplantation have confirmed the fact that this increased iodide uptake is intrinsic to the thyroid. Paradoxically, the gland produces less effective thyroid hormone than do normal thyroid glands, suggesting that there is a basic genetic defect in organic iodinization (Sundick *et al.*, 1979). Although the exact nature of this lesion has not yet been defined, it results in the development of an autonomous thyroid; that is, one not suppressible by thyroxin (Livezey and Sundick, 1980).

Each of these three genetic effects in the OS sorts independently (Rose *et al.*, 1976). The birds that are most prone to the development of severe autoimmune disease inherit all three defects; that is, the *B* locus immune-response gene, the abnormality in thymic maturation, and the intrinsic abnormality in the thyroid gland. Each of these three defects has a different specificity. The immune-response gene probably determines response to one or a small number of antigenic determinants on the thyroglobulin molecule. The defect in thymic maturation seems to predispose to the development of a number of self-directed immunological reactions. The thyroid abnormality pertains to the function of the thyroid gland. One can understand, therefore, why the OS chicken may develop a number of autoantibodies, the major autoantibody is directed to thyroglobulin and the only significant disease is thyroiditis.

C. Experimental Autoimmune Thyroiditis in Mice

The third model for the genetic investigation of thyroiditis is experimentally induced thyroiditis. The most informative animal for these in-

vestigations has been the mouse, where a large number of inbred strains is available. Early investigations showed great differences in the responses of various strains of mice to a standardized stimulus of murine thyroglobulin plus a suitable adjuvant (Vladutiu and Rose, 1975). These responses were not due to susceptibility to the action of adjuvant itself, since several different adjuvants could be used (Esquivel *et al.*, 1977). Genetic analysis of the strains showed that the responses were predictable on the basis of their *H-2* genotype. Certain alleles, such as *H-2k*, *H-2s*, and *H-2q*, are consistently associated with vigorous responses to thyroglobulin, as evidenced by a prompt production of thyroglobulin autoantibody and the early appearance of severe lesions. Other *H-2* genotypes, such as *H-2b* or *H-2d*, are characterized by the later production of autoantibodies and very mild or even insignificant lesions in the thyroid gland.

The basis of this *H-2*-associated response was studied by means of cell transfers (Vladutiu and Rose, 1975). It was found that T cells alone account for the genetic difference. The source of B cells was irrelevant in the genetically determined response of mice to mouse thyroglobulin. Moreover, T cells are a requisite for the development of these diseases. On the other hand, T-cell transfers were much more effective when carried out with thymectomized recipients, suggesting that there is a normal T-suppressor cell in some strains of mice.

Further genetic analyses were carried out using intra-*H-2* recombinants. The major gene controlling responsiveness to thyroglobulin is located in the *I-A* subregion (Beisel *et al.*, 1982). This region corresponds roughly to the *D* region of human *HLA*. Animals with *k* or *s* alleles at *I-A* are basically good responders, whereas those with *b* or *d* alleles are poor responders. However, there was a difference in thyroid lesions in intra-*H-2* recombinants, depending on genes located to the right of the *I-A* region. Further investigation showed that a second gene within *H-2* influences the development of experimental thyroiditis in mice. This gene was localized at the *D* region of *H-2*. Its major effect was not on antibody production but on severity of thyroid infiltration. In animals with *k* or *s* at *I-A*, some alleles, such as *k* and *f*, produced severe disease, whereas *b*, *d*, and *q* were associated with milder lesions. Interestingly, the *f* allele at *I-A* is responsible for a poor response, whereas *q* at *I-A* provides a good response. Therefore, it is the combination between the gene at *I-A* and the gene at *D* that determines the final picture of antibody production and disease.

In order to explain these results, we assume that there are two levels of MHC genetic control (Rose *et al.*, 1981). The first involves the induction of immune response and is encoded primarily in class II genes of the *I-A* region. The second level of genetic control determines the development of

the effector response as assessed by pathological lesions. This genetic control is endowed in the class I genes of the *D* region. It appears that the class I genes of *K* region also affect the severity of thyroiditis lesions. As shown in Fig. 2, several different populations of T cells can be hypothesized as involved in the genetic control of experimental thyroiditis in the mouse. The class II genes of the *I-A* region code for the proliferative response to thyroglobulin of helper/inducer T cells and can be measured directly by such a proliferative response *in vitro* (Christadoss *et al.*, 1978; Okayasu *et al.*, 1981). These helper/inducer cells bear the Lyt-1 antigen. It is possible to demonstrate an effector T cell *in vitro* in the mouse (Creemers *et al.*, 1983). The experiments require labeling of mouse thyroid cultures with ^{111}indium. Lymphocytes from immunized mice stimulated *in vivo* and then *in vitro* with mouse thyroglobulin produce a significant cytotoxic effect on these targets. The effect is *H-2*-restricted. Moreover, antisera to the appropriate K or D determinants abrogate the cytotoxic effect in part, and a combination of the two antisera completely inhibits the cytotoxic reaction. The reaction is also inhibited by polyvalent rabbit antisera to thyroglobulin, suggesting that thyroglobulin itself acts as the surface determinant on the thyroid epithelial cells. The cytotoxic reaction can be prevented by treatment of the lymphocyte suspen-

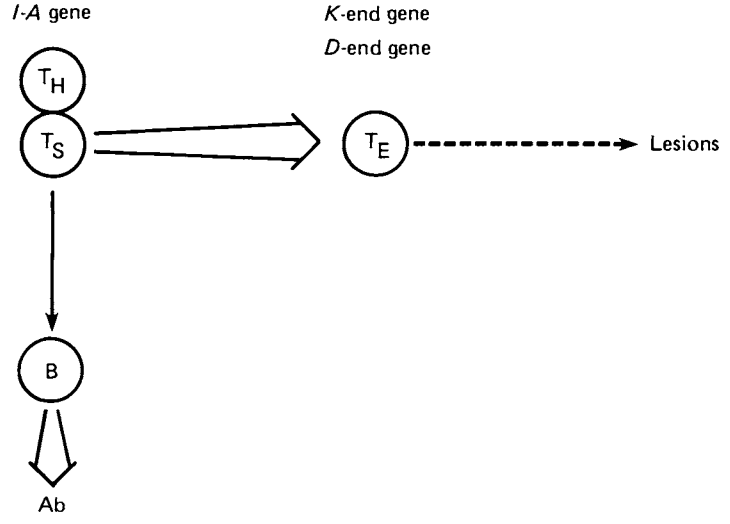


FIG. 2. Proposed schema for genetic control of the immune responses of mice to murine thyroglobulin. T_H, Helper/inducer T cell; T_S, suppressor T cell; T_E, effector/cytotoxic T cell; B, B cell; and Ab, antibody.

sion with Lyt-2 antisera, suggesting that a Lyt-2^+ cell is responsible for cytotoxicity. However, treatment of the suspension at an earlier time with Lyt-1 antisera prevents the development of the Lyt-2 cytotoxic cells, pointing to an Lyt-1^+ helper/inducer cell.

Antigen-specific suppressor T cells can be induced in mice by injection of soluble antigen, that is, thyroglobulin, without any adjuvant (Kong *et al.*, 1982). The suppressors can be transferred with thymus or spleen suspensions.

III. MULTISYSTEM AUTOIMMUNE DISEASES

A. Introduction

The multisystem non-organ-specific autoimmune diseases, including the connective tissue diseases, have long been thought to show genetic linkages, but the nature and degree of such linkages could not be specified by reason of the difficulties in precise nosologic characterization of the entities under consideration. As a group, the multisystem autoimmune diseases show relatively weak patterns of inheritance and the tendency for disease overlap within patients, or within family members, is less prominent than that seen in the thyroiditis-related autoimmune diseases. The multisystem autoimmune diseases are a heterogeneous group and some, such as Sjögren's syndrome, might even be more appropriately classified as organ specific: Their characterization as multisystem may depend more on the propensity for immune complexes to be generated by the major target autoantigens of the disease than on multiple target tissues. Particular examples used in this discussion to illustrate aspects of inheritance include systemic lupus erythematosus (SLE), Sjögren's syndrome (SS), chronic active hepatitis (CAH), and rheumatoid arthritis. The major genetic determinants to be considered will include female sex; the *HLA* system with comment on phenotypes, haplotypes, and extended haplotypes; immunoglobulin allotypes, Gm and Km; *HLA*-Gm interactions; and familial predisposition of undefined nature.

For a few of the diseases in the multisystem group there are models in animals, either spontaneously occurring or experimentally induced; spontaneously occurring models are exemplified by the lupuslike diseases in various mouse strains, which are genetically quite complex, and experimentally induced models derived by immunization with an autoantigen in Freund's complete adjuvant, for example, rheumatoid arthritis and myasthenia gravis, resemble the experimental organ-specific diseases in having strong dependence on genes of the *H-2* complex.

B. Female Sex

Almost all of the multisystem autoimmune diseases occur predominantly in females, the cited sex ratios (F:M) being 6:1 for SLE, 9:1 for Sjögren's syndrome, 8:1 for autoimmune CAH, and 2:1 for rheumatoid arthritis. There is substantial evidence in humans and other species for a generally augmented humoral immune responsiveness in females (Mackay *et al.*, 1977), exemplified by higher immunoglobulin levels, especially IgM, greater immunological responses to antigenic challenge, and a higher background frequency among females in normal populations of various autoantibodies. This effect may be explained by regulatory genes that influence the expression of genes which determine levels of immunoglobulin production. For example, a gene locus on the X chromosome was postulated to explain differences in the antibody response to dextran in mice (Blomberg *et al.*, 1972), and a Y-linked regulatory locus has been postulated (Brewerton, 1984).

Detailed studies on models of SLE in mice, including castration and hormone supplementation, point very strongly to the female predisposition to autoimmunity being due to modulating effects of hormones on immune responses. In human SLE, and in lupus in NZ mouse strains, estrogens have potentiating effects whereas testosterone is protective (Talal *et al.*, 1980). Of interest, if males have abnormal conversion of androgens to estrogenic metabolites, the protective effect of androgens is attenuated (Lahita *et al.*, 1982). On the other hand, there are other models of autoimmunity in mice, for example, BXSB, in which males have accelerated disease which, in this case, is due to a Y-chromosome-linked factor that is dependent on autosomal genes but not sex hormones; this may have some analogy with a rare inherited type of SLE in males, expressed in sons and fathers (Lahita *et al.*, 1983).

C. *HLA*

Although it is readily understandable that MHC-encoded cell-surface products could influence the propensity to develop immune-mediated disease, the manner whereby this occurs is still unclear. Six points can be made. (1) Multiple and different effects must be operative, since *HLA*-associated diseases may have some clinical similarities yet differ quite markedly in their MHC markers and immunological expression, for example, ankylosing spondylitis (*HLA-B27*) and rheumatoid arthritis (*HLA-DR4*). (2) The disease association is usually not with a single *HLA* specificity but rather with a chromosomal set of MHC genes that may include *HLA* specificities, complement alleles, and other genes which together

constitute a supratype, or extended haplotype—where among these the actual disease susceptibility gene is located is unknown. (3) *HLA* effects on predisposition to disease are generally not strong since the relative risk for SLE conferred by *HLA-B8* or *-DR3* is only of the order of 3 to 4, and for Sjögren's syndrome or autoimmune CAH of the order of 10 to 15; however, the strength of association may be attenuated by unknown degrees of heterogeneity among disease entities and so, as diseases become more precisely classified, genetic associations with specific subsets will become stronger (Mackay, 1983; Stastny *et al.*, 1983). (4) There may be more than one *HLA*-linked determinant involved (Scholz and Albert, 1983); the additional determinants may confer increased susceptibility, for example, DR3 and DR4 in insulin-dependent diabetes mellitus; or may confer protection, as suggested for certain *HLA* specificities in SLE (Whittingham *et al.*, 1983). (5) Different *HLA* specificities may confer susceptibility to the same disease among different ethnic groups. (6) In analysis of pedigrees, the *HLA* specificity under consideration may not segregate with the occurrence of disease in family members, implicating non-MHC as well as MHC genes in susceptibility (Reveille *et al.*, 1983). An analog of this is provided by spontaneous insulin-dependent diabetes mellitus in BB rats in which there are operative at least two genes, one determining T lymphopenia and not linked to the rat MHC and the other influencing susceptibility to diabetes and closely linked to the rat MHC (Jackson *et al.*, 1984).

HLA associations with disease will be clarified by defining extended haplotypes. There was early recognition that certain HLA associations with disease were secondary to linkage disequilibrium—a greater than expected association of one HLA specificity with another. For example, autoimmune CAH was found by studies of families to be strongly associated with *HLA-A1, -B8, -DR3*, present as a haplotype on one chromosome (Mackay and Tait, 1980). More complex extended haplotypes have been identified for several diseases, involving complement alleles and other MHC-associated loci, notably in insulin-independent diabetes mellitus (Raum *et al.*, 1984) and rheumatoid arthritis (Dawkins *et al.*, 1983). It has been suggested (Alper *et al.*, 1982) that these gene clusters are frozen in the genome because their components fail to cross over during meiosis: The disease susceptibility gene is assumed to be incorporated with the gene cluster of the extended haplotype. However, as indicated by Shoenfeld and Schwartz (1984), this attractive concept does not take account of the immune functions specified by the MHC genes.

There is no single overall explanation to explain HLA associations with disease: Multiple mechanisms must be operative. For some examples

such as the association of HLA-B27 and ankylosing spondylitis, there is appeal in the concept of molecular mimicry in which the HLA specificity is cross-reactive with microbial antigens. For *B8/DR3*-associated autoimmune diseases, there may be an associated general immune hyperactivity (or weakness of suppression), according to data of Ambinder *et al.* (1982), as well as antigen-specific effects. The latter could operate via an *Ir* gene effect, which specifies responsiveness to an autoantigen (or an extrinsic antigen which is cross-reactive with an autoantigen), or by an immune-suppressor gene. For example, Sazazuki and colleagues (1983) addressed immunological aspects of MHC-linked susceptibility by studies on immune responses *in vitro* to streptococcal cell wall (SCW) antigen, and high and low responders could be associated with HLA haplotypes as defined by family studies. A good case was made for the existence of HLA-associated immune-suppressor (*Is*) genes to which were attributed "crucial roles in the pathogenesis of autoimmune disease and allergy." The concept developed was that the genetic defect in autoimmunity is a homozygous state for the lack of an *Is* gene for the relevant autoantigen.

D. Complement System

Genetic polymorphism is frequent among components of the complement system, the best studied being those (C2, C4, and factor B) coded by genes in the MHC between *HLA-B* and *DR* (Hobart *et al.*, 1984). There are two common alleles for factor B, at least four alleles for C2, and there are complex polymorphisms for C4. Alleles of C2, C4, and factor B are found to be included in the extended haplotypes associated with insulin-dependent diabetes mellitus, rheumatoid arthritis, and multiple sclerosis, but probably have no influence *per se* on susceptibility to disease.

The disease that is associated particularly with inherited complement deficiencies is SLE; this disease or an analogous syndrome has been found in association with deficiencies of almost every complement component (Walport *et al.*, 1984), particularly C2 and also C4, C5, C6, C7, and C8. It was shown in a family study by Fielder *et al.* (1983) that 83% of 29 patients with SLE had one or more null or silent alleles for C4A, C4B, or C2, thus resulting in a failure of synthesis of the corresponding complement gene product; it was inferred that partial deficiencies of a complement component in heterozygous states conferred susceptibility to SLE. It is uncertain how complement deficiency predisposes to SLE, but facilitation or prolongation of viral infections that provoke SLE is one possibility.

E. Inheritance of *V* Genes for Autoantigens

The mammalian immune system has a huge capacity for generating diversity: for antibodies, 10^6 to 10^8 different combining sites; and for T cell receptors, an as-yet-unknown number. This diversity is created by multiple gene segments in the germ line, somatic recombination of these segments, somatic mutation, and association of different V_H and V_L polypeptides during immunoglobulin assembly. Despite this potential for random diversity among antibody binding sites, there is reason to believe that there may be a preferred or dominant expression of particular *V* region genes in the B-cell repertoire (Manser *et al.*, 1984). Accordingly, inheritance of autoantibody reactivity by B cells could be explained by preferred expression of a germ line *V* gene. However, on the basis that a germ line gene coding for self should be selected against, the occurrence of self-reactivity in B cells is more comfortably explained by postcombinatorial mutations or by failure of immunoregulatory processes. A theory of *V* gene inheritance has been developed by Adams and recently applied to the inheritance of insulin-dependent diabetes mellitus (Adams *et al.*, 1984).

A investigative approach to the activity of *V* genes for autoantigens could be made by enzymatic digestion of the DNA of blood leukocytes and by searching for disease-associated DNA fragments by hybridization with *V* region probes. A simpler, but less direct, approach described in the next section is the examination for associations of disease with allotypic variants on the constant chains of immunoglobulin molecules, on the basis of linkage disequilibrium between *V* and *C* region genes.

F. Association of Gm with Antibody Responses and Disease

Possible relationships between Gm phenotypes and immune responsiveness or disease have been keenly investigated. The claims (reviewed below) for positive, albeit weak, associations can be balanced by the unenthusiastic comments of Steinberg (1984).

Wells *et al.* (1971) measured levels of natural antibody to flagellin, and of antibody following a primary challenge with 5 μg of monomeric flagellin. The finding was of a high number of Gm(1;21)-positive subjects among the groups with high natural antibody and with high titers of antibody following primary challenge, in contrast to low titers in Gm(3;5)-positive subjects. These and subsequent observations by Mackay *et al.* (1975) suggested that the high-responder Gm phenotype might confer a selective

survival advantage. Nevo (1974) reported that subjects homozygous for Gm(3;5) were poor antibody responders to *Salmonella typhi* antigens compared with heterozygotes and, among subjects with typhoid fever, the relative frequency of homozygotes for Gm(3;5) was increased; thus, susceptibility to typhoid could be determined in part by the Gm phenotype of the subject. The basis for associations between antibody response and Gm allotypes would presumably be linkage disequilibrium between genes for the constant region of Ig molecules and variable region (*V*) genes that specify the structure of the antibody-combining site of the immunoglobulin heavy chain.

In regard to disease associations with Gm allotypes, one of the early observations was a weak association (RR,2.3) of Gm(1;21) with CAH, discussed below in the context of HLA interactions. In Graves' disease, an association between *Gm(3;23;5)* and *Gm(3;5)* haplotypes and the presence of thyroid-stimulating autoantibodies was reported by Farid (1977), but the validity of these findings was questioned by Propert (1984) on the basis of the statistical analysis used in the study. A very strong association with Gm(1,2,21) was ascertained for antiglomerular basement membrane antibody (anti-GBM) and for the associated glomerulonephritis by Rees *et al.* (1984); this increase was attributed entirely to presumed heterozygotes with phenotype Gm(1,2,21;3,5,11), and heterozygotes at *Gm* loci had higher titers of anti-GMB irrespective of the presence of Gm(1,2,21). In another study on autoantibodies, Mentnech *et al.* (1980) found a significantly higher titer of anti-IgG antibodies, including rheumatoid factors, in Caucasian coal miners with pneumoconiosis who were homozygous for Gm, Gm(3;5,11) than in miners who were heterozygous for Gm, Gm(1,2,3,17;5,11,21). In celiac disease, only patients with the phenotype Gm(23) had substantial levels of antibody to a pure preparation of gliadin, a derivative of gluten that excites an immunopathological response in the intestine. Weiss and colleagues (1983) used a sensitive radioimmunoassay to detect antibody in 30 Caucasian patients with a celiac disease who had been maintained on a gluten-free diet for a period of 1.5 to 20 years, and found that none of 8 patients negative for Gm(23) had raised levels of antibody to gliadin.

Other disease associations with Gm allotypes, derived from studies on Caucasians of northern European origin, include SLE and multiple sclerosis (Whittingham *et al.*, 1984a) and insulin-dependent diabetes mellitus (Schernthaner and Mayr, 1984). Among Japanese with myasthenia gravis, Nakao *et al.* (1980) found that the frequency of the Gm(1,2;21) phenotype was significantly increased among patients with a high titer of autoantibodies to the acetylcholine receptor, and Nakao *et al.* (1981) reported a significant difference among Japanese insulin-dependent dia-

betics in the frequency of certain Gm specificities in patients who had developed antibodies to insulin as compared with those who had not.

There are similar reports of disease associations for the Km allotypes, formerly referred to as Inv, a structural variant on the κ light chain encoded by a gene on chromosome 2. Vogel *et al.* (1971) found an association between ABO blood groups and Km(1) specificity in Thai patients with leprosy and Brachtel *et al.* (1979) found that patients with the atopic diseases atopic dermatitis, hay fever, allergic rhinitis, bronchial asthma, and acute urticaria, which result from an IgE response to allergens, had an increased frequency of Km(1). These studies did not clearly define the association of Km(1) with an antibody response to a specific antigen, whereas Pandey *et al.* (1979), in a study on 20 white children and 33 black children immunized with a number of vaccines, found differences related to ethnic background and Km allotype in responses to *Haemophilus influenzae* type polyribose phosphate vaccine and meningococcus C polysaccharide vaccine. Whittingham *et al.* (1984b) investigated patients with autoantibody to the extractable nuclear antigen La (SS-B), which is associated with primary Sjögren's syndrome; there was a highly significant association between the Km(1) phenotype and anti-La; this is of particular interest because no association was found between Km(1) and autoantibody to another structurally related extractable nuclear antigen, (U1)RNP, which is characteristic of mixed connective disease. Other disease associations with Km(1) have not as yet been identified.

G. Effect of *HLA*–Gm Interaction on Antibody Response and Disease

In a study designed to analyze the effect of Gm types on antibody responses, Whittingham *et al.* (1980) immunized 200 unrelated healthy adult Caucasian volunteers with 1 μg of monomeric flagellin; measured IgM and IgG antibodies to flagellin serologically before and 2 weeks after immunization, that is, at the peak of the primary response; and performed *HLA* and Gm typing to examine associations of these genetic markers with the antibody response. Of the 200 subjects, 125 responded to this low dose and 100 developed IgG antibodies to flagellin. The mean log titer of serum IgG antibody was higher in females than in males and was higher in subjects heterozygous for Gm allotypes than in homozygotes. Analysis using appropriate statistical models indicated that the magnitude of the antibody response in subjects with particular Gm phenotypes depended on the phenotype determined by the *HLA-B* locus (Whittingham *et al.*, 1984a). The data pointed to interactive effects between *Gm* and *HLA* loci, so that the joint effects of particular HLA and Gm phenotypes were

substantially greater (or less) than the sum of the mean effects of the two loci considered separately.

There are few other studies in which an interactive effect of *HLA* and Gm on the magnitude of an antibody response has been examined. No such effect was apparent in the antibody response to gliadin in gluten-sensitive enteropathy (Weiss *et al.*, 1983), although the statistical methods differed from those used in flagellin study; thus, more than 70% of the patients with gluten-sensitive enteropathy were *HLA-B8* and/or *-DR3*, yet IgG antibody to gliadin was found in affected patients with Gm(23), irrespective of the presence or absence of *HLA-B8* and/or -DR3. In mice two genetic loci, *H-2* on chromosome 17 and the immunoglobulin heavy chain allotype locus *Igh* on chromosome 12 were found to have controlling influences on the antibody response to gliadin (Kagnoff, 1982), although their interactive effect was not determined.

The influence of interactive effects of Gm and *HLA* on susceptibility to development of autoimmune disease was established for autoimmune CAH in a study in Melbourne on 50 patients and 180 healthy controls (Whittingham *et al.*, 1981). The relative risk for autoimmune CAH was 11.6 for patients who were *HLA-B8*, 11.7 for patients who were *HLA-DR3*, and 2.3 for patients who were Gm(1,2); 45% of the patients who were *HLA-B8* had the phenotype Gm(1,2) but none of the patients who were negative for *HLA-B8* were Gm(1,2), in sharp contrast to the distribution of Gm(1,2) relative to the presence or absence of *HLA-B8* in the 180 healthy controls, 24 and 18%, respectively. Applying statistical techniques, Whittingham *et al.* (1984a) found that the relative risk for disease was lowest in patients negative for *HLA-B8* yet positive for Gm(1,2) and, relative to this low risk group, the risk was increased 39 times for subjects positive for both *HLA-B8* and Gm(1,2), 15 times for subjects positive for *HLA-B8* but not Gm(1,2), and 2 times for subjects negative for both phenotypes. Thus, in the presence of *HLA-B8*, immunoglobulin genes presumed to be in linkage disequilibrium with the Gm(1,2) phenotype substantially augmented the risk of development of autoimmune CAH; and, in contrast, in the absence of *HLA-B8*, these same Gm-linked genes appeared to be inactive.

Interactive effect of *HLA* and Gm have been sought in other autoimmune diseases. For rheumatoid arthritis, Propert *et al.* (1982) reported on a study on 45 Caucasian patients with classical disease in whom the determinants involved were *HLA-DR4* and Gm(1,2,3;5). The risk for disease was increased 21.7 times in patients with both *HLA-DR4* and Gm(1,2,3;5) and 5.8 times in patients with *HLA-DR4* but not Gm(1,2,3;5), in comparison with those who lacked both the -DR4 and Gm(1,2,3;5) phenotypes. There are several other diseases in which predisposition is

conferred by both HLA and immunoglobulin allotypes, examples including myasthenia gravis, thyrotoxicosis, Sjögren's syndrome, and multiple sclerosis (Whittingham *et al.*, 1984a); in these studies interactive effects of *HLA* and Gm were not implicated, although among insulin-dependent diabetes mellitus a Gm phenotype (1−,2−,3+,5+) showed significant heterogeneity according to subdivision of disease by *HLA-DR* type (Schernthaner and Mayr, 1984). Note can be taken of an interesting study among Japanese by Uno *et al.* (1981), who were attempting to identify genes predisposing to susceptibility to Graves' disease. In a study of 30 families in whom there were two or more affected relatives with disease, all affected siblings bar one shared the disease-associated HLA and Gm phenotypes; siblings with disease-associated phenotypes who did not have disease were not tested for thyroid antibodies, and so latent disease in these was not excluded.

H. Familial Predisposition of Undefined Basis: Background Genes

Genetic analysis of the multisystem diseases is complicated, for some at least, by an unknown degree of diagnostic heterogeneity, pertaining almost certainly to RA and probably also to SLE. This may account for difficulties with family studies.

Information on familial predisposition to SLE is now substantial and has been well reviewed, for example, by Reveille *et al.* (1983). Data from Estes and Christian (1971), Dubois (1974), and Reveille *et al.* (1983) indicate that a first or second degree relative of a propositus with SLE would have a 7–12% risk of developing SLE, and an even higher risk for serologic abnormalities.

However, the most impressive evidence for genetic determinants of SLE is the concordance for both disease and serologic abnormalities in monozygotic twins. Block *et al.* (1975) reviewed twin pairs with concordance for disease in 70%, considerably in excess of the frequency for dizygotic twins and sib pairs. Since known genetic determinants (see below) such as female gender, HLA phenotypes (B8, DR3), and deficiencies of complement components would explain only a small component of the genetic risk for SLE, there are obviously multiple background genes which, acting with environmental influences, determine susceptibility to SLE; similar consideration would presumably apply to other multisystem, autoimmune diseases.

The nature of the background non-MHC genetic determinants of autoimmunity is still a mystery, and will require detailed analysis of autoimmune-prone families. In one such study by Grennan *et al.* (1984) on inheritance of rheumatoid arthritis, it was found that autoimmune thyroid

disease was significantly more frequent in those families in which rheumatoid arthritis segregated with a *DR4*-negative haplotype, suggesting that genes independent of *DR4* may predispose to both diseases. As rheumatoid arthritis–thyroiditis haplotype sharing by appropriate sibling pairs was not increased, this suggests that part of this genetic background in common is determined by non-*HLA*-linked genes.

I. Murine Lupus

The recognition of autoimmune disease in inbred New Zealand (NZ) mice in the late 1950s held out much promise for an understanding of the genetic basis of autoimmune disease, but this has yet to be realized, even 25 years later. Knight and Adams (1982) reviewed their studies on NZ mice and attributed three genetic influences to the occurrence of lupus nephritis in B/W hybrids, with one being closely linked to the MHC and two to the occurrence of autoimmune hemolytic anemia in NZB, with neither being linked to the MHC. Other authors (Theofilopoulos and Dixon, 1981; A. D. Steinberg, 1984) emphasize the genetic complexity of these murine models, with at least six autosomal genes being implicated by A. D. Steinberg (1984) in the disease of NZB mice, and multiple genes in the B/W hybrid. It is of interest that genetic interference with the final pathway is possible by introduction of the *xid* gene, which inhibits the development of a subset of splenic B cells on which depend responses to certain antigens, including autoantigens. Thus there appears to be an interplay among multiple genetic influences and environmental determinants in predisposition to SLE in various models in mice, exemplified by the NZB strain and the B/W hybrids, and the MRL-lpr/lpr strains. There is no reason to believe that the situation is any less complex in human SLE.

IV. CONCLUDING REMARKS: FUTURE PROSPECTS

It is now obvious that the development of human autoimmune disease is determined to a great extent by inheritance. The most commonly identified genetic factors are those associated with female gender and the major histocompatibility complex. Of the several ways in which *HLA* might play a role in autoimmune disease, the most likely is that *HLA* incorporates an immune-response gene similar to the murine *Ir* gene. Thus individuals who inherit certain *HLA* haplotypes are probably better responders to particular antigenic determinants of self-antigens. However, as the experiments with murine thyroiditis have shown, HMC genes can have several different effects. The usual explanation is that the actual gene responsible

for increased susceptibility is merely associated with the HLA determinant by linkage disequilibrium. Therefore, the HLA determinant serves only as the marker for the nearby disease susceptibility gene. Thus far, no such disease susceptibility gene has been identified, either in experimental animals or in human populations. Another explanation is that there is an abnormality in the *HLA* gene itself. To find such an abnormality, it is necessary to compare DNA restriction maps of DNA from diseased and normal individuals, and it may be that restriction fragment analysis will demonstrate an abnormal sequence in those individuals who actually develop disease; this approach has proven to be very promising in insulin-dependent diabetes mellitus (Dausset and Cohen, 1984).

The genetic susceptibility to the initiation of immune response may differ from control of the effector arm of the immune response, which results in the actual production of disease. It is well known that the occurrence of autoantibodies is much more common than the development of lesions, suggesting that additional levels of genetic control are required for the development of disease above those required for the production of autoantibodies, and experimental studies support this idea.

A genetic influence on immunoregulation is probably conferred by the thymus. In a few diseases there may be a decrease in nonspecific suppressor cells such as those induced by Concanavalin A, but in most autoimmune diseases only certain populations of suppressor T cells are diminished, that is, the populations that are specific for the antigens involved in the particular autoimmune response. General decreases in numbers or activity of nonspecific suppressor cells is not a prominent feature of autoimmunity, although such decreases may be evident only during exacerbations of disease, as claimed in multiple sclerosis (Compston, 1983), or particularly at the onset of disease, as reported for insulin-dependent diabetes mellitus (Buschard *et al.*, 1983). In most cases, however, it will be necessary to develop measures for quantitative analysis of activity of antigen-specific suppressor cells in order to gauge the immune status of the patient.

A further type of genetic abnormality relates to the target organ. In the case of thyroiditis in particular, it appears that some underlying defect in the function, and perhaps the anatomy, of the thyroid gland is necessary for the development of thyroid disease. The same could apply to other tissues affected by organ-specific autoimmune diseases, for example, stomach. These target organ defects may themselves be responsible for disease and may explain why patients with nonautoimmune disease, for example, of the thyroid, are encountered in families in which autoimmune thyroid disease is common.

An overview is present in Fig. 3. It is based on the experimental evi-

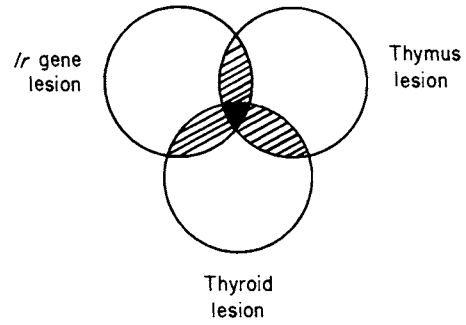


FIG. 3. Theoretical construct of the genetic predisposition to autoimmune thyroiditis. Reprinted with permission from Rose and Burek (1985).

dence that the development of autoimmune disease represents the conjunction of several independent genetic events, for which the evidence is good in both thyroiditis and murine SLE. In the case of thyroiditis, these events include one or more immunoregulatory genes, an abnormality in maturation of the thymus, and a disturbance in the function of the thyroid gland. Individuals who inherit all three of these defects are the ones most likely to develop autoimmune disease. If only two defects are inherited, the frequency of disease might be considerably less. However, the intervention of developmental and environmental factors may shift the balance to development of disease. Such factors as sex hormones, thyroid infection, or even stress may be significant contributors. In this respect, it is important to point out that even genetically resistant mice and rats can be induced to produce a full picture of thyroiditis, if a strong enough antigenic stimulus is given. In the case of experimental thyroiditis, that stimulus is provided by thyroglobulin combined with Freund's adjuvant and, in the case of experimental autoimmune encephalomyelitis, the use of pertussigen will induce disease in resistant mouse strains (Munoz and Mackay, 1984).

The concept just presented does not presuppose that there are disease susceptibility genes as such. It is the conjunction of several abnormalities in addition to the *HLA*-determined effect that is responsible for the development of disease in some, but not all, individuals with a particular HLA phenotype. Also, the autoimmune disease in family members does not always clearly segregate with *HLA* haplotypes. This picture seems to fit rather closely the clinical concept of an autoimmune diathesis since, in families predisposed to the development of autoimmunity, not every member develops autoimmune disease, and different autoimmune diseases can develop in various members of the same family and at different

ages. The differences are attributed to *HLA* and non-*HLA* genetic effects, the occurrence of various target organ differences in various members of the family, and age and sex which come into play because of their role in altering thymus function, and to environmental influences.

In many ways the concept of an autoimmune diathesis is similar to the familiar principle of atopy as a predisposition to the development of immediate hypersensitivities or allergies. It is well known that the development of allergy is more common in some families than others. However, the particular form that allergy may take varies from person to person within the same family. Moreover, there is a strong age-related factor in the development of allergy.

The role of the major histocompatibility complex in the development of autoimmunity may have important implications as far as the prognosis and treatment of autoimmune disease are concerned, particularly if restriction fragment analysis of DNA can sharpen estimates of risk. First, within predisposed families, it may be possible to select the individuals at greatest risk and, under such circumstances, to intervene at an earlier time. In the case of thyroiditis, prognosis is feasible even if there is not a strong linkage disequilibrium with a particular *HLA* haplotype; rather, the sharing of any haplotype with the affected proband would be significant. The presence of other genetic markers, such as Gm, could add to the prognostic value of *HLA*. All of these measures may make it possible to identify the individuals at risk for autoimmune disease much more effectively than was done in the past.

As far as treatment is concerned, one logical strategy is to attack the *HLA*-determined factors. Recent experiments in mice have shown the effectiveness of anti-Ia sera in abrogating genetically controlled immunological responses. Two experimentally produced autoimmune diseases, experimental autoimmune encephalomyelitis (EAE) and experimental myasthenia gravis, can be prevented in mice by concurrent injection of the appropriate anti-Ia sera. In the case of recurrent EAE, one can even prevent exacerbation of the disease by injection of the anti-Ia reagent. In animals predisposed to genetically determined autoimmune disease, such as murine lupus in NZB/NZW mice, the occurrence of disease can be delayed by the injection of the appropriate anti-Ia antisera. It should be noted that the anti-Ia antiserum must be directed to the particular allotypic determinant responsible for good response. In heterozygous mice, one can demonstrate that only anti-Ia antisera directed to the allele associated with a good response are effective. Since most humans are heterozygous, this finding suggests that it may be possible to reduce autoimmune susceptibility without general reduction of the immune response. However, a note of caution must be inserted. A recent study has shown

that injections of anti-Ia antisera in mice produced a general depletion of B cells in spleen and lymph node; this finding could limit the general application of anti-Ia therapy for the treatment of autoimmune disease.

REFERENCES

Adams, D. D., Knight, J. G., White, P., Adams, Y. J., McCall, J., Horrocks, R., and van Loghem, E. (1984). *Lancet* **1,** 420–424.

Alper, C. A., Awdeh, Z. L., Raum, D. D., and Yunis, E. J. (1982). *Clin. Immunol. Immunopathol.* **24,** 276–285.

Ambinder, J. M., Chioraggi, N., Sibofsky, A., Fotino, M., and Kunkel, H. G. (1982). *Clin. Immunol. Immunopathol.* **23,** 269–274.

Bacon, L. D., Sundick, R. S., and Rose, N. R. (1977). *In* "Avian Immunology" (A. A. Benedict, ed.), pp. 309–315. Plenum, New York.

Bacon, L. D., Polley, C. R., Cole, R. K., and Rose, N. R. (1981). *Immunogenetics* **12,** 339.

Beisel, K. W., David, C. S., Giraldo, A. A., Kong, Y. M., and Rose, N. R. (1982). *Immunogenetics* **15,** 427.

Bigazzi, P. E., and Rose, N. R. (1975). *Prog. Allergy* **19,** 245.

Block, S. R., Winfield, J. B., Lockshin, M. D., D'Angelo, W. A., and Christian, C. L. (1975). *Am. J. Med.* **59,** 533–552.

Blomberg, G., Geckeler, W. R., and Weigert, M. (1972). *Science* **177,** 178–180.

Brachtel, R., Walter, H., Beck, W., and Hilling, M. (1979). *Hum. Genet.* **49,** 337–348.

Brewerton, D. A. (1984). *Lancet* **2,** 799–802.

Burek, C. L., Hoffman, W. H., and Rose, N. R. (1982). *Clin. Immunol. Immunopathol.* **25,** 395.

Burek, C. L., Rose, N. R., Najar, G. M., Gimelfarb, A., Zmijewski, C. M., Polesky, H. F., and Hoffman, W. H. (1984). In "Immunogenetics" (G. S. Panayi and C. S. David, eds.), pp. 207–233. Butterworth, London.

Buschard, K., Röpke, C., Madsbad, S., Mehlsen, J., Sprenson, T. B., and Rygaard, J. (1983). *J. Clin. Lab. Immunol.* **10,** 127–131.

Christadoss, P., Kong, Y. M., El Rehewy, M., Rose, N. R., and David, C. S. (1978). *In* "Control of Autoimmune Response" (N. R. Rose and P. E. Bigazzi, eds.), pp. 445–453. Elsevier, Amsterdam.

Cole, R. K. (1966). *Genetics* **54,** 1021.

Compston, A. (1983). *J. Neurol. Neurosurg. Psychiatry* **46,** 105–114.

Creemers, P., Rose, N. R., and Kong, Y. M. (1983). *J. Exp. Med.* **154,** 559.

Dausset, J., and Cohen, D. (1984). *Clin. Immunol. Allergy* **4,** 581–592.

Dawkins, R. L., Christiansen, F. T., Kay, P. H., Garlepp, M., McCluskey, J., Hollingsworth, P. N., and Zilko, P. J. (1983). *Immunol. Rev.* **70,** 5–22.

Doherty, P. C., and Zinkernagel, R. M. (1975). *J. Exp. Med.* **141,** 502–507.

Dubois, E. L. (1974). *In* "Lupus Erythematosus" (E. L. Dubois, ed.), Chapter 7. Univ. of California Press, Los Angeles.

Esquivel, P. S., Rose, N. R., and Kong, Y. M. (1977). *J. Exp. Med.* **145,** 1250.

Estes, D., and Christian, C. L. (1971). *Medicine* **50,** 85–95.

Farid, N. R., Newton, R. M., Noel, E. P., and Marshall, W. H. (1977). *J. Immunogenet.* **4,** 429–432.

Farid, N. R., Moens, H., Bear, J., and Bernard, J. M. (1979). *Tissue Antigens* **13,** 342.

Farid, N., Sampson, L., Moens, H., and Bernard, J. (1981). *Tissue Antigens* **17,** 265.

Fielder, A. H. L., Walport, M. J., Batchelor, J. R., Rynes, R. I., Block, C. M., Dodi, I. A., and Hughes, G. R. V. (1983). *Br. Med. J.* **286,** 425–428.

Grennan, D. M., Dyer, P., Dods, W., Read, A., Haeney, M., Clague, R., and Harris, R. (1984). *Q. J. Med.* **53,** 479–485.

Hall, R., and Stanburg, J. B. (1967). *Clin. Exp. Immunol.* **2,** 719–725.

Hall, R., Owen, S. G., and Smart, G. A. (1960). *Lancet* **2,** 187.

Hobart, M. J., Walport, M. J., and Lachmann, P. J. (1984). *Clin. Immunol. Allergy* **4,** 647–664.

Irvine, J. (1965). *N. Engl. J. Med.* **273,** 432–438.

Jackson, R. A., Buse, J. B., Rifai, R., Pelletier, D., Milford, E. L., Carpenter, C. B., Eisenbath, G. S., and William, R. M. (1984). *J. Exp. Med.* **159,** 1629–1636.

Jakobisiak, M., Sundick, R. S., Bacon, L. D., and Rose, N. R. (1976). *Proc. Nat. Acad. Sci. U.S.A.* **73,** 2877.

Kagnoff, M. F. (1982). *Nature* **296,** 158–160.

Khoury, E. L., Bottazzo, G. F., Carvalho, L. C., Wick, G., and Roitt, I. M. (1982). *Clin. Exp. Immunol.* **49,** 273.

Kite, J. H., Tyler, J., and Pascak, J. (1979). *In Immune Pathology* (F. Milgrom and B. Albini, eds.), pp. 96–100. Karger, Basel.

Knight, J. G., and Adams, D. D. (1982). *In* "Receptors, Antibodies and Disease" Ciba Foundation Symposium 90, pp. 35–56.

Kong, Y. M., Okayasu, I., Giraldo, A. A., Beisel, K. W., Sundick, R. S., Rose, N. R., Daivd, C. S., Audibert, F., and Chedid, L. (1982). *Ann. N.Y. Acad. Sci.* **392,** 191.

Lahita, R. G., Bradlow, L., Fishman, J., and Kunkel, H. G. (1982). *Arthritis Rheum.* **25,** 843–846.

Lahita, R. G., Chiorazzi, N., Gibofsky, A., Winchester, R. J., and Kunkel, H. G. (1983). *Arthritis Rheum.* **26,** 39–44.

Livezey, M. D., and Sundick, R. S. (1980). *Gen. Comp. Endocrinol.* **41,** 243.

Livezey, M. D., Sundick, R. S., and Rose, N. R. (1981). *J. Immunol.* **127,** 1469.

McDevitt, H. O., and Benacerraf, B. (1969). *Adv. Immunol.* **11,** 31–74.

Mackay, I. R. (1983). *Recent Adv. Liver Dis.* **1,** 1–23.

Mackay, I. R., Wells, J. V., and Fudenberg, H. H. (1975). *Clin. Immunol. Immunopathol.* **3,** 408–411.

Mackay, I. R., Whittingham, S., and Tait, B. D. (1977).

Manser, T., Huang, S.-Y., and Gefter, M. L. (1984). *Science* **226,** 1283–1288.

Mentnech, M. S., Pearson, D. J., Elliott, J. A., Taylor, G., and Major, P. C. (1980). *Immunopathology* **17,** 274–279.

Munoz, J. J., and Mackay, I. R. (1984). *J. Neuroimmunol.* **7,** 91–96.

Nakao, Y., Matsumoto, H., Miyazaki, T., Nishitani, H., Ota, K., Fujita, T., and Tsuji, K. (1980). *Lancet* **1,** 677–680.

Nakao, Y., Matsumoto, H., Miyazaki, T., Mizuno, N., Arima, N., Wakisaka, A., Okimoto, K., Akazawa, Y., Touji, K., and Fujita, T. (1981). *N. Engl. J. Med.* **304,** 407–409.

Nevo, S. (1974). *In* "Protides Biological Fluids" p. 649. Pergamon, Oxford.

Okayasu, I., Kong, Y. M., David, C. S., and Rose, N. R. (1981). *Cell. Immunol.* **61,** 32.

Pandey, J. P., Fudenberg, H. H., Virella, G., Kyong, C. U., Loadholt, C. B., Galbraith, R. M., Gotschlich, E. C., and Parke, J. C., Jr. (1979). *Lancet* **1,** 190–192.

Propert, D. N. (1984). *In* "Detection of Immune-Associated Genetic Markers of Human Disease" (M. J. Simons and B. D. Tait, eds.), Volume 7, p. 70. Livingstone, Edinburgh.

Propert, D. N., Kay, P., McCluskey, J., Zilko, P. J., and Mathews, J. D. (1982). *In* "Immunogenetics in Rheumatology: D-Penicillamine and Musculoskeletal Diseases" R. L. Dawkins, P. J. Zilko, and F. T. Christiansen, eds.), pp. 127–129. Exerpta Medica, Amsterdam.

Raum, D., Awdeh, Z., Yunis, E. J., Alper, C. A., and Gabbay, K. H. (1984). *J. Clin. Invest.* **74,** 449–454.
Rees, A. J., Demaine, A. G., and Welsh, K. I. (1984). *Hum. Immunol.* **10,** 213–220.
Reveille, J. D., Bias, W. B., Winkelstein, J. A., Provost, T. T., Dorsh, C. A., and Arnett, F. C. (1983). *Medicine (Baltimore)* **62,** 21–35.
Roitt, I. M. (1984). *Triangle* **23,** 67–76.
Roitt, I. M., and Doniach, D. (1967). *Clin. Exp. Immunol.* **2,** 727–736.
Rose, N. R., and Burek, C. L. (1985). *Ann. Allergy* **54,** 1–7.
Rose, N. R., Bacon, L. D., and Sundick, R. S. (1976). *Transplant. Rev.* **31,** 264.
Rose, N. R., Kong, Y. M., and Sundick, R. S. (1980). *Clin. Exp. Immunol.* **39,** 545–550.
Rose, N. R., Kong, Y. M., Okayasu, I., Giraldo, A., Beisel, K., and Sundick, R. S. (1981). *Immunol. Rev.* **55,** 299–314.
Sasazuki, T., Nishimura, Y., Muto, M., and Ohta, N. (1983). *Immunol. Rev.* **70,** 51–75.
Schernthaner, G., and Mayr, W. R. (1984). *Metabolism* **33,** 833–836.
Scholz, S., and Albert, E. (1983). *Immunol. Rev.* **70,** 77–88.
Shoenfeld, Y., and Schwartz, R. S. (1984). *N. Engl. J. Med.* **311,** 1019–1029.
Spielman, R. S., Baker, L., and Zmijewski, C. M. (1981). *In* "HLA in Endocrine and Metabolic Disorders" (N. R. Farid, ed.), pp. 37–67. Academic Press, New York.
Stastny, P., Ball, E. J., Dry, P. J., and Nunez, G. (1983). *Immunol. Rev.* **70,** 113–153.
Steinberg, A. D. (1984). *Ann. Intern. Med.* **100,** 714–727.
Steinberg, A. G. (1984). *In* "Molecular Immunology: A Textbook" Chapter 11. Dekker, New York.
Sundick, R. S., and Wick, G. (1974). *Clin. Exp. Immunol.* **18,** 127–139.
Sundick, R. S., and Wick, G. (1976). *J. Immunol.* **116,** 1319.
Sundick, R. S., Bagchi, N., Livezey, M. D., Brown, T. R., and Mack, R. E. (1979). *Endocrinology* **105,** 493.
Talal, N., Roubinian, J. R., Shear, H., Hom, J. T., and Miyasaka, N. (1980). *In* "Immunology 80: Progress in Immunology IV" (M. Fougereau and J. Dausset, eds.), pp. 889–905. Academic Press, New York.
Theofilopoulos, A. N., and Dixon, F. J. (1981). *Immunol. Rev.* **55,** 121–154.
Uno, H., Sasazuki, T., Hajime, T., and Matsumoto, H. (1981). *Nature* **292,** 768–770.
Vladutiu, A. O., and Rose, N. R. (1971). *Science* **174,** 1137.
Vladutiu, A. O., and Rose, N. R. (1975). *Cell Immunol.* **17,** 106.
Vogel, F., Kruger, J., and Chakravartti, M. R. et al. (1971). *Hum. Genet.* **12,** 284–301.
Walport, M. J., Fielder, A. H. L., and Batchelor, J. R. (1984). *In* "Immunogenetics " (G. S. Panyai and C. S. David, eds.), Chapter 7. Butterworth, London.
Weiss, J. B., Austin, R. K., and Schanfield, M. S. (1983). *J. Clin. Invest.* **72,** 96–101.
Wells, J. V., Fudenberg, H. H., and Mackay, I. R. (1971). *J. Immunol.* **107,** 1505–1511.
Whittingham, S., Mathews, J. D., Schanfield, M. S., Matthews, J. V., Tait, B. D., Morris, J. P., and Mackay, I. R. (1980). *Clin. Exp. Immunol.* **40,** 8–15.
Whittingham, S., Mathews, J. D., Schanfield, M. S., Tait, B. D., and Mackay, I. R. (1981). *Clin. Exp. Immunol.* **43,** 80–86.
Whittingham, S., Mathews, J. D., Schanfield, M. S., Tait, B. D., and Mackay, I. R. (1983). *Tissue Antigens* **21,** 50–57.
Whittingham, S., Mackay, I. R., and Mathews, J. D. (1984a). *Clin. Immunol. Allergy* **4,** 623–640.
Whittingham, S., Propert, D. N., and Mackay, I. R. (1984b). *Immunogenetics* **19,** 295–299.
Wick, G. (1970). *Clin. Exp. Immunol.* **7,** 187.
Witebsky, E., and Rose, N. R. (1956). *J. Immunol.* **76,** 408.
Witebsky, E., Rose, N. R., Terplan, K., Paine, J. R., and Egan, R. W. (1957). JAMA, *J. Am. Med. Assoc.* **164,** 1439.

CHAPTER 2

Systemic Lupus Erythematosus*

P. BRENT FERRELL
George Washington University Medical Center
Washington, D.C.

ENG M. TAN
W. M. Keck Autoimmune Disease Center
Research Institute of Scripps Clinic
Scripps Clinic and Research Foundation
La Jolla, California

* This is publication number 3511BCR from the Research Institute of Scripps Clinic, La Jolla, California. Supported in part by NIH grants AM32063 and AI10036.

I. INTRODUCTION

A. Historical Background

The use of the term "lupus" to describe various ulcerative conditions of the skin was popular for many years prior to the more restricted application of the term "lupus erythemateaux" by Cazenave (Talbott, 1974). Later physicians recognized the systemic involvement, especially William Osler (Talbott, 1974) who wrote " . . . a disease of unknown etiology with polymorphic skin lesions—hyperemia, oedema, and hemorrhage—arthritis occasionally and a variable number of visceral manifestations, of which the most important are gastrointestinal crises, endocarditis, pericarditis, acute nephritis, and hemorrhage from the mucous surfaces." However, it was not until the discovery of the lupus erythematosus (LE) cell phenomenon in 1948 by Hargraves at the Mayo Clinic that the disease began to be appreciated as an autoimmune phenomenon. Later work by Holman and Kunkel (1957), Miescher (1957), and Seligmann (1975) gave real importance to this phenomenon. These investigators were able to show that serum from patients with systemic lupus erythematosus (SLE) very frequently contained not only antibodies responsible for the LE cell phenomenon but, more importantly, antibodies directed to their own deoxyribonucleic acid (DNA). With the discovery of anti-DNA antibodies, a new era in the field of human autoimmune diseases began. Subsequently, it was shown that patients with SLE demonstrated multiple defects in both humoral and cellular immunity.

B. Importance of Autoantibodies

Autoantibodies, especially antinuclear antibodies (ANA), have played an important role in furthering our understanding of the pathogenesis of SLE. Despite continued debate over the existence of "ANA-negative" SLE, it remains clear that autoantibodies and ANA are closely related to the expression of the disease in both humans and animal models. This nearly constant relationship between disease and the presence of ANA has been more strongly emphasized in new American Rheumatism Association (ARA) classification criteria for SLE (Tan *et al.*, 1982). The new criteria include not only the presence of a positive ANA but also the more SLE-specific ANA, anti-Sm (anti-Smith), and anti-DNA. Accordingly, this review of SLE, although general in scope, will emphasize the immunobiology of ANA in SLE and the clinical usefulness of ANA profiles in the differential diagnosis of SLE. In addition, this chapter will highlight

important clinical, pathogenetic, and immunological aspects of human SLE as well as recent work with experimental animal models of SLE.

II. GENERAL DESCRIPTION

A. Classification Criteria

In the early years of investigation of SLE, it became readily apparent that an effective study of an autoimmune diathesis that could present with such a wide spectrum of clinical and serological abnormalities would first require a set of classification criteria. In 1971 (Cohen *et al.*, 1971), a committee of the ARA devised criteria for SLE (Table I) to more readily permit comparison of studies done by different investigators. It is worth emphasizing that these criteria were never intended to be diagnostic criteria, although they are widely used as such. In 1982 (Tan *et al.*, 1982), with the advent of more sensitive, specific, and reliable tests for ANA and the realization that some items of the early clinical criteria were not sufficiently discriminating, the 1971 set of criteria was revised. These new criteria (Table II) eliminated infrequent and less discriminating clinical signs such as Raynaud's phenomenon and alopecia, and added important laboratory findings such as positive fluorescent ANA and anti-DNA or anti-Sm. To be classified as having SLE, one must have at least four of the

TABLE I

Preliminary ARA Classification Criteria (1971) for SLE[a]

Facial erythema (butterfly rash)
Discoid lupus
Raynaud's phenomenon
Alopecia
Photosensitivity
Oral or nasopharyngeal ulceration
Arthritis (nondeforming)
LE Cell
Chronic false-positive STS
Profuse proteinuria (>3.5g/day)
Cellular casts
Pleuritis and/or pericarditis
Psychosis and/or convulsions
Hemolytic anemia/leukopenia/thrombocytopenia

[a] From Cohen *et al.* (1971).

TABLE II
Revised ARA Classification Criteria (1982) for SLE

Malar rash
Discoid rash
Photosensitivity
Oral ulcers
Arthritis (nondeforming polyarthritis)
Serositis (pleuritis and/or pericarditis)
Renal disorder (proteinuria >0.5 g/day or cellular casts)
Neurologic disorder (psychosis and/or seizures)
Hematologic disorder (leukopenia or lymphopenia/hemolytic anemia/thrombocytopenia)
Immunologic disorder (anti-DNA/anti-Sm/LE cell/false-positive STS)
Antinuclear antibody

11 criteria, either at one time or serially. These new criteria were shown to have a 96% sensitivity and specificity for SLE (Tan *et al.*, 1982).

B. Epidemiology

Systemic lupus erythematosus is a disease primarily of women, especially of young black women. Prevalence rates in one study (Siegel and Lee, 1975) were 55.8 per 100,000 for black women and 3.3 per 100,000 for black men. Corresponding rates for white women and white men were 16.8 and 2.9. The overall annual incidence of SLE in the 15-year study ending in 1973 was 2.0 per 100,000. Earlier studies (Kurland *et al.*, 1969; Siegel *et al.*, 1970) showed rates approximately twice as high. A later study by Fessel (1974) of the Kaiser Foundation Health Plan group from 1965 to 1973 showed an incidence of 7.6 per 100,000, using the 1971 ARA criteria. Prevalence rates were nearly three times as high for the Kaiser group as for the New York study (Siegel and Lee, 1975). In both studies incidence rates rose nearly two- or threefold over the periods of study. These increases probably reflected improved diagnostic capabilities (e.g., LE cell test), but also may have been due to an increased attack rate (Hahn, 1980).

C. Etiology

1. Viruses

It is generally assumed that SLE is an autoimmune disease associated with various definite immune defects. Neither the initiation nor the actual sequence of events leading to these immune defects is known. It is felt,

however, that some combination of environmental, genetic, and host factors (e.g., hormones) must be present for full expression of the disease. Foremost among the possible candidates for environmental agents that may trigger the disease are viruses. Spurred by the discoveries in 1969 of viruslike structures in kidney biopsies from SLE patients, interest in a viral etiology for SLE increased (Hurd *et al.*, 1969). Although it was shown that on closer inspection these tubuloreticular structures were not truly paramyxovirus nucleocapsids (Pincus *et al.*, 1970), attempts to show increased antiviral antibody titers in SLE patients met with more success (Phillips and Christian, 1973). In fact, SLE patients very often had elevated levels of antibodies to various common viruses such as measles and rubella viruses. Because of broad-spectrum antiviral reactivity, these results were interpreted as merely a reflection of polyclonal B-cell reactivity (Pincus, 1982). Attempts at viral isolation and direct demonstration of viruses by electron microscopy were uniformly unsuccessful.

Type C retroviruses are animal viruses with unique capabilities for complex interactions with host cells. It is possible that these interactions, inadequately controlled in the susceptible host; might be related to the autoimmune phenomenon occurring in SLE. One experimental model for SLE, the New Zealand B/W F_1 hybrid mouse (NZB/W), has been intensively studied with respect to the possible role of type C retrovirus. Viral particles may be seen early in the life of NZB/W mice, with a subsequent rise in antibody titers to viral antigens correlating with immune complex nephritis (Izui *et al.*, 1979). On the other hand, separation of viral infection and nephritis can be achieved with backcross matings with normal SWR mice (Datta *et al.*, 1978). These F_1 mice show an increased viral output without nephritis. Additionally, Dixon *et al.* (1971) showed that more of these glomerulonephritis-causing antibodies were directed to deoxyribonucleoprotein (DNP) than to the gp70 protein and other viral antigens. At present the role of viruses in both murine and human lupus still remains an intriguing but unproven hypothesis.

2. Drugs

A number of drugs have now been shown to cause a drug-induced lupus syndrome (DILE) (Harmon and Portanova, 1982). Drug-induced lupus syndrome is in many ways similar to idiopathic SLE, but differs in the population at risk, lack of renal and CNS disease, and the relatively limited variety of ANA. Patients with DILE are most often older, white, and male, with predominantly polyarthritis and/or serositis. ANA are usually antihistone but may include anti-single-strand DNA (ssDNA) antibodies (Fritzler and Tan, 1978). HLA-DRw4 has been shown to be increased in hydralazine DILE patients (Batchelor *et al.*, 1980).

Liver acetyltransferase enzymes detoxify drugs by acetylation. Enzyme levels are genetically controlled and are phenotypically expressed in humans as fast or slow acetylators. Acetylator phenotype is important in the expression of ANA in patients on hydralazine and procainamide. A fast acetylator phenotype with cumulative hydralazine ingestion of >400 g is not associated with the production of ANA, while 60% of persons with a slow acetylator phenotype will produce ANA at this level (Perry *et al.*, 1970). Increased frequencies of procainamide-induced ANA are also associated with a slow acetylator phenotype (Woosley *et al.*, 1978). The mechanisms by which these drugs can induce ANA were reviewed by Harmon and Portanova (1982). The reasons only some of these patients with drug-induced ANA develop clinical lupus remain unknown.

3. Environmental Factors

It has long been known that sunlight is closely related to some cases of SLE. Certain patients give dramatic histories of having symptoms appearing *de novo* after a sunburn, while others relate personal histories of photosensitivity to sunlight that are more than the usual experiences of a normal population. In the skin rash of human SLE, a common and almost characteristic finding is the presence of stippled deposits of immunoglobulin and complement at the dermal–epidermal junction. A very attractive hypothesis for this finding is that these deposits comprise antigen–antibody complexes, which are precipitated at this region. The pathogenic mechanism might be considered to arise from the destruction of epithelial cells that release nuclear antigens into the dermis of the skin, come into contact with antinuclear antibody at the dermal–epidermal junction, and are precipitated as immune complexes. This hypothesis and the proposed pathogenic mechanism were tested in an experimental model using ultraviolet light (UVL)-irradiated mice (Tan *et al.*, 1976). Ultraviolet light irradiation of DNA has been shown to induce a photochemical reaction in DNA, causing the formation of thymine dimers. These thymine dimers are highly immunogenic and cause the production of antibodies specific only for thymine dimers. In the experimental model, mice were first immunized with thymine dimers so that circulating antidimer antibodies were known to be present. These animals were then whole-body irradiated with UVL, which caused a photochemical induction of thymine dimer in the nuclei of epithelial cells. The epithelial cells have the capacity to repair thymine dimers by a mechanism of excision of the thymine dimer and extrusion of the dimers across the cell membrane.

In this model, it was shown that immune complexes of thymine dimer and antibody to dimer were present as stippled deposits in the dermal–

epidermal junction of mice recovering from whole-body UVL irradiation. An extension of this work has been performed on humans to show that similar types of thymine dimers were induced in the nuclei of epithelial cells after irradiation of isolated patches of skin (Tan and Stoughton, 1969). A reasonable conclusion from these experiments is that a similar mechanism for the formation of immune complex deposition at dermal–epidermal junction could occur in the skin of SLE patients who already have preexisting circulating ANA. It might be important to remember that thymine dimers might not be the antigen of importance as far as the lesions of human SLE skin are concerned. UVL causes transformation of other intranuclear macromolecules such as nuclear proteins and DNA– or RNA–protein complexes, and these other macromolecules may be of greater importance in pathogenicity.

The effect of sunlight is of greater than just theoretical interest. It has been shown that UVL of wave lengths of 300 nm can be effective in the photochemical transformation to thymine dimer. Studies in photobiology have shown that UVL of wave lengths in the region of 300 nm are detected on the earth's surface (Koller, 1965).

D. Animal Models

Animal models of SLE have been studied intensively from both an immunological and a viral standpoint. Currently, there are at least three separate mouse strains that express features of an autoimmune disease very similar to human SLE (Dixon, 1982).

The NZB/W mouse strain was first developed by Helyer and Howie in 1963 by mating the NZB autoimmune hemolytic mouse with a normal NZW mouse. Almost 100% of the hybrid animals developed a rapidly progressive and lethal glomerulonephritis associated with antinuclear antibodies. The ANA that were produced included antibodies to double strand DNA (dsDNA) and single strand DNA (ssDNA), as well as to nuclear proteins and to nuclear and transfer RNA (Talal *et al.*, 1974). The antibodies that appeared to have the most pathological importance were anti dsDNA, since they were eluted from sections of diseased kidneys in high concentrations (Dixon *et al.*, 1971). Animals developed disease and died at a mean of 439 days and 280 days for males and females, respectively (Howie and Helyer, 1968). The marked difference in survival time between sexes pointed to an important aspect that has held true for the more recently developed lupuslike models: the fact that within the strain endogenous modifying factors exist that affect the severity of disease expression. In the NZB/W mouse, female sex hormones produce more rapid onset of disease. In another murine lupus model, the MRL strain,

lupus develops early in life in the MRL/lpr mouse and later in the MRL/n mouse. The factor responsible for this difference in onset is the *lpr* (or *lymphoproliferative*) gene (Murphy and Roths, 1979). Like estrogens in NZB/W disease, the *lpr* gene is clearly an accelerating factor in MRL mice. These mice also make a wide variety of autoantibodies including anti-DNA and anti-Sm (40%), a marker for human lupus (Eisenberg *et al.*, 1978). Like human SLE, anti-Sm-negative and -positive mice do not differ in disease severity. The third murine lupus model is the BXSB mouse (Murphy and Roths, 1979); again, early and late disease onset can be seen. In contradistinction to the NZB/W, the accelerating factor predisposing to early disease in the BXSB appears to be linked to the Y chromosome, making the male mouse in BXSB disease more prone to early death (Theofilopoulos and Dixon, 1981).

After many years of intensive investigation into the etiopathogenesis of murine lupus, Dixon (1982) suggested that some generalizations can be made that are pertinent to our understanding of human SLE. First, no single gene or group of genes [e.g., the major histocompatibility complex (MHC)] has been defined that can account for the autoimmune diathesis of SLE. Rather, the genetic endowment predisposing to autoimmune disease is a diverse one, probably coming from within as well as outside the MHC region. Second, both endogenous (*lpr* gene in MRL mice) and exogenous (neonatal lymphocytic chorimeningitis in all strains) factors can modify disease expression. Finally, B-cell hyperactivity, autoantibody formation, and immune complex pathology are consistent findings in all forms of murine lupus.

III. CLINICAL MANIFESTATIONS

As a multisystemic disease, SLE may present with widespread involvement of various organs and tissues. Ordinarily, however, its initial presentation is limited in scope with additional areas of involvement occurring as the disease progresses. Fever, weight loss, and increasing fatigue together or separately may herald the onset of SLE. Arthralgias and arthritis are the most common forms of clinical expression, symmetrically affecting the small joints of the hands, wrists, and knees (Rothfield, 1982). Deforming arthritis is unusual, and erosive changes are quite rare (Russel *et al.*, 1981). Synovial fluid findings from swollen joints reveal leukocyte counts <3000, in fluid that is characteristically clear to only slightly cloudy with good viscosity (Pekin and Zvaifler, 1970). The next most common manifestation of SLE is cutaneous disease. Cutaneous lesions may range from

the erythematous "butterfly rash," which is evanescent and nonscarring, to the disfiguring discoid lesion. Other skin changes include (1) erythematous maculopapular, often pruritic rash, (2) lupus profundus (nonsupporative panniculitis), (3) vasculitic nodules with a tendency to ulcerate, (4) annular, polycyclic lesions (Gilliam and Sontheimer, 1982). These latter skin lesions are often associated with a positive serum test for anti-SSA/Ro antibody.

Mucous membrane lesions occur in the form of oral and nasal ulcers (Ropes, 1976). These are painless and usually heal with corticosteroid therapy. Serositis may appear as pleuritis with or without effusions or as pericarditis. Pericarditis is usually detected as a friction rub without significant pericardial fluid (Ropes, 1976). Lung involvement in SLE is relatively uncommon, but may take the form of either an acute pneumonitis with transient infiltrates or as diffuse interstitial pneumonitis. Pulmonary hypertension and severe pulmonary hemorrhage have also been described (Matthay *et al.*, 1974). Cardiac manifestations other then pericarditis include myocarditis and the verrucous endocarditis of Libman and Sacks (1924); however, the latter condition is rarely observed at present.

As the above summary indicates, SLE may present clinically in many different ways and in different degrees of severity. However, the involvement of two organ systems, the kidney and the nervous system, portends a poor prognosis. Renal disease in SLE may be divided into mesangial, membranous, focal, and diffuse proliferative nephritis (Baldwin *et al.*, 1970; McCluskey, 1982). Mesangial nephritis is associated with relatively few if any urinary abnormalities but has on occasion been noted to progress to a diffuse proliferative lesion. Membranous nephritis is often associated with nephrotic syndrome, although hematuria is infrequent. This type generally carries a better prognosis than the proliferative types. In focal proliferative lupus nephritis, $<50\%$ of glomeruli are affected by segmental proliferation. Urinalyses are abnormal with respect to both protein and red blood cells, but significant renal insufficiency is uncommon. On the other hand, diffuse proliferative glomerulonephritis is associated with the entire spectrum of urine abnormalities, significant renal failure, and/or nephrotic syndrome.

In addition to renal disease, involvement of the nervous system in SLE may herald a significant change in disease progression. Besides peripheral neuropathy, patients may manifest a long list of neuropsychiatric problems including paralysis, seizures, headaches, and organic as well as nonorganic syndromes (Feinglass *et al.*, 1976). Psychotic states secondary to lupus are notoriously difficult to diagnose and treat. This problem is best illustrated when one considers the patient with severe lupus nephritis on

high doses of corticosteroids who develops psychosis. Diagnostic possibilities range from active lupus cerebritis and corticosteroid-induced psychosis to brain abscess in an immunosuppressed host.

Hematological disturbances, very common in SLE, have been reviewed by Budman and Steinberg (1977). Anemia is present in more than one-half of cases, the primary cause being bone marrow suppression related to chronic disease. Other causes for anemia such as autoimmune hemolytic anemia should be sought. Autoimmune hemolytic states can precede the other manifestations of SLE by a number of years. Direct Coombs' tests are often positive (Boehner *et al.*, 1968), but are not usually associated with a hemolytic anemia. Anemia may also be secondary to chronic blood loss. Bleeding may occur in association with thrombocytopenic states. Thrombocytopenia is associated with IgG antiplatelet antibodies and is present in ~15% of patients (Budman and Steinberg, 1977). Like autoimmune hemolytic anemia, autoimmune thrombocytopenic purpura can also predate the onset of typical SLE by many years. In contrast to thrombocytopenia, bleeding does not usually occur with the presence of a lupus anticoagulant. Antibodies to clotting factors have been noted in lupus very uncommonly, but when present are directed to one of the following: VII, IX, XI, or XII (Rick and Hoyer, 1975). More commonly, prolonged partial thromboplastin times are seen secondary to the inhibitory action of anti-phospholipid in patients' sera. Enhanced thrombosis rather than bleeding is the rule with this latter lupus anticoagulant. Because of cross-reactivity with cardiolipin, lupus anticoagulant may be correlated with biological false-positive tests for syphilis (Laurell and Nilsson, 1957).

Given the long list of possible clinical manifestations, one can easily see the rationale for devising the previously listed classification criteria. Other diagnostic possibilities often considered during the evaluation of patients shown ultimately to have SLE include subacute bacterial endocarditis, rheumatoid arthritis, rheumatic fever, leukemia, lymphoma, secondary syphilis, and serum sickness. It may take a number of years actually to confirm a diagnosis of SLE, because separate organ system involvement may proceed very slowly. For this reason, performance of serological tests may yield an early and specific diagnosis.

IV. HISTOPATHOLOGY

The histopathology of SLE has been described in comprehensive reviews (Cruickshank, 1974) and will not be discussed in detail here. How-

ever, certain features of histopathology of the kidney, skin, and brain will be mentioned.

Renal involvement is a major cause of morbidity and mortality in SLE. The pathological features have been elucidated by renal biopsies performed on patients with varying severity of illnesses and in some instances by serial biopsies performed on the same patient. A classification of lupus nephritis has been proposed by Andres *et al.* (1970). (1) Mild (focal) lupus nephritis is characterized by segmental proliferation of some glomeruler tufts while other tufts appear normal. Usually <50% of the glomeruli are involved, and they exhibit segmental proliferation, that is, hypercellularity with some areas of necrosis of the involved segment. Immunoglobulins and C3 can be demonstrated in the mesangium of all glomeruli, and scattered areas of fine granular fluorescence may be present among the capillary loops especially in areas of proliferation. (2) Severe (diffuse) proliferative lupus nephritis is characterized by abnormalities of >50% of the total area of the glomerular tufts. Interstitial infiltrates of mononuclear cells are seen, and epithelial crescents are common as are sclerotic glomeruli. Deposits of immunoglobulins and C3 appear as granules or irregular lumps along the peripheral capillary well and mesangium. (3) Membraneous nephritis is characterized by thickening of the basement membrane when observed by light microscopy. No proliferation is observed, although there may be slight irregular increases in mesangial cells and matrix. By immunofluorescence, immune complex deposits are diffuse and located in the subepithelial region and mesangium. In recent years, it has been shown that some patients with these different classifications of nephritis can be seen to transform from the mild or focal nephritis to the diffuse and proliferative form; in patients who have responded to treatment, the reverse transition has also been observed. Finally (4), an additional classification, mesangial (minimal) lupus nephritis, has been described, in which the pathological abnormalities are limited to the mesangium. The diagnosis is made mostly by studies by immunofluorescence and electron microscopy that show deposits of IgG and C3 in the mesangium and occasionally along capillary walls.

A second organ that has yielded considerable information regarding the pathogenesis of SLE is the skin. With immunofluorescence methods, it has been clearly demonstrated that the "liquefaction necrosis" of the basement membrane zone seen by light microscopy consists of stippled or small lumpy deposits of immunoglobulin and complement at the dermal–epidermal junction (Burnham *et al.*, 1963; Tan *et al.*, 1966). Some authors have referred to this finding as the "lupus band," but this is an inaccurate description because this phenomenon at the dermal–epidermal junction is

not restricted to SLE and may occur in diseases such as rosacea, mixed connective tissue disease, and urticaria. Recently, properdin and factor B (Schrager and Rothfield, 1976) have also been demonstrated at the dermal–epidermal junction, which brings up the possibility that nonimmunological factors that can activate the alternative pathway of complement–polysaccharides, endotoxins, and other substances—may be involved in pathogenesis.

The pathology of central nervous system (CNS) disease in SLE is not well understood. A patient may have had seizures, hemiparesis, or hemiplegia without postmortem evidence of cerebral vasculitis or brain damage detectable by light microscopy. However, by immunofluorescence microscopy, IgG and IgM have been found in the choroid plexus of the brain together with complement components (Atkins *et al.*, 1972; Lampert and Oldstone, 1974). Such immunoreactants can also be found in the walls of small arterioles, capillaries, and venules in the brain. The search for immune complexes that might play a significant role in the CNS disease of SLE deserves extension, since CNS disease is becoming a more frequently observed complication of the disease.

V. IMMUNOLOGY

A. Immunological Aberrations

1. Humoral Immunity

It has been demonstrated that the serum factor in SLE patients responsible for the LE cell phenomenon noted by Hargraves (1948) is antibody directed to nucleoprotein or histone (Holman and Deicher, 1959; Tan *et al.*, 1982). This hallmark discovery by Hargraves and the usefulness of the LE cell test have been overshadowed by the advent of the fluorescent antinuclear antibody test and the characterization of numerous antinuclear antibodies during the last 20 years. In addition to non-tissue-specific autoantibodies such as ANA and anticytoplasmic antibodies, tissue-specific autoantibodies are also seen in SLE (Table III). These latter antibodies include those directed to circulating erythrocytes, lymphocytes, and platelets as well as those directed to tissue-specific antigens of the thyroid, liver, stomach, adrenal gland, and muscle.

Certainly the most intensively studied and best characterized of these types of autoantibodies are the antinuclear antibodies (ANA). Antinuclear antibodies occurring in SLE are quite diverse, but may be broken down into broad groups according to their specificities for various nuclear

TABLE III
Autoantibodies in SLE

To nuclear antigens
DNA
Nucleoprotein
Histones
Nonhistone (acidic) proteins
To cytoplasmic antigens
RNA
Ribosomes and other RNA–protein complexes
Cytoplasmic proteins
To clotting factors
Lupus anticoagulants
To red cell antigens
To white cell antigens
T-Lymphocyte cell-surface antigens
B-Lymphocyte cell-surface antigens
To platelet antigens
To other tissue-specific (thyroid, liver, muscle, stomach, adrenal) antigens

macromolecules (Table IV). The first demonstration of anti-DNA antibodies came about as investigators were dissecting out the reactive moiety accounting for the LE cell phenomenon (Robbins *et al.*, 1957). This moiety eventually proved to be 1 *M* NaCl soluble nucleoprotein (DNP) (Holman and Deicher, 1957). As a result of these studies, however, interest grew in developing methods to detect anti-DNA in SLE sera. Methods employing purified native DNA were developed including double immunodiffusion, complement fixation, hemagglutination, and counterimmunoelectrophoresis (CIE). By immunodiffusion it appeared that anti-DNA antibodies were of at least two different specificities. One specificity was felt to be directed to the phosphodiester backbone of DNA, which would account for antibody reactive with both single strand and double strand DNA. The other large group of anti-DNA antibodies had specificity for the purine and pyrimidine bases exposed only on single strand DNA. This would account for those antibodies reactive only with single strand DNA. A third but relatively rare type of anti-DNA antibody has been seen with specificity only for the helical structure of DNA. This third type could not be absorbed by single strand DNA or synthetic homopolymers (Gilliam *et al.*, 1980).

Anti-DNA antibodies rapidly became a serological hallmark of SLE. Tan and others (1966) demonstrated the close relationship of changes in anti-DNA antibodies and free serum DNA with clinical manifestations of SLE (Fig. 1). In the clinical course depicted here, significant levels of

TABLE IV
Autoantibodies to Nuclear Antigens (ANA) in SLE

Antibody specificity	Clinical characteristics
Double strand DNA (dsDNA)	
Antigenic determinant present in both dsDNA and single strand (ss) DNA.	60–70% of patients with SLE. When in high titer, practically a diagnostic marker
Single strand DNA (ssDNA)	
Antigenic determinant related to exposed purines and pyrimidines	60–70% of patients with SLE. However, present in other diseases, including nonrheumatic diseases
Histones	
Antigenic determinants in all subclasses: H1, H2A, H2B, H3, H4, and H2A/H2B, H3/H4 complexes	70% of patients with SLE. >95% of patients with procainamide- and hydralazine-induced LE
Nonhistone antigens	
Sm antigen	
Antigenic determinant is protein(s) complexed to five species of small nuclear RNA (snRNA)	30–40% of patients with SLE. Diagnostic marker
U1 RNP	
Antigenic determinant is protein(s) complexed to U1-RNA	35–45% of patients with SLE. >95% of patients with mixed connective tissue disease
SS-A/Ro	
Antigenic determinant is 61-kdalton protein complexed to RNAs	30–40% of patients with SLE. 60–70% of patients with Sjögren's syndrome. Related to neonatal lupus
SS-B/La	
Antigenic determinant is 43-kdalton protein complexed to RNAs	15% of patients with SLE. 45–60% of patients with Sjögren's syndrome
PCNA (proliferating cell nuclear antigen)	3% of patients with SLE
Determinant is 33-kdalton protein	

anti-DNA are seen in a period of disease quiescence without free DNA being present. Subsequently a flare of activity is associated with disappearance of antibody and appearance of free antigen. This sequence of events suggested that immune complexes of DNA and antibody were being formed during flares of disease activity and was the first evidence that circulating ANA might be involved in pathogenesis, such as in tissue deposition of immune complexes. Other investigators confirmed this by demonstrating the presence of anti-DNA in cryoprecipitates (Davis *et al.*, 1978) and acid eluates of kidney sections from SLE patients (Koffler, 1974). It is clear from these and other studies that a unique relationship

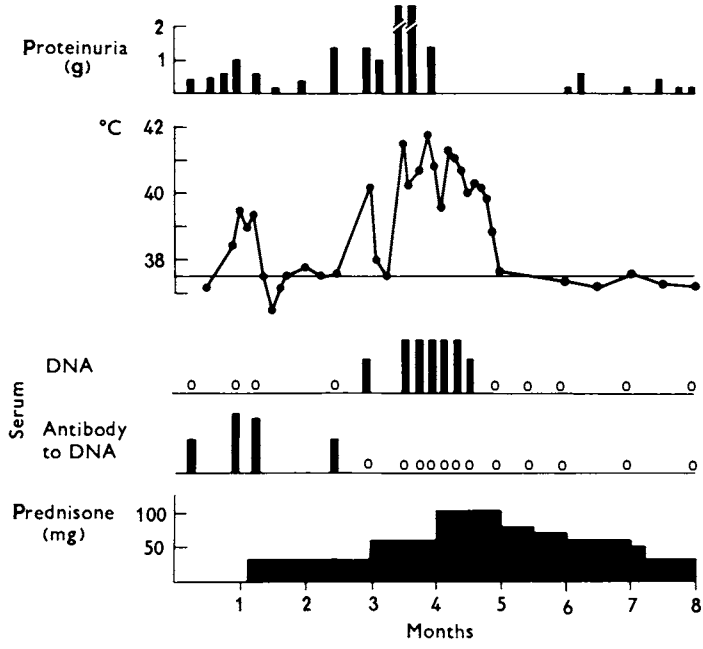


FIG. 1. A patient with SLE initially showing DNA antibody in four consecutive serum samples. Exacerbation of illness with high fever and increased proteinuria coincided with appearance of DNA antigen in serum and disappearance of antibody. (From Tan *et al.*, 1966.)

exists between anti-DNA and SLE. One group of investigators has been quite successful in showing that in individual patients sequential elevations and depressions in anti-DNA levels may be valuable indicators of impending disease flares (Swaak *et al.*, 1979, 1982). Other investigators have attempted to correlate disease activity with potentially pathogenic properties of the anti-DNA itself. Such properties include complement-fixing ability, immunoglobulin class or subclass, and avidity or ability to precipitate. It is generally held that activity and nephritis are most often correlated with IgG complement-fixing antibodies (Hahn, 1982). From these studies has also come the finding that high titers of anti-DNA antibodies are nearly always found only in SLE. This point is emphasized in the graph in Fig. 2. Although other systemic rheumatic diseases may express anti-DNA antibodies, high titers are almost exclusively associated with SLE.

Turning from antibodies reactive with DNA to antibodies directed to the basic nuclear proteins, histones, we note that approximately two-

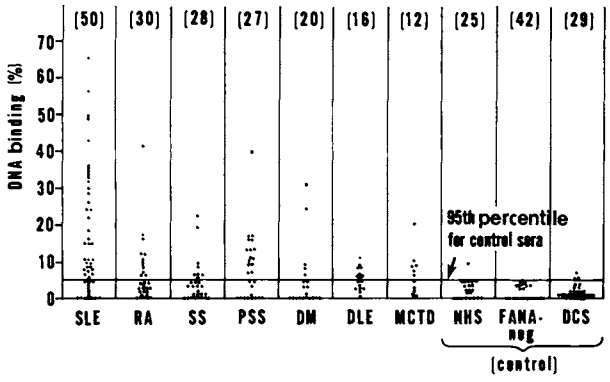


FIG. 2. Antibody to native DNA was measured by a radioimmunoassay method. The numbers in parentheses indicate the number of patients tested. With few exceptions, high binding for native DNA was present in the SLE group. SLE, systemic lupus erythematosus; RA, rheumatoid arthritis; SS, Sjögren's syndrome; PSS, progressive systemic sclerosis; DM, dermatomyositis; DLE, discoid lupus; MCTD, mixed connective tissue disease; NHS, normal human serum; FANA-neg, serum negative for fluorescent antinuclear antibody; DSC, disease control serum. (From Notman *et al.*, 1975.)

thirds of patients with idiopathic SLE show antihistone reactivity (Fritzler and Tan, 1978). Antihistone antibodies are much more common in procainamide-induced lupus, occurring in 96% of patients when their sera are examined by the histone reconstitution ANA test. In hydralazine-induced lupus, the next most common form of drug-induced SLE, antibodies to histones also occur but are not detected by the histone reconstitution test. These antibodies are directed to histones H3 and H4 and must be detected by an enzyme-linked immunosorbent assay (ELISA) using purified histone fractions (Portanova *et al.*, 1982). Use of this latter assay showed that procainamide-induced antihistone antibodies were predominantly directed to H2A–H2B complexes (Table V). In distinguishing idiopathic from drug-induced SLE by serological means, it is also important to remember that the latter disorder often displays a restricted diversity of ANA compared with the heterogeneity of the ANA response in the former (Table VI).

Like antibodies to dsDNA, antibodies to a nonhistone protein named Sm carry high specificity for SLE. Anti-Sm antibodies occur in ~30% of SLE patients, and may be considered a serological marker for this disease (Fig. 3). It is not surprising in view of its diagnostic significance that the Sm–anti-Sm system has evoked so much interest since its discovery in 1966 by Tan and Kunkel. In immunofluorescence, anti-Sm gives a speckled staining pattern not significantly different from that of antinuclear

TABLE V

Types of Antihistone Antibodies Induced by Procainamide (Pr) and Hydralazine (HY)

Antibody class	Drug	Antibodies to			
		H2A	H2B	H2A/H2B	H3
IgM	Pr	0	0	++	0
	Hy	++	+/−	+/−	++
IgG	Pr	++	+/−	+++	+/−
	Hy	0	0	+/−	+/−

ribonucleoprotein (nRNP) but may be distinguished from anti-nRNP by double immunodiffusion. Recently, Lerner and Steitz (1979), have made interesting new observations on the nature of autoantigens such as Sm and nRNP. By using anti-Sm as a reagent to immunoprecipitate ^{32}P- and ^{35}S-labeled tissue culture cell extracts, these investigators have been able to show that certain small nuclear RNAs (U1, U2, U4, U5, and U6) are complexed with Sm antigen in protein–RNA particles (Fig. 4). Some of these small nuclear RNA species such as U1 RNA are thought to be involved in posttranscriptional processing of heterogeneous nuclear RNA to messenger RNA. Yang *et al.* (1981) demonstrated that anti-Sm could inhibit this splicing function of small RNA and suggested that the U1 RNA–protein particle was involved in splicing of early RNA transcripts.

Another nonhistone autoantigen that also reacts with SLE sera is nuclear RNP (Sharp *et al.*, 1972). Like Sm, nRNP also associates with small

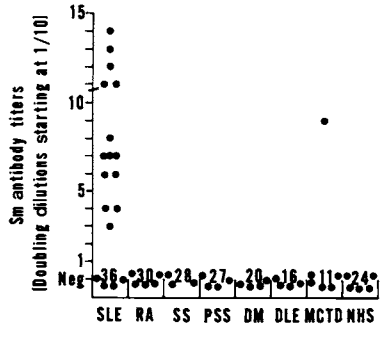


FIG. 3. Antibody to Sm antigen was determined by hemagglutination in different rheumatic diseases. With the exception of one patient with mixed connective disease, all other sera with anti-Sm were from patients with SLE. (From Notman *et al.*, 1975.)

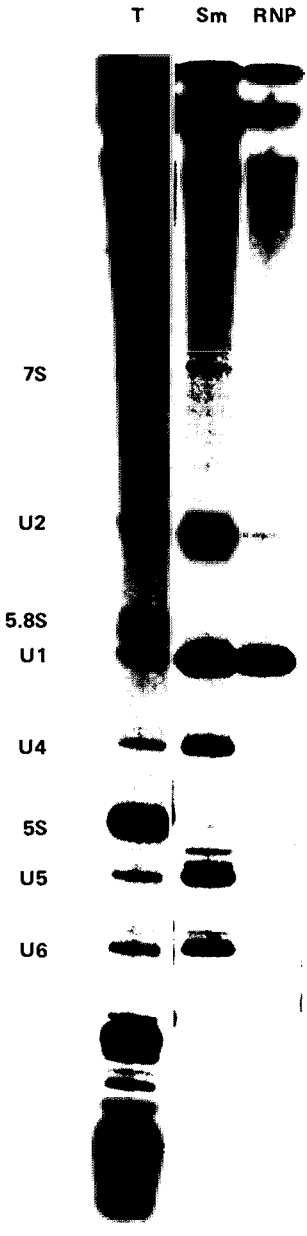


FIG. 4. Profiles of small nuclear RNAs immunoprecipitated by anti-Sm and antinuclear RNP sera. HeLa cells were labeled with ^{32}P and cell extract reacted with sera. The immunoprecipitates were solubilized and run on polyacrylamide gels to identify precipitated RNAs. Left lane represents total cellular RNA. Middle lane shows five major RNA bands precipitated by anti-Sm serum: U1, U2, U4, U5, and U6 RNAs. Right lane shows that only U1 RNA is precipitated by anti-RNP serum.

TABLE VI
Antibody Profiles in Idiopathic SLE and Drug-Induced LE

	Antibody to				
	Histones	Sm	nRNP	dsDNA	ssDNA
Idiopathic SLE (20)[a]	7	9	3	7	5
Drug-LE (23)[a]	22	0	0	0	3

[a] Numbers in parentheses indicate numbers of patients.

nuclear RNA but unlike Sm is highly specific in that it always associates with only U1 RNA (Lerner and Steitz, 1979). For this reason anti-nRNP has come to be known as anti-U1 RNP. Anti-U1 RNP may be seen as a nonspecific speckled pattern of immunofluorescent staining on nuclear substrates but may be specifically identified by immunodiffusion or CIE. High titers of anti-U1 RNP in the absence of other ANA, especially anti-DNA, are indicative of a clinically distinct entity, mixed connective tissue disease (MCTD) (Tan, 1982). The mere presence of anti-U1 RNP lacks the specificity of anti-Sm, as can be seen in Fig. 5.

A final group of ANA in SLE directed to nonhistone nuclear proteins

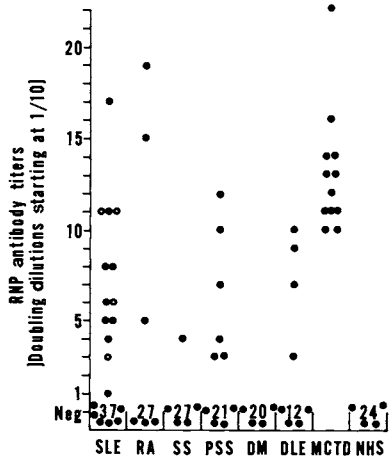


FIG. 5. Antibody to nRNP was determined by hemagglutination with the same sera as in Fig. 5. Note that all patients with MCTD had uniformly high titers. However, antibody was also present in other diseases, sometimes in relatively high titers. (From Notman *et al.*, 1975.)

are the Sjögren's syndrome antibodies, anti-SSA/Ro and anti-SSB/La. First described by Alspaugh and Tan (1975, 1976) in a study of precipitating ANA in Sjögren's syndrome, anti-SSA and anti-SSB were later shown to be identical (Alspaugh and Maddison, 1979) to two previously described antibodies, anti-Ro (Clark *et al.*, 1969) and anti-La (Mattioli and Reichlin, 1974). Anti-SSA/Ro occurs in 30–40% of patients with SLE but in almost 70% of Sjögren's sicca patients. Anti-SSB/La is less frequent in SLE, appearing in only 15%, but three to four times more frequently in Sjögren's syndrome. From a clinical standpoint, recent investigations have revealed very interesting disease associations of the anti-SSA/Ro antibody. In a 2-year study of three infant–mother pairs with SLE, Franco *et al.* (1981) showed that all the infants had anti-SSA/Ro at time of birth and that the antibody was no longer detectable at 6 months postnatally. This finding was consistent with transplacental passage of ANA. More important was the observation of complete heart block (CHB) in one of the infants. Similar observations were made in an extension of the initial study (Scott *et al.*, 1983), confirming the association of anti-SSA/Ro with CHB.

Screening of pregnant lupus patients for anti-SSA/Ro may provide valuable information with regard to the possible occurrence of the neonatal lupus syndrome and life-threatening CHB in the fetus. The frequent occurrence of anti-SSA/Ro and/or anti-SSB/La has also been discovered in another dermatological subset of SLE, subacute cutaneous lupus (Gilliam and Sontheimer, 1982). This disease is characterized by extensive nonscarring psoriaform or annular, polycyclic lesions, sometimes accompanied by various musculoskeletal complaints. Interestingly, it has also been found that the expression of anti-SSA/Ro in these patients is closely linked to the presence of the HLA-DRw3 antigen (Maddison, 1982). In one study (Bell and Maddison, 1980) all 10 patients with SLE and anti-Ro were HLA-DRw3 positive, whereas only 28% of anti-Ro-negative patients were HLA-DRw3 positive. This latter figure was almost identical to the prevalence of DRw3 in a control population.

2. Cellular Immunity

Clearly, the most reproducible abnormality of cellular immune function in SLE is the marked B-cell hyperreactivity (Kunkel, 1980). Despite overall reduced numbers of both T and B cells, assays for plaque-forming cells performed on peripheral blood lymphocytes (PBL) from SLE patients show excessive numbers of such cells when compared to normal PBL (Decker *et al.*, 1979). Pokeweed mitogen stimulation seems to have little enhancing effect on these immunoglobulin (Ig)-producing cells. Response

to stimulation with chemical haptens *in vitro* is also enhanced with production of specific antihapten IgM antibodies. Without such stimulation, however, the Ig-producing B cells present in SLE PBL tend to be IgG producers rather than IgM producers. Initial attempts to measure T-cell function in SLE consisted of skin tests with common antigens for a delayed-type hypersensitivity reaction. These crude assays showed a diminished reactivity to common antigens that was also present with *in vitro* assays of antigen responsiveness (Horwitz, 1972). More recently, studies of T-cell function in SLE have centered around the apparent loss of suppressor-cell activity and its correlation with elevated levels of lymphocytotoxic antibodies and increased disease activity (Williams, 1982). Also exciting are the findings of ANA cross-reactive with lymphocyte membranes or reactive with membrane-bound nuclear antigens such as DNA or histones (Searles *et al.*, 1979; Rekvig and Hannestad, 1979). Other areas of defective cellular immune function include work showing diminished autologous mixed lymphocyte reactivity (Sakane *et al.*, 1978) and diminished natural killer cell activity (Gato *et al.*, 1980). The significance of these findings has yet to be resolved.

B. Mechanisms of Damage

A major mechanism of tissue injury in SLE is mediated by immune complexes, and the damage can be caused either by immune complexes formed exogenously and trapped in various tissues or by immune complexes formed endogenously *in situ* in tissue by the combination of fixed antigen and circulating antibody (Koffler *et al.*, 1971; McCluskey, 1982). The latter (if fixed antigen is the problem) may be either a structural part of the tissue or an antigen deposited and avidly bound by that tissue. In the renal lesion of SLE, immune complexes are seen in various locations throughout the glomerulus as well as in the tubulointerstitial areas of the kidney as a whole. Activation of complement by the classical pathway is supported by the presence of C1q, C4, and C3 in these areas also. Although it appears clear that damage is correlated with immune complex deposition and complement activation, little is known about the actual way in which immune complexes are localized. There is evidence that complexes are actually trapped initially in the mesangial area, but with chronic deposition may exceed the removal capacity of the mesangium and spill over into subendothelial and subepithelial locations (Kunkel, 1980). In support of this, it has been shown that subendothelial deposits are very uncommonly seen without coexistent mesangial deposits (Hill *et al.*, 1978). Little doubt exists concerning the composition of the majority of immune complex components in lupus nephritis. As shown by the work

of Koffler and associates (1974), high concentrations of anti-DNA can be eluted from kidney specimens and direct staining of renal tissue with fluorescent-tagged anti-DNA shows tissue-bound DNA (Andres *et al.*, 1970). Izui (1977) has evidence that suggests that some of these complexes may be formed *in situ*. Formation *in situ* may be the primary mechanism for the occurrence of membranous lupus glomerulonephritis (McCluskey, 1982). In addition to anti-DNA, other ANA have been eluted from the renal tissue of patients with lupus nephritis. These include antibodies to Sm, nRNP, SSA/Ro, and ssDNA (Winfield *et al.*, 1975; Maddison and Reichlin, 1979).

C. Immunogenetics

In clinical studies of the relationship of genetic factors to SLE, the strongest evidence for the importance of such factors has come from studies of SLE in twin pairs. Although studies of families (Siegel and Lee, 1968; Masi, 1968) and of household contacts (Cleland *et al.*, 1978) provided somewhat disparate conclusions, studies of mono- and dizygotic twin pairs with SLE emphasized the importance of genetic factors in the etiology of SLE (Block *et al.*, 1975). One especially convincing report described the onset of SLE in twins who were separated at 16 months of age and yet both developed the disease only 1 month apart some 13 years later (Block *et al.*, 1975). Overall the concordance rate in monozygotic twins is 70%, suggesting some role still exists for environmental factors. On a more basic level, evidence is also available for the influence of genetic factors on the expression of disease. For instance, the prevalence of ANA and hypergammaglobulinemia in first-degree relatives of the twins studied above was shown to be increased 28 and 33%, respectively. Concordance rates for the same serological factors in the twins were 71 and 87%, respectively. In a study by Miller and Schwartz (1979), a defect in suppressor-cell function was demonstrated to be more common in lupus relatives than in controls.

Early reports of histocompatibility typing were not able to provide strong evidence for a common HLA genotype (Goldberg *et al.*, 1973; Grumet *et al.*, 1971; Stastny, 1972; Arnett *et al.*, 1972; Waters *et al.*, 1971; Nies *et al.*, 1974). More recently, the work of Gibofsky *et al.* (1978) provided evidence for a link between SLE and the combined B-cell alloantigens HLA-DRw2 and -DRw3. Other investigators suggested that the risk of SLE increases some 16 times if both these alloantigens and reactivity of B cells to alloantiserum Ia-715 are present (Reinertsen *et al.*, 1980). Further support for the role of DRw3 has come from studies at the Mayo Clinic that showed that the presence of this B-cell alloantigen is

closely associated with the ability of patients to make antibodies to native DNA (Griffing *et al.*, 1980). These data suggested that control for autoantibody production may be related to an immune-response gene, but that the disease expression ultimately depends on other factors, both genetic and environmental, as well.

D. Laboratory Diagnosis

Although various important ANA in SLE have been discussed, little mention has been made of methods for their detection. The most sensitive test for ANA is the indirect immunofluorescent technique using a polyvalent fluorescein-tagged anti-Ig. For many years tissue sections from either mouse or rat kidney have been used as substrate for the detection of ANA. More recently, various tissue culture cell lines, especially the HEp2 cell line, have come into common usage (Fig. 6). These latter substrates tend to be more appropriate for some nuclear antigens such as SSA/Ro due to higher concentrations of these antigens in HEp2 cell lines (Harmon *et al.*, 1984). In addition, larger nuclei and nucleoli allow for earlier recognition of some patterns, and significant numbers (2–4%) of mitotic cells allow detection of antibodies to chromosomal binding proteins. Without the use of such substrates, antibodies to centromere (Moroi *et al.*, 1980) proliferating cell nuclear antigen (Miyachi *et al.*, 1978) and rheumatoid arthritis nuclear antigen (Alspaugh and Tan, 1975, 1976) would not have been recognized.

Once a positive fluorescent ANA has been detected, further investigation should continue to define the specificity of the ANA. A number of ANA may be detected by the Ouchterlony immunodiffusion technique. Using protype operationally monospecific antisera (some of which are now available from the Centers for Disease Control in Atlanta) (Tan *et al.*, 1982), antibodies to Sm, U1 RNP, SSA/Ro, and SSB/La may be specifically identified when placed against appropriate tissue or cell extracts. Antibodies to DNA may also be detected by this technique, but only when present in high titer. For this reason other techniques such as the Farr assay and the Millipore filter binding assay (Ginsberg and Kaiser, 1973) have become available. If *Crithidia luciliae* is used as substrate in an immunofluorescent assay, then anti-dsDNA may be specifically detected by reaction with the organism's kinetoplast (Aarden *et al.*, 1975). Antihistone antibodies (as previously alluded to) may be detected in procainamide-induced SLE by a reconstitution assay developed by Tan and co-workers (1976). An improvement using ELISA has been developed by Rubin *et al.* (1982) that increases the sensitivity of detection and facilitates identification of antibody to histones H1, H2A, H2B, H3, and H4.

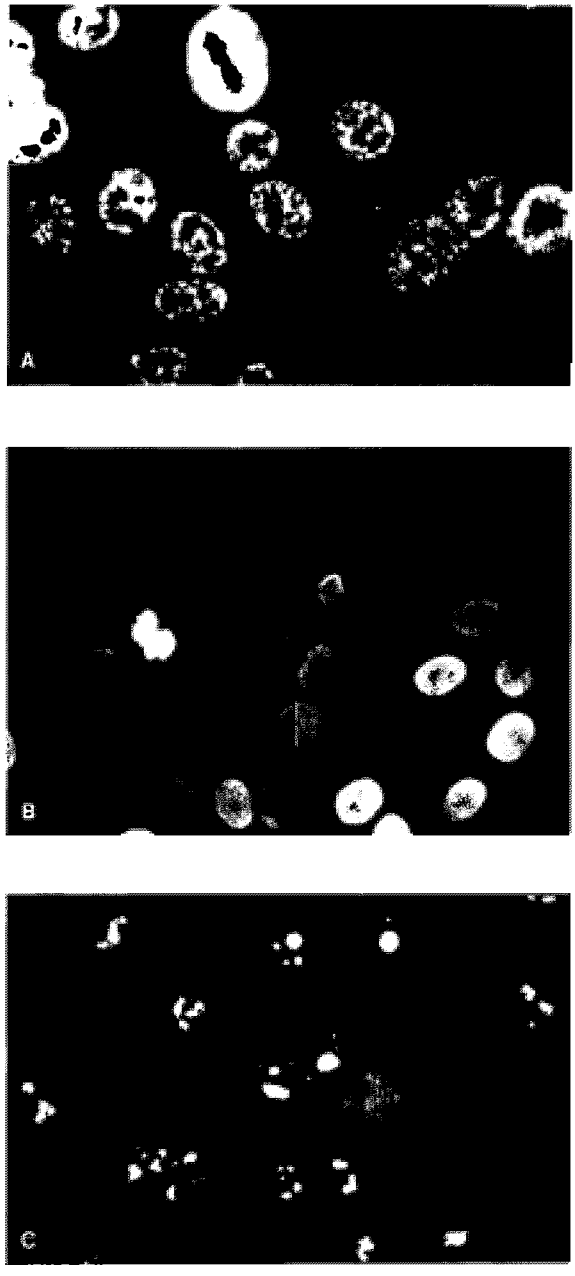


FIG. 6. (A) Speckled nuclear staining produced by serum with antibody to nuclear RNP (U1 RNP) on the HEP-2 cell line. Note that nucleoli were completely unstained (900×). (B) Homogeneous nuclear staining produced by serum from a patient with SLE in indirect immunofluorescence on HEP-2. (C) A pattern of nucleolar staining demonstrated by sera from some patients with scleroderma on HEP-2.

TABLE VII
ANA Profiles in Certain Autoimmune Diseases

Antibodies to	SLE	MCTD	Sjögren's syndrome	Scleroderma	Dermato-polymyositis
dsDNA	60–70	<5[a]	<5	<5	<5
ssDNA	60–70	10–20	10–20	10–20	10–20
Histones	70	<5	<5	<5	<5
Sm	30–40	<5	<5	<5	<5
U1-RNP	35–45	95–100	<5	20	<5
SS-A/Ro	30–40	<5	60–70	<5	<5
SS-B/La	15	<5	45–60	<5	<5
Scl-70	<5	<5	<5	20–30	<5
Centromere/ kinetochore	<5	<5	<5	25–30	<5
Nucleolar antigen	<5	<5	5–10	50–60	5–10

[a] <5% frequency is used to denote the fact that the antibody is rarely observed. (All data are percentages.)

Armed with these various serological techniques, one can better approach the diagnoses of connective tissue disorders. As seen in Table VII, profiles of ANA can be most useful when trying to distinguish among these disorders. For example, SLE is characterized by a number of different ANA, whereas scleroderma, Sjögren's syndrome, dermatomyositis, and MCTD exhibit more narrow specificities. Diagnosis can often be assured in some cases with presence of a single marker antibody such as anti-Sm in SLE or anti-Scl-70 in scleroderma.

VI. TREATMENT

Treatment of SLE should be carried out in a logical fashion, matching increasingly more severe clinical manifestations with more potent, although unfortunately more toxic, strategies. The mainstay of treatment is undoubtedly corticosteroids, but it is not appropriate to use these without a global therapeutic strategy. As discussed extensively in earlier sections of this chapter, SLE is a prototypic autoimmune disease with many measurable immune abnormalities manifested in part as ANAs, immune complexes, and complement consumption. While laboratory tests should not be used as dictators of therapy, they should be used to supplement clinical observations and as guides to therapy. In this context, titers of ANA, anti-DNA, immune complexes, and complement are most useful.

In the patient with mild arthralgias or arthritis, minimal constitutional symptoms and mild skin lesions associated with a positive ANA but no other serological abnormalities, treatment with aspirin, another nonsteroidal antiinflammatory drug, or hydroxychloroquin may be sufficient to control the disease. If symptoms are not controlled by these measures or there is evidence of worsening serological abnormalities, treatment with corticosteroids is indicated. Prednisone produces a prompt response, usually within 24 to 48 h. Doses should be tapered as rapidly as possible as long as symptoms can be suppressed. Alternate-day corticosteroids may be adequate for some persons. More severe manifestations such as renal or CNS involvement ordinarily indicate higher doses of corticosteroids (1 mg/kg). Life-threatening manifestations not controlled by corticosteroids call for the use of cytotoxic drugs such as cyclophosphamide and azathioprine (Cameron *et al.*, 1970; Cade *et al.*, 1973; Shelp *et al.*, 1971). Apheresis procedures aimed at removing circulating antibodies and immune complexes are new modalities of treatment, but the final evaluation of their therapeutic efficacy is not available (Huston *et al.*, 1981; Wei *et al.*, 1981). However, with our increased understanding of the pathophysiology of SLE and with the development of rational approaches to therapy, significant advances have already been made in control of disease activity and in long-term survival of patients suffering from this disease (Decker, 1982).

REFERENCES

Aarden, L. A., deGroot, E. R., and Feltkamp, T. E. W. (1975). *Ann. N. Y. Acad. Sci.* **254,** 505.

Alspaugh, M. A., and Tan, E. M. (1975). *J. Clin. Invest.* **55,** 1067.

Alspaugh, M. A., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 711.

Alspaugh, M. A., and Maddison, P. (1979). *Arthritis Rheum.* **22,** 796.

Andres, R. A., Accini, L., Berser, S. M., Christian, C. L., Cinotti, G. A., Erlanger, B. F., Hsu, K. C., and Seegal, B. C. (1970). *J. Clin. Invest.* **49,** 2106.

Arnett, F. C., Bias, W. B., and Shulman, L. E. (1972). *Arthritis Rheum.* **15,** 428.

Atkins, C. J., Kondo, J. J., Quismorio, F. P., and Friou, G. J. (1972). *Ann. Intern. Med.* **76,** 65.

Baldwin, D. S., Lowenstein, J., Rothfield, N. F., Gallo, G., and McCluskey, R. T. (1970). *Ann. Intern. Med.* **73,** 929.

Batchelor, J. R., Welsh, K. I., Tinoro, R. M., Dollery, C. T., Hughes, G. R. U., Bernstein, R., Ryan, P., Narsh, P. F., Aber, G. M., Bing, R. F., and Russell, G. I. (1980). *Lancet* **i,** 1107.

Bell, D. A., and Maddison, P. J. (1980). *Arthritis Rheum.* **23,** 1268.

Block, S. R., Winfield, J. B., Lockshin, M. D., *et al.* (1975). *Ann. J. Med.* **59,** 533.

Boehner, R. F., Altmann, J. C., Gorman, J. G., Ferbought, M., and Scudder, J. (1968). *Ann. Intern. Med.* **68,** 19.

Budman, D. R., and Steinberg, A. D. (1977). *Ann. Intern. Med.* **86,** 220.
Burnham, T. K., Nebbitt, T. R., and Fine, G. (1963). *J. Invest. Dermatol.* **41,** 451.
Cade, R., Spooner, G., Schleurs, E., *et al.* (1973). *Nephron* **10,** 37.
Cameron, J. S., Boulton-Jones, M., Robinson, R., and Ogg, C. (1970). *Lancet* **ii,** 846.
Clark, G., Reichlin, M., and Tomasi, T. B. (1969). *J. Immunol.* **102,** 117.
Cleland, L. G., Bell, D. A., Willians, M., *et al.* (1978). *Arthritis Rheum.* **21,** 183.
Cohen, A. S., Reynolds, W. E., Franklin, E. C., Kulka, J. P., Ropes, M. W., Shulman, L. E., and Wallace, S. L. (1971). *Arthritis Rheum.* **21,** 643.
Cruickshank, B. (1974). *In* "Lupus Erythematosus" (E. L. Dubois, ed.) Univ. of California Press, Los Angeles.
Datta, S. K., Manny, N., Andrzejewski, C., Andre-Schwartz, J., and Schwarz, R. S. (1978). *J. Exp. Med.* **147,** 854.
Davis, J. B., Godfrey, S. M., and Winfield, J. B. (1978). *Arthritis Rheum.* **21,** 17.
Decker, J. L. (1982). *Arthritis Rheum.* **25,** 891.
Decker, J. L., Steinberg, A. D., Reinertsen, J. L., Plotz, P. H., Balow, J. E., and Klippel, J. H. (1979). *Ann. Intern. Med.* **91,** 587.
Dixon, F. J. (1982). *Hosp. Practice March* p. 63.
Dixon, F. J., Dedstone, M. B. A., and Tonietti, G. (1971). *J. Exp. Med.* **134,** 65S.
Eisenberg, R. A., Tan, E. M., and Dixon, F. J. (1978). *J. Exp. Med.* **147,** 582.
Feinglass, E. J., Arnett, F. C., Dorsch, C. A., Zizic, T. M., and Stevens, M. C. (1976). *Medicine (Baltimore)* **55,** 323.
Fessel, W. J. (1974). *Arch Int. Med* **134,** 1027.
Franco, H. L., Weston, W. L., Peebles, C. P., Forstot, L. J., and Phanuphak, P. (1981) *J. Am. Acad. Dermatol.* **4,** 67.
Fritzler, M. J., and Tan, E. M. (1978). *J. Clin. Invest.* **62,** 560.
Gato, M., Tanimoto, K., and Horuichi, Y. (1980). *Arthritis Rheum.* **23,** 1274.
Gibofsky, A., Winchester, R. J., and Patarryo, M. (1978). *J. Exp. Med.* **146,** 1725.
Gilliam, H., Lang, D., and Lo Spalluto, J. J. (1980). *J. Immunol.* **125,** 874.
Gilliam, J. N., and Sontheimer, R. D. (1982). *Clin. Rheum. Dis.* **8,** 207.
Ginsberg, B., and Kaiser, H. (1973). *Arthritis Rheum.* **16,** 199.
Goldberg, M. A., Arnett, F. C., Bias, W. B., and Shulman, L. E. (1973). *Arthritis Rheum.* **16,** 546.
Griffing, W. L., Moore, S. B., Luthra, H. S., McKenna, C. H., and Fathman, C. G. (1980). *J. Exp. Med.* **752,** 319S.
Grumet, F. C., Cookell, A., Bodmer, J. G., Bodmer, W. F., and McDevitt, H. O. (1971). *N. Engl. J. Med.* **285,** 193.
Hahn, B. (1980). *In* "Clinical Immunology" (C. W. Parker, ed.), Vol. I, p. 584. Saunders, Philadelphia, Pennsylvania.
Hahn, B. (1982). *Arthritis Rheum.* **25,** 747.
Hargraves, M. M., Richmond, H., and Morton, R (1948). *Mayo Clin. Proc.* **23,** 25.
Harmon, C. H., and Portanova, J. P. (1982). *Clin. Rheum. Dis.* **8,** 121.
Harmon, C. H., Deng, J. S., Peebles, C. L., and Tan, E. M. (1984). *Arthritis Rheum.* **27,** 166.
Hill, G. S., Hingoars, N., Tron, F., and Bach, J. F. (1978). *Am. J. Med.* **64,** 65.
Holman, H. R., and Deicher, H. R. (1959). *J. Clin. Invest.* **38,** 2059.
Holman, H. R., and Kunkel, H. G. (1957). *Science* **126,** 162.
Horwitz, D. A. (1972). *Arthritis Rheum.* **15,** 353.
Howie, J. B., and Helyer, B. J. (1968). *Adv. Immunol.* **9,** 215.
Hurd, E. R., Eigenbrodt, E., Ziff, M., and Strunk, S. W. (1969). *Arthritis Rheum.* **12,** 541.
Huston, D. P., Wei, N., Klewa, E. N., *et al.* (1981). *Arthritis Rheum.* **24,** S92.

Izui, S., Lambert, P. H., Fournie, G. J., Turler, H., and Miescher, P. A. (1977) *J. Exp. Med.* **145,** 1115.
Izui, S., McConahey, P. J., Theofilopoulos, A. N., and Dixon, F. J. (1979). *J. Exp. Med.* **149,** 1099.
Koffler, D. (1974). *Annu. Rev. Med.* **25,** 149.
Koffler, D., Agnello, V., Thoburn, R., and Kunkel, H. G. (1971). *J. Exp. Med.* **134,** 169S.
Koller, L. R. (1965). *In* "Ultraviolet Radiation." Wiley, New York.
Kunkel, H. G. (1980). *Hosp. Pract.*
Kurland, L. T., Hauser, W. A., Ferguson, R. H., and Holley, K. E. (1969). *Mayo Clin. Proc.* **44,** 649.
Lampert, P. W., and Oldstone, M. B. A. (1974). *Virchows Arch. A: Pathol. Anat. Histol.* **363,** 21.
Laurell, A. B., and Nilsson, I. M. (1957). *J. Lab. Clin. Med.* **52,** 1871.
Lerner, M. R., and Steitz, J. A. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **71,** 685.
Libman, E. and Sacks, B. (1924). *Arch Int. Med.* **33,** 701.
McCluskey, R. T. (1982). *Arthritis Rheum.* **25,** 867.
Maddison, P. J. (1982). *Clin. Rheum. Dis.* **8,** 105.
Maddison, P. J., and Reichlin, M. (1979). *Arthritis Rheum.* **22,** 858.
Masi, A. T. (1968). *In* "Population Studies of the Rheumatic Diseases" (P. H. Bennett and P. H. N. Wood, eds.), p. 267. Excerpta Med. Found., Amsterdam.
Matthay, R. A., Schwartz, M. I., Petty, T. L., Stanford, R. E., Ceupta, R. C., Sohn, S. A., and Steigerwald, J. C. (1974). *Medicine (Baltimore)* 397.
Mattioli, M., and Reichlin, M. (1974). *Arthritis Rheum.* **17,** 421.
Miescher, P., and Strassle, R. (1957). *Vox Sanguinis* **2,** 283.
Miller, K. B., and Schwartz, R. S. (1979). *N. Eng J. Med.* **301,** 803.
Miyachi, K., Fritzler, M. J., and Tan, E. M. (1978). *J. Immunol.* **121,** 2228.
Moroi, Y., Peebles, C. P., Fritzler, M. J., Steigerwald, J. C., and Tan, E. M. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 627.
Murphy, E. D., and Roths, J. B. (1979). *In* "Genetic Control of Autoimmune Disease" (N. R. Rose, P. E. Bigassi, and N. L. Warner, eds.) Elsevier North-Holland, New York.
Nies, K. M., Brown, J. C., Dubois, E. L., Quismorio, F. P., Friou, G. J., and Terasaki, P. I. (1974). *Arthritis Rheum.* **17,** 397.
Notman *et al.* (1975). *Ann. Intern. Med.* **83,** 464.
Pekin, T. J., and Zvaifler, N. J. (1970). *Arthritis Rheum.* **13,** 777.
Perry, H. M., Tan, E. M., Carmody, S., and Nakamoto, A. (1970). *J. Lab. Clin. Med.* **76,** 114.
Phillips, P. E., and Christian, C. L. (1973). *Ann. Rheum. Dis.* **32,** 450.
Pincus, T. (1982). *Arthritis Rheum.* **25,** 847.
Pincus, T., Blacklow, N. R., and Grumley, P. M. (1970). *Lancet* **ii,** 1058.
Portanova, J. P., Rubin, R. L., Joslin, F. G., and Tan, E. M. (1982). *Clin. Immunol. Immunopathol.* **25,** 67.
Reinertsen, J. L., Klippel, J. H., Johnson, A. H., Steinberg, A. D., Decker, J. L., and Mann, D. L. (1980). *N. Engl. J. Med.* **299,** 515.
Rekvig, O. P., and Hannestad, K. (1979). *J. Immunol.* **123,** 2673.
Rick, M. E., and Hoyer, L. W. (1975). *Clin. Rheum. Dis.* **1,** 583.
Robbins, W. C., Holman, H. R., Deicher, H. R., and Kunkel, H. G. (1957). *Proc. Soc. Exp. Biol. Med.* **96,** 575.
Ropes, M. W. (1976). "Systemic Lupus Erythematosus." Harvard Univ Press, Porters, Massachusetts.
Rothfield, N. (1982). *In* "Textbook of Rheumatology" (W. N. Kelley, E. D. Harris, S. Ruddy, and C. Sledge, eds.). Saunders, Philadelphia, Pennsylvania.

Rubin, R. L., Joslin, F. G., and Tan, E. M. (1982). *Arth. Rheum.* **25,** 779.
Russel, A. S., Percz, J. S., Rigal, W. M., and Wilson, G. L. (1981). *Ann. Rheum. Dis.* **33,** 204.
Sakane, T., Steinberg, A. D., and Green, I. (1978). *Proc. Natl. Acad. Sci. U.S.A.* **75,** 3464.
Schrager, M. A., and Rothfield (1976). *J. Clin. Invest.* **57,** 212.
Scott, J. S., Maddison, P. J., Taylor, P. V., Escher, E., Scott, O., and Skinner, R. P. (1983). *N. Engl. J. Med.* **309,** 209.
Searles, R. P., Messner, R. P., and Bankhurst, A. D. (1979). *Clin. Immunol. Immunopathol.* **14,** 292.
Seligmann, M. (1975). *C. R. Hebd. Seances Acad. Sci. (Paris)* **245,** 243.
Sharp, G. C., Irwin, W. S., Tan, E. M., Gould, R. C., and Holman, H. R. (1972). *Am. J. Med.* **52,** 148.
Shelp, W. D., Bloodworth, J. M. P., and Bieselbach, R. E. (1971). *Arch. Intern. Med.* **128,** 566.
Siegel, M., and Lee, S. L. (1968). *In* "Population Studies of the Rheumatic Diseases" (P. H. Bennett and P. H. N. Wood, eds.), p. 245. Exceptia Med. Found., Amsterdam.
Siegel, M., and Lee, S. L. (1975). *Semin. Arthritis Rheum.* **3,** 1.
Siegel, M., Holley, H. L., and Lee, S. L. (1970). *Arthritis Rheum.* **13,** 802.
Stastny, P. (1972). *Arthritis Rheum.* **15,** 455.
Swaak, A. J. G., Aarden, L. A., Status Van Eps, L. W., and Feltkamp, T. E. W. (1979). *Arthritis Rheum.* **22,** 226.
Swaak, A. J. G., Groenwold, J., Parder, L. A., Status Van Eps, L. W., and Feltkamp, T. E. W. (1982). *Ann. Rheum. Dis.* **41,** 388.
Talal, N., and Steinberg, A. D. (1974). *Curr. Top. Microbiol. Immunol.* **64,** 79.
Talbott, J. H. (1974). *In* "Lupus Erythematosus" (E. L. Dubois, ed.), p. 1. Univ. of Southern California Press, Los Angeles, California.
Tan, E. M. (1982). *Adv. Immunol.* **33,** 167.
Tan, E. M., and Kunkel, H. G. (1966). *Arthritis Rheum.* **9,** 37.
Tan, E. M., and Stoughton, R. B. (1969). *J. Invest. Dermatol.* **52,** 537.
Tan, E. M. *et al.* (1966). *J. Clin. Invest.* **45,** 1732.
Tan, E. M., Robinson, J., and Robitaille, P. (1976). *Scand. J. Immunol.* **5,** 811.
Tan, E. M., Cohen, A. S., Fries, J. F., Masi, A. T., McShane, D. J., Rothfield, N. F., Schaller, J. G., Talal, N., Winchester, R. J. (1982). *Arthritis Rheum.* **25,** 1271.
Theofilopoulos, A. N., and Dixon, F. J. (1981). *Adv. Immunol.* **55,** 179.
Waters, H., Konrad, P., and Walford, R. L. (1971). *Tissue Antigens* **1,** 68.
Wei, N., Huston, D. P., Klippel, J. H. *et al.* (1981). *Arthritis Rheum.* **24,** S107.
Williams, R. C. (1982). *Arthritis Rheum.* **25,** 810.
Winfield, J. B., Koffler, D., and Kunkel, H. G. (1975). *Arthritis Rheum.* **18,** 531.
Woosley, R. L., Drayer, D. E., Reidenberg, M. M., Nies, A. S., Corr, K., and Oates, J. A. (1978). *N. Engl. J. Med.* **298,** 1157.
Yang, V. W., Lerner, M. R., Steitz, J. A., and Flint, S. J. (1981). *Proc. Natl. Acad. Sci. (USA)* **78,** 1371.

CHAPTER 3

Autoimmune Aspects of Rheumatoid Arthritis

MORRIS ZIFF

Department of Internal Medicine
The University of Texas Health Science Center at Dallas
Dallas, Texas

I. INTRODUCTION

In considering the entity of adult rheumatoid arthritis (RA), one must keep in mind the possibility that the disease which has traditionally gone by this name may be composed of more than one condition. The reason

for this caution is the observation that patients with seronegative (rheumatoid factor-negative) RA, aside from their milder clinical course and low incidence of systemic complications, do not demonstrate the aggregation of the HLA-DRw4 tissue marker that is characteristic of seropositive disease (Doubloug *et al.*, 1980; Alarcon *et al.*, 1982). The information presented in this chapter pertains largely to seropositive RA.

Rheumatoid arthritis shares a group of autoantibodies with the other connective tissue diseases. However, like the other members of the group, it is characterized by a selective increase in frequency and titer of one of these autoantibodies. Most characteristic of RA are the anti-IgG autoantibodies or rheumatoid factors (RFs). A minority of patients also have antinuclear and anti-single-stranded DNA antibodies. It should be recognized, however, that autoantibody formation and immune complex phenomena constitute only one aspect of a series of immunological reactions that take place in the rheumatoid joint. Perhaps more basic to rheumatoid synovitis is the cellular immune response in the synovial membrane itself, which leads to the synthesis of immunoglobulins and lymphokines that mediate the immunologically stimulated inflammation that takes place.

II. HLA-D REGION AND THE GENETICS BACKGROUND

Because of the tendency of RA to aggregate in certain families, there has long been an impression that there is a genetic basis for its development. Although population studies have been contradictory (Bunim *et al.*, 1964; Masi and Shulman, 1965; Lawrence, 1970), familial aggregation of the rheumatoid factor in the serum of asymptomatic relatives has been demonstrated (Ziff *et al.*, 1958; Ball and Lawrence, 1961), and HLA typing has established that there is a genetic basis for the disease. HLA-Dw4 was found in 59% of seropositive Caucasian patients with RA and in 16% of normal controls (Stastny, 1977).

HLA-D locus typing, as carried out above, is performed using typing cells in a mixed lymphocyte reaction. *HLA-DR* locus typing utilizes specific alloantibody to induce complement-mediated cytotoxicity directed at host B cells. The B-cell alloantigen HLA-DR4 was found in 65% of seropositive patients with RA and in 28% of normal controls (Stastny, 1981). There was no association with the *HLA-A* or *HLA-B* loci. The aggregation of the *HLA-D* locus does not extend to juvenile RA, emphasizing again the differences between the adult and juvenile diseases.

III. SERUM AUTOANTIBODIES

A. The Rheumatoid Factors

A role for autoimmunity in RA was first indicated by the presence of RF or anti-IgG in the serum of rheumatoid patients. Antiimmunoglobulin activity was first noted by Waaler (1940) and Rose *et al.* (1948), who found that the serum of high proportion of patients with RA agglutinated sheep erythrocytes sensitized with rabbit antibody. Although it is now known that anti-IgG antibodies occur in patients with a variety of conditions other than RA, the term rheumatoid factor has continued to be used for this type of autoantibody. The RFs are, in fact, a group of autoantibodies that are directed against the Fc fragment of the heavy chain of IgG (Fig. 1). The antigenic sites in the Fc fragment against which IgM-RF is directed have been identified, and these are in the second ($C\gamma2$) and third ($C\gamma3$) constant homology regions that constitute the Fc fragment (Natvig *et al.*, 1972).

B. Multiplicity of Rheumatoid Factors

For many years, it was believed that there was only one RF and that this was present only in the IgM class. This supposition arose because the detection and assay of the RFs was carried out by agglutination methods such as the latex fixation and sensitized sheep cell agglutination (SSCA) procedures, which are sensitive in the detection of IgM immunoglobulin antibodies. IgG-RF was first demonstrated in certain sera when intermediate complexes of a molecular weight between IgG and IgM, which

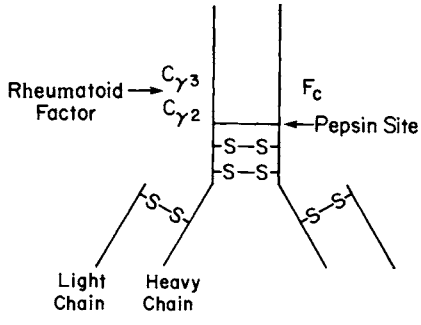


FIG. 1. The antigenic sites in the Fc fragment against which IgM-RF is directed are in the $C\gamma2$ and $C\gamma3$ constant homology regions that constitute the Fc fragment.

dissociated in acid to IgG molecules, were found on ultracentrifugation of certain sera (Kunkel *et al.*, 1961). Subsequently, IgA-RF was isolated (Taylor and Abraham, 1973), and recently IgE-RF (Zuraw *et al.*, 1981) and IgD-RF (Pope *et al.*, 1982) were demonstrated in the serum of a majority of patients with seropositive RA.

Not only can RFs exist in the different immunoglobulin classes, but they also combine with IgG of various species. RF has been widely assayed by the SSCA method, which employs rabbit antibody to sensitize the sheep cell. A factor reacting with human IgG and another reacting with both human and rabbit IgG have been described (LoSpalluto and Ziff, 1959; Williams and Kunkel, 1963; Milgrom *et al.*, 1962). As a general rule, the sera of RA patients show positive reactions in both the latex fixation test, which employs human IgG for sensitization, and the SSCA test, which employs rabbit IgG. The human gamma-globulin-sensitized latex agglutination test was positive in 88% of patients with definite and classic RA, and in 73% of patients with possible or probable disease (Waller *et al.*, 1961). The rabbit gamma globulin SSCA test was less sensitive, the corresponding figures being 72% in definite or classical patients and 36% in those possible or probable. Values in healthy donors were of the order of 3–4% with both procedures.

Sera of patients with hypergammaglobulinemic states like viral hepatitis and chronic liver disease (Bonomo *et al.*, 1963), syphilis (Peltier, 1959), and sarcoidosis (Kunkel *et al.*, 1958), which are frequently seropositive, tend to react mainly in test systems employing human gamma globulin for sensitization. It should be pointed out, however, that sera of patients with pulmonary fibrosis (Tomasi *et al.*, 1962), leprosy (Cathcart *et al.*, 1961), and subacute bacterial endocarditis (Williams and Kunkel, 1962) tend to react in both the latex and SSCA procedures.

C. Specificity of Rheumatoid Factors

Since the RFs appear to function as autoantibodies against IgG, it has been commonly assumed that they are produced as a result of autoimmunization by the patient's own IgG in altered form, this alteration resulting perhaps from reaction of the IgG with an exogenous antigen. There are a number of reports demonstrating that heat-aggregated IgG has the capacity to stimulate peripheral blood mononuclear cells (PBMC) of patients with RA to undergo blastic transformation (Pisko *et al.*, 1982), stimulate the synthesis of immunoglobulin (Yamasaki and Ziff, 1977), and generate the lymphokine MIF (Kinsella, 1974). However, these may represent nonspecific polyclonal responses to aggregated gamma globulin.

The association of seropositivity with chronic inflammatory states and with chronic infections suggests that the RFs are induced by a chronic immunizing process. RF has, in fact, has been induced in the rabbit by prolonged immunization with *Escherichia coli* (Abruzzo and Christian, 1961). A striking example of such a process in man is the development of RFs in the course of subacute bacterial endocarditis (Williams and Kunkel, 1962). In this condition, IgG-RF and IgM-RF appear in the serum, reaching a peak following the development of maximum levels of circulating immune complexes and suggesting that the IgG in the immune complexes stimulates the synthesis of the RFs. With treatment, the titers of the RFs decrease, but more slowly than do the levels of the complexes (Carson *et al.*, 1978).

Certain IgM-RFs show serological specificity for genetically defined determinants on IgG that are not present on the host IgG but are present on IgG of normal individuals (Fudenberg and Kunkel, 1961). This and the fact that rheumatoid patients have RFs that react also with the IgG of other species indicate that RFs have the capacity to cross-react relatively widely. How wide this cross-reactivity may be is emphasized by the recent observation of Hannestad (1978) that some RFs also react with nuclear antigens. Hannestad observed that certain purified RFs gave positive antinuclear antibody reactions. Approximately 40% of sera from 62 cases of seropositive classical RA contained such cross-reacting RFs. Thus, many patients have a RF subset that reacts with both the Fc portion of IgG and an antigenic site on the histone moiety of the nucleoprotein. Williams (1979) has suggested that the characteristic that the cross-reacting antigens may have in common is a configurational one, that is, a common structural shape, and Johnson and Faulk (1976) summarized various concepts for the formation of RFs. However, as with other autoantibodies, the true mechanism is at present not established.

D. Solid Phase Immunoassays for Rheumatoid Factor

The availability of solid phase immunoassays has created a new dimension in the measurement of autoantibodies. These assays have been extensively applied to the measurement of the RFs. IgD- and IgE-RFs have recently been demonstrated using this technique. By radioimmunoaasay, Wernick *et al.* (1981) found that the concentration of IgG-RF (439 $\pm$ 755 μg/ml) was greater than that of IgM-RF (175 $\pm$ 221 μg/ml). Values in normal adults were 4.1 $\pm$ 2.3 μg/ml for IgG-RF and 1.3 $\pm$ 0.96 μg/ml for IgM-RF. Normal values were observed in almost all patients with juvenile RA.

E. Other Autoantibodies

The presence of a number of autoantibodies in the serum suggests that there is polyclonal activation of B cells in RA. Antinuclear antibody (ANA) and anti-single-strand DNA (anti-ssDNA) are commonly present. Antinuclear antibody, measured by indirect immunofluorescence, was present in 35–40% of patients with RA (Garcia-de la Torre and Miranda-Mendez, 1982; Notman *et al.*, 1975). The staining pattern is usually diffuse, indicating that the antibody is directed against the histone of deoxyribonucleoprotein. This antibody is also responsible for the positive LE cell reactions that are observed in RA patients. Positive LE cell tests were seen in only 11% of patients (Fallet and Ziff, 1958).

The role of collagen in the rheumatic diseases has recently become a subject of interest. This arose with the work of Steffen *et al.* (1973), who observed that 45% of rheumatoid sera showed positive reactions with human collagen using a globulin consumption test. Michaeli and Fudenberg (1974) found collagen antibody in 60% of RA patients and only 9% of controls by a hemagglutination method. The occurrence of anticollagen antibody was independent of the presence or titer of RF. Moreover, patients with RA who had the highest levels of collagen antibodies had severe erosive disease (Clague *et al.*, 1979). Antibody to types I, II, and III collagens has been demonstrated (Andriopoulos *et al.*, 1975). However, immunofluorescent staining of rheumatoid synovial tissue with collagen antigens showed the presence of cells producing antibody only to type II collagen, which is restricted to hyaline cartilage. It has been argued from this finding that the cartilage serves as a local source of antigen to stimulate the synthesis of autoantibody in the synovial membrane. It is of interest that collagen–anticollagen immune complexes have been demonstrated in the cells of rheumatoid synovial effusions (Steffen *et al.*, 1974).

Cellular immunity to collagen has also been demonstrated in RA. The PBMC of most patients responded in culture to stimulation by types II and III collagen with the production of MIF (Trentham *et al.*, 1978). In contrast, cells from patients with other forms of arthritis did not respond. Also, it has been possible to induce arthritis in rats by injecting them intradermally with type II collagen, but not with other types (Trentham *et al.*, 1977). The PBMC of the arthritic rats responded to rat collagen with proliferation in culture, indicating that autoimmunity to collagen had been induced.

There are objections to the notion that immunity to collagen may be responsible for rheumatoid synovitis. About 40% of patients with RA do not have collagen antibody. Furthermore, collagen antibodies have been

observed in patients with ankylosing spondylitis (Clague *et al.*, 1979), leprosy (McAdam *et al.*, 1978), and relapsing polychondritis (Foidart *et al.*, 1978). It seems likely, therefore, that collagen antibody formation is a secondary phenomenon in the pathogenesis of RA, although the possibility that immunity to collagen plays a role in maintaining the synovitis cannot be ruled out.

Two other autoantibodies may be mentioned. A serum antiperinuclear factor has been identified in RA patients (Nienhuis and Mandema, 1964; Marmont *et al.*, 1967). The antiperinuclear factors are immunoglobulins that react with the keratohyaline granules of human oral mucosal cells. They have been demonstrated by immunofluorescent staining in 51% of patients with RA, rarely in SLE, and in isolated cases of Sjögren's syndrome and hyperglobulinemic purpura. Also, IgM antibodies against intermediate filaments (IMF) of the cytoskeleton have been demonstrated by immunofluorescent staining to be present in the serum of 80% of patients with seropositive RA (Osung *et al.*, 1982). These antibodies are also present in the synovial fluid. Intermediate-filament antibody has been identified with lesser frequency in patients with ankylosing spondylitis and osteoarthritis and in normal subjects.

IV. IMMUNE COMPLEX PHASE OF RHEUMATOID SYNOVITIS

A. Circulating Immune Complexes

Immune complexes occur in the serum of RA patients. It has been reported that they may be detected before it is possible to diagnose the disease (Jones *et al.*, 1981). Opinions have differed about the value of quantitating immune complex levels in RA serum. Although Roberts-Thomson *et al.* (1980) were able to discriminate 79% of rheumatoid patients from healthy blood donors by the Clq binding assay, others (McDougal *et al.*, 1982) found measurements of immune complex levels of little value in the diagnosis or management of RA. The various procedures used correlated poorly with each other and failed to reflect the activity of the disease.

Although nonspecific measurements of immune complex levels in the circulation have not yielded valuable information, investigation of the composition, properties, and complement-fixing activity of RF complexes has been of considerable interest. Immune complex levels in rheumatoid sera are correlated with IgG-RF and IgM-RF levels (Pope *et al.*, 1981).

Serum from patients with RA contains a complex of IgG with IgM-RF, which has a sedimentation coefficient of 22 S. This complex can be frequently detected in high-titer sera by ultracentrifugation (Franklin *et al.*, 1957). It is of low affinity and appears to have no important role in rheumatoid inflammation. It is possible, however, that this type of complex is the source of the cryoprecipitates of mixed cryoglobulinemia (LoSpalluto *et al.*, 1961).

IgG-RF tends to react preferentially with itself rather than with other molecules of autologous IgG to form dimers and larger molecules (Pope *et al.*, 1974). These, if present in sufficient concentration, may be seen on ultracentrifugation as intermediate complexes with sedimentation coefficients of 11 to 18 S (Kunkel *et al.*, 1961). Increased levels of these intermediate complexes have been observed in association with rheumatoid pulmonary interstital fibrosis (Tomasi *et al.*, 1962), Felty's syndrome (Andreis *et al.*, 1978), and the hyperviscosity syndrome that may accompany RA (Jasin *et al.*, 1970). Self-associated IgG-RF complexes have the capacity to react with IgM-RF to form molecules containing IgG-RF, IgM-RF, and C3. Such complexes have been phagocytosed from the serum of patients with Felty's syndrome on incubation with normal polymorphonuclear cells (Hurd *et al.*, 1977) to form stainable inclusions in these cells.

B. Immune Complexes in the Synovial Effusion

Although total hemolytic complement activity in RA serum is normal or slightly elevated, rheumatoid synovial effusions are uniquely characterized by low levels of complement (C′) components (Hedberg, 1964; Pekin and Zvaifler, 1964; Ruddy and Austen, 1970). The reasons for this decrease are several, but in the last analysis the trigger for the activation of C′ is the presence of RF complexes. IgG-RF tends to self-associate to form dimers and even larger molecules (Pope *et al.*, 1974, 1975), with the composition $(IgG\text{-}RF)_2$ and $(IgG\text{-}RF)_n$, respectively. These complexes dissociate at acid pH to IgG. They can be precipitated by Clq (Winchester *et al.*, 1970), and are able to fix C′ directly (Brown *et al.*, 1980). Indeed, Winchester and co-workers found that the decrease in C′ levels in synovial effusions was proportional to the level of the IgG-RF complexes in the effusion (Winchester *et al.*, 1970). The IgG-RF complexes can, in turn, react with IgM-RF and form larger complexes with the composition $(IgG)_n$-IgM-RF. These are also capable of activating C′ (Taylor-Upsahl *et al.*, 1977; Bianco *et al.*, 1974; Sabharwal *et al.*, 1982). As a result of the C′ activation, chemotactic factors like C5a, which attract large numbers of polymorphonuclear cells into the effusion, are generated.

Generation of C3a and C5a yields anaphylotoxic activity leading to the

accumulation of fluid in the joint space. The generation of C3b activates the alternative complement pathway (Ruddy and Austen, 1973). Moreover, the binding of C3 to the RF complexes yields molecules with the composition $(IgG)_n$-IgM-RF-C3, which are phagocytosed to yield inclusions in the polymorphonuclear cells of the synovial effusion and the phagocytic type A cells of the synovial lining layer (Fig. 2). These inclusions can be demonstrated to stain positively for IgG, IgM, and C3 by the immunofluorescent technique (Kinsella *et al.*, 1970; Tursi *et al.*, 1970; Hurd *et al.*, 1970). The uptake of these complexes leads to (1) release of lysosomal enzymes in the joint cavity that also participate in the activation of C3 and C5; (2) release of oxygen radicals and hydrogen peroxide, which leads to local cell injury and secondary inflammation; (3) synthesis and release from the granulocytes and lining cells of proinflammatory prostaglandins of the E series; and (4) synthesis and release of powerful leukocyte chemotactic factors like 5-hydroxyeicosatetraenoic acid (5-HETE) and leukotriene B4, which are produced in the cell by the lipoxygenase catalyzed oxidation of arachidonic acid.

It is clear from the above that the synovial effusion phase of rheumatoid synovitis is essentially an immune complex-induced exudative response that is similar to that which occurs in immune complex-induced vasculitis or in an Arthus reaction in the skin. Instead of taking place in a vessel wall, it occurs in the joint space, which is in free communication with the perivascular interstitial space surrounding the synovial vasculature. In this sense, RA partakes of the characteristics of an immune complex disease.

There is agreement that the levels of immune complexes are higher in the synovial fluid than in the serum (Halla *et al.*, 1979; Roberts-Thomson *et al.*, 1980). Cecere *et al.* (1982), moreover, have compared the IgG-RF, IgM-RF, and IgA-RF levels in the serum and synovial fluid with the levels

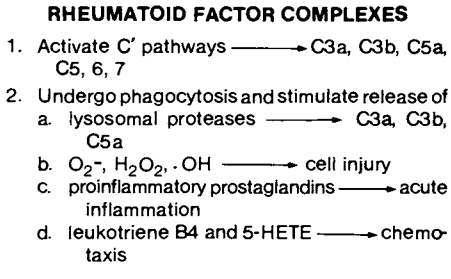


FIG. 2. Summary of pathways of acute inflammation initiated by rheumatoid factor complexes in the synovial fluid and extracellular space of the synovial membrane.

of the corresponding immunoglobulin classes in each of these sites. The percentages of IgM-RF and IgA-RF in their respective immunoglobulin classes were greater in the synovial fluid than in the serum. However, the percentage of IgG-RF relative to total IgG in the synovial fluid was not. This was interpreted to mean that although significant contributions of the three RFs are made to the synovial fluid by synovial membrane synthesis, this is not apparent in the case of IgG-RF because of preferential uptake of IgG-RF complexes relative to IgG by the rheumatoid inflammatory cells.

Male and Roitt (1981) isolated the C3-containing immune complexes of the synovial fluid of rheumatoid patients on conglutinin columns, and determined that they contained only immunoglobulins (presumably IgG-RF and IgM-RF) and activated components of the complement system. This result suggested that antigens other than IgG were not involved in the immune response in the synovial membrane at least to the extent of forming immune complexes in amounts sufficient to be detected. It should be added, however, that collagen–anticollagen immune complexes have been demonstrated in rheumatoid synovial cells by the staining of collagen and immunoglobulin in synovial cell inclusions (Steffen *et al.*, 1974). Also, cryoprecipitable complexes containing immunoglobulins, DNA, and antinuclear antibodies have been found in synovial effusions (Marcus and Townes, 1968; Barnett *et al.*, 1970; Marcus and Townes, 1971).

C. Immune Complexes in Rheumatoid Synovial Membrane

Evidence for the presence of immune complexes in rheumatoid synovial tissue is based mainly on fluorescent antibody staining of synovial cells. Positive staining has been considered evidence of the phagocytosis of immune complexes by these cells. Kinsella *et al.* (1969) demonstrated an IgG-C3 complex in the type A phagocytic synovial lining cells on the basis of combined staining of IgG and C3 in the cytoplasm of RA synovial cell suspensions following trypsin digestion of the synovial tissue; Hurd *et al.* (1971) also stained inclusions made up of IgG, IgM, and C3 in synovial lining cells in suspension. Tursi *et al.* (1970) made similar observations by staining intact synovial tissue sections. These immunofluorescent staining experiments indicated that the phagocytic or type A synovial lining cells phagocytose $(IgG\text{-}RF)_n$-C3 complexes and IgG-IgM-RF-C3 complexes. Munthe and Natvig (1971) have eluted similar complexes from rheumatoid synovial tissue. Finally, extensive staining of IgG, IgM, and C3 in the interstitial space of the synovium (Brandt *et al.*, 1968) suggested that RF complexes are deposited in the interstitium in large quantities. In view of

the extensive synthesis of immunoglobulins and RFs in the rheumatoid synovial membrane (Smiley *et al.*, 1968), it appears likely that these intracellular and extracellular complexes are synthesized locally.

D. Immune Complexes in Cartilage

Rheumatoid cartilage has been examined by the immunofluorescent technique at the light microscopic level (Cooke *et al.*, 1975) and by the immunoperoxidase technique at the electron microscopic level (Ishikawa *et al.*, 1975). By both methods, deposits of IgG, IgM, and RF have been stained in the superficial portions of the articular cartilage and meniscus, suggesting that RF-containing immune complexes are deposited in these tissues. Their location in the superficial portion of the cartilage suggests that the complexes or their constituents have diffused into the cartilage from the synovial space. These complexes may stimulate the growth of the pannus across the cartilage. It is also possible that they release antigen that stimulates the immunocompetent cells of the synovial membrane to propagate the synovial inflammatory reaction. The latter possibility is of interest in relation to the observation that joints that have been replaced by a mechanical prosthesis and are, therefore, free of cartilage, tend to be less subject than intact joints to exacerbation of rheumatoid synovitis (Lance, 1982).

V. CELLULAR PHASE OF RHEUMATOID SYNOVITIS

A. Cellular Changes

The sublining or deep layer of the rheumatoid synovial membrane is the site of a chronic inflammatory reaction. The mononuclear cells that infiltrate this layer are participants in an immune response in which T and B lymphocytes undergo blastic transformation, lymphokines and a monokine are secreted, and immunoglobulin is synthesized. As observed in the electron microscope, small lymphocytes emigrate from the postcapillary venules to form perivascular collections that consist mainly of these cells. These collections have been referred to as lymphocyte-rich areas (Kobayashi and Ziff, 1973; Ishikawa and Ziff, 1976); such areas may assume a nodular configuration. In neighboring regions, which have been called transitional areas, both T and B lymphocytes undergo blastic transformation and macrophages collect in increased numbers. These areas appear to be the sites of immunological stimulation of both T and B cells. In still a

third type of region, the cells are composed mainly of plasma cells. Such areas have been designated plasma-cell-rich areas.

In addition to the electron microscopic data, evidence of T-cell transformation in the sublining layer is seen in the presence of MIF and a blastogenic factor (Stastny *et al.*, 1975a,b) in rheumatoid synovial effusions. Interleukin-1 (IL-1), a macrophage secretory product, has also been demonstrated (Fontana *et al.*, 1982). These polypeptides diffuse into the synovial fluid through the easily permeable synovial lining layer.

The lymphocytes of the rheumatoid synovial effusion are predominantly T cells as in the blood. When the results of five studies were averaged, 66% of the lymphocytes in the blood and 71% of the lymphocytes in the synovial fluid were of the T-cell variety. Moreover, measurements made on collagenase digests of rheumatoid synovial tissue have demonstrated that the lymphocytes in the synovial membrane are also mainly T cells; the mean value of three studies was 77% (Van Boxel and Paget, 1975; Abrahamsen *et al.*, 1975; Bankhurst *et al.*, 1976). Thus the lymphocytes in all three compartments, blood, synovial effusion, and synovial membrane, are mainly T lymphocytes.

The proportions of helper and suppressor T lymphocytes in the three compartments has been a subject of great interest. These have been identified both by monoclonal antibody staining and by functional testing. Monoclonal antibodies that selectively react with human T cells (OKT3), helper/inducer cells (OKT4), and suppressor/cytotoxic cells (OKT8) have made it possible to differentiate these T-cell populations. An increased ratio of T4 to T8 cells in the peripheral blood of rheumatoid patients has been reported (Raeman *et al.*, 1981; Veys *et al.*, 1982a; Fox *et al.*, 1982). In contrast, a number of groups have observed decreased (Fox *et al.*, 1982; Veys *et al.*, 1982b; Lydyard *et al.*, 1982) or normal (Forre *et al.*, 1982) ratios of T4 to T8 cells, as compared to the blood, in rheumatoid synovial fluid (Fox *et al.*, 1982; Veys *et al.*, 1982b; Lydyard *et al.*, 1982) and synovial tissue digests (Burmester *et al.*, 1981). Utilizing either immunofluorescent (Janossy *et al.*, 1981; Duke *et al.*, 1982) or immunohistochemical staining techniques (Klareskog *et al.*, 1982; Meijer *et al.*, 1982) to stain intact synovial tissue, markedly increased numbers of T4 cells have been observed in the tissue. Most of these cells were described as being in close contact with HLA-DR or Ia-positive interdigitating dendritic cells, which are believed to present antigen to the surrounding T4 cells and in this way help to propagate the synovial inflammatory response.

Kurosaka and Ziff (1983) studied the distribution of lymphocyte subsets in the rheumatoid synovial membrane by an electron microscopic im-

munoperoxidase staining technique. In this technique, individual cells are identified from the staining of their cell membranes by peroxidase-labeled monoclonal antibody. At the same time, the cells with which the lymphocytes are in close contact are identified on the basis of their ultrastructure. It was observed that the percentage of T4 and T8 cells in a given area varied with the type of area examined. In the lymphocyte-rich collections, which consisted predominantly of small lymphocytes, most of the lymphocytes were T4 cells. The T4:T8 ratio was 3.0. In the transitional areas, where blastic transformation takes place, plasma cells are present, and macrophagelike cells are found in increased numbers, the T4/T8 ratio was 0.8. It appeared from this that in regions of immunological activity (transitional areas), the percentage of T8 or suppressor/cytotoxic cells rose as if in response to the immunological stimulation taking place. This work emphasized that the T4:T8 ratio in rheumatoid synovial tissue is dependent on the region selected for staining.

Functional measurements of suppressor activity in the T-cell population isolated from rheumatoid synovial tissue have disclosed an almost complete lack of suppressor activity (Chattopadhyay *et al.*, 1979; Romain *et al.*, 1982). The discrepancy between the positive staining of suppressor cells in synovial tissue or digests and the absence of suppressor activity in cultures of synovial T cells has not been explained.

B. Immunoglobulin Synthesis in the Synovial Tissue

The rheumatoid synovium removed at the time of surgery carries on extensive immunoglobulin synthesis. In experiments of Smiley *et al.* (1968), the level of synthesis in synovial pieces cultured *in vitro* was greater than that of spleen or lymph node tissue. The major portion of the immunoglobulin synthesized was IgG, with lesser amounts of IgA and IgM. On the basis of kinetic studies, Sliwinski and Zvaifler (1970) calculated that between 12 and 26% of the IgG present in the synovial fluid had been produced locally in the synovial membrane. Wernick *et al.* (1985) measured the synthesis of IgG-RF and IgM-RF by solid phase radioimmunoassay in synovial mononuclear cell suspensions obtained from seropositive patients. Both RFs were synthesized, and in approximately equal quantities. IgG-RF accounted for 5.4 ± 2.4% of the IgG secreted, and IgM-RF for 30.1 ± 13.8% of the IgM secreted. These were larger fractions of the total immunoglobulin in each class than found in the serum, and in the case of IgM-RF, than synthesized by the peripheral blood mononuclear cells, suggesting that there was a selective enrichment of RF synthesis in the synovial tissue over that synthesized elsewhere in the body.

This evidence of selective RF synthesis in rheumatoid synovial tissue suggested that locally synthesized RFs may make a significant contribution to the RF complexes in the synovial effusion.

C. Epstein–Barr Virus and Rheumatoid Arthritis

There has been much recent interest (Ferrell *et al.*, 1981; Depper and Zvaifler, 1981; Catalano *et al.*, 1979) in the possibility that Epstein–Barr virus (EBV) has an etiological role in RA. This interest began with the observation of Alspaugh and Tan (1976) that rheumatoid patients had a high frequency of precipitating antibody to a nuclear antigen present only in EBV-transformed B lymphocytes. This antibody, designated the rheumatoid arthritis precipitin (RAP), was subsequently also demonstrated by immunofluorescent staining of EBV-infected B-cell lines such as the WI-L_2 and Raji cell lines. The nuclei of such cell lines were stained by a majority of rheumatoid sera, indicating the presence of antibody to a unique antigen called the rheumatoid arthritis associated nuclear antigen (RANA). Opinion on the elevation of titers to EBV antigens other than RANA has been divided, but serological data do not support the possibility that rheumatoid patients have a different exposure experience to EBV than the rest of the population (Ferrell and Tan, 1981).

Additional evidence for an association between EBV and RA is the observation of Bardwick *et al.* (1980) that PBMC from rheumatoid patients spontaneously establish permanent cell lines more frequently and rapidly than do lymphocytes from control individuals. Investigation showed that the earlier "takeoff" of rheumatoid B cells to form a cell line was due to a deficiency of a T-cell population, which ordinarily slows the rate of formation of cell lines. This observation pointed to a T-cell-related inhibition of EBV stimulation of B cells that was deficient in RA.

Tosato *et al.* (1981) also provided evidence that T-cell control of the B-cell response to EBV is deficient in RA. Lymphocytes from EBV-immune rheumatoid patients failed to develop a T-suppressor-cell population after 12 days in the EBV-infected cultures of PBMC, while lymphocytes from normal individuals suppressed the number of ISC formed at this time. Finally, Hasler *et al.* (1982) offered evidence that the T-cell suppressor of EBV-induced B-cell proliferation alluded to above exerts this suppression by secreting γ-interferon. It appears likely, in view of the evidence of diminished T-cell function observed in RA, that decreased amounts of γ-interferon are produced by the T cells in EBV-infected cultures of rheumatoid peripheral blood mononuclear cells, leading to an increased B-cell response to the virus. Thus, the EBV abnormalities observed in RA are in all likelihood a secondary phenomenon.

VI. INITIATION OF RHEUMATOID SYNOVITIS

The initiation of rheumatoid synovial inflammation presumably takes place at the level of the postcapillary venule of the synovial membrane. The stimulus, whatever it may be, activates the emigration of mononuclear cells from the postcapillary venule. In some areas, these mononuclear cells collect to form quiescent nodules of small lymphoyctes, which maintain the approximate ratios of the T- and B-cell populations in the blood. In other perivascular areas, an immunological reaction appears to occur with the formation of blast cells and plasma cells and the accumulation of macrophagelike cells. The nature of the stimulus for these changes is not known. Evidence of an extrinsic antigen is not at hand. Analysis of the immune complexes of the synovial effusion (Male and Roitt, 1981) has detected only immunoglobulins and complement. This leaves the possibility that IgG is the responsible autoantigen. The major argument in favor of this possibility is the almost invariable association of classical RA with the RFs. However, RFs are present in a number of chronic inflammatory states other than RA and it is not likely that an autoimmune response to autologous IgG is the initial stimulus in all of them. It also does not appear likely that EBV is the causative agent. The evidence to date indicates that although RA patients have a decreased capacity to suppress the EBV stimulation of B cells, this is very likely a result of the rheumatoid process rather than a cause of it (Hasler *et al.*, 1982).

The possibility remains that neither extrinsic antigen nor autoantigen is involved in the causation of RA but that this disease is a consequence of an autologous mixed lymphocyte reaction (AMLR) (Weksler *et al.*, 1981). In the AMLR, proliferation of T cells *in vitro* takes place as a consequence of stimulation by autologous non-T cells when a sufficiently elevated ratio of the adherent cell fraction to the T-lymphocyte fraction is established. Such a reaction, if it occurred in the synovium, would be sufficient to trigger the series of cell interactions which takes place in the rheumatoid synovium. If the concentration of macrophages or dendritic cells in a given perivascular collection of mononuclear cells became sufficiently high, conditions would be ripe for the development of an AMLR. In the RA synovium, conditions favoring an AMLR would exist in the macrophage-rich transitional areas. On the other hand, in the absence of significant numbers of macrophages or dendritic cells, perivascular lymphocyte collections would grow in a passive manner to form inactive populations of small lymphocytes, the lymphocyte-rich collections. Although it is not clear what abnormality is responsible for the initial emigration of small lymphocytes from the postcapillary venule, it is possible that the immunological reactivity seen in the transitional areas is a conse-

quence of the increased ratio of macrophages or dendritic cells to lymphocytes permitted by the endothelial cells to emigrate to these areas from the blood.

VII. CONCLUSION

Rheumatoid synovitis (Fig. 3) consists of two simultaneous reactions: a chronic inflammatory reaction that is present in the sublining layer of the synovium, and a more acute exudative reaction that takes place in the synovial fluid. In the chronic inflammatory phase, T cells, B cells, and macrophagelike cells interact. Initially, a mixed population of small lymphocytes emigrates from the postcapillary venule (Fig. 4). It consists mainly of T cells and the majority of these, as in the blood, are T4 cells (Kurosaka and Ziff, 1985). They form lymphocyte-rich collections in which T4 lymphocytes are predominant (T4/T8 = 2.9). In neighboring perivascular areas that have been referred to as transitional areas, the T cells appear to undergo blastic transformation, lymphokines are released, and perhaps as a consequence, macrophages are attracted into the area.

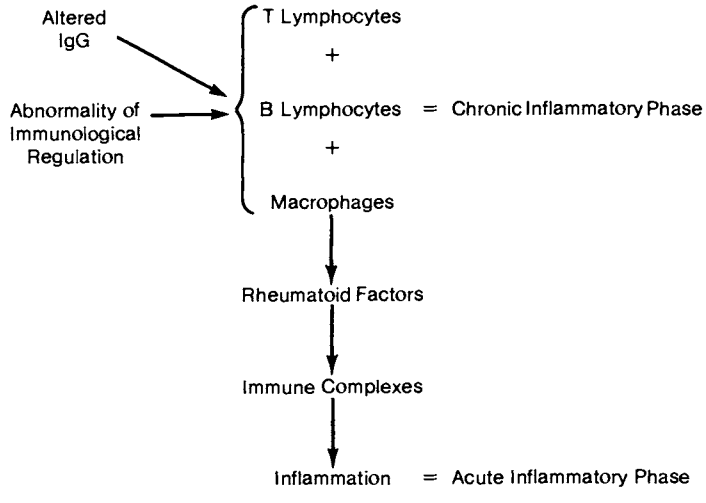


FIG. 3. The interaction of the mononuclear cells of the rheumatoid synovial membrane (chronic inflammatory phase) leads to the production of rheumatoid factor immune complexes, perhaps requiring the participation of altered IgG. The resulting immune complex-induced acute, exudative reaction, manifested by polymorphonuclear cells and lowered complement levels in the synovial effusion, constitutes the acute inflammatory phase of rheumatoid synovitis.

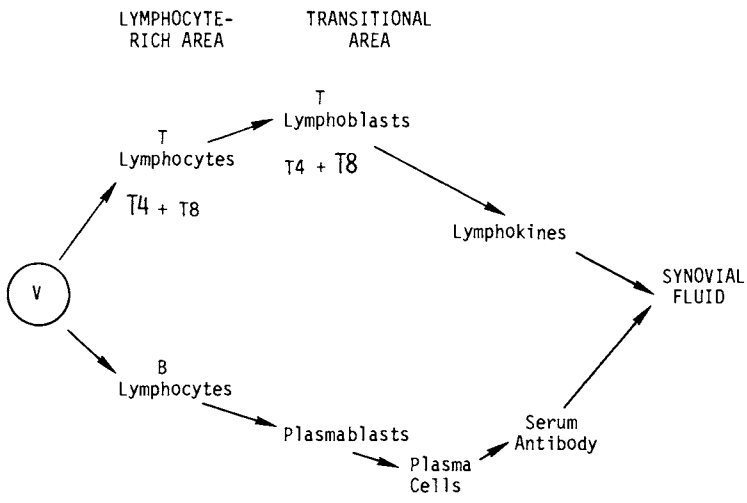


FIG. 4. Changes in the T- and B-lymphocyte populations following emigration from the postcapillary venule of the rheumatoid synovial membrane with production of lymphokines and antibodies.

Presumably, as a result of proliferation of T8 supressor/cytotoxic lymphocytes, the T4/T8 ratio in the transitional areas falls (T4/T8 = 0.8). The lymphokines liberated by the activated T cells diffuse into the synovial fluid where a blastogenic factor and MIF have been identified. The macrophages also secrete the monokine IL-1, which has also been identified in the synovial fluid.

The B lymphocytes, which make up 10–20% of the early lymphocyte-rich collections, undergo blastic transformation in the transitional areas and become plasma cells. The plasma cells synthesize large amounts of immunoglobulin and significant amounts of rheumatoid factors. IgG-RF and IgM-RF are synthesized in approximately equal quantities (Wernick, *et al.*, 1985) and are liberated into the synovial fluid where they participate in the formation of the immune complexes that are responsible for the immune complex or exudative phase of rheumatoid synovitis. In this phase, large numbers of polymorphonuclear cells are attracted into the synovial effusion. In this setting, the chronic inflammatory reaction in the sublining layer appears to act as a factory for the production of the immunoglobulins of the immune complexes that are responsible for the exudative phase.

The stimulus that initiates the abnormal synovial response is as yet not known. The presence in the serum of a number of the autoantibodies that characterize the connective tissue diseases as a group suggests that a

defect in immunoregulation is, at least in part, responsible for the abnormal synovial immune response. The specific clinical configuration of rheumatoid disease and its particular association with anti-IgG autoantibody formation may constitute only the pattern in which individuals with a "rheumatoid" genetic background respond to a defect in immunoregulation.

REFERENCES

Abrahamsen, T. G., Frøland, S. S., Natvig, J. B., and Pahle, J. (1975). *Scand. J. Immunol.* **4,** 823–830.

Abruzzo, J. L., and Christian, C. L. (1961). *J. Exp. Med.* **114,** 791–806.

4, 823–830. Abruzzo, J. L., and Christian, C. L. (1961). *J. Exp. Med.* **114,** 791–806.

Alarcon, G. S., Koopman, W. J., Acton, R. T., and Barger, B. O. (1982). *Arthritis Rheum.* **25,** 502–507.

Alspaugh, M. A., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 711–719.

Andreis, M., Hurd, E. R., LoSpalluto, J. J., and Ziff, M. (1978). *Arthritis Rheum.* **21,** 310–315.

Andriopoulos, N. A., Miller, E. J., and Mestecky, J. (1975). *Fed. Proc, Fed. Am. Soc. Exp. Biol.* **34,** 940.

Ball, J., and Lawrence, J. S. (1961). *Ann. Rheum. Dis.* **20,** 235–243.

Bankhurst, A. D., Husby, G., and Williams, R. C., Jr. (1976). *Arthritis Rheum.* **19,** 555–562.

Bardwick, P. A., Bluestein, H. G., Zvaifler, N. J., Depper, J. M., and Seegmiller, J. E. (1980). *Arthritis Rheum.* **23,** 626–632.

Barnett, E. V., Bluestone, R., and Cracciola, A. (1970). *Ann. Intern. Med.* **73,** 95–107.

Bianco, N. E., Dobkin, L. W., and Schur, P. H. (1974). *Clin. Exp. Immunol.* **17,** 91–101.

Bonomo, L., LoSpalluto, J. J., and Ziff, M. (1963). *Arthritis Rheum.* **6,** 104–114.

Brandt, K. D., Cathcart, E. S., and Cohen, A. S. (1968). *J. Lab. Clin. Med.* **72,** 631–647.

Brown, P. B., Nardella, F. A., and Mannik, M. (1982). *Arthritis Rheum.* **25,** 1101–1107.

Bunim, J. J., Burch, T. A., and O'Brien, W. M. (1964). *Bull. Rheum. Dis.* **15,** 349–350.

Burmester, G. R., Yu, D. T. Y., Irani, A.-M, Kunkel, H. G., and Winchester, R. J. (1981). *Arthritis Rheum.* **24,** 1370–1376.

Carson, D. A., Bayer, A. S., Eisenberg, R. A., Lawrance, S., and Theofilopoulos, A. (1978). *Clin. Exp. Immunol.* **31,** 100–103.

Catalano, M. A., Carson, D. A., Slovin, S. F., Richman, D., and Vaughan, J. H. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76,** 5825–5828.

Cathcart, E. S., Williams, R. C., Ross, H., and Calkins, E. (1961). *Am. J. Med.* **31,** 758–765.

Cecere, F., Lessard, J., McDuffy, D., and Pope, R. M. (1982). *Arthritis Rheum.* **25,** 1307–1314.

Chattopadhyay, C., Chattopadhyay, H., Michaelsen, T. E., Mellbye, O. J., and Natvig, J. B. (1979). *Scand. J. Immunol.* **10,** 309–316.

Clague, R. B., Shaw, M. J., and Lennox, P. J. L. (1979). *Arthritis Rheum.* **22,** 598.

Cooke, T. D., Hurd, E. R., Jasin, H. J., Bienenstock, J., and Ziff, M. (1975). *Arthritis Rheum.* **18,** 541–551.

Depper, J. M., and Zvaifler, N. J. (1981). *Arthritis Rheum.* **24,** 755–761.

Doubloug, J. H., Forre, O., Kass, E., and Thorsby, E. (1980). *Arthritis Rheum.* **23,** 309–313.

Duke, O., Panayi, G. S., Janossy, G., and Poulter, L. W. (1982). *Clin. Exp. Immunol.* **49,** 22–30.
Fallet, G., and Ziff, M. (1958). *Arthritis Rheum.* **1,** 70–76.
Ferrell, P. B., and Tan, E. M. (1981). *Springer Semin. Immunopathol.* **4,** 181–191.
Ferrell, P. B., Aitcheson, C. T., Pearson, G. R., and Tan, E. M. (1981). *J. Clin. Invest.* **67,** 681–687.
Foidart, J.-M., Shigeto, A., Martin, G. R., Zizic, T. M., Barnett, E. V., Lawley, T. J., and Katz, S. I. (1978). *N. Engl. J. Med.* **299,** 1203–1207.
Fontana, A., Hengartner, H., Weber, E., Fehr, K., Grob, P. J., and Cohen, G. (1982). *Rheumatol. Int.* **2,** 49–53.
Forre, O., Thoen, J., Dobloug, J. H., Egeland, T., Kvien, T. K., Mellbye, O. J., and Natvig, J. B. (1982). *Scand. J. Immunol.* **15,** 221–226.
Fox, R. I., Fong, S., Sabharwal, N., Carstens, S. A., Kung, P. C., and Vaughan, J. H. (1982). *J. Immunol.* **128,** 351–354.
Franklin, E. C., Holman, H. R., Muller-Eberhard, H. J., and Kunkel, H. G. (1957). *J. Exp. Med.* **105,** 425–438.
Fudenberg, H. H., and Kunkel, H. G. (1961). *J. Exp. Med.* **114,** 257–278.
Garcia-de la Torre, I., and Miranda-Mendez, L. (1982). *J. Rheumatol.* **9,** 603–609.
Halla, J. T., Volanakis, J. E., and Schrohenloher, R. E. (1979). *Arthritis Rheum.* **22,** 440–448.
Hannestad, K. (1978). *Scand. J. Immunol.* **7,** 127–136.
Hasler, F., Bluestein, H. G., and Zvaifler, N. J. (1982). *Arthritis Rheum.* **24,** S51.
Hedberg, H. (1964). *Acta Rheum. Scand.* **10,** 109–127.
Hurd, E. R., LoSpalluto, J. J., and Ziff, M. (1970). *Arthritis Rheum.* **13,** 724–733.
Hurd, E. R., Kinsella, T. D., and Ziff, M. (1971). *J. Exp. Med.* **134,** 296s–305s.
Hurd, E. R., Andreis, M., and Ziff, M. (1977). *Clin. Exp. Immunol.* **28,** 413–425.
Ishikawa, H., and Ziff, M. (1976). *Arthritis Rheum.* **19,** 1–14.
Ishikawa, K., Smiley, J. D., and Ziff, M. (1975). *Arthritis Rheum.* **18,** 563–576.
Janossy, G., Panayi, G., Duke, O., Bofill, M., Poulter, L. W., and Goldstein, G. (1981). *Lancet* **ii,** 839–841.
Jasin, H. E., LoSpalluto, J. J., and Ziff, M. (1970). *Am. J. Med.* **49,** 484–493.
Johnson, P. M., and Faulk, W. P. (1976). *Clin. Immunol. Immunopathol.* **6,** 414–430.
Jones, E. V., Jacoby, R. K., Wallington, T., and Holt, P. (1981). *Clin. Exp. Immunol.* **44,** 512–521.
Kinsella, T. D. (1974). *J. Clin. Invest.* **53,** 1108–1114.
Kinsella, T. D., Baum, J., and Ziff, M. (1969). *Clin. Exp. Immunol.* **4,** 265–271.
Kinsella, T. D., Baum, J., and Ziff, M., (1970). *Arthritis Rheum.* **13,** 734–753.
Klareskog, L., Forsum, U., Wigren, A., and Wigzell, H. (1982). *Scand. J. Immunol.* **15,** 501–508.
Kobayashi, I., and Ziff, M. (1973). *Arthritis Rheum.* **16,** 471–486.
Kunkel, H. G., Simon, H. J., and Fudenberg, H. H. (1958). *Arthritis Rheum.* **1,** 289–296.
Kunkel, H. G., Muller-Eberhard, H. J., Fudenberg, H. H., and Tomasi, T. B. (1961). *J. Clin. Invest.* **40,** 117–129.
Kurosaka, M., and Ziff, M. (1983). *J. Exp. Med.* **158,** 1191–1210.
Lance, E. M. (1982). *Arthritis Rheum.* **25,** S9.
Lawrence, J. S. (1970). *Ann. Rheum. Dis.* **29,** 357–379.
LoSpalluto, J. J., and Ziff, M. (1959). *J. Exp. Med.* **110,** 169–186.
LoSpalluto, J. J., Miller, W. E., Dorward, B., and Ziff, M. (1961). *Am. J. Med.* **32,** 142–147.
Lydyard, P. M., Hanglow, A., Hartley, I., Young, A., and Roitt, I. M. (1982). *Lancet* **i,** 799.

McAdam, K. P. W., Fudenberg, H. H., and Michaeli, D. (1978). *Clin. Immunol. Immunopathol.* **9,** 16–21.

McDougal, J. S., Hubbard, M., McDuffie, F. C., Strobel, P. L., Smith, S. J., Bass, N., Goldman, J. A., Hartman, S., Myerson, G., Miller, S., Morales, R., and Wilson, C. H., Jr. (1982). *Arthritis Rheum.* **25,** 1156–1166.

Male, D. K., and Roitt, I. M. (1981). *Clin. Exp. Immunol.* **46,** 521–529.

Marcus, R. L., and Townes, A. S. (1968). *Arthritis Rheum.* **11,** 497–498.

Marcus, R. L., and Townes, A. S. (1971). *J. Immunol.* **106,** 1499–1506.

Marmont, A. M., Damasio, E. E., Bertorelli, C., and Rossi, F. (1967). *Arthritis Rheum.* **10,** 117–128.

Masi, A. T., and Shulman, L. E. (1965). *Ann. Rheum. Dis.* **8,** 418–425.

Meijer, C. J. L. M., de Graaff-Reitsma, C. B., Lafeber, G. J. M., and Cats, A. (1982). *J. Rheumatol.* **9,** 359–365.

Michaeli, D., and Fudenberg, H. H. (1974). *Clin. Immunol. Immunopathol.* **2,** 153–159.

Milgrom, F., Witebsky, E., Goldstein, R., and Loza, U. (1962). *J. Am. Med. Assoc.* **181,** 476–484.

Munthe, E., and Natvig, J. B. (1971). *Clin. Exp. Immunol.* **8,** 249–262.

Natvig, J. N., Gaarder, P. I., and Turner, M. W. (1972). *Clin. Exp. Immunol.* **12,** 177–184.

Nienhuis, R. L. F., and Mandema, E. (1964). *Ann. Rheum. Dis.* **23,** 302–305.

Notman, D. D., Kurata, N., and Tan, E. M. (1975). *Ann. Intern. Med.* **83,** 464–469.

Osung, O. A., Chandra, M., and Holborow, E. J. (1982). *Ann. Rheum. Dis.* **41,** 69–73.

Pekin, T. J., and Zvaifler, N. J. (1964). *J. Clin. Invest.* **43,** 1372–1382.

Peltier, A. (1959). *Arthritis Rheum.* **2,** 1–7.

Pisko, E. J., Pruitt, N. L., and Turner, R. A. (1982). *Arthritis Rheum.* **25,** 1108–1116.

Pope, R. M., Teller, D. C., and Mannik, M. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71,** 517–521.

Pope, R. M., Teller, D. C., and Mannik, M. (1975). *J. Immunol.* **115,** 365–373.

Pope, R. M., Yoshimoya, S., and McDuffy, S. J. (1981). *Clin. Exp. Immunol.* **46,** 259–267.

Pope, R. M., Keightly, R., and McDuffy, S. (1982). *J. Immunol.* **128,** 1860–1863.

Raeman, F., DeCock, W., DeBeukelaar, T., DeCree, J., and Verhaegen, H. (1981). *Clin. Exp. Immunol.* **45,** 475–479.

Roberts-Thomson, P. J., Neoh, S. H., Bradley, J., and Milazzo, S. C. (1980). *Ann. Rheum. Dis.* **39,** 438–444.

Romain, P. L., Burmester, G. R., Enlow, R. W., and Winchester, R. J. (1982). *Rheumatol. Int.* **2,** 121–127.

Rose, H. M., Ragan, C., Pearce, E., and Lipman, M. (1948). *Proc. Soc. Exp. Biol. Med.* **68,** 1–6.

Ruddy, S., and Austen, K. F. (1970). *Arthritis Rheum.* **13,** 713–623.

Ruddy, S., and Austen, K. F. (1973). *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **32,** 134–137.

Sabharwal, U. K., Vaughan, J. H., Fong, S., Bennett, P. H., Carson, D. A., and Curd, J. G. (1982). *Arthritis Rheum.* **25,** 161–167.

Sliwinski, A. J., and Zvaifler, N. J. (1970). *J. Lab. Clin. Med.* **76,** 304–310.

Smiley, J. D., Sachs, C., and Ziff, M. (1968). *J. Clin. Invest.* **47,** 624–632.

Stastny, P. (1981). *Adv. Inflamm. Res.* **3,** 41–48.

Stastny, P., Rosenthal, M., Andreis, M., Cooke, D., and Ziff, M. (1975a). *Ann. N. Y. Acad. Sci.* **256,** 117–131.

Stastny, P., Rosenthal, M., Andreis, M., and Ziff, M. (1975b). *Arthritis Rheum.* **18,** 237–243.

Steffen, C., Ludwig, H., Thumb, N., Frank, O., Eberl, R., and Tausch, S. (1973). *Klin. Wochenschr.* **41,** 222–229.

Steffen, C., Ludwig, H., and Knapp, W. (1974). *Z. Immun.-Forsch.* **147,** 229–235.
Taylor, H., and Abraham, G. N. (1973). *Clin. Exp. Immunol.* **13,** 529–536.
Taylor-Upsahl, M. M., Johnson, P. M., Mellbye, O. J., and Natvig, J. B. (1977). *Clin. Exp. Immunol.* **28,** 204–211.
Tomasi, T. B., Fudenberg, H. H., and Finby, H. H. (1962). *Am. J. Med.* **33,** 243–248.
Tosato, G., Steinberg, A. D., and Blaese, R. (1981). *N. Engl. J. Med.* **305,** 1238–1243.
Trentham, D. E., Townes, A. S., and Kang, A. H. (1977). *J. Exp. Med.* **146,** 857–868.
Trentham, D. E., Dynesius, R. A., Rocklin, R. E., and David, J. R. (1978). *N. Engl. J. Med.* **299,** 1327–1332.
Tursi, A., Trizio, D., and Bonomo, L. (1970). *Clin. Exp. Immunol.* **6,** 767–772.
Van Boxel, J. A., and Paget, S. A. (1975). *N. Engl. J. Med.* **293,** 517–520.
Veys, E. M., Hermanns, P., Schindler, J., Kung, P. C., Goldstein, G., Symoens, J., and Wauwe, J. V. (1982a). *J. Rheumatol.* **9,** 25–29.
Veys, E. M., Hermanns, P., Verbruggen, G., Schindler, J., and Goldstein, G. (1982b). *Lancet* **i,** 225–226.
Waaler, E. (1940). *Acta Pathol. Microbiol. Scand.* **17,** 172–188.
Waller, M. V., Decker, B., Toone, E. C., and Irby, R. (1961). *Arthritis Rheum.* **6,** 578–591.
Weksler, M. E., Moody, C. E., and Kozak, R. W. (1981). *Adv. Immunol.* **31,** 271–312.
Wernick, R. M., LoSpalluto, J. J., Fink, C. R., and Ziff, M. (1981). *Arthritis Rheum.* **24,** 1501–1511.
Wernick, R. M., Lipsky, P. E., Marban-Arcos, E., Maliakkal, John J., Edelbaum, D., and Ziff, M. (1985). *Arthritis Rheum.* **28,** 742–752.
Williams, R. C. (1979). *Am. J. Med.* **67,** 179–191.
Williams, R. C., and Kunkel, H. G. (1962). *J. Clin. Invest.* **41,** 666–675.
Williams, R. C., and Kunkel, H. G. (1963). *Arthritis Rheum.* **6,** 665–675.
Winchester, R., Agnello, V., and Kunkel, H. G. (1970). *Clin. Exp. Immunol.* **6,** 689–706.
Yamasaki, K., and Ziff, M. (1977). *Arthritis Rheum.* **20,** 679–684.
Ziff, M., Schmid, F. R., Lewis, A. J., and Tanner, M. (1968). *Arthritis Rheum.* **1,** 392–399.
Zuraw, B. L., O'Hair, C. H., Vaughan, J. H., Mathison, D. A., Curd, C. G., and Katz, D. H. (1981). *J. Clin. Invest.* **68,** 1610–1613.

CHAPTER 4

Mixed Connective Tissue Disease

GORDON C. SHARP
BERNHARD H. SINGSEN
Departments of Child Health, Medicine and Pathology
University of Missouri Health Sciences Center
Columbia, Missouri

Any one who has seen a half a dozen examples of common lupus and lupus erythematosus is able with ease to distinguish the one from the other, . . . but let him wait awhile and see more, and he will find before long that there are examples of mixed forms of disease which it is impossible to denote correctly without employing hybrid names or qualifying adjectives. (Hutchinson, 1880)

I. HISTORICAL DEVELOPMENT

Patients with overlap syndromes of various types, who do not fit traditional classifications, have been recognized for many years (Dubois, 1974; Sharp, 1975). In the early and mid-1960s at Stanford University, the initial serological evaluations were made by Sharp *et al.* of some interesting patients who had overlapping clinical features of systemic lupus erythematosus (SLE), progressive systemic sclerosis (PSS), and polymyositis. Their characteristic finding was that they had persistent high titers of antibody to whole calf thymus nuclei by complement fixation, both when the disease was active and when in remission. This distinguished them from typical SLE patients, whose antibody titer would diminish as the disease became inactive. It was believed that these Stanford patients might represent a distinct clinical entity, and hence they were referred to as mixed connective tissue disease (MCTD), rather than by the more general term overlap syndrome (Sharp *et al.*, 1972).

It was hypothesized that the circulating antinuclear antibody in these patients might be unique. In order to try to clarify its nature, a very sensitive passive hemagglutination test was developed in which various nuclear antigens were adsorbed to tannic acid-treated sheep erythrocytes. These studies of Sharp *et al.* (1972) revealed that the MCTD patients all had very high hemagglutination titers (1:1,000 to 1:10,000,000) of antibodies directed against an extractable nuclear antigen (ENA) that had been described earlier by Holman (1965). Analysis of sera from other rheumatic diseases showed that ENA antibodies were present, usually at lower levels, in ~50% of patients with typical SLE and were rarely found in other diseases.

Since ENA was a nuclear preparation that was only partially purified, it seemed reasonable to postulate that it might contain more than a single antigenic specificity, and that ENA antibodies occurring in MCTD might be different from those detected in SLE. In order to investigate this possibility, enzyme digestion studies were pursued in which ENA-coated erythrocytes treated with ribonuclease (RNase), trypsin, or deoxyribonuclease were reacted with MCTD and SLE sera. It was found that digestion with RNase and trypsin eliminated or greatly reduced the hemagglutination reaction to ENA in all cases of MCTD, but had little effect on ENA antibody titers in SLE (Sharp *et al.*, 1972).

Tan then performed fluorescent antinuclear antibody (FANA) analyses on these MCTD sera, and found that they produced speckled reactions in high titer (Sharp *et al.*, 1972). Rim or patchy FANA patterns were occasionally seen at lower dilutions, but at higher dilutions all reactions were speckled. RNase treatment of the tissue sections eliminated the speckled

reactions of MCTD sera. These findings suggested that the circulating antibody in MCTD had a specificity for a nuclear ribonucleoprotein (RNP) antigen.

Immunodiffusion analyses in several laboratories during the early 1970s (Northway and Tan, 1972; Reichlin and Mattioli, 1972; Parker, 1973) showed more precisely that ENA contains at least two immunologically distinct antigenic components, the RNase- and trypsin-sensitive RNP antigen and the RNase- and trypsin-resistant Sm antigen; the latter had been described previously by Tan and Kunkel (1966). The immunodiffusion test, using well-characterized reference sera, provides a more definitive identification of RNP and Sm antibodies, but the hemagglutination test with titers before and after RNase continues to be highly useful because it is much more sensitive and quantitative. As an example, RNP antibody with a hemagglutination titer of $<1:1000$ is rarely detectable by immunodiffusion. The hemagglutination test for RNP antibody is also 100- to 1000-fold more sensitive than the FANA test; thus, the FANA may be inadequate as a screening test early in the course of MCTD when the antibody titer may be at a low level. Counterimmunoelectrophoresis may also, in some cases, be able to demonstrate RNP or Sm antibodies that have not been detected by other techniques (Kurata and Tan, 1976; Bresnihan *et al.*, 1977).

During the past decade, it has become progressively more clear that the serological pattern of very high titers of RNP antibodies and no Sm antibodies is more frequently associated with MCTD, is infrequent in SLE and PSS, and is very rare in other rheumatic diseases (Sharp *et al.*, 1972; Northway and Tan, 1972; Parker, 1973; Notman *et al.*, 1975; Sharp *et al.*, 1976; Farber and Bole, 1976; Singsen *et al.*, 1977; Tan *et al.*, 1980; Sharp, 1982). The typical serological pattern in MCTD shows a high-titer, speckled FANA (often $>1:1{,}000$), and very high titers (frequently $>1:1{,}000{,}000$) of antibody to RNase-sensitive ENA by hemagglutination and RNP antibody by immunodiffusion. Sm antibodies and high titers of antibody to native DNA are infrequent in MCTD, and their appearance usually correlates with a severe flare of SLE-like disease (Esther *et al.*, 1981; Grant *et al.*, 1981; Frank *et al.*, 1983). Sm antibodies, usually in association with lower titers of RNP antibodies, are most commonly associated with SLE (Notman *et al.*, 1975; Sharp *et al.*, 1976).

Patients with clinical features of MCTD will rarely be negative for RNP antibody (Sharp *et al.*, 1976), and occasionally they may transiently have negative tests for FANA and RNP antibody early in their course. Subsequently these tests will become positive, usually following corticosteroid therapy (Alarcón-Segovia, 1979; Sharp *et al.*, 1972; Grant *et al.*, 1981). This serological evolution may relate to the presence of circulating im-

mune complexes, often detected by Raji cell or Clq binding assays (Halla *et al.*, 1978; Frank *et al.*, 1983), which during active MCTD bind the RNP antibody and impede its detection. Once present, high titers of free circulating RNP antibody usually persist both during periods of active and inactive disease, but long-term studies have disclosed that in some patients RNP antibody levels may fall or become undetectable after 5 to 11 years (Sullivan *et al.*, 1984). This phenomenon is more often seen in patients in prolonged remission and less frequently in association with active disease. Other recent reports also emphasize that ENA antibody titers and FANA patterns, as well as other serological tests, may change with time in MCTD (Nimelstein *et al.*, 1980; Grant *et al.*, 1981).

The initial description of typical clinical characteristics in MCTD included Raynaud's phenomenon, swollen hands with a sausagelike appearance of the fingers, polyarthralgia or arthritis, esophageal hypomotility, inflammatory myopathy, lupuslike skin rashes, serositis, leukopenia, and anemia. The infrequency of renal disease, responsiveness to corticosteroids, and a relatively good prognosis were suggested by that first report (Sharp *et al.*, 1972). Over the past decade it has emerged that renal disease may occur (although progressive renal failure is rare), and that pulmonary disease is very common. Pulmonary hypertension, associated with proliferative vascular lesions, is also a serious complication in some patients, and therefore the prognosis is not always favorable (Nimelstein *et al.*, 1980; Singsen *et al.*, 1980a; Grant *et al.*, 1981; Sharp, 1981; Sullivan *et al.*, 1984).

II. THE CLINICAL CONTROVERSY OF MIXED CONNECTIVE TISSUE DISEASE (MCTD)

The status of MCTD as a distinct rheumatic disease syndrome remains controversial, at least partially because no single laboratory, roentgenographic, or clinical feature is pathognomonic. Indeed, one might view the delay in the development of diagnostic criteria as recognition that no illness of evolution, which may encompass many overlapping features of "classic" rheumatic disorders, will ever occupy a comfortable niche as a defined entity. And perhaps it should not. Perhaps the value of mixed connective tissue disease lies in the investigative, clinical, and academic "tension" that it creates, spurring both its opponents and proponents toward better understanding of fundamental immunopathological mechanisms.

Historically, descriptions of overlapping features of rheumatic disorders have been with us for a century. Despite the furious pace of techno-

logical advances in immunology during the past 30 years, the diagnosis of rheumatic syndromes is still based almost exclusively on clinical phenomena, with only limited knowledge of etiology and pathogenesis. Thus, one sees myopathic features at clinical or histological levels of observation in problems as diverse as localized and progressive systemic sclerosis (PSS), rheumatoid arthritis, JRA, SLE, polymyositis, and a variety of postinfectious reactive arthropathies. The same can be said for synovitis in these and in almost all other rheumatic disorders. Similarly, the ubiquity of pulmonary involvement, albeit most commonly quite mild, is now recognized in the majority of rheumatic syndromes in both children and adults.

As suggested by Christian (1979), nosological frustration ensues when arbitrary criteria for disease classification collide with the real world. Terms such as "lupoderma," "sclerodermatomyositis," and "rupus" already seem out of step with today's sophistication, and yet to call something an "overlap syndrome" or "undifferentiated connective tissue disease" is perhaps only a more linguistically sophisticated manner of stating that we do not yet understand things adequately. Thus, one cannot help but suspect that substitution of the word "mixed" for "undifferentiated" is not momentous and that possibly MCTD as a concept will have only historical interest in another short decade.

One of the greatest disservices we may do to ourselves is to develop classification criteria for diseases, such as those proposed by the American Rheumatism Association (ARA) for rheumatoid arthritis, SLE, and now scleroderma, without adequately protecting the practitioner from those criteria. Beyond aids to memory of salient disease features, such criteria should have no diagnostic utility; rather, they ought to be employed mainly for the purposes of uniform disease classification for immunobiological, clinical, and treatment investigations.

Given this, is the call for diagnostic criteria for MCTD, as suggested by Alarcón-Segovia (1981) and others, a timely and helpful step? Perhaps not quite yet. Despite being born into a rich family, MCTD is only 13 years old; and like most preadolescents, it is perhaps only dimly aware of its future directions and ultimate place in the adult world of rheumatic diseases, some of which have been recognized since antiquity. It is instructive to reflect that in 1976 Reichlin persuasively argued that MCTD was merely a subset of SLE, and yet by 1980 LeRoy and associates were convinced that MCTD patients frequently developed scleroderma. In a similar fashion, early rushes of literature quickly disspelled the notions that MCTD was without significant renal disease or erosive arthritis (Bennett and Spargo, 1977; Halla and Hardin, 1978).

But criteria of another sort may yet be helpful. The distinguishing fea-

tures of lymphocyte types, numbers, and functions that quite clearly separate MCTD and its four component classic disorders is fascinating new information (Alarcón-Segovia, 1981). And one begins to suspect that our progressive unraveling of the precise nature of the nuclear acidic protein antigens will have far-reaching "diagnostic" consequences as well. But the number of potential laboratory discriminators is becoming overwhelming, and it is clear that "pattern recognition" will be required.

Fortunately, the computer is at hand. An "artificial intelligence" model at the University of Missouri can now with a standard clinical and laboratory data base, differentiate MCTD and 24 other rheumatic disorders with 95% accuracy when compared with an experienced rheumatologist (Lindberg *et al.*, 1982, 1983). Since our understanding suggests that patterns of immunological behavior will continue to become diagnostically restrictive, the challenge will lie in our ability to provide the rapid and cost-effective performance of many laboratory tests. The computer will then be able to give us diagnostic probabilities, although the worrisome implications of deintellectualizing the art of medical thinking will need to be carefully addressed.

Progressively, it has become clear that MCTD is a disease of evolution. However, the sequential development of clinical features has caused problems for both "lumpers" and "splitters." On the one hand, 60% of MCTD patients can be diagnosed as "classic rheumatoid arthritis" by ARA criteria, and an additional 25% can be categorized as "probable rheumatoid arthritis" (Bennett and O'Connell, 1978). Conversely, the majority of Sharp's original 25 patients with MCTD also satisfy the current preliminary criteria for SLE. Several reports suggest that ultimately the largest percentage of MCTD patients progress to a clinical state difficult to differentiate from progressive systemic sclerosis (Nimelstein *et al.*, 1980; Singsen *et al.*, 1980a).

Serological findings may evolve sequentially as well. The latter has been demonstrated in rheumatoid factor-positive, FANA-negative children with "systemic onset JRA," who then developed high-titer RNP antibodies, subsequently had the transient presence of anti-Sm antibodies, and then developed MCTD (Singsen *et al.*, 1977). Grant *et al.* (1981) also reported the evolution of laboratory features, including 4 of 23 MCTD patients with initially absent FANA, 3 who had anti-Sm antibodies transiently present at 3 years after disease onset, and a progressive increase in the number of patients with abnormal binding of dsDNA at their 5-year follow-up. The frequency with which FANA could change to or from a speckled pattern was also notable. Seventy percent of the MCTD group was rheumatoid factor positive at some time during the disease course.

Since even "classical" rheumatic syndromes are seen to share common features with each other as they develop over time, it is particularly dangerous to describe or compare illnesses which are studied largely by cross-sectional, "point-in-time," or retrospective methods. As a new illness, MCTD is particularly liable to the criticism of absent prospective studies. Thus as one example the paucity of histological and postmortem studies of MCTD testifies to our lack of understanding of the natural history of the illness. To some extent, this fact blunts the arguments from both sides—pro and con—as to whether MCTD is a separate entity. Indeed, it is precisely this awareness that has led us to direct our research energies away from clinical description and toward three more important areas: (a) prospective studies of organ function and long-term outcome; (b) development of possible animal models; and (c) detailed biochemical studies of the composition and function of the antigens that react with antibodies commonly associated with MCTD.

Much has been written and misunderstood since the inital description of MCTD in 1972 that implies that the presence of anti-RNP antibodies in high titer arbitrarily defines MCTD. It is important to distinguish between that which is necessary and that which is enough. Most investigators are comfortable with the rare but striking patient who has features of all four components of MCTD but who has no RNP antibodies; they usually appear shortly. Much more distressing are the clinicians and occasional investigators who accept these RNP antibodies as a diagnostic *sine qua non*. Admittedly, the patient who demonstrates RNP antibodies has a higher probability of evolving into MCTD, but the diagnosis should be reserved for those patients who have demonstrated both clinical and serological features of a truly mixed pattern.

Is MCTD a diagnosis that should be actively sought, or can the diagnosis be allowed to develop naturally while treating its component parts? Bennett (1982) suggests that " . . . it behooves the alert physician to consider the possibility of MCTD in a diverse array of clinical situations," but then notes that such disparate things as fever of unknown origin, JRA, aseptic meningitis, unexplained serositis or arthritis, and myocarditis are by themselves adequate to prompt consideration of MCTD. In our opinion, such a listing may encourage overdiagnosis of MCTD and further confuse an already difficult terminology.

The most important reason for differentiating subsets of rheumatic illness is that they foster structured thought and research and allow investigators to produce sharply focused new information regarding management, outcome, etiology, and pathogenesis. It is evident that the original association of RNP antibodies with MCTD has spurred unique advances in our understanding of nuclear acidic protein antigens. Our derivative

knowledge of their biochemistry, and its impact on cellular and humoral immunity, may help to clarify the pathogenesis of illness as disparate as Sjögren's syndrome, SLE, dermatomyositis, and scleroderma.

III. AUTOIMMUNE IMMUNOPATHOLOGICAL MECHANISMS

A. Elucidation of Other Distinct Nuclear Antigens

Elucidation of the RNP and Sm nuclear antigens and the correlation of autoantibodies to these antigens with MCTD and SLE have stimulated further investigation of this type, resulting in the identification of other nuclear antigens that appear to have clinical importance. Some of these antigens and the rheumatic diseases with which they are correlated are shown in Fig. 1. The findings of new precipitin bands on Ouchterlony analysis, which were immunologically distinct from RNP, Sm, and other previously described systems, has led to the delineation of the MA and Su antigens associated with SLE (Winn *et al.*, 1979a; Treadwell *et al.*, 1984) and the Jo-1, PM-1, and Ku antigens found mainly in polymyositis and a polymyositis–scleroderma overlap syndrome (Nishikai and Reichlin, 1980; Wolfe *et al.*, 1977a; Mimori *et al.*, 1981). The SS-A (Ro) and SS-B (La) (Ha) antigens discovered independently in several laboratories have been primarily associated with Sjögren's syndrome and SLE (Alspaugh and Tan, 1975; Alspaugh *et al.*, 1976; Scopelitis *et al.*, 1980; Alspaugh and Maddison, 1979; Akizuki *et al.*, 1977). The rheumatoid arthritis nuclear antigen (RANA) system is most frequently seen and occurs in highest titers in rheumatoid arthritis (Alspaugh and Tan, 1976). Another distinct precipitin reaction, the Scl-70 system, has been found mainly in PSS as has the nucleolar antibody demonstrated in the FANA test, while the centromere antibody most clearly defined with FANA tests employing cultured cells is most frequently seen in the CREST variant of PSS (Tan *et al.*, 1980).

B. Correlation of Antinuclear Antibody Pattern and Clinical Manifestations

Evidence is emerging that suggests some of these circulating autoantibodies are correlated with predictable clinical manifestations and therefore may serve as markers for rheumatic disease subsets. Numerous published reports have documented the strong correlation of antibodies to native DNA and active SLE and the role of DNA–anti-DNA complexes

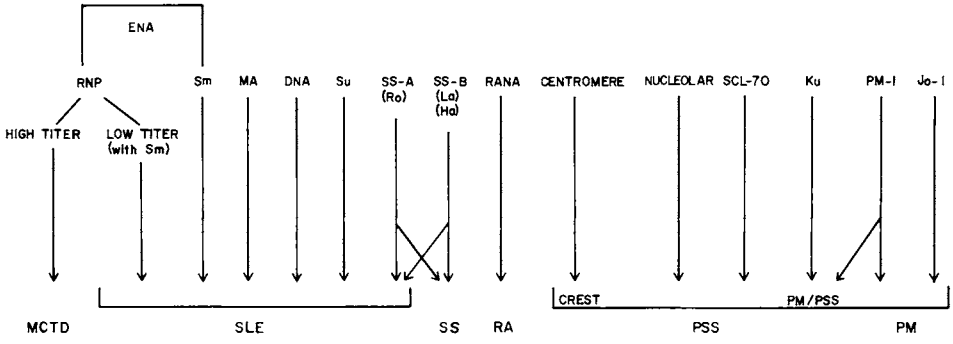


FIG. 1. Various nuclear antigens, to which circulating antibodies have been found in the rheumatic diseases, are listed at the top of the diagram. Arrows are drawn from each antigen to the disease listed below in which the system is most frequently observed. ENA, extractable nuclear antigen consisting of ribonucleoprotein (RNP) and Sm components; Sm, MA, Su, Ro, La, Ha, Ku, and Jo-1 are antigen designates by various authors, taken from the names of the prototype patients; SS-A and SS-B are systems associated with Sjögren's syndrome (SS); RANA, rheumatoid arthritis (RA)-associated nuclear antigen. Scl-70 and PM-1 designations refer to their association with scleroderma (PSS) and polymyositis (PM). The PM-1 and Ku systems are most closely related to a polymyositis–scleroderma overlap syndrome (PM/PSS). Ro and La were originally reported as cytoplasmic systems but are immunologically identical to the nuclear SS systems. Antibodies to the centromere and nucleolar antigens are identified by the FANA method, and are associated with the CREST variant of PSS and systemic sclerosis, respectively. MCTD, mixed connective tissue disease; SLE, systemic lupus erythematosus.

in SLE nephritis (Tan *et al.*, 1966; Koffler *et al.*, 1967). Antibody to the MA antigen is much less frequently detected but seems to be found exclusively in SLE, and appears to identify lupus patients who have very severe disease (Winn *et al.*, 1979a). Studies by Wasicek and Reichlin (1982) indicate that the occurrence of anti-Ro (SS-A) antibodies in the absence of anti-La (SS-B) antibodies in lupus patients is also associated with a high risk of having DNA antibodies and severe nephritis. In contrast, SLE patients with antibodies to both Ro and La (Wasicek and Reichlin, 1982) and with Sm and RNP antibodies (Winn *et al.*, 1979a,b; Sharp, 1982; Reichlin, 1979) may have a lower frequency of serious renal disease and a more favorable prognosis. Antibodies to RNP in high titers in the absence of other autoantibodies are rarely associated with progressive renal disease, but instead are very frequently correlated with Raynaud's phenomenon, swollen hands, sclerodactyly, esophageal hypomotility, abnormal pulmonary diffusing capacity, and inflammatory myositis (Sharp, 1979, 1981; Sullivan *et al.*, 1984).

It appears that patients with the CREST subset of scleroderma (who may have anticentromere antibodies) and MCTD patients with RNP anti-

bodies are at greater risk for developing pulmonary hypertension and proliferative vascular lesions (with minimal fibrosis) than are patients with classical PSS, SLE, and polymyositis (Salerni *et al.*, 1977; Sharp, 1981; Sullivan *et al.*, 1984). Pulmonary hypertension, when it occurs in PSS, is more apt to be associated with pathological findings of substantial fibrosis in addition to vascular changes (Salerni *et al.*, 1977).

The PM-1 and Ku antinuclear autoantibodies are most frequently correlated with a polymyositis–scleroderma overlap syndrome (Wolfe *et al.*, 1977a; Mimori *et al.*, 1981), although the PM-1 system also occurs less frequently in typical polymyositis. The Jo-1 system, on the other hand, is usually associated with classical polymyositis (Nishikai and Reichlin, 1980).

C. Immunopathological Findings in MCTD

1. B-Cell Hyperactivity

The numerous immunological aberrations that have been identified in MCTD strongly suggest that immune injury mechanisms are involved in the pathogenesis of the disease. The facts that extremely high titers of RNP antibody persist for many years and that patients often have marked polyclonal hypergammaglobulinemia are indicative of B-cell hyperactivity. Although the actual amount of RNP antibody has not been determined by direct measurement, using indirect methods, Maddison and Reichlin (1977) showed that RNP antibody consitituted 33% of the total immunoglobulin in one case.

2. Immune Regulation

In order to try to elucidate the abnormalities of immune regulation that may give rise to the autoimmune phenomena in MCTD and other rheumatic diseases, Alarcón-Segovia and colleagues studied the numbers and functions of helper T cells (T_H), suppressor T cells (T_S), and postthymic precursor T cells in these diseases (Palacios *et al.*, 1981; Alarcón-Segovia *et al.*, 1981; Alarcón-Segovia, 1981; Alarcón-Segovia and Palacios, 1981). Their findings, illustrated in Fig. 2, suggest that the immunoregulatory T-cell circuits in MCTD may be adequate both in postthymic precursor cells and in serum thymic factor, but that there is a defect at the level of the helper T cells either in their signaling to precursor cells for feedback inhibition or in the reception of suppressor signals from suppressor T cells (Alarcón-Segovia and Palacios, 1981). There is a defect in the suppressor T cells that could be due to a decreased stimulus from feedback inhibition

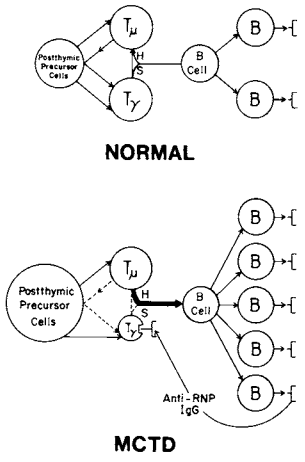


FIG. 2. Schematic representation of immunoregulation in normal subjects and patients with MCTD. Normal postthymic precursor T cells generate T_μ and T_γ cells, which are in equilibrium providing a balanced helper (H) and suppressor (S) function and a normal B cell response in the healthy subject. T_μ cells normally produce a feedback inhibition signal that acts upon the postthymic precursor T cells. This leads to suppressor signals from T_γ cells that help to maintain normal regulation. In the MCTD patient, the pool of postthymic precursor T cells is increased; these cells are qualitatively normal and generate adequate numbers of T_μ cells. Although numerically adequate, the T_μ cells are functionally defective either in their signaling to precursor T cells or in the reception of suppressor signals from the T_γ cells. Decreased generation of T_γ cells and the consequent loss of suppression causes B-cell activation by uncontrolled T_μ helper cells. This results in production of high levels of IgG autoantibodies to RNP. Penetration of RNP antibodies into T_γ cells through their Fc receptors may further accentuate the loss of T_γ cells and abrogate their suppressor function, thus leading to an autoperpetuating phenomenon. (Modified from Alarcón-Segovia and Palacios, 1981.)

or a result of the effect of anti-RNP antibodies, which apparently can penetrate living mononuclear cells through Fc receptors (Alarcón-Segovia *et al.*, 1978, 1979a). These abnormalities of immunoregulatory T-cell circuits in MCTD are different from those found in SLE, Sjögren's syndrome, PSS, and rheumatoid arthritis (Alarcón-Segovia, 1981), supporting the concept that MCTD is a distinct entity (Table I).

3. RNP Antibody Penetration of Living Mononuclear Cells

Alarcón-Segovia's report of RNP antibody penetrating living mononuclear cells is of considerable interest, since there have been few instances in which antibodies have been shown to penetrate living cells. Possibly there is something unique about RNP antibodies, perhaps their enormous

TABLE I

Comparison of Lymphocyte Studies in MCTD with Other Connective Tissue Diseases[a]

	T-Cell numbers				T-Cell functions				
					Suppressor		Generation of suppression	Feed back inhibition	
Disease	Total	T_γ[b]	T_μ	T_p	Spontaneous	Con-A			Help
MCTD	↓	↓	N	↑	↓	↓	↓	↓	PN
SLE	↓	↓	N	↓	N or ↓	↓	↓	↓	PN
PSS	N	N	↓	N	N	N	N	N	↑
RA	N	N	N	N	N	N	N	N	N
SS	N or ↓	N	N	↓	N	N	N	N	N

[a] Modified from Alarcón-Segovia (1981).

[b] Abbreviations and symbols: T_γ, T cells with receptors for Fc portion of IgG; T_μ, T cells with receptors for Fc portion of IgM; T_p, postthymic precursor T cells (autologous rosette forming T cells); Con-A, concanavalin-A-induced suppression; MCTD, mixed connective tissue disease; SLE, systemic lupus erythematosus; PSS, progressive systemic sclerosis; RA, rheumatoid arthritis; SS, primary Sjögren's syndrome; ↓, decreased; ↑, increased; N, normal; PN, probably normal.

quantity compared to other antibodies, that would permit this to occur. In this regard, it is pertinent that there have been several reports of direct speckled nuclear fluorescence noted in tissues from MCTD patients when they were incubated with fluoresceinated anti-human gamma globulin (Alarcón-Segovia *et al.*, 1979b; Gilliam and Prystowsky, 1977; Mundy *et al.*, 1981). There has been controversy as to whether these findings are the result of *in vivo* penetration of cells by RNP antibody or whether the observations are due to an artifact in which high levels of circulating RNP antibodies enter the cells during the processing of nonviable tissue.

4. Other Immunopathological Abnormalities

Other immunological abnormalities occurring in MCTD include reduced serum complement levels in ~25% of patients (Wolfe *et al.*, 1977b), and circulating immune complexes during active disease (Halla *et al.*, 1978; Frank *et al.*, 1983). Fishbein and associates (1977) described the presence of circulating RNP antigen and antibody. Deposition of IgG, IgM, and complement have been observed within muscle fibers, in the sarcolemmal–basement membrane region, in vascular wells, and along the glomerular basement membrane in MCTD (Oxenhandler *et al.*, 1977; Singsen *et al.*, 1978).

Hamburger and associates recently studied 18 patients with MCTD, most of whom had active disease, and found that only 4 patients had defective reticuloendothelial system (RES) Fc-specific immune clearance as measured by clearance of IgG-sensitized erythrocytes, although 17 had circulating immune complexes by Raji cell and Clq binding assays (Frank *et al.*, 1983). The 4 patients with defective clearance had an illness that more resembled typical SLE with antibodies to Sm and/or DNA, higher levels of immune complexes, glomerulonephritis, and severe skin lesions. Renal disease was absent in the 14 patients with normal RES clearance. It is possible that the maintenance of normal RES function in patients with more classic MCTD may enable these patients to clear immune complexes efficiently, precluding serious renal damage. On the other hand, it would seem that the lungs are more susceptible to injury, possibly because of the different nature of the circulating immune complexes in MCTD, or possibly because the pulmonary disease is due to other mechanisms.

In addition to the immunoregulatory abnormalities already described, there is widespread pathological evidence of lymphocytic and plasma cell infiltration of numerous human tissues in MCTD (Singsen *et al.*, 1980a). Lymph nodes are frequently enlarged, sometimes to enormous proportions, and muscle, spleen, salivary glands, and synovium are particularly affected.

D. Purification and Biochemical Characterization of RNP and Sm Antigens

Numerous laboratories are actively pursuing purification and biochemical characterization of RNP, Sm, and other nuclear antigens. In earlier studies Northway and Tan (1972) emphasized the immunological distinctness of these antigens, while Mattioli and Reichlin (1973) showed their close relationship. The subsequent observations of Takano *et al.* (1981), consistent with both of these concepts, indicated that RNP and Sm are distinct antigenic sites on the same macromolecular complex. Biochemical investigations from various laboratories showed that the whole complex containing RNP and Sm activities consists of ~8 polypeptides and the 165-nucleotide RNA designated U1 RNA (Lerner *et al.*, 1981a; Hardin *et al.*, 1982; Takano *et al.*, 1981; White *et al.*, 1981). Polypeptides of ~68,000 and lower molecular weights (referred to as A and C peptides) are associated with the RNP antigenic determinant, while the Sm antigenic determinant is associated with lower molecular weight polypep-

tides, referred to as B/B′, D, and E peptides (Takano *et al.*, 1981; Pettersson *et al.*, 1984).

Purification of the RNP and Sm antigens will facilitate studies of their role in the rheumatic diseases and furthermore may contribute to our understanding of basic cell biology, since Lerner *et al.* recently indicated that small nuclear RNPs may have a key role in the processing of hnRNA to mature mRNA (Lerner *et al.*, 1980; Yang *et al.*, 1981). If RNP antibodies can penetrate living cells, this might result in inhibition of one or more of the specific steps in the RNA-processing pathway.

When compared in evolutionarily distinct organisms, the RNP U1 RNA has a remarkably stable 165-nucleotide-long sequence (Branlant *et al.*, 1980). The human and calf sequences are identical (Agris *et al.*, 1984), and their sequences are only 2 nucleotides different from the rat. Rat, calf, and man have 7 nucleotide differences from that of chicken, and the secondary structure of this sequence has also been determined (Branlant *et al.*, 1980, 1981; Epstein *et al.*, 1981). Agris and colleagues (1984) have shown that RNP antigenicity is predominantly restricted to one-third of the RNP-RNA nucleotide sequence, numbers 40–90, plus two large molecular weight peptides. Complementing this work, Liautard *et al.* (1983) found that the Sm-specific peptides protected the separate 3′, uridine-rich terminus of the U1 RNA. Protein protection of the middle and 3′ terminus of the RNA leaves the 5′ end open for the hypothesized functions of RNP in the processing of messenger RNA (Lerner *et al.*, 1980; Rogers and Wall, 1980).

E. MRL Mouse Autoimmune Disease: An Experimental Animal Model of MCTD?

The study of nuclear acidic protein antigens is enhanced by the availability of the nearly congenic mouse strains MRL/Mp-lpr/lpr (MRL/l) and MRL/Mp-+/+ (MRL/n). The autoimmune disease in MRL/l mice is characterized by massive lymphoproliferation, high levels of circulating immune complexes, arthritis, arteritis, and the production of antibodies to DNA, RNP, and Sm (Eisenberg *et al.*, 1978; Singsen *et al.*, 1980b; Pisetsky *et al.*, 1980,1981; Billings *et al.*, 1982). The fact that these mice produce RNP antibodies, have some similarities to MCTD in their disease, and may have similar immunoregulatory T-cell circuit abnormalities (Alarcón-Segovia and Palacios, 1981) suggests that the MRL mouse may be an animal model of MCTD. The MRL mice, since they produce RNP and Sm antibodies, are also being used for construction of hybridomas to generate monoclonal antibodies to these antigens (Lerner *et al.*, 1981b; Billings *et al.*, 1982; Wise *et al.*, 1982).

F. Vascular Lesions in MCTD and MRL Mice

A significant complication in MCTD may be the proliferative vascular lesion associated with pulmonary hypertension. Using connective tissue stains on lung biopsies of MCTD patients, Hurst *et al.* (1980) noted proliferation of young collagen throughout the intimal area of pulmonary arterioles. Biochemical analysis of these specimens showed that synthesis of type III collagen in MCTD lung is selectively increased, resulting in an abnormal type I:III ratio. Ultrastructural analysis of these vascular lesions also shows a disordered array of collagen fibers, an abundance of intimal ground substance, and perforations of the basal lamina. Similar lesions have also been observed by Hurst *et al.* (1980) in MRL mice. Thus, investigation of this animal model may provide information regarding the pathogenesis of vascular complications in MCTD.

IV. MAJOR CLINICAL CHARACTERISTICS OF MCTD

A. General Features of the Adult Patient

Patients with mixed connective tissue disease range in age from 4 to 80 years, with a mean of 37 years; ~80% of patients are female (Sharp, 1981). MCTD is observed as frequently as scleroderma, but is more common than polymyositis and less frequent than SLE. Typical clinical features of MCTD (Table II) include Raynaud's phenomenon, pulmonary disease, polyarthritis, swollen hands or sclerodactyly, inflammatory muscle disease, and esophageal hypomotility (Sullivan *et al.*, 1984). This long-term, prospective study demonstrated that lymphadenopathy, alopecia, malar rash, serositis, cardiac and renal disease, anemia, and leukopenia are less frequent findings. The increased frequency of esophageal hypomotility, myositis, and pulmonary disease in this report, as compared to some previous investigations, is probably the result of a prospective, systematic, and long-term collection of data in this series.

In some MCTD patients the typical overlapping disease pattern is fully expressed when the patients first come to medical attention. However, as physicians have more fully appreciated the association of RNP antibody with MCTD and as testing for RNP antibody has become more widely available, larger numbers of patients have been identified in an early phase of disease that may or may not become MCTD. During this time minimal symptoms, such as Raynaud's phenomenon, arthralgias, myalgias, and/or swollen hands may be insufficient to make a definitive diagno-

TABLE II
Clinical and Laboratory Characteristics of 34 Patients with MCTD[a]

Characteristic	Number	Percentage
Clinical		
Raynaud's phenomenon	31	91
Pulmonary disease	29	85
Polyarthritis	29	85
Swollen hands/sclerodactyly	29	85
Inflammatory myositis	27	79
Esophageal hypomotility	25	74
Lymphadenopathy	17	50
Alopecia	13	41
Pleuritis	12	35
Malar rash	10	29
Cardiac disease	9	26
Renal disease	9[b]	26[b]
Anemia	8	24
Leukopenia	7	21
Diffuse scleroderma	7	21
Sjögren's syndrome	4	12
Trigeminal neuropathy	2	6
Laboratory		
Positive FANA	34	100
Positive RNP antibody	34	100
Positive rheumatoid agglutinins	20	59
Hypergammaglobulinemia	18	53
Hypocomplementemia	11	32
Positive LE cell test	6	18

[a] From Sullivan *et al.* (1984).
[b] Only 6 patients or 18% had clinical evidence of renal disease.

sis. LeRoy and associates (1980) refer to this constellation as an undifferentiated connective tissue syndrome. Long-term follow-up reveals that in some patients this mild syndrome may persist for years; in others, it may progress to PSS; while yet in others, it may develop additional manifestations, resulting in a diagnosis of MCTD (Gilliam *et al.*, 1977; Grant *et al.*, 1981; Hench *et al.*, 1975; Singen *et al.*, 1977). In the study by Sullivan *et al.* (1984), 60% of their patients initially had more limited clinical involvement and were thought to have rheumatoid arthritis, PSS, SLE, polymyositis, or an undifferentiated connective tissue disease (Table III). By the time of their most recent medical evaluation, 91% had demonstrated typical overlapping features of MCTD, while 9% maintained their undifferentiated status.

TABLE III
Transitions in Mixed Connective Tissue Disease over Longitudinal Study[a]

Disease classification	Number classified at initial medical evaluation	Number classified at latest medical evaluation
Mixed connective tissue disease (MCTD)	14	31
Undifferentiated connective tissue Disease	7	3
Systemic lupus erythematosus	6	0
Rheumatoid arthritis or JRA	4	0
Progressive systemic sclerosis	2	0
Polymyositis	1	0

[a] Sullivan *et al.* (1984).

B. Specific Organ Involvement

1. Skin

The skin in MCTD may be taut, thickened, or shiny, and over time such sclerodermatous changes can become extensive; however, widespread involvement with tightly bound-down skin and contractures seems uncommon (Sharp *et al.*, 1972, 1976). Two-thirds of MCTD patients have swollen hands accompanied by a tapered or sausagelike appearance of the fingers (Farber and Bole, 1976; Parker, 1973; Sharp *et al.*, 1972). Other skin manifestations include lupuslike malar eruptions, chronic discoid lesions, diffuse vasculitic changes, the violaceous heliotrope eyelids and erythema over the knuckles that resemble dermatomyositis, periungual telangiectasia, "squared" telangiectasia over the face and hands, areas of hypopigmentation and hyperpigmentation, and diffuse nonscarring alopecia (Gilliam *et al.*, 1977).

2. Joints

Three-fourths of MCTD patients have arthritis, and almost all have polyarthralgias (Sharp, 1981; Bennett *et al.*, 1978; Halla and Hardin, 1978). The arthritis is often nondeforming, but may be suggestive of rheumatoid arthritis by its erosive and disabling capabilities (Alarcón-Segovia and Uribe-Uribe, 1979; Ramos-Niembro *et al.*, 1979; Singsen *et al.*, 1977; Sharp, 1981). Rheumatoidlike changes were present in 30–35% of patients in some investigations, but these have been primarily limited to involvement of the hands and wrists (Halla and Hardin, 1978). Subcutaneous

nodules have occasionally been found in patients with MCTD (Hench *et al.*, 1975; Sharp *et al.*, 1972).

3. Muscles

Proximal muscle weakness, with or without tenderness, is frequent and may be severe in MCTD (Oxenhandler *et al.*, 1977; Parker, 1973; Sharp *et al.*, 1976). Markedly elevated serum levels of creatine phosphokinase, SGOT, and/or aldolase, myopathic changes on electromyography, and muscle histology with inflammatory changes similar to polymyositis may all be observed in both adults and children (Singsen *et al.*, 1978, 1980a).

4. Esophagus

Esophageal abnormalities and swallowing difficulties are frequent in MCTD (Farber and Bole, 1976; Sharp, 1975; Winn *et al.*, 1976). Systematic studies of 35 consecutive MCTD patients by cine-esophagram and/or manometry revealed dysfunction in 80%; 70% of these had no historical symptoms suggesting esophageal abnormalities (Winn *et al.*, 1976). Decreased amplitude of peristalsis in the distal two-thirds of the esophagus and decreased upper and lower esophageal sphincter pressures are characteristic abnormalities. The severity of measured esophageal dysfunction appears to correlate with disease duration, but not necessarily with expression of symptoms (Sharp, 1981; Singsen *et al.*, 1977; Wolfe *et al.*, 1977b).

5. Heart

Pericarditis may be the only clinical difficulty, and is the most frequent cardiac abnormality reported in MCTD (Singsen *et al.*, 1977; Alpert *et al.*, 1983; Oetgen *et al.*, 1983). The pericarditis is usually corticosteroid responsive. Cardiac involvement appears to be less common than pulmonary disease in adult MCTD, but may be frequent in children (Singsen *et al.*, 1977; Singsen, 1981; Hepburn, 1981). We have prospectively studied 37 adults with MCTD and found that pericarditis and/or pericardial effusion occurred in 10 patients; pulmonary hypertension was present in 10 of 15 patients who underwent cardiac catheterization; and mitral valve prolapse occurred in 9 patients. On cardiovascular examination 78% had abnormalities including murmurs, gallop rhythm, pericarditis, mitral valve prolapse, pulmonary hypertension, and edema. Electrocardiographic abnormalities were noted in 70% and included chamber enlargement, arrhythmias, conduction defects, and ST-T wave abnormalities (Al-

pert *et al.*, 1983). It is not yet clear what prognostic significance this frequent cardiac involvement will have.

6. Lungs

Pulmonary dysfunction is usually clinically silent in the early phases of MCTD, and may go undetected unless detailed evaluations are performed. Although not appreciated initially (Sharp *et al.*, 1972), a 1976 prospective study (Harmon *et al.*) reported that 80% of MCTD patients had evidence of pulmonary disease, including 69% who were asymptomatic. Serial follow-up of those patients who were treated with corticosteroids showed significant clinical and physiological improvement in 12 of 14. Subsequently, however, it has become apparent that pulmonary involvement in MCTD may seriously compromise function, leading to exertional dyspnea and/or pulmonary hypertension, particularly in the later stages of disease (Jones *et al.*, 1978; Rosenberg *et al.*, 1979; Esther *et al.*, 1981; Wiener-Kronish *et al.*, 1981; Sullivan *et al.*, 1984).

We have just completed analysis of the clinical, laboratory, and pathological findings from a prospective, longitudinal evaluation of 34 patients with MCTD (Sullivan *et al.*, 1984). The most frequent serious problem was pulmonary hypertension, with associated proliferative vascular lesions in the lungs and other organs. The most common clinical finding was dyspnea, followed by pleuritic pain and bibasilar rales. Twenty-nine patients had evidence of pulmonary involvement (85%); included in these were 8 of 11 asymptomatic patients (73%) who also had abnormal pulmonary function tests and/or chest films. Single-breath diffusing capacity (DLCO) was abnormal in 73%, vital capacity was reduced in 33%, total lung capacity was low in 41%, and FEV_1 was abnormal in 17%; resting hypoxemia was present in 21% of patients. Thirty percent had chest roentgenographic abnormalities consisting of small irregular opacities that involved predominantly the bases and middle regions. These findings confirmed that pulmonary involvement in MCTD may not only be common, but also may be clinically inapparent until far advanced.

Right heart catheterization with cardiac output determination was also performed on 15 of these patients. Ten had elevations in pulmonary vascular resistance (PVR), and 10 had increased pulmonary artery pressure, but pulmonary wedge pressure was abnormal in only 1 patient. The clinical findings of this study suggested that MCTD patients with features similar to PSS are more likely to develop pulmonary hypertension. It is of additional interest that nailfold capillaroscopy, performed by Dr. H. Maricq in 3 of these patients, showed severe capillary loop changes before pulmonary symptoms were present. All three subsequently devel-

oped pulmonary vascular disease; thus, nailfold capillary microscopy may be helpful in determining which MCTD patients will develop pulmonary hypertension.

7. Kidney

Longer follow-up now suggests that renal disease is more frequent in MCTD than was thought initially. When isolated case reports are excluded, recent studies give a combined incidence of 28% renal involvement, including children (Rao *et al.*, 1976; Bresnihan *et al.*, 1977; Bennett and Spargo, 1977; Singsen *et al.*, 1977). Glomerular deposition of immune complexes may be observed, although the patterns differ from SLE (Fuller *et al.*, 1977; Baldassare *et al.*, 1976). Patients with MCTD occasionally die of progressive renal failure (Bennett and Spargo, 1977).

In our longitudinal study of MCTD, there has been clinical evidence of renal disease in 6 patients, manifested by proteinuria and/or hematuria; 1 developed the nephrotic syndrome (Sullivan *et al.*, 1984). Two renal biopsies showed focal glomerulonephritis. In all of these patients, the renal abnormalities responded well to corticosteroid therapy and progressive renal failure did not occur. At autopsy, 3 patients who never exhibited clinical manifestations of renal disease had proliferative vascular lesions in the kidneys similar to those noted in other organs; the glomeruli were normal in 2 of these patients and mesangial thickening and hypercellularity were noted in the third. In a report by Singsen *et al.* (1980a), 7 of 15 children with MCTD had clinical or histological evidence of renal disease, and Grant *et al.* (1981) noted 1 patient who died from "scleroderma kidney." Observation of comparable levels of circulating immune complexes in MCTD and SLE may be important (Parker and Marion, 1977; Halla *et al.*, 1978), but the clinical and histological findings suggest that vascular lesions may represent a more serious problem than immune complex nephritis in MCTD.

8. Nervous System

Significant neurological deficits are observed in only 10% of patients with MCTD, but in some of these cases it may be the primary clinical concern (Bennett *et al.*, 1978; Sharp, 1975; Wolfe *et al.*, 1977b). Trigeminal neuropathy appears to occur much more frequently in MCTD than in other rheumatic disorders (Sharp *et al.*, 1976), but usually this is of minor clinical consequence. Other problems noted have included aseptic meningitis (Bernstein, 1980), seizures, peripheral neuropathy, organic mental syndromes, "vascular headaches," and cerebral infarction or hemorrhage (Cryer and Kissane, 1978).

9. Miscellaneous Clinical Features

Sjögren's syndrome appears to occur frequently in MCTD (Alarcón-Segovia, 1976; Fraga *et al.*, 1978; Sharp, 1975). Hashimoto's thyroiditis and persistent hoarseness are occasionally observed in children and adults (Wolfe *et al.*, 1977b; Singsen *et al.*, 1980a), and fever and lymphadenopathy occur in about one-third of MCTD patients (Gilliam *et al.*, 1977; Sharp, 1981). Lymph nodes may be massively enlarged and suggest a lymphoma, but when biopsied they reveal only lymphoid hyperplasia. Hepatomegaly and splenomegaly may occur, but serious liver function disturbances are uncommon. The intestinal tract may also be involved, with resultant hypomotility, pseudosacculation, dilatation, malabsorption, sclerosis, and perforation, much as is seen with scleroderma (Cooke and Lurie, 1977; Davis *et al.*, 1978; Samach *et al.*, 1978). In one report, three patients with MCTD had extensive gastrointestinal changes similar to those found in PSS (Norman *et al.*, 1978); this has also been observed in children (Singsen *et al.*, 1977). Several serological features including hypergammaglobulinemia should also be emphasized, and up to 60% of patients will manifest rheumatoid factor, often at very high titers (Sharp, 1981).

C. Pediatric Discriminations

Since 1973 there have been at least 12 reports of pediatric patients with MCTD (Sanders *et al.*, 1973; Singsen *et al.*, 1977). These reports encompass 39 children, with an additional 12 others (defined as $\leq$16 years old) briefly mentioned in other investigations, but in such fashion that features and treatment responses could not be separated from the larger groups of adults. The incidence and prevalence of pediatric MCTD is unknown. In one large pediatric rheumatology referral service, MCTD was one-tenth as frequent as SLE and occurred approximately once for every 100 cases of JRA (Singsen *et al.*, 1977). The median age of onset for the 39 reported children with MCTD was 10.8 years (range, 4–16 years); there have been 29 girls and 10 boys (Peskett *et al.*, 1978; Hepburn, 1981; Singsen, 1981).

MCTD in children and adults is largely similar, although there are several distinguishing pediatric features. (1) Dermatomyositis and SLE-like rashes have been more commonly observed in children (Singsen *et al.*, 1977; Baldassare *et al.*, 1976). (2) Clinical and histological evidence of renal involvement appears more common in childhood MCTD (Oetgen *et al.*, 1981; Jones *et al.*, 1978). In our experience, 5 of 16 pediatric developed significant renal disease, including 2 who progressed to chronic hemodialysis. (3) Thrombocytopenia continues to be observed in ~25% of

children with MCTD (Singsen *et al.*, 1977). (4) Sjögren's syndrome, most often observed as recurrent episodes of "mumps," is quite frequent in children with MCTD (Fraga *et al.*, 1978). (5) There is a higher frequency of cardiac disease, especially pericarditis, reported in childhood MCTD (Silver *et al.*, 1976; Singsen *et al.*, 1977; Jones *et al.*, 1978).

Several additional aspects of pediatric MCTD deserve comment. Deforming but relatively painless arthritis and a high frequency of rheumatoid factor have remained prevalent since their first descriptions in childhood MCTD (Baldassare *et al.*, 1976; Oetgen *et al.*, 1981; Singsen *et al.*, 1977). Since the degree of synovitis is mild to moderate, and complaints of discomfort are not usually marked, responses to treatment may appear better than they actually are. Erosive disease in these children has been infrequent, but flexion contractures and deformity are common.

Since Sjögren's syndrome is generally rare in children with rheumatic disease (Athreya *et al.*, 1977), it is infrequently or incompletely searched for in pediatric MCTD. At least half of the children with MCTD have salivary gland involvement and/or keratoconjunctivitis sicca (Fraga *et al.*, 1978; Singsen *et al.*, 1980a). This suggests that clinical salivary gland evaluation, Schirmer's test, and lip biopsy should be considered in all children with possible MCTD.

In the initial report of Sanders and associates (1973), the 9-year-old patient had absent serum IgA; one other child without serum IgA has also been noted (Peskett *et al.*, 1978). Selective deficiency of serum IgA has also been observed in JRA (Petty *et al.*, 1977); it should continue to be searched for in chidren with MCTD.

An infant with congenital complete heart block (CCHB) has been reported born to a mother with MCTD (Nolan *et al.*, 1979); the child also developed pericardial tamponade and a discoid LE rash. CCHB is now a well-recognized complication of several maternal rheumatic or "pre-rheumatic" illnesses, most commonly SLE (McCue *et al.*, 1977).

In children as in adults, MCTD is frequently a sequential disease (Hench *et al.*, 1975; Grant *et al.*, 1981; Singsen *et al.*, 1977; Sullivan *et al.*, 1984). It is rare for a child to present with all of the overlapping features of MCTD. Rather, the majority manifest Raynaud's phenomenon in association with polyarthritis. Progressively, these patients then develop other disease features of MCTD, but not in any particular order. Serological findings may also change over time, commensurate with the clinical situation. Thus, the absence of FANA and the presence of rheumatoid factor may be associated with early polyarthropathy, while anti-Sm antibodies may appear along with renal disease, hypocomplementemia, and other SLE features (Hepburn, 1981; Singsen *et al.*, 1977). A speckled FANA

will first develop several years after other nuclear patterns are recognized in some of these children.

It should be emphasized that many of the features of pediatric MCTD are intially clinically silent and may remain undetected for long periods of time unless carefully sought. The limited verbal skills, memory, and experiential understanding of children may magnify these findings. These silent or asymptomatic features may include: (1) pulmonary function (PFT) abnormalities, which are found prior to historical complaints or roentgenographic change; (2) esophageal dysfunction; (3) minimal alterations in bowel habits, perhaps related to early gastrointestinal tract alterations; (4) early sclerodermatous skin changes, particularly around the mouth or at the distal digits; (5) salivary gland disease with xerostomia or keratoconjunctivitis sicca; and (6) minimal muscle involvement with mild atrophy, moderate serum muscle enzyme elevations, and perhaps EMG changes, but with no significant proximal weakness. Children most commonly do not complain of dyspnea or shortness of breath. Thus, PFTs should be part of the routine evaluation of any child in whom MCTD is suspected. Pulmonary hypertension and idiopathic fibrosis must be recognized as potentially life threatening in pediatric MCTD (Jones *et al.*, 1978; Rosenberg *et al.*, 1979; Singsen *et al.*, 1977).

D. Histopathology

Isolated biopsy information from early MCTD patients demonstrated skin alterations including increased dermal collagen and edema (Sharp *et al.*, 1972). Other cases with immunofluorescent analysis of skin and muscle demonstrated deposits of IgG or IgM within normal-appearing blood vessels, within normal muscle fibers, around or on sarcoplasmic basement membranes, or within perimysial connective tissues (Oxenhandler *et al.*, 1977; Singsen *et al.*, 1978). Histochemical analysis of these tissues showed perifascicular atrophy and type I fiber predominance.

Lung tissues from early MCTD patients revealed vascular intimal proliferation and medial hypertrophy, interstitial mononuclear infiltrates, and infrequent fibrosis in a few instances (Harmon *et al.*, 1976; Jones *et al.*, 1978; Kennedy, 1976). Early descriptions of MCTD mentioned vascular changes in several of the other tissues available (Cooke and Lurie, 1977; Fuller *et al.*, 1977), but recognition of the diffuse vasculopathic nature of MCTD had not yet developed. This lack of appreciation of the unique vascular changes of MCTD perhaps related to the frequency with which similar alterations are associated with systemic or pulmonary hypertension or with natural aging.

The first comprehensive histological investigation of MCTD reviewed material from three autopsies and five additional renal biopsies (Singsen *et al.*, 1980a). The study group included 15 pediatric MCTD cases with a median age at disease onset of 10.7 years; the three children who died had had disease for 5.4 years prior to histological evaluation. The ages, short disease durations, absence of systemic or pulmonary hypertension, and lack of corticosteroid treatment in young patients are relevant because in older patients all would be variables that could cause or reduce the observed vasculopathy.

Widespread proliferative vascular lesions were the most prominent histopathological feature in these young patients. All three children had diffuse vascular diease; in combination, 31 of 58 organs (53%) had intimal vascular change, and 9 organs (16%) revealed medial vessel wall thickening (Fig. 3). The vascular abnormalities occurred both in large vessels such as the aorta and renal artery and within small myocardial arterioles. The severity and frequency of this obliterative or obstructive vasculopathy in organs without overt evidence of clinical involvement was striking; these included intestinal tract, kidney, aorta, coronary vessels, and myocardium (Singsen *et al.*, 1980a). Inflammatory infiltration of vessels was not a prominent feature. Similar proliferative vascular lesions have now been described in MCTD in adults (Esther *et al.*, 1981; Wiener-Kronish *et al.*, 1981; Sullivan *et al.*, 1984). Thus, the presence of pulmonary hypertension in MCTD appears to be related to the proliferative vascular abnormalities and marked luminal vessel narrowing, but without significant interstitial fibrosis. In marked contrast is the uniform finding of Salerni *et al.* (1977) of pulmonary fibrosis in a series of 350 PSS patients.

The microvasculature in scleroderma has been implicated as one of its primary sites of injury (Norton and Nardo, 1970; Campbell and LeRoy, 1975). Several features of the histology of MCTD, however, appear to distinguish it from scleroderma: (1) there is limited capillary disease in MCTD, in contrast to the frequency of capillary change described in progressive systemic sclerosis; (2) large vessel involvement is a new and striking aspect of MCTD patients that is said to be absent in PSS; (3) the significant fibrosis that is commonly observed in scleroderma lung, heart, and gastrointestinal tract is absent in MCTD; and (4) in contrast to scleroderma, the rarity of fibrinoid change in MCTD appears noteworthy (Campbell and LeRoy, 1975; Singsen *et al.*, 1980a).

It should be noted, however, that widely varying degrees of attention have been directed toward tissue studies of both MCTD and scleroderma. Neither has had longitudinal histological evaluations, and each has so few pathological investigations that it is possible that further study would demonstrate them to be either dramatically similar or dissimilar. The

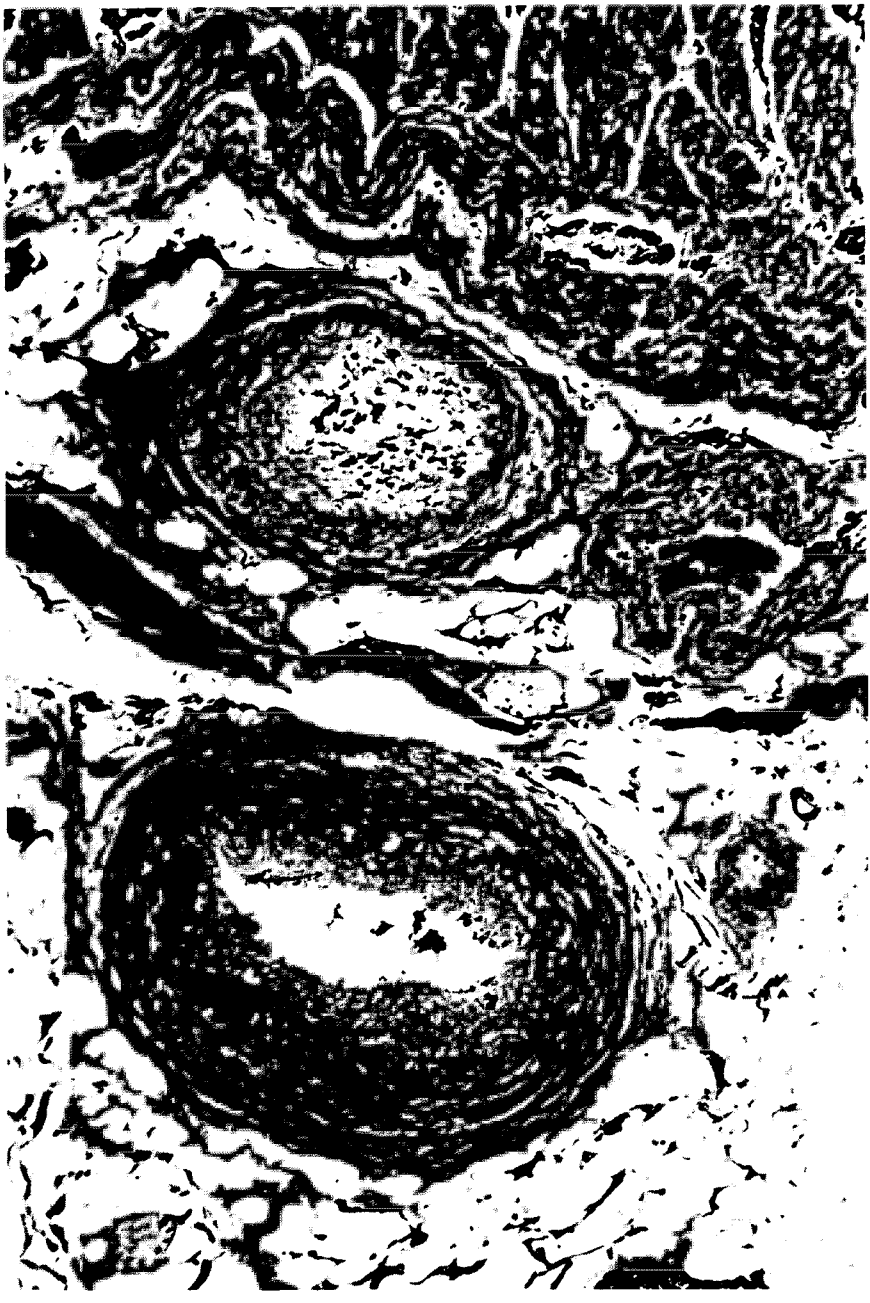


FIG. 3. Two medium-sized arteries from the terminal ilium of a 15-year-old female with MCTD. These arteries demonstrate both intimal and medial proliferation and the luminal narrowing that are characteristic of MCTD.

vascular lesions in either illness will be underestimated if they are not diffusely present or if they are obscured by late atrophy and fibrosis. Since the early serological and clinical characteristics of scleroderma and MCTD can be so markedly different, it is important that we critically describe the histopathological similarities and differences so that points of convergence and divergence can be clearly recognized.

The widespread large and small vessel changes of MCTD are of concern because vascular compromise may lead to hemorrhage (Harmon *et al.*, 1976) and to organ failure (Cooke and Lurie, 1977; Jones *et al.*, 1978; Singsen *et al.*, 1980a). Until now, the most likely organs to be affected by vascular compromise have been the kidneys, lungs, and gastrointestinal tract (Norman and Fleischmann, 1978; Samach *et al.*, 1978).

Initial investigations of MCTD suggested a low frequency of renal involvement, but evidence of membranous and/or diffuse glomerulonephritis may be found in up to 28% of patients in more recent reports (Baldassare *et al.*, 1976; Bennett and Spargo, 1977; Singsen *et al.*, 1980a). Immunofluorescent studies of some cases showed granular deposits of IgG, C3, and C4 in the glomerular basement membrane, and electron microscopy demonstrated electron-dense deposits in subepithelial and deep mesangial regions (Fuller *et al.*, 1977; Jones *et al.*, 1978; Rao *et al.*, 1976). Singsen and co-workers (1980a) described 6 of 15 children who had clinical or histological evidence of renal involvement. Five of six renal tissues demonstrated marked renal vascular sclerosis (87%), two had membranous nephropathy, and the remaining four all had diffuse proliferative glomerulonephritis. Two of these latter have required chronic hemodialysis, and two others have developed severe systemic hypertension. In the series of 308 patients with MCTD reported by Wolfe and associates (1977b), the 20 renal biopsies included 5 that were normal, 5 with mesangial hypercellularity, 5 with focal glomerulitis, 3 with membranous glomerulonephritis, 1 with membranoproliferative glomerulonephritis, and 1 with severe intimal proliferative vascular changes.

Gastrointestinal alterations in MCTD may include atrophy and fibrosis, proliferative vasculopathy, widespread dysmotility, or pneumatosis cystoides intestinalis (Cryer and Kissane, 1978; Norman and Fleischmann, 1978; Samach *et al.*, 1978). These changes may be caused by vascular insufficiency, fibrosis, or muscle layer involvement. A distinctive replacement of the inner and outer muscle layers of the esophagus, pylorus, and colon with some form of hyaline material has been observed in MCTD (Singsen *et al.*, 1980a). These changes are quite distinct from the type of atrophy or fibrosis that has been associated with PSS.

The synovial histopathology of MCTD has been briefly noted in two studies (Bennett *et al.*, 1978; Singsen *et al.*, 1980a). Both described

chronic inflammatory cell infiltrates and cartilage destruction that appeared quite similar to that found with rheumatoid arthritis; one rheumatoid nodule in MCTD was typical of those found in rheumatoid arthritis (Ramos-Niembro *et al.*, 1979).

Other pathological alterations of MCTD that may be unique include the prominent plasmacytic inflammatory infiltrations in salivary glands, liver, and GI tract and the widespread inflammatory myopathy (Fig. 4). Pericarditis and cardiac myopathy have been found in small numbers of children and adults reported with MCTD (Singsen *et al.*, 1980a; Whitlow *et al.*, 1978; Alpert *et al.*, 1983), but their prevalence is not yet known.

Histopathological investigation of MCTD remains at a nascent stage. There is recent evidence suggestive of a widespread but largely silent vasculopathy, inflammatory cardiac and skeletal myopathy, membranoproliferative nephropathy, and widespread pulmonary and gastrointestinal involvement in MCTD. Most of these changes imply a potential for morbidity and mortality greater than that originally observed. Prospective histopathological studies with pre- and posttreatment observations are necessary to determine whether current expectations of MCTD morbidity and mortality are overly optimistic.

E. Differential Diagnosis

The differential features of MCTD versus the other four classical rheumatic syndromes are outlined in Table IV. A total of only 12 clinical and 9 laboratory discriminators is employed in this simplified schematic. It is important to recall that in most reports regarding these illnesses relatively insensitive tests are commonly employed, particularly when investigating esophageal, pulmonary, or muscle function. Thus, the original report of MCTD (Sharp *et al.*, 1972) did not systemically assess these organ systems, and the follow-up by Nimelstein *et al.* (1980) of the same patients largely reassessed only symptoms. Sullivan and associates (1984) have now confirmed that significant functional abnormalities in these systems may precede clinically apparent disease. Therefore, in a patient with RNP antibody and a presumptive diagnosis of MCTD, detailed pulmonary, esophageal, and muscle evaluations may be indicated even in the absence of clinical symptoms.

F. Treatment and Prognosis

Treatment recommendations specifically for MCTD are largely based upon anecdotal information. There are no controlled studies with long-term follow-up, and thus no comparisons of various treatment methodolo-

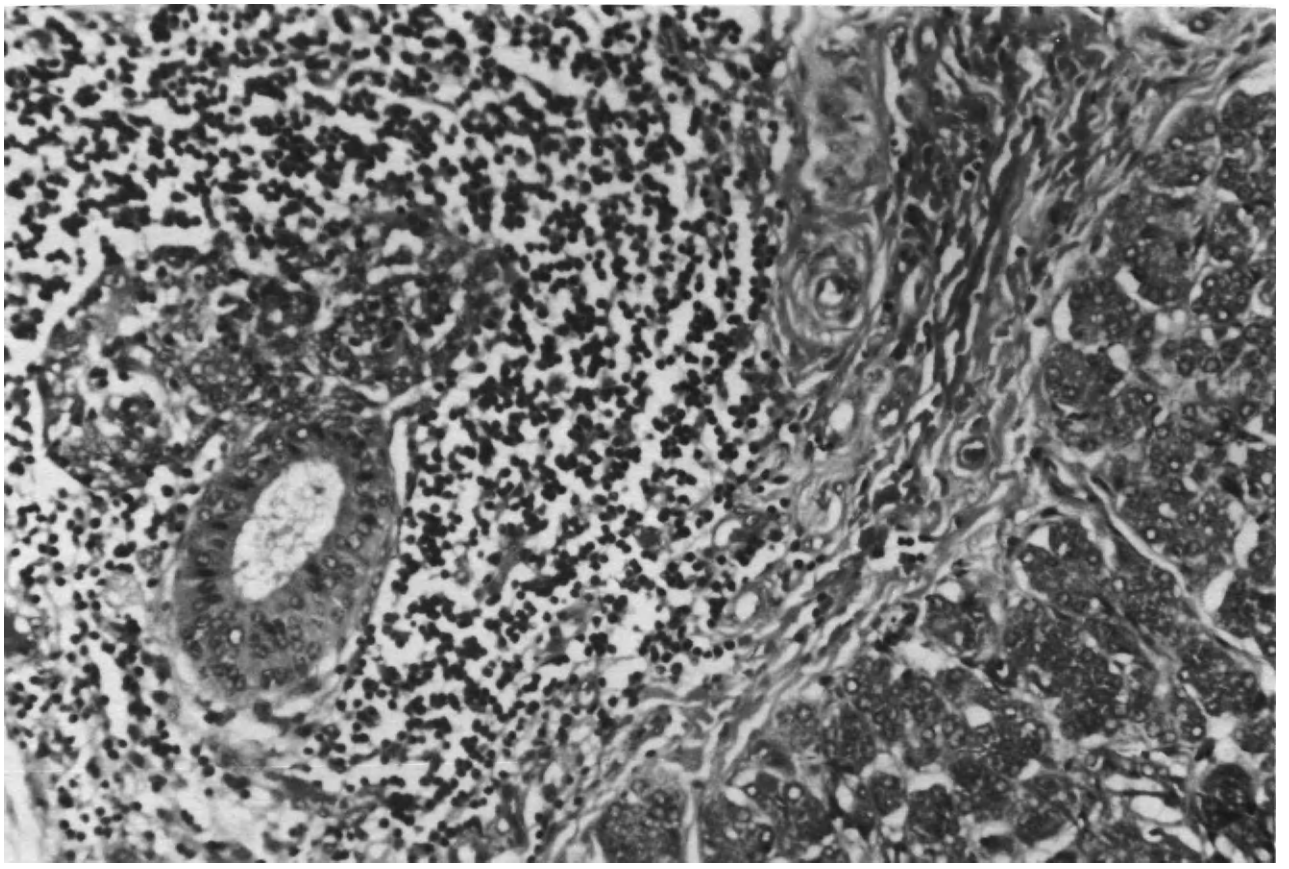


FIG. 4. Plasma cells and lymphocytes infiltrating a parotid gland. There is metaplasia of ductal epithelial cells (myoepithelial islands) in this 14-year-old girl with MCTD.

TABLE IV

Differential Diagnosis of Mixed Connective Tissue Disease

	Mixed connective tissue disease	Systemic lupus erythematosus	Scleroderma	Rheumatoid arthritis	Polymyositis
Raynaud's phenomenon	++++	+	++++	Rare	+
Swollen hands	+++	Rare	+++	Rare	Rare
Polyarthritis	++++	+++	+	++++	+
Esophageal hypomotility	+++	+	+++	0	+
Pulmonary disease	+++	+	++	+	+
Myositis	+++	Rare	+	+	++++
Lymphadenopathy	+++	++	Rare	+	Rare
Leukopenia	++	+++	Rare	Rare	Rare
Thrombocytopenia	+[a]	++	Rare	Rare	Rare
Serious renal disease	+	+++	++	Rare	Rare
Serious CNS disease	+	+++	Rare	Rare	Rare
Diffuse skin sclerosis	+	Rare	++++	0	+
Hypergammaglobulinemia	++++	+++	+	++	+
High titer antibody to RNP	++++	+	Rare	0	Rare
Antibody to native DNA	+	++++	+	Rare	Rare
Antibody to Sm	Rare	+++	0	0	0
Antibody to Scl-70	Rare	Rare	++	Rare	Rare
Antibody to PM-1	Rare	Rare	++[b]	0	+++[b]
Antibody to Ku	Rare	Rare	++[b]	0	+++[b]
Rheumatoid factor	++	+	Rare	+++	Rare
Hypocomplementemia	+	+++	+	+	Rare

[a] More frequent in pediatric MCTD.

[b] Usually a polymyositis–scleroderma overlap.

gies are available. Additionally, since MCTD is an evolutionary disease, very different treatments may be appropriate for the same patient over the duration of several years. For instance, children with MCTD have been diagnosed who initially responded to intramuscular gold for systemic and articular JRA-like features, then improved with corticosteroids when dermatomyositis ensued, and finally required immunosuppressives for significant SLE-like renal involvement (Singsen *et al.*, 1977).

The original descriptions of MCTD (Sharp *et al.*, 1972; Sharp, 1975) suggested that there was a high degree of corticosteroid responsiveness, and that patients who did not respond to corticosteroids in high doses sometimes improved after the addition of cyclophosphamide. Scleroderma-like features of MCTD appeared least likely to respond. However, more extended observations have not supported a uniformly optimistic outlook. In our long-term study (Sullivan *et al.*, 1984), approximately two-thirds of patients had a favorable response to treatment and had mild disease or were in remission at the time of the most recent observation. That MCTD can have a good long-term prognosis is indicated by 13 of 34 (38%) who had sustained inactive disease, including 10 who were off all therapy for 1 to 8 years (mean, 5 years). As previously noted, inflammatory changes such as rashes, arthritis, serositis, myositis, and nephritis were most responsive to corticosteroids, while sclerodermalike involvement, esophageal hypomotility, and pulmonary disease were less likely to respond.

In contrast to the segment of our patients who had a favorable course, 36% have had more severe, less steroid-responsive disease including four patients who died. A major complication of MCTD that is clearly emerging from this and other studies is pulmonary hypertension and widespread proliferative vascular lesions. Pulmonary hypertension in MCTD may or may not respond to combinations of corticosteroids and cytotoxic agents, but no controlled studies have been performed and thus the most effective treatment remains uncertain.

It is useful to consider therapy of MCTD in terms of its patterns of organ system involvement. Arthritis is a frequent feature of MCTD that may be present well before other overlapping disease suggests MCTD (Halla and Hardin, 1978). As recognition of the erosive and deforming potential of MCTD has increased, so has the early addition of remittive agents such as gold and penicillamine to antiinflammatory treatment. In some settings arthritis has been amenable to nonsteroidal antiinflammatory drugs (Bennett and O'Connell, 1978), but others have achieved little success with these medications (Ramos-Niembro *et al.*, 1979). Ibuprofen-related aseptic meningitis in MCTD, perhaps on a hypersensitivity basis, has been described in one case and may suggest caution in the use of this

agent (Bernstein, 1980). When corticosteroids have been used to treat other MCTD manifestations, many physicians have observed a good response of the arthritis, at least during the early inflammatory phase. However, it seems unlikely that the long-term outcome for MCTD arthritis is different when treated with corticosteroids versus other forms of therapy. The erosions and deformities of joints in longstanding MCTD are indications for a comprehensive program similar to that used in rheumatoid arthritis, including education, rest, and physical therapy.

Usually anemia, leukopenia, rash, serositis, fever, and lymphadenopathy will improve within days to weeks of beginning corticosteroids for MCTD (Sharp, 1981). Inflammatory myositis also will usually respond, but may require higher doses (1–2 mg/kg) of corticosteroids for longer periods of time.

The treatment responses of other organ systems involved in MCTD are more difficult to characterize. Pulmonary disease may significantly improve following oral corticosteroids in some cases (Harmon *et al.*, 1976), but whether the improvement reported is predominantly pleural or interstitial and whether it was assessed by clinical, roentgenographical, or pulmonary function criteria is not always evident from available information. Pericarditis and pleuritis in MCTD apparently respond well to corticosteroids (Davis *et al.*, 1978), but interstitial and restrictive lung disease may (Hepburn, 1981) or may not (Cryer and Kissane, 1978). Pulmonary hypertension in MCTD also may (Rosenberg *et al.*, 1979; Wiener-Kronish *et al.*, 1981) or may not (Jones *et al.*, 1978) improve with combinations of corticosteroids and cytotoxic agents. Raynaud's phenomenon and sausage fingers may be reduced by employing corticosteroids (Ellman *et al.*, 1981; Fraga *et al.*, 1978).

The therapeutic response of features of Sjögren's syndrome is not well characterized in MCTD. Some adults and children with salivary gland involvement definitely improve with corticosteroids, but long-term evaluations are not available (Alarcón-Segovia, 1976; Fraga *et al.*, 1978).

It is also not clear how often the gastrointestinal aspects of MCTD will respond to treatment. One investigation revealed improved esophageal motility in 8 of 14 patients following corticosteroid treatment for other disease manifestations (Winn *et al.*, 1976). Several other specific reports of intestinal involvement in MCTD exist, but indications of therapeutic response are absent (Samach *et al.*, 1978; Norman and Fleischmann, 1978; Cooke and Lurie, 1977).

Renal disease in MCTD may be an indication for high-dose corticosteroids, possibly in association with immunosuppressive agents. Kitridou and associates (1980) noted corticosteroid responsiveness in three of six patients, but their report had no details of dosage, treatment duration,

TABLE V
Prognosis in Five Series of MCTD Patients with Sequential Evaluations

Number studied	Deaths	Mean duration of disease (years)	Reference
100	4	6	Sharp *et al.* (1976)
15	4	6.8	Singsen *et al.* (1977, 1980a)
22	8	12	Nimelstein *et al.* (1980)
23	5	—[a]	Grant *et al.* (1981)
34	4	11	Sullivan *et al.* (1984)
194	25 (13%)		

[a] Not stated for whole series; mean duration was 11 years in those who died from MCTD.

or relation of outcome to type of renal lesion. Cytotoxic agents have been utilized in small numbers of patients, but without a uniform outcome experience (Bennett and Spargo, 1977; Rao *et al.*, 1976; Singsen *et al.*, 1977).

Initial prognostic estimations for MCTD suggested a good long-term outcome (Sharp, 1972). A later multicenter study reported that only 4% of 100 MCTD patients had died (Sharp *et al.*, 1976); the mean duration of illness in these adults was 6 years, with a range of 6 months to 22 years. Concern for a possibly more severe prognosis in MCTD first appeared in the pediatric literature (Singsen *et al.*, 1977, 1980a). Particularly worrisome were the frequencies of cardiac involvement (64%), life-threatening thrombocytopenia (43%), and renal disease (47%). Four of 15 children (27%) had succumbed after a mean disease duration of only 5.4 years. The immediate causes of death in these children included infection in 3 and cerebral hemorrhage in 1. The infections were rapidly fatal, caused by encapsulated organisms (pneumococcal in 2 and meningococcal in 1), and were not related to significant corticosteroid doses.

The overall prognosis in MCTD can be projected by reviewing the statistics in five recent series with long-term, sequential evaluation (Table V). It is interesting that in the report by Nimelstein *et al.* (1980), six of the eight deaths were not directly related to rheumatic disease. Pulmonary disease, pulmonary hypertension, and congestive heart failure were major factors in four of the deaths in the series of Grant and associates (1981); the fifth death was due to squamous cell carcinoma of the lung. In the study of Sullivan *et al.* (1984), four deaths were related to MCTD; significant factors included pulmonary hypertension, proliferative vascular lesions, and cor pulmonale in three, pericardial tamponade in one, perfo-

rated bowel in two, and disseminated sepsis in three. The overall mortality in these five studies was 13%, with the mean duration of disease varying from 6 to 12 years.

The collected observations of mobidity and mortality that are developing over time suggest that modification of the early predictions of a generally favorable outcome for MCTD is necessary. In both children and adults, it is becoming clear that the prognosis for MCTD is generally similar to that of SLE, but somewhat better than for scleroderma. Our knowledge of morbidity, outcome, and ideal therapy for MCTD all remain in flux. Treatment protocols, longitudinal prospective assessments, and information sharing regarding this illness will all promote improved future care for MCTD.

V. SUMMARY

Mixed connective tissue disease (MCTD) is a syndrome characterized by overlapping clinical features and high titers of circulating antibody to nuclear RNP. The debate over whether MCTD constitutes a distinct clinical entity still has not been resolved, but a number of recent studies support this possibility. Features that appear to distinguish MCTD from SLE, PSS, and other rheumatic diseases include (1) the serological pattern of extremely high titers of RNP antibody, usually in the absence of significant titers of other antibodies; (2) abnormalities of immunoregulatory T-cell circuits in MCTD that differ from the defects found in other rheumatic diseases; (3) normal RES clearance of immune complexes in most MCTD patients, in contrast to SLE; (4) the high frequency in MCTD of pulmonary hypertension and proliferative vascular lesions (possibly unique) in the lungs and other organs, virtually without fibrosis; and (5) other pathological alterations that may be unique for MCTD (e.g., widespread plasmacytic and lymphocytic infiltration in many tissues, and the distinctive replacement of the inner and outer muscle layers of the esophagus with hyaline material).

In our view, however, the present status of MCTD as a distinct entity is not the most critical issue. As these patients came to be recognized by their clinical and serological characteristics and reports of more long-term observation began to appear, certain repetitive patterns emerged: (1) MCTD tends to evolve gradually, and during the early phases, functional tests and biopsies of tissues often reveal abnormalities even prior to the appearance of clinical symptoms; (2) the limited (e.g., undifferentiated) connective tissue disease that is present initially frequently progresses over time to become more widespread and typical of MCTD, and transi-

tions in clinical and serological patterns are observed during longitudinal study; (3) certain organ systems (e.g., lung, esophagus, muscle) are at a high risk to become involved; (4) many of the disease manifestations are initially responsive to corticosteroids, but the sclerodermalike features, being most resistant to treatment, are apt to be the predominant clinical findings later in the course of MCTD; (5) while the prognosis is relatively favorable in many MCTD patients, a significant percentage will develop serious and sometimes fatal complications, particularly pulmonary hypertension and proliferative vascular lesions.

When it has its onset in childhood, MCTD appears to be a more severe disease. Systematic, long-term studies of adults are however beginning to reveal similar types of clinical features and pathological lesions. The paucity of reports of MCTD patients (children and adults) who are exclusively cared for by their private physicians and never come to tertiary referral centers might suggest that a significant segment of the MCTD population with milder disease is not represented in the medical literature.

The most important result of the continuing controversy of mixed connective tissue disease is the stimulus that it has provided to further investigation. The numerous immunopathological abnormalities demonstrated in MCTD indicate that immune injury mechanisms may play an important role in this autoimmune disease. Further elucidation of these mechanisms may lead to new and more selective treatments for certain rheumatic diseases. It seems likely that a variety of genetic, environmental, immunological, hormonal, and other factors will determine whether patients will continue to have a prolonged benign undifferentiated connective tissue disease, develop major organ system involvement of typical MCTD, evolve into more lupus- or sclerodermalike disease, or develop serious proliferative vascular complications. The MRL mouse animal model, which may be analogous to human MCTD, may permit incisive investigation of some of these factors in ways which would not be possible in patients; already, the MRL mouse has provided a means of producing monoclonal antibodies to RNP, Sm, and other nuclear antigens.

Finally, an unexpected but remarkable consequence of the investigations stimulated by MCTD is the demonstration that RNP, Sm, and other nuclear acidic protein antigens may have important roles in basic cell biology, such as in the processing of messenger RNA. Thus, just as the monoclonal proteins occurring in multiple myeloma facilitated subsequent elucidation of immunoglobulin structure and function, the autoantibodies to RNP, Sm, and other nuclear antigens occurring in MCTD and other rheumatic diseases have stimulated the purification, biochemical characterization, and study of the biological role of these nuclear antigens.

REFERENCES

Agris, P. F., KiKuchi, Y., Gross, H. J., Takano, M., and Sharp, G. C. (1984). *Immunol. Commun.* **13**(2), 137–149.

Akizuki, M., Powers, R., and Holman, H. R. (1977). *J. Clin. Invest.* **59,** 264–272.

Alarcón-Segovia, D. (1976). *J. Rheumatol.* **3,** 191–195.

Alarcón-Segovia, D. (1979). *J. Rheumatol.* **6,** 694–699.

Alarcón-Segovia, D. (1981). *J. Rheumatol.* **8,** 535–540.

Alarcón-Segovia, D., and Palacios, R. (1981). *Arthritis Rheum.* **24,** 1486–1494.

Alarcón-Segovia, D., and Uribe-Uribe, O. (1979). *Arthritis Rheum.* **22,** 1013–1018.

Alarcón-Segovia, D., Ruiz-Arguelles, A., and Fishbein, E. (1978). *Nature (London)* **271,** 67–69.

Alarcón-Segovia, D., Ruiz-Arguelles, A., and Llorente, L. (1979a). *J. Immunol.* **122,** 1855–1863.

Alarcón-Segovia, D., Ruiz-Arguelles, A., and Fishbein, E. (1979b). *Clin. Exp. Immunol.* **35,** 364–375.

Alarcón-Segovia, D., Palacios, R., and Ibanez de Kasep, G. (1981). *J. Clin. Lab. Immunol.* **5,** 143–148.

Alpert, M. A., Goldberg, S., Singsen, B. H., Durham, J. B., Sharp, G. C., Ahmad, M., Hurst, D. P., Madigan, N. P., and Sullivan, W. D. (1983). *Circulation.* **68,** 1182–1193.

Alspaugh, M., and Maddison, P. (1979). *Arthritis Rheum.* **22,** 796–798.

Alspaugh, M. A., and Tan, E. M. (1975). *J. Clin. Invest.* **55,** 1067–1073.

Alspaugh, M. A., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 711–719.

Alspaugh, M. A., Talal, N., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 216–222.

Athreya, B. H., Norman, M. E., Myers, A. R., and South, M. A. (1977). *Pediatrics* **59,** 931–938.

Baldassare, A., Weiss, T., Auclair, R., and Zuckner, J. (1976). *Arthritis Rheum.* **19,** 788.

Bennett, R. M. (1982). *Intern. Med. Special.* **3,** 40–53.

Bennett, R. M., and O'Connell, D. J. (1978). *Ann. Rheum. Dis.* **37,** 397–403.

Bennett, R. M., and Spargo, B. H. (1977). *Am. J. Med.* **63,** 534–541.

Bennett, R. M., Bong, D. M., and Spargo, B. H. (1978). *Am. J. Med.* **65,** 955–962.

Bernstein, R. F. (1980). *Ann. Intern. Med.* **92,** 206–207.

Billings, P. B., Allen, R. W., Jensen, F. C., and Hoch, S. O. (1982). *J. Immunol.* **128,** 1176–1180.

Branlant, C., Krol, A., Ebel, J., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J., and Jeanteur, P. (1980). *Nucleic Acids Res.* **8,** 4143–4154.

Branlant, C., Krol, A., and Ebel, J. (1981). *Nucleic Acids Res.* **9,** 841–856.

Bresnihan, B., Bunn, C., Snaith, M. L., and Hughes, G. R. V. (1977). *Br. Med. J.* **1,** 610–611.

Campbell, P. M., and LeRoy, E. C. (1975). *Semin. Arthritis Rheum.* **4,** 351–368.

Christian, C. L. (1979). *Sci. Pract. Clin. Med.* **4,** 154–158.

Cooke, C. L., and Lurie, H. I. (1977). *Arthritis Rheum.* **20,** 1421–1427.

Cryer, P. F., and Kissane, J. M., eds. (1978). *Am. J. Med.* **65,** 833–842.

Davis, J. D., Parker, M. D., and Turner, R. A. (1978). *J. Rheumatol.* **5,** 96–98.

Dubois, E. L. (1974). *In* "Lupus Erythematosus" (E. L. Dubois, ed.), pp. 464–476. Univ. of Southern California Press, Los Angeles.

Eisenberg, R. A., Tan, E. M., and Dixon, F. J. (1978). *J. Exp. Med.* **147,** 582–587.

Ellman, M. H., Pachman, L., and Medof, M. E. (1981). *J. Rheumatol.* **8,** 632–634.

Epstein, P., Reddy, R., and Busch, H. (1981). *Biochemistry* **78,** 1562–1566.

Esther, J. H., Sharp, G. C., Agia, G., and Hurst, D. J. (1981). *Arthritis Rheum.* **24,** S105.

Farber, S. J., and Bole, G. G. (1976). *Arch. Intern. Med.* **136,** 425–431.
Fishbein, E., Alarcón-Segovia, D., and Ramos-Niembro, F. (1977). *Proc. XIV Int. Congr. Rheumatol., San Francisco,*
Fraga, A., Gudino, J., Ramos-Niembro, F., and Alarcón-Segovia, D. (1978). *Am. J. Dis. Child.* **132,** 263–265.
Frank, M. M., Lawley, T. J., Hamburger, M. I., and Brown, E. J. (1983). *Ann. Intern. Med.* **98,** 206–218.
Fuller, T. J., Richman, A. V., Auerbach, D., Alexander, R. W., Lottenberg, R., and Longley, S. (1977). *Am. J. Med.* **62,** 761–764.
Gilliam, J. N., and Prystowsky, S. D. (1977). *Arch. Dermatol.* **113,** 538–587.
Grant, K. D., Adams, L. E., and Hess, E. V. (1981). *J. Rheumatol.* **8,** 587–598.
Halla, J. T., and Hardin, J. G. (1978). *Arthritis Rheum.* **21,** 497–503.
Halla, J. T., Schrohenloher, R. E., Hardin, J. G., and Volanakis, J. E. (1978). *Arthritis Rheum.* **21,** 562–563.
Hardin, J. A., Rahn, D. R., Shen, C., Lerner, M. R., and Steitz, J. A. (1982). *J. Clin. Invest.* **70,** 141–147.
Harmon, C., Wolfe, J. F., Lillard, S., Held, C., Cordon, R., and Sharp, G. C. (1976). *Arthritis Rheum.* **19,** 801.
Hench, P. K., Edgington, T. S., and Tan, E. M. (1975). *Arthritis Rheum.* **18,** 404.
Hepburn, B. (1981). *J. Rheumatol.* **8,** 635–638.
Holman, H. R. (1965). *Ann. N. Y. Acad. Sci.* **124,** 800–806.
Hurst, D. J., Baker, W. M., and Gilbert, G. (1980). *Clin. Res.* **28,** 744.
Hutchinson, J. (1880). *Br. Med. J.* **1,** 650–652.
Jones, M. B., Osterholm, R. K., Wilson, R. B., Martin, F. H., Commers, J. R., and Bachmayer, J. D. (1978). *Am. J. Med.* **65,** 855–863.
Kennedy, J. M. (1976). *Ariz. Med.* **33,** 192–193.
Kitridou, R. C., Akmal, M., Ehresmann, G. R., Quisomorio, F. P., and Massry, S. (1980). *Arthritis Rheum.* **23,** 704.
Koffler, D., Schur, P. H., and Kunkel, H. G. (1967). *J. Exp. Med.* **126,** 607–624.
Kurata, N., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 574–580.
Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., and Steitz, J. A. (1980). *Nature (London)* **283,** 220–224.
Lerner, M. R., Boyle, J. A., Hardin, J. A., and Steitz, J. A. (1981a). *Science* **211,** 400–402.
Lerner, E., Lerner, M., Janeway, C., and Steitz, J. (1981b). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 2737–2741.
LeRoy, E. C., Maricq, H. R., and Kahaleh, M. B. (1980). *Arthritis Rheum.* **23,** 341–343.
Liautard, J. P., Sri-Widada, J., Brunel, C., and Jeanteur, P. (1983). *J. Mol. Biol.* **162,** 623–643.
Lindberg, D. A. B., Kingsland, L. C., III, Roeseler, G. C., Kay, D. R., and Sharp, G. C. (1982). *Proc. 1st AMIA Congr. Med. Informa., San Francisco, California,* pp. 299–303.
Lindberg, D. A. B., Sharp, G. C., Kay, D. R., Kingland, L. C., III, Roeseler, G., Kulikowski, C. A., and Weiss, S. M. (1983). *Mobius* **3,** 30–36.
Maddison, P. J., and Reichlin, M. (1977). *Arthritis Rheum.* **20,** 819–824.
Mattioli, M., and Reichlin, M. (1973). *J. Immunol.* **110,** 1318–1324.
McCue, C. M., Mantakas, M. E., Tingelstad, J. B., and Ruddy, S. (1977). *Circulation* **56,** 82–90.
Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S., and Homma, M. (1981). *J. Clin. Invest.* **68,** 611–620.
Munday, T. M., Landing, B. H., Hanson, V., Bernstein, B., and King, K. (1981). *Arthritis Rheum.* **24,** S101.

Nimelstein, S. H., Brody, S., McShane, D., and Holman, H. R. (1980). *Medicine (Baltimore)* **59,** 239–248.
Nishikai, M., and Reichlin, M. (1980). *Mol. Immunol.* **17,** 1129–1141.
Nolan, R. J., Shulman, S. T., and Victorica, B. E. (1979). *J. Pediatr. (St. Louis)* **95,** 420–422.
Norman, D. A., and Fleischmann, R. M. (1978). *Arthritis Rheum.* **21,** 811–819.
Northway, J. D., and Tan, E. M. (1972). *Clin. Immunol. Immunopathol.* **1,** 140–154.
Norton, W. L., and Nardo, J. M. (1970). *Ann. Intern. Med.* **731,** 317–324.
Notman, D. D., Kurata, N., and Tan, E. M. (1975). *Ann. Intern. Med.* **83,** 464–469.
Oetgen, W. J., Boice, J. A., and Lawless, O. J. (1981). *Pediatrics* **67,** 333–337.
Oetgen, W. J., Mutter, M. L., Lawless, O. J., and Davia, J. E. (1983). *CHEST* **2,** 185–188.
Oxenhandler, R., Hart, M., Corman, L., Sharp, G. C., and Adelstein, E. (1977). *Arthritis Rheum.* **20,** 985–988.
Palacios, R., Alarcón-Segovia, D., Llorente, L., Ruiz-Arguelles, A., and Fishbein, E. (1981). *J. Clin. Lab. Immunol.* **5,** 71–80.
Parker, M. D. (1973). *J. Lab. Clin. Med.* **82,** 769–775.
Parker, M. D., and Marion, T. (1977). *Arthritis Rheum.* **20,** 130.
Peskett, S. A., Ansell, B. M., Fizzman, P., and Howard, A. (1978). *Rheum. Rehab.* **17,** 245–248.
Pettersson, I., Hinterberger, M., Mimori, T., Gottlieb, E., and Steitz, J. A. (1984). *J. Biol. Chem.* **259,** 5907–5914.
Petty, R. E., Cassidy, J. T., and Sullivan, D. B. (1977). *Arthritis Rheum.* **20,** 260–267.
Pisetsky, D. S., McCarty, G. A., and Peters, D. V. (1980). *J. Exp. Med.* **152,** 1302–1310.
Pisetsky, D. S., Stewart, P., Roths, J. B., and Murphy, E. D. (1981). *J. Supramol. Struct. Cell Biochem., Suppl.* **5,** 89.
Ramos-Niembro, F., Alarcón-Segovia, D., and Hernandez-Ortiz, J. (1979). *Arthritis Rheum.* **20,** 43–51.
Rao, K. V., Berkseth, R. O., Crosson, J. T., Raij, L., and Shapiro, F. L. (1976). *Ann. Intern. Med.* **84,** 174–176.
Reichlin, M. (1976). *In* "Modern Topics in Rheumatology" (E. R. V. Hughes, ed.), pp. 157–162. Yearbook Medical Publ., Chicago.
Reichlin, M. (1979). *In* "Immunopathology of the Skin" (E. H. Beutner, ed.), 2nd ed., pp. 330–331. Wiley, New York.
Reichlin, M., and Mattioli, M. (1972). *N. Engl. J. Med.* **286,** 908–911.
Rogers, J., and Wall, R. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 1877–1879.
Rosenberg, A. M., Petty, R. E., Cumming, G. R., and Koehler, B. E. (1979). *J. Rheumatol.* **6,** 700–704.
Salerni, R., Rodnan, G. P., Leon, D. F., and Shover, J. A. (1977). *Ann. Intern. Med.* **86,** 394–399.
Samach, M., Brandt, L. J., and Bernstein, L. H. (1978). *Am. J. Gastroenterol* **69,** 494–500.
Sanders, D. Y., Huntley, C. C., and Sharp, G. C. (1973). *J. Pediatr. (St. Louis)* **83,** 642–645.
Scopelitis, E., Biundo, J. J., and Alspaugh, M. A. (1980). *Arthritis Rheum.* **23,** 287–293.
Sharp, G. C. (1975). *Clin. Rheum. Dis.* **1,** 561–572.
Sharp, G. C. (1979). *In* "Arthritis and Allied Conditions" (D. J. McCarty, ed.), 9th ed., pp. 737–741. Lea & Febiger, Philadelphia, Pennsylvania.
Sharp, G. C. (1981). *In* "Textbook of Rheumatology" (W. N. Kelly, E. D. Harris, Jr., and S. Ruddy, eds.), Vol. II, pp. 1151–1161. Saunders, Philadelphia, Pennsylvania.
Sharp, G. C. (1982). *Am. J. Kid. Dis.* **2,** 201–205.
Sharp, G. C., Irvin, W. S., Tan, E. M., Gould, R. G., and Holman, H. R. (1972). *Am. J. Med.* **52,** 148–159.

Sharp, G. C., Irvin, W. S., May, M. C., Holman, H. R., McDuffie, F. C., Hess, E. V., and Schmid, F. R. (1976). *N. Engl. J. Med.* **295,** 1149–1154.

Silver, T. M., Farber, S. J., Bole, G. G., and Martel, W. (1976). *Radiology* **120,** 269–275.

Singsen, B. H. (1981). *In* "Practice of Pediatrics" (V. C. Kelley, ed.), Vol. 3, pp. 1–17. Harper & Row, Philadelphia, Pennsylvania.

Singsen, B. H., Bernstein, B. H., Kornreich, H. K., King, K. K., and Hanson, V. (1977). *J. Pediatr. (St. Louis)* **90,** 893–900.

Singsen, B. H., Landing, B., Wolfe, J. F., Bernstein, B., Oxenhandler, R. W., Sharp, G. C., and Hanson, V. (1978). *Arthritis Rheum.* **21,** 593.

Singsen, B. H., Swanson, V. L., Bernstein, B. H., Heuser, E. T., Hanson, V., and Landing, B. H. (1980a). *Am. J. Med.* **68,** 710–717.

Singsen, B. H., Olive, P. M., Luger, A. M., Benage, D. D., Takano, M., Walker, S. E., and Sharp, G. C. (1980b). *Arthritis Rheum.* **23,** 747.

Sullivan, W. D., Hurst, D. J., Harmon, C. E., Ester, J. H., Agia, G. A., Maltby, J. D., Lillard, S. B., Held, C. N., Wolfe, J. F., Sunderrajan, E. V., Maricq, H. R., and Sharp, G. C. (1984). *Medicine (Baltimore)* **63,** 92–107.

Takano, M., Golden, S. S., Sharp, G. C., and Agris, P. F. (1981). *Biochemistry* **21,** 5929–5935.

Tan, E. M., and Kunkel, H. G. (1966). *J. Immunol.* **96,** 464–471.

Tan, E. M., Schur, P. H., Carr, R. I., and Kunkel, H. G. (1966). *J. Clin. Invest.* **45,** 1732–1740.

Tan, E. M., Rodnan, G. P., Garcia, I., Moroi, Y., Fritzler, M. J., and Peebles, C. (1980). *Arthritis Rheum.* **23,** 617–625.

Treadwell, E. L., Alspaugh, M. A., and Sharp, G. C. (1984). *Arthritis Rheum.* **27,** 1263–1271.

Wasicek, C. A., and Reichlin, M. (1982). *J. Clin. Invest.* **69,** 835–843.

White, P. J., Gardner, W. D., and Hoch, S. O. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 626–630.

Whitlow, P. L., Gilliam, J. N., and Chubick, A. (1978). *Am. J. Med.* **65,** 955–962.

Wiener-Kronish, J. P., Solinger, A. M., Warnock, M. L., Churg, A., Ordonez, N., and Golden, J. A. (1981). *Am. Rev. Respir. Dis.* **124,** 499–503.

Winn, D., Gerhardt, D., Winship, D., and Sharp, G. C. (1976). *Clin. Res.* **24,** 545A.

Winn, D. M., Wolfe, J. F., Harmon, D., and Sharp, G. C. (1979a). *J. Clin. Invest.* **64,** 820–823.

Winn, D. M., Wolfe, J. F., Lindberg, D. A. B., Fristoe, F. H., Kingsland, L., and Sharp, G. C. (1979b). *Arthritis Rheum.* **22,** 1334–1337.

Wise, K. S., Takano, M., and Sharp, G. C. (1982). *Arthritis Rheum.* **25,** S16.

Wolfe, J. F., Adelstein, E., and Sharp, G. C. (1977a). *J. Clin. Invest.* **59,** 176–178.

Wolfe, J. F., Kingsland, L., Lindberg, D., and Sharp, G. C. (1977b). *Clin. Res.* **25,** 488A.

Yang, V. W., Lerner, M. R., Steitz, J. A., and Flint, S. J. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 1371–1375.

CHAPTER 5

Scleroderma

DONATO ALARCÓN-SEGOVIA
Department of Immunology and Rheumatology
Instituto Nacional de la Nutrición "Salvador Zubirán"
México, D.F., México

I. INTRODUCTION

Scleroderma is not a single disease, but a group of diseases whose main feature is sclerosis of the skin due to collagen deposition in the superficial and deep dermis. Diseases with this clinicopathological feature are often pathogenetically related, and in most of them, an immunological derangement seems to be involved.

Although the clinical definition of scleroderma dates back to the mid-nineteenth century (Benedek and Rodnan, 1982), it was not until it was included among the collagen diseases with systemic lupus erythematosus by Klemperer (1950) that interest in these conditions was aroused. The acumen of Klemperer became evident when antinuclear antibodies were

found in both diseases [see review by Jablonska *et al.* (1974) and Alarcón-Segovia (1975)].

The description of mixed connective tissue disease (MCTD) by Sharp *et al.* (1972) as a condition in which features previously considered characteristic of systemic lupus erythematosus, polymyositis, and rheumatoid arthritis as well as scleroderma converge in the presence of high titers of antinuclear ribonucleoprotein antibody called attention to the aforementioned statement that scleroderma is a manifestation of various diseases rather than a disease proper. The term scleroderma, however, is frequently given, with the connotation of disease entity, to the diffuse form of scleroderma that affects multiple internal organs in addition to affecting the skin. Because of this, Goetz (1945) proposed that this form of diffuse scleroderma be called progressive systemic sclerosis (PSS), a term that we will use here for this particular condition.

II. CLASSIFICATION

Diseases and syndromes in which scleroderma is a manifestation are many, and they vary from the localized to the systemic. A classification of scleroderma is presented in Table I.

Most forms of scleroderma are idiopathic whether primary or associated with or secondary to other diseases. In some of those considered under this heading, a mechanism for the development of scleroderma has been partially unraveled. This may be the case in the sclerodermalike skin changes of juvenile diabetes mellitus (Lundbaek, 1957; Garza-Elizondo *et al.*, 1983), where nonenzymatic glycosylation of collagen may alter the packing, cross-linking, and turnover of collagen. Likewise, the sclerodermalike changes in the carcinoid syndrome seem to depend on the fibrosing properties of serotonin (Zarafonetis *et al.*, 1958) and have been found to occur after L-5-OH-tryptophan administration (Sternberg *et al.*, 1980). Other forms of scleroderma, herein called reactive, may result from contact with chemical agents, may occur following infection, or may result from immunological phenomena like the graft-versus-host reaction (Spielvogel *et al.*, 1977). Primary scleroderma may be localized either as morphea or as linear scleroderma; the latter includes the *en coup de sabre* form that causes facial hemiatrophy. Two main categories conform primary PSS; the diffuse scleroderma form in which scleroderma is found proximal to the metacarpophalangeal joints whether acral (acrosclerosis) or generalized, and the CREST syndrome, an acronym derived from Calcinosis, Raynaud's phenomenon, Esophageal dysfunction, Sclerodactily, and Telangectasia (Winterbauer, 1964), in which, at least in the early

Table I
Classification of Scleroderma

I. Idiopathic Scleroderma
  A. Localized, primary
    1. Morphea
    2. Linear scleroderma
  B. Progressive systemic sclerosis, primary
    1. Diffuse scleroderma
    2. CREST syndrome
  C. Scleroderma in MCTD
  D. Scleroderma in overlap with other diseases
    1. Immunological
      a. Sclerodermatomyositis
      b. Primary biliary cirrhosis
      c. Amyloidosis
    2. Metabolic
      a. Porphyria
      b. Insulin-dependent diabetes mellitus
      c. Scleromyxedema
    3. Tumoral
      a. Carcinoid syndrome (and L-5-hydroxytryptophan administration)
      b. Bronchoalveolar carcinoma
    4. Inherited
      a. Progeria and Werner's syndrome
      b. Phenylketonuria
      c. Brandywine triracial isolate
    5. Other
      a. Fasciitis with eosinophilia
II. Reactive Scleroderma
  A. Chemically induced scleroderma
    1. Polyvinyl chloride-related
    2. Bleomycin-induced
    3. Pentazocine-induced fibrosis
    4. Epidemic toxic denatured rapeseed oil syndrome
    5. Silicone-induced (human adjuvant) disease
  B. Postinfections
    1. Scleroderma adultorum of Buschke
  C. Immunological
    1. Graft-versus-host reaction

stages, scleroderma is limited to the fingers and face. There is evidence both for and against the notion of the identity of the two. Their immunological differences and similarities are discussed in Section V.

Scleroderma may be a prominent component of MCTD (Sharp *et al.*, 1972), and in long-term follow-up, may remain as its main manifestation (Nimelstein *et al.*, 1980). There is, however, ample evidence that the

immunological disturbance that takes place in MCTD is quite different from that occurring in PSS (Alarcón-Segovia, 1983).

Scleroderma may occur in overlap with other diseases but whether these associations constitute actual distinct entities, as seems to be the case with MCTD, or are the result of common pathogenetic mechanisms or mere coincidence, remains to be determined. Of particular interest to the readers of this volume are the sclerodermatomyositis, in which a peculiar antigen–antibody system, the Ku–anti-Ku, has been described (Mimori *et al.*, 1981) and the association of scleroderma and primary biliary cirrhosis (Reynolds *et al.*, 1971), an occurrence too common to be coincidental. Like L-5-hydroxytryptophan, when used in conjunction with carbidopa (Sternberg *et al.*, 1980) other drugs and chemicals can result in scleroderma. Included among these drugs are pentazocine and bleomycin (Finch *et al.*, 1980). Workers in the plastics industry in contact with polyvinyl chloride (Maricq *et al.*, 1976), women subjected to cosmetic surgery with silicone implants (Kumagai *et al.*, 1984), and Spaniards who consumed adulterated rapeseed oil (Olmedo-Garzóu *et al.*, 1982) have developed diseases with features of scleroderma.

Chronic graft-versus-host reactions may yield changes akin to those of scleroderma (Spielvogel *et al.*, 1977), and may constitute a model of scleroderma (Jaffe and Claman, 1983). Other models of scleroderma include the spontaneous inherited disease of White Leghorn chickens (Gershwin *et al.*, 1981). Each of these may in turn be models of different diseases manifesting themselves with scleroderma.

III. CLINICAL PRESENTATIONS

Primary progressive systemic sclerosis is a connective tissue disease that is characterized clinically by induration and thickening of the skin that may be diffuse or limited to distal portions (acrosclerosis), by abnormalities of the vasculature, both at the small vessels causing digital infarcts and at larger vessels causing Raynaud's phenomenon, and by organ involvement that is characterized by dysmotility and distention of the gastrointestinal tract, by pulmonary involvement causing fibrosis, by cardiac involvement causing heart failure, and by renal involvement causing hypertension and/or renal failure. Patients with progressive systemic sclerosis (PSS) develop a characteristic facies due to skin tightening (Fig. 1). A variant of PSS is the CREST syndrome where the five manifestations forming the acronym occur with a more protracted course. Scleroderma in the CREST syndrome is confined to the fingers and the face, and

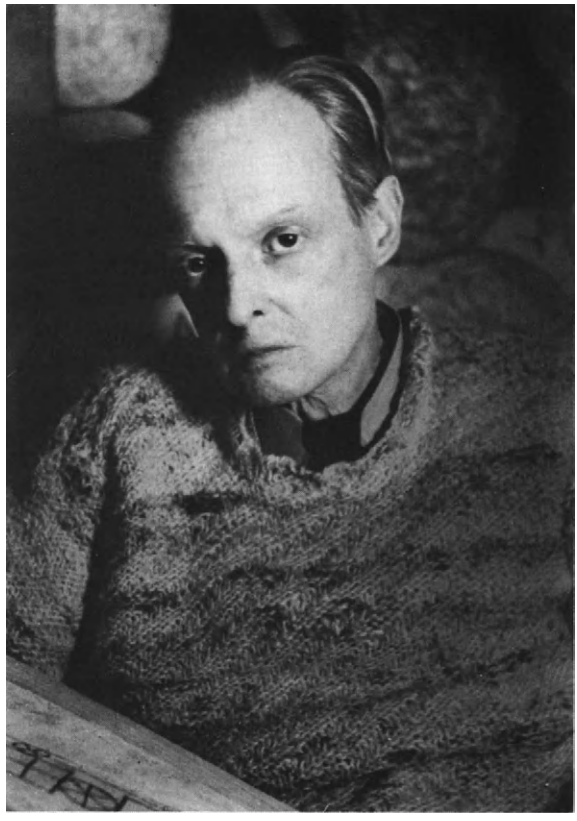


FIG. 1. Facial features of scleroderma as developed by Paul Klee. In 1935, Klee developed progressive systemic sclerosis and was to die of its pulmonary complications in 1940. This photograph of him was taken in December, 1939, 5 months before he died. Despite being severely ill and often bed-ridden, he continued to have a remarkable pictorial output that has given contemporary painters innumerable guidelines. (Reproduced with the permission of Felix Klee, to whom I am grateful).

internal manifestations other than the esophageal usually take many years to appear but are not by any means absent. Indeed, severe pulmonary hypertension or renal involvement may occur also in CREST syndrome patients (Steen *et al.*, 1984).

The course of PSS is usually but not always progressive, and the progression may be slow both in the diffuse and the CREST forms. Excellent reviews of the clinical features of systemic sclerosis are those of Rodnan (1978) and LeRoy (1981). Preliminary criteria for the classification of PSS

have been drawn (Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapuetic Criteria Committee, 1980).

Mixed connective tissue disease has scleroderma features most often confined to the hands where it can coexist with edema of the hands, a cardinal feature of this disease (Alarcón-Segovia, 1981). These patients, most of whom are females, also have features previously considered characteristic of systemic lupus erythematosus, including butterfly rash and frontal hair shortening; features of dermatomyositis, including muscle weakness and heliotrope lid rash; a peripheral arthritis only second in its destructive potential to rheumatoid arthritis (Ramos-Niembro *et al.*, 1979); Raynaud's phenomenon; and Sjögren's syndrome (Alarcón-Segovia, 1976). This disease could be distinguished as an entity because of the presence of high serum titers of anti-nRNP antibody.

IV. HISTOPATHOLOGY

Involved skin has increased compact collagen in the reticular dermis that extends inwards below the skin adnexae well into the subcutis. The epidermis shows loss of rete pegs, atrophy of dermal appendages, and hyalinization of arterioles. Lymphocyte accumulation may be present, consisting mostly of T lymphocytes (Kondo *et al.*, 1976).

Arteries show intimal thickening that may be responsible for both the Raynaud's phenomenon and the digital infarcts peripherally and the disease of internal organs such as the lungs, heart, and kidneys (D'Angelo *et al.*, 1969). In the synovium, there may be lymphocytic infiltration as well as fibrosis that may become quite thick.

Histological changes in the esophagus usually follow the distribution found in esophageal motility studies, with more marked changes in the lower portions. There may be thinning of the mucosa and increased collagen deposition in the submucosa and lamina propria with scarring of the muscularis. There may be anteriolar intimal thickening. Similar changes are found in the intestinal tract. Increased collagen has been found surrounding Brunner's glands as a sign of scleroderma on peroral small bowel biopsy (Rosson and Yesner, 1965).

Pulmonary findings are mainly those of an intense interstitial and alveolar fibrosis with intimal thickening of small pulmonary vessels. There is no necessary correlation between the degree of interstitial fibrosis and the extent of pulmonary vascular changes (Rodnan, 1978). Renal lesions consist mainly of intimal hyperplasia of intralobular and smaller arteries with fibrinoid necrosis of arterioles and glomerular tufts.

Adventitial fibrosis may distinguish the findings in PSS from those made in malignant hypertension of other causes. There are deposits of IgM, complement and, antinuclear antibodies found in immunohistochemical and elution studies (McCoy *et al.*, 1976).

At autopsy, interstitial myocardial fibrosis may be found in irregular patches that may extend to replace myocardial fibers and involve the conduction system (Ridolfi *et al.*, 1976). Small coronary arteries may show intimal thickening but larger coronary arteries are usually spared. Salivary glands are frequently involved, both with lymphocytic infiltration and ductular changes identical to those of Sjögren's syndrome and by peri- and intraglandular collagen deposition (Alarcón-Segovia *et al.*, 1974).

V. IMMUNOLOGY OF SCLERODERMA

A. Immunological Features of PSS

As compared to other connective tissue diseases such as SLE, Sjogren's syndrome, or RA, PSS has few clinical features to permit conceiving an immunological pathogenesis. More often than not the vascular lesions show intimal proliferation rather than necrotizing vasculitis. This type of lesion is also responsible for the renal involvement of late onset and is associated with severe hypertension, which is rather rare in SLE (Herrera-Acosta *et al.*, 1978). In the skin and other organs, the predominant lesions are due to collagen deposition rather than to a flagrant inflammatory process. From the laboratory standpoint, serum complement levels are normal, antinuclear antibodies are not so readily apparent as they are in SLE, and serum immunoglobulins are not grossly abnormal. Thus, clues to the immunological nature of PSS were less readily apparent and took longer to become apparent.

Lymphocytic infiltrate was noticed in the dermis of PSS patients, and when studied by newer techniques these were found to be mostly formed by T lymphocytes (Kondo *et al.*, 1976) Elution and immunocytochemical studies of scleroderma kidneys revealed the presence of IgM in glomeruli and eluted antinuclear antibodies from them (McCoy *et al.*, 1976). A large proportion of patients with PSS, either with diffuse scleroderma or the CREST syndrome, have evidence of Sjögren's syndrome (Alarcón-Segovia *et al.*, 1974; Cippoletti *et al.*, 1977) with lymphocytic infiltrate of salivary glands. Both primary Sjögren's syndrome and Sjögren's syndrome associated with other diseases have prominent immunological features; all the diseases they are associated with are immunological, and

most of them are autoimmune. These include primary biliary cirrhosis (Alarcón-Segovia *et al.*, 1973), an autoimmune disease that associates not uncommonly with CREST syndrome (Reynolds *et al.*, 1971).

Antinuclear antibodies showing fine speckles (Kleinsmith *et al.*, 1982) or nucleolar staining (Miyawaki and Ritchie, 1977) in immunofluorescent studies were described in scleroderma patients (CREST and diffuse scleroderma, respectively) and soon other antigen–antibody systems were found to occur predominantly in PSS patients such as the uracil-specific anti-RNA (Alarcón-Segovia *et al.*, 1975), the anticentromere (Fritzler and Kinsella, 1980), and the anti-Scl-70 antibodies (Douvas *et al.*, 1979). The finding of high titers of anti-nRNP antibodies in patients with scleroderma, as well as features previously considered characteristic of SLE, RA, and dermatomyositis (Sharp *et al.*, 1972), made it also evident that scleroderma as a disease manifestation could occur as a result of autoimmunity. This notion was confirmed when scleroderma features appeared in both human (Spielvogel *et al.*, 1977) and experimental (Jaffe and Claman, 1983) graft-versus-host disease.

B. Immunological Derangements

1. Lymphocyte Subpopulations

Patients with PSS have been found to have normal proportions of T cells with receptors for the Fc portion of IgG (T_γ cells), but decreased T cells with receptors for the Fc portion of IgM (T_μ cells), in two of three studies (Gupta *et al.*, 1979; Alarcón-Segovia *et al.*, 1981; Inoshita *et al.*, 1981). T Cells with the capacity of forming rosettes with autologous erythrocytes (T_{ar} cells), a T cell found to have functions of postthymic precursor T cells (Palacios *et al.*, 1981 and to be a responding cell in the autologous mixed lymphocyte reaction (Palacios *et al.*, 1980; Tomonari *et al.*, 1980; Fournier and Charreire, 1982), are found in normal proportions in patients with PSS (Alarcón-Segovia *et al.*, 1981). T-Cell subpopulations, as detected with monoclonal antibodies of the OKT series, are quite variable in PSS (Keystone *et al.*, 1982; Melendro *et al.*, 1983). In one study of 22 patients with PSS not receiving any treatment and their carefully matched controls, the only statistically significant difference was a decreased proportion of T3 cells (Melendro *et al.*, 1983). Other findings are shown in Table II.

The decreased number of T cells is not associated with decreased B cells (Horwitz *et al.*, 1977) but B-cell subpopulations have not yet been

Table II

T-Cell Subpopulations As Determined with Monoclonal Antibodies of the OKT Series in 22 Untreated PSS Patients and their Age-and-sex Matched Controls

	T-cell subpopulations		
	T3	T4	T8
Normal[a]	8	9	11
Decreased[a]	14	8	9
Increased[a]	0	5	2
Mean ± ISD			
PSS	41.4 ± 9.4[b]	34.1 ± 7.8	20.0 ± 8.2
Controls	51.1 ± 6.4	35.5 ± 6.8	21.7 ± 5.8

[a] As compared to the 95% confidence limits of their age-and-sex-matched controls.

[b] Significance versus controls, $p < .01$.

studied in PSS. Decreased proportions of T3 lymphocytes may be due to their homing in tissue lesions (Fleishmajer *et al.*, 1977).

2. T-Cell-Function Studies

Concanavalin A-induced suppressor-cell function has been found to be normal in patients with PSS, using different indicator systems (Alarcón-Segovia *et al.*, 1981; Krakauer *et al.*, 1981; Krawitt *et al.*, 1982; Segond *et al.*, 1982). Suppressor-cell function spontaneously expanded in culture has also been found to be normal in PSS (Alarcón-Segovia *et al.*, 1981). In three different simultaneous studies, helper-cell function was studied with allogeneic combinations of PSS and normal cells and found to be increased (Alarcón-Segovia *et al.*, 1981; Krakauer *et al.*, 1981; Inoshita *et al.*, 1981).

The function of postthymic precursor T cells, identified in man as T_{ar} cells, was found to be normal in nine untreated patients with PSS (Alarcón-Segovia *et al.*, 1981). Feedback inhibition, studied by coculturing T_{γ} and B cells, T_{μ} cells, or T_{ar} and B cells in the presence of pokeweed mitogen (Palacios *et al.*, 1981), was found to be normal in patients with PSS (Alarcón-Segovia *et al.*, 1981).

A suppressor monocyte was found to be responsible for a defect of primary *in vitro* antibody response to trinitrophenyl coupled to polyacrylamide beads found in PSS patients (Segond *et al.*, 1982).

3. Autologous and Allogeneic Mixed Lymphocyte Reactions

Normal human T lymphocytes proliferate when cocultured *in vitro* with inactivated non-T cells from the same donor, a phenomenon called autologous mixed lymphocyte reaction (AMLR) (Opelz *et al.*, 1975), which has both memory and specificity (Weksler and Kozak, 1977) and may thus reflect the complex interaction of the immunoregulatory circuits in self-recognition. The reaction of T cells to *HLA-Dr*-nonidentical allogeneic non-T cells and AMLR have different cellular bases and significance (Palacios *et al.*, 1980).

Proliferation in AMLR in PSS was found to be normal in one study (Morse and Bodi, 1982). However, in another investigation where day-to-day kinetics of response were studied in 17 untreated PSS patients, an early peak proliferative response was noticed on the fourth day in 9 patients and on the fifth day in another 4 (Laffón *et al.*, 1983). As this early response might be anamnestic, kinetic studies lasting 12 days were conducted, and these revealed a second proliferative response on the ninth or tenth day (Alcocer-Varela *et al.*, 1984a). Two possibilities were considered to explain these findings; on the one hand, that autoreactivity between T and non-T cells might have occurred *in vivo* in PSS patients and be recalled in the AMLR, or on the other hand, that alterations of immunoregulatory cells or their products might permit this earlier activation in the system (Alcocer-Varela *et al.*, 1984a).

Response in allogeneic mixed lymphocyte cultures (MLC) of PSS T cells and normal B cells showed no significant differences from those stemming from normal cocultures (Morse and Bodi, 1982; Alcocer-Varela *et al.*, 1984a). However, when MLC of PSS non-T cells and normal T cells were done, proliferation was significantly less than that of normal non-T cells and PSS T cells (Morse and Bodi, 1982). This could indicate either that PSS non-T cells are poor stimulators in MLC or that PSS T cells are better responders than normal T cells.

4. Production of, and Response to, Soluble Mediators

T Lymphocytes from patients with PSS produce the T-cell growth factor interleukin-2 (IL-2) adequately (Alcocer-Varela *et al.*, 1984b). They also respond normally *in vitro* to IL-2 of an exogenous source. The production of interleukin-1 upon stimulation of monocytes with lipopolysaccharide S is normal. However, unstimulated monocytes from some patients with early PSS release IL-1 activity spontaneously into their supernatants (Alcocer-Varela *et al.*, 1985). Response by T lymphocytes

from PSS patients to IL-1 was found defective (Alcocer-Varela *et al.*, 1985). Both of these findings could indicate *in vivo* preactivation. These findings are of particular interest since both monocytes (Diegelmann *et al.*, 1982) and their soluble products (Diegelmann *et al.*, 1982) including IL-1 (Schmidt *et al.*, 1982) induce fibroblast proliferation as well as collagenase production (Postlethwaite *et al.*, 1983). Indeed, cells from the immune system, including T lymphocytes, seem to actively participate in the regulation of fibrogenesis mainly through the production of soluble factors (Jiménez *et al.*, 1979; Korn *et al.*, 1980; Postlethwaite *et al.*, 1984).

Two other factors have been reported to be present in serum from patients with PSS. One is a factor that is cytotoxic to human endothelial cells (Kahaleh *et al.*, 1979; Cohen *et al.*, 1983), which although originally thought to be unique to PSS patients has been found in other connective tissue diseases as well (Shanahan and Korn, 1982). Another serum factor is the one described by Emerit (1976) as capable of inducing chromosomal aberrations that she found in most PSS patients, but this is still largely unconfirmed.

5. Interactions between Mononuclear Cells and Fibroblasts

Mononuclear cells are in close association with fibroblasts and the collagen they deposit in excess in PSS. This is noticeable in the early stages of the disease (Fleishmajer *et al.*, 1977; Jiménez, 1983), is quite similar to what occurs in the chronic graft-versus-host disease where scleroderma and hyperpigmentation develop (Jiménez, 1983), and may concur with vascular endothelial changes.

Most likely the cells of the immune system play important regulatory roles on fibrogenesis, probably through soluble mediators that include both stimulators (Johnson and Ziff, 1976; Schmidt *et al.*, 1982; Keyser *et al.*, 1980; Postlethwaite *et al.*, 1984) and inhibitors (Jiménez *et al.*, 1979; Neilson *et al.*, 1982; Postlethwaite *et al.*, 1984). They can also have both effects depending on their concentration (Diegelmann *et al.*, 1982). The various soluble factors produced by cells of the immune system that influence fibroblast functions are presented in Table III.

Mixed mononuclear leukocyte–fibroblast cultures have been done in an attempt to better understand their interactions (Korn, 1981; Hibbs *et al.*, 1983). Both autologous and allogeneic cocultures stimulate both collagen and noncollagen protein production (Hibbs *et al.*, 1983).

We have recently examined this model in our laboratory with cells from PSS patients and normal age-matched controls (González-Amaro *et al.*,

Table III
Soluble Factors Produced by Cells of the Immune System That Have Regulatory Effects on Fibroblast Function

Factor	Producer cell	Effect on fibroblast function	Reference
IL-1	Monocyte	Stimulates collagenase production	Mizel *et al.* (1981); Postlethwaite *et al.* (1983)
		Stimulates fibroblast proliferation	Schmidt *et al.* (1982)
16 Kdalton	MNL[a]	Stimulates growth of subconfluent fibroblasts	Postlethwaite and Kang (1983)
22 Kdalton	T Cells	Chemotactic for fibroblasts	Postlethwaite *et al.* (1976)
55 Kdalton	MNL[b]	Inhibits fibroblast collagen production	Postlethwaite *et al.* (1984)
60 Kdalton	T Cells[c]	Stimulates fibroblast proliferation	Postlethwaite and Kang (1983)
>100 Kdalton	T Cells[d]	Stimulates collagen fibroblast production	Postlethwaite *et al.* (1984)

[a] May be identical to IL-1.
[b] MNL, mononuclear leukocytes; may be the same factor described by Jiménez *et al.* (1979).
[c] May be the same factor described by Walil and Gately (1983).
[d] May be the factor described by Johnson and Ziff (1976).

1984). Supernatants recovered from fibroblast–MNC autologous cocultures from scleroderma patients induced higher thymidine and [^{14}C]proline incorporation than did those from normal subjects. Supernatants from those cocultures also induced higher thymidine incorporation by autologous fibroblasts from PSS patients than did those from AMLR. The increased thymidine incorporation was mostly by fibroblasts, as determined in experiments in which one of the two cells types was treated with mitomycin C.

6. Differences in Immunoregulation between PSS and MCTD

Despite the presence of scleroderma in both PSS and MCTD, the findings in both of these diseases from the standpoint of immune regulation are markedly different (Alarcón-Segovia, 1983). These are shown in Table IV. Both also differ from other connective tissue diseases (Alarcón-Segovia, 1983).

Table IV
Comparison between Cellular Immunological Findings in the Peripheral Blood of PSS and MCTD Patients[a]

Study	PSS	MCTD
Total lymphocytes	Normal	Decreased
T_γ Cells	Normal	Decreased
T_μ Cells	Decreased	Normal
T_{ar} Cells	Normal	Increased
T3 Cells	Decreased	Normal
T4 Cells	Normal	Normal
T8 Cells	Normal	Decreased
Suppressor function	Normal	Decreased
Feedback inhibition	Normal	Decreased
Generation of suppression	Normal	Decreased
Helper function	Increased	Normal
AMLR	Early response	Early response
Production of IL-2	Normal	Decreased[b]
Response to IL-2	Normal	Decreased
Production of IL-1	Normal	Decreased
Response to IL-1	Decreased	—

[a] Modified from Alarcón-Segovia (1983).
[b] With PHA stimulus; normal in AMLR.

C. Mechanism of Damage

The solution of the riddle of the pathogenesis of PSS may have begun. The peculiar association of vascular endothelial damage, increased collagen deposition, and peculiar antinuclear antibodies may now be linked.

Increased helper T-cell function (Alarcón-Segovia *et al.*, 1981; Krakauer *et al.*, 1981; Inoshita *et al.*, 1981) might be a primary event. Some of those helper T cells might locate in tissues proximal to fibroblasts, perhaps attracted by tissue-associated antigens such as collagen to which they may be committed, or they may be able themselves to further attract lymphocytes by means of the soluble fibroblast chemotactic factors they produce (Postlethwaite *et al.*, 1976). This homing in tissues might explain the decrease in T_μ and T3 cells, the former having been shown to have predominant helper activity (Moretta *et al.*, 1977). Increased helper function could result in increased production of lymphokines, with predominance of those favoring fibroblast proliferation and collagen production as well as activation of clones of B cells that produce autoantibodies.

A cytotoxic factor to endothelial cells (Kahaleh *et al.*, 1979) could also be a byproduct of helper T-cell activation and be responsible for endothelial cell damage and intimal proliferation. Monocytes could participate in

the response to lymphokines by themselves producing IL-1 and perhaps other monokines. IL-1 would, in turn, stimulate fibroblast proliferation and increased collagenase production. Such stimulation, if collagen or a fraction from collagen or laminin (Huffstutter *et al.*, 1984) was the antigen or antigens, might contribute to the vicious circle of T-cell activation, lymphokine-monokine production/fibrogenesis, and production of antinuclear antibodies. These might also contribute to counteract suppressor cells, perhaps by penetrating into them and causing their deletion (Alarcón-Segovia *et al.*, 1979b).

D. Immunogenetics

Different studies have shown different HLA associations in PSS, probably reflecting the ethnic variations of each study. In general, association with the strongly linked *A1*, *B8*, and *DR3* haplotypes or with some of the phenotypes that compose it have been found in British, German, and Dutch patients (Fraundenberg *et al.*, 1978; Hughes *et al.*, 1978; Kallenberg *et al.*, 1981; Welsh, 1983). Raised *DR5* phenotypes have been found in both Canadian and British patients (Gladman *et al.*, 1981; Welsh, 1983), particularly in those with CREST syndrome (Gladman *et al.*, 1981).

Conversely, in white scleroderma patients from the United States no association was found with either the *A1*, *B8*, or *DR3* haplotype or their components or *DR5* (Whiteside *et al.*, 1983). The strongest association was found between *DR1* and CREST syndrome patients having anticentromere antibodies. These findings were also made in the British study (Welsh, 1983).

It seems that no clear-cut association has been found between PSS and *HLA* antigens, but that this may reflect both the heterogeneity of the diseases that cause scleroderma as well as of the ethnic background of the populations studied.

E. Serology

Antibodies to nuclear antigens (ANA) have been found in the serum of patients with scleroderma (Rothfield and Rodnan, 1968; Jablonska *et al.*, 1974; Alarcón-Segovia, 1976). Immunofluorescent tests for ANA using serum from PSS patients yield three main patterns of fluorescence (Kleinsmith *et al.*, 1982); a fine speckled pattern that is associated with CREST syndrome, a thready pattern that is found mainly in patients with either diffuse scleroderma or acrosclerosis but having pulmonary involve-

ment, and a nucleolar pattern that is associated with diffuse scleroderma (Miyawaki and Ritchie, 1974; Kleinsmith *et al.*, 1982).

The antibody giving the fine speckled pattern probably corresponds to the anticentromere antibody (Kleinsmith *et al.*, 1982; McCarty *et al.*, 1983) that has been found in patients with the CREST syndrome (Fritzler and Kinsella, 1980; Tuffanelli *et al.*, 1983). The antigen giving the thready pattern has not been identified, but this could be due to reaction with the Scl-70 antigen (Maddison, 1983) (see below), and the nucleolar pattern probably corresponds to the uracil-specific anti-RNA antibodies (Miyawaki and Ritchie, 1973) that are found in most PSS patients with early active diffuse scleroderma (Alarcón-Segovia *et al.*, 1975; Heinzerling *et al.*, 1980; Fishbein and Alarcón-Segovia, 1985).

Failure to detect the uracil-specific anti-RNA antibodies in these patients (Cattogio and Maddison, 1981) seems to be due to different patient populations and to low assay sensitivity, since the initial observations using counterimmunoelectrophoresis (Alarcón-Segovia *et al.*, 1975) have been confirmed with a sensitive radioimmunoassay using synthetic polyuridylic acid (poly-U) (Heinzerling *et al.*, 1980), and more recently by an ELISA method also using poly-U (Fishbein and Alarcón-Segovia, 1985). That the antibodies to RNA found in PSS are specific to uracil was shown by their reactivity with protein-conjugated uridine and uridine-containing dinucleotides (Alarcón-Segovia *et al.*, 1975) and with poly-U (Alarcón-Segovia and Fishbein, 1975), but not with polyadenylic-polyuridylic acid (poly A-poly U) where the bases are not exposed, thus ruling out interaction with the sugar–phosphate moiety (Alarcón-Segovia and Fishbein, 1975; Fishbein and Alarcón-Segovia, 1985). These antibodies were also inhibited with uracil (Alarcón-Segovia *et al.*, 1975).

Antibodies to a chromosomal antigen initially called Scl-1 and later Scl-70 (Douvas *et al.*, 1979) are found in ~20% of patients with various forms of PSS (23% in diffuse scleroderma; 13% in CREST syndrome) and in few other instances (Tan *et al.*, 1980; Cattoggio *et al.*, 1983). They seem, therefore, to be less sensitive, albeit probably specific for scleroderma.

Antibodies to centrioles have been observed in four patients within the scleroderma spectrum (Tuffanelli *et al.*, 1983), in one with Raynaud's phenomenon and telangiectasia, one with CREST syndrome, and two with diffuse scleroderma.

Patients with localized forms of scleroderma also frequently have serum ANA (Rodnan *et al.*, 1977). These are more easily detected when using HeLa cells for their study (73% versus 50% on mouse kidney sections) (Takehara *et al.*, 1983). Thus 2 of 4 patients with morphea, 6 of 6 patients with generalized morphea, and 8 of 12 patients with linear sclero-

derma were found to have ANA with various patterns of fluorescence (homogeneous speckled and nucleolar) in a recent study (Takehara *et al.*, 1983).

The association of antibodies to nuclear RNP with MCTD (Sharp *et al.*, 1972) permitted the definition and characterization of this disease to the point that at present it can be suspected and, I believe, diagnosed in their absence (Alarcón-Segovia, 1981). Most patients with MCTD, however, have high titers of antibodies to RNP, and, if originally absent, these antibodies usually appear either spontaneously or after the initiation of corticosteroid treatment (Alarcón-Segovia, 1979b). High titers of antibodies to RNP have a high specificity for MCTD. In a study of 2800 sera from 289 SLE patients, we found 155 patients with antibodies to extractable nuclear antigens, 40 patients with antibodies to RNP, and 21 with these antibodies at high titers, 7 of whom fulfilled a predefined clinical criteria for the diagnosis of MCTD. All the other 14 had Raynaud's phenomenon, and included among them were all SLE patients with myositis and all with calcinosis.

It is therefore likely that these 14 patients do have MCTD although they do not fulfill the proposed criteria (Alarcón-Segovia, 1981). Although most if not all extractable nuclear antigens associate with RNA, for clinical purposes the main ribonuclease-sensitive antigen that gives high titers in hemagglutination assays is nuclear RNP, and this allows its differentiation from another major constituent of ENA, the Sm antigen (Tan and Kunkel, 1966). Antibodies to the Sm antigen are most frequently found in SLE, albeit in a small proportion of patients (Barada *et al.*, 1981).

Another interesting nonhistone protein antigen–antibody system is the Ku–anti-Ku (Mimori *et al.*, 1981), which is found to have high specificity for the scleroderma–polymyositis overlap syndrome, thus supporting the contention that such patients are distinct from those with MCTD (Alarcón-Segovia, 1981).

Antibodies to collagen have been detected in patients with PSS (Mackel *et al.*, 1982). These are directed to interstitial (type I) and basement membrane (type IV) collagen. Both were found to correlate with the presence of pulmonary functional abnormalities (Mackel *et al.*, 1982). These antibodies did not seem to cross-react among themselves, and those to type IV collagen could be unique to PSS whereas those to type I collagen are also found in RA (Menzel *et al.*, 1978). Antibodies to insulin receptors have also been described in at least one patient with PSS (Weinstein *et al.*, 1980).

Immune complexes have been detected in PSS (Pisko *et al.*, 1979; Dan *et al.*, 1981; Seibold *et al.*, 1982; Hughes *et al.*, 1983; Fishbein *et al.*, 1985)

using various methods: Raji cell assay (Pisko *et al.*, 1979; Seibold *et al.*, 1982; Hughes *et al.*, 1983), Clq binding (Dan *et al.*, 1981; Seibold *et al.*, 1982; Hughes *et al.*, 1983; Fishbein *et al.*, 1985), inhibition of antibody-dependent cellular cytotoxicity (Hughes *et al.*, 1983; Fishbein *et al.*, 1985), and agarose gel electrophoresis (Seibold *et al.*, 1982).

Presence of immune complexes in PSS has varied from 9 (Chia *et al.*, 1979) to 82% (Seibold *et al.*, 1982). The immune complexes seemed to be heterogeneous (Fishbein *et al.*, 1984), and to correlate with organ involvement (Seibold *et al.*, 1982; Hughes *et al.*, 1983) and with the presence of antinuclear antibodies (Seibold *et al.*, 1982) or rheumatoid factor (Hughes *et al.*, 1983). They appeared not to be related to the presence of antibodies to collagen (Mackel *et al.*, 1982).

VI. TREATMENT

In the past few years several advances have been made in the treatment of PSS. These include those made in intervention on collagen metabolism and those made in the treatment of several of the complications of PSS.

A. INTERVENTION IN COLLAGEN METABOLISM

Two main drugs have been proposed for the purpose of decreasing collagen deposition in scleroderma. D-Penicillamine has been shown to inhibit the intramolecular and intermolecular cross-links of collagen causing increasing proportion of soluble collagen in the skin (Nimni and Bavetta, 1965; Uitto *et al.*, 1970). D-Penicillamine also has effects on immune functions that might be beneficial in PSS (Chwalinska-Sadowska and Baum, 1976).

Results of trials with this drug have been controversial, and no well-controlled study has shown its effectiveness without question (Nassonova and Ivanova, 1979). A recent study suggested a beneficial effect (Steen *et al.*, 1982). Toxicity, however, is high, and withdrawal is often necessary.

Colchicine has also been used to intervene in the metabolism of collagen in PSS because of its effects in the disruption of microtubule formation that prevent the extrusion of collagen from the fibroblasts (Ehrlich and Bornshtein, 1972; Diegelmann and Peterkofsky, 1972; Fernández-Madrid *et al.*, 1974). It has also been found to increase collagen degradation *in vitro*, apparently by enhancing the action of collagenase (Harris and Krane, 1971).

As with D-penicillamine, conflicting results have been reported on the

effect of colchicine in the treatment of PSS (Alarcón-Segovia, 1979). However, a long-term evaluation of colchicine in the management of PSS and localized scleroderma showed improvement in most patients (Alarcón-Segovia *et al.*, 1979a). Those patients who had <5 years duration of disease when placed on treatment, and those who received larger total doses of colchicine, improved the most. As opposed to D-penicillamine, colchicine, which was given at the maximum tolerated dose, had practically no untoward effects.

B. Treatment of Complications of PSS

Renal failure and malignant hypertension in PSS had proven to be uniformly fatal or to require renal dialysis, nephrectomy, and transplantation until the advent of angiotensin-converting enzyme inhibitors (López-Ovejero *et al.*, 1979). Results, however, have not been uniformly successful (Whitman *et al.*, 1982; Thurm and Alexander, 1984).

Treatment of Raynaud's phenomenon in PSS has been attempted with a number of medications, but has been shown to be regularly effective until the advent of the calcium channel blocker, nifedipine, which has been shown to improve Raynaud's phenomenon in several controlled studies (Smith and McKendry, 1982; Sauza *et al.*, 1984) and to improve finger ulcerations (Jaffe, 1982).

Nifedipine may also have beneficial effects on pulmonary hypertension in patients with PSS (Ocken *et al.*, 1983). Caution should be exerted with the use of this medication, however, since it may have untoward effects on the lower esophageal sphincter pressure (Kahan *et al.*, 1984). Recently, prazocin has also been found beneficial for Raynaud's phenomenon in PSS (Surwit *et al.*, 1984).

The esophageal aperistalsis of patients with PSS may be a difficult problem to deal with and may be extremely bothersome and deleterious to the patient. Metoclopramide has been found to induce the appearance of a pressure zone at the sphincteric area in PSS patients and to cause the appearance of pressure waves in 5 of 11 patients with aperistalsis of the esophagus due to PSS (Ramírez-Mata *et al.*, 1977).

C. Immunosuppressive Agents

Despite the growing evidence that PSS may be due to an immunological derangement, immunosuppressive agents and corticosteroids have not been shown to be of any significant benefit in the treatment of this disease (Steigerwald, 1979).

VII. CONCLUDING REMARKS

Scleroderma is a manifestation of various diseases, many of them autoimmune. The primary form of scleroderma, PSS, remained for many years as the most elusive and puzzling of the autoimmune connective tissue diseases.

Although some evidence of autoimmunity had become apparent in PSS, it remained difficult to understand how it could relate to increased collagen deposition and to intimal proliferation. New knowledge in the regulatory actions of cells and their products from the immune system, and on the functions of fibroblasts, including collagen production as well as the description of circulating factors possibly produced by those same cells, has given clues to the possible pathogenetic mechanisms whereby autoimmunity and increased fibrogenesis might be related. At the same time aberrations in immunoregulatory T-cell circuits in PSS began to be unraveled and to be distinguished from those found in the other connective tissue diseases, including MCTD where scleroderma also occurs.

A number of antinuclear antibodies have been found present either primarily or uniquely in the serum of patients with scleroderma, either as PSS or as its CREST syndrome variety or in overlap, and either as MCTD or as sclerodermatomyositis. The meaning of these antibodies as a cause or as a result of immunodysregulation will have to be understood before the entire picture on the pathogenesis of scleroderma can be drawn. Antibodies to type I and type IV collagen may be indicative of a role of collagen as a triggering antigen, a possibility that might close the circle of the pathogenesis or scleroderma. Laminin could also be implicated.

These advances in the knowledge of the pathogenetic mechanisms in PSS might seem futile were it not that they permit to foresee more rational therapeutic approaches. In the meantime, however, it has also become possible to offer something to the patient with PSS. Intervention on collagen metabolism has become available, and despite controversial results, this has been received enthusiastically. Various modalities of treatment for the renal, vascular, pulmonary, and esophageal complications of PSS are also available.

VIII. SUMMARY

Scleroderma is not a disease, but a manifestation of various diseases in most of which an immunological derangement is likely to participate. The primary forms of scleroderma, localized and systemic, are diseases that are also probably related. The systemic form, progressive systemic sclerosis (PSS), has in turn two variants: the diffuse and the CREST

syndrome (acronym for *C*alcinosis, *R*aynaud's phenomenon, *E*sophageal involvement, *S*clerodactyly, and *T*elangectasia).

Scleroderma may also occur with features previously considered characteristic of systemic lupus erythematosus, dermatomyositis, rheumatoid arthritis, and Sjögren's syndrome in the disease entity called mixed connective tissue disease (MCTD) that associates with high titers of antibody to nuclear ribonucleoprotein. An association of scleroderma and dermatomyositis has also been found to have an antinuclear antibody called anti-Ku.

Patients with PSS have decreased T cells with receptors for the Fc portion of IgM and decreased T cells in their peripheral blood, perhaps because these cells locate in tissues in the vicinity of fibroblasts. Suppressor-cell function and feedback inhibition function are normal in patients with PSS, but they do seem to have increased helper function that could account for an early proliferative response in the autologous mixed lymphocyte reaction. Preactivation of monocytes is also apparent from the spontaneous release of mononuclear cells in culture of the monokine interleukin-1, while the production of and the response to the lymphokine interleukin-2 is normal in PSS.

Mononuclear cells and fibroblasts seem to have strong interactions for regulatory purposes. These could be deranged in PSS, and thereby explain the increased fibrogenesis and collagen deposition characteristic of this disease.

The immunoregulatory derangement in PSS seems to be quite different from that which occurs in mixed connective tissue disease and in other connective tissue diseases.

No clear-cut *HLA* associations has been found in PSS patients, but this may reflect ethnical variations as well as disease heterogeneity.

Patients with PSS have been found to have various characteristic autoantibodies. These include antinucleolar antibodies that are probably directed to nucleolar RNA and may be the same as the uracil-specific anti-RNA antibodies, as antibodies to the chromosomal centromere that are found primarily in patients with the CREST syndrome, antibodies to a nucleolar antigen called Scl-70, and antibodies to type I and IV collagen. Circulating immune complexes can also be detected in the serum of patients with PSS.

New avenues of treatment have been uncovered by means of intervention on collagen metabolism, which has been tried with either D-penicillamine or colchicine. New forms of treatment for various complications of PSS have also become available to make the situation less hopeless. Indeed, advances in knowledge about PSS suggest a brighter future for patients with this puzzling disease.

REFERENCES

Alarcón-Segovia, D. (1975). *In* "Modern Topics in Rheumatology" (G. R. V. Hughes, ed.), pp. 152–156. Heinemann Medical Books, London.

Alarcón-Segovia, D. (1976). *J. Rheumatol.* **3,** 191–195.

Alarcón-Segovia, D. (1979a). *Clin. Rheum. Dis.* **5,** 263–302.

Alarcón-Segovia, D. (1979b). *J. Rheumatol.* **6,** 694–699.

Alarcón-Segovia, D. (1981). *J. Rheumatol.* **8,** 535–540.

Alarcón-Segovia, D. (1983). *Semin. Arthritis Rheum.* **13,** 114–120.

Alarcón-Segovia, D., and Fishbein, E. (1975). *J. Immunol.* **115,** 28–31.

Alarcón-Segovia, D., Díaz-Jouanen, E., and Fishbein, E. (1973). *Ann. Intern. Med.* **79,** 31–36.

Alarcón-Segovia, D., Ibáñez, G., Hernández-Ortíz, J., Velázquez-Forero, F., and Gonzalez-Jiménez, Y. (1974). *Am. J. Med.* **57,** 78–85.

Alarcón-Segovia, D., Fishbein, E., García-Ortigoza, E., and Estrada-Parra, S. (1975). *Lancet* **1,** 363–365.

Alarcón-Segovia, D., Ramos-Niembro, F., Ibáñez de Kasep, G., Alcocer-Varela, J., and Pérez-Tamayo, R. (1979a). *J. Rheumatol.* **6,** 705–112.

Alarcón-Segovia, D., Ruíz-Arguelles, A., and Llorente, L. (1979b). *J. Immunol.* **122,** 1855–1863.

Alarcón-Segovia, D., Palacios, R., and Ibáñez de Kasep, G. (1981). *J. Clin. Lab. Immunol.* **5,** 143–148.

Alcocer-Varela, J., Martínez-Cordero, E., and Alarcón-Segovia D. (1985). *Clin. Exp. Immunol.* **59,** 666–672.

Alcocer-Varela, J., Laffón, A., and Alarcón-Segovia, D. (1984b). *Rheumatol. Int.* **4,** 39–44.

Alcocer-Varela, J., Martínez-Cordero, E., and Alarcón-Segovia D. (1984c). (Submitted for publication.)

Barada, F. C., Andrews, B. B., Davis, J. S., IV, and Taylor, R. P. (1981). *Arthritis Rheum.* **24,** 1236–1244.

Benedek, T. G., and Rodnan, G. P. (1982). *Semin. Arthritis Rheum.* **12,** 52–67.

Cattogio, L. J., and Maddison, P. J. (1981). *Rev. Esp. Rheumatol. (Lett.)* **8,** 88.

Cattogio, L. J., Skinner, R. P., and Maddison, P. J. (1983). *Rheumatol. Int.* **3,** 19–21.

Chia, D., Barnett, E. U., Yamagata, J., Knutson, D., Restivo, C., and Furst, D. (1979). *Clin. Exp. Immunol.* **37,** 399–407.

Chwalinska-Sadowska, H., and Baum, J. (1976). *J. Clin. Invest.* **58,** 871–879.

Cippoletti, J. F., Buckingham, R. B., Barnes, E. L., Peel, R. L., Malimoad, K., Cignetti, F. E., Pierce, J. M., Rabin, B. S., and Rodnan, G. P. (1977). *Am. Intern. Med.* **87,** 535–541.

Cohen, S., Johnson, A. R., and Hurd, E. (1983). *Arthritis Rheum.* **26,** 170–178.

Dan, P. C., Kahaleh, M. B., and Sagebiel, R. W. (1981). *Arthritis Rheum.* **24,** 1128–1136.

D'Angelo, W. A., Fries, J. F., Masi, A. T., and Shulman, L. E. (1969). *Am. J. Med.* **46,** 428–440.

Diegelmann, R. F., and Peterkofsky, B. (1972). *Proc. Natl. Acad. Sci. (U.S.A.)* **69,** 892–897.

Diegelmann, R. F., Cohen, I. K., and Kaplan, A. M. (1982). *Proc. Soc. Exp. Biol. Med.* **169,** 445–451.

Douvas, A., Achten, M., and Tan, E. M. (1979). *J. Biol. Chem.* **254,** 10514–10522.

Ehrlich, P., and Bornstein, P. (1972). *Nature (London), New Biol.* **238,** 257–260.

Emerit, I. (1976). *Dermatologica* **155,** 145–156.

Fernández-Madrid, F., Noonas, S., and Riddle, J. (1974). *J. Rheumatol.* **1,** 82. (Abstr.)

Finch, W. R., Rodnau, G. P., Buckingham, R. B., Prince, R. K., and Winkelstein, A. (1980). *J. Rheumatol.* **7,** 651–659.
Fishbein, E., and Alarcón-Segovia, D. (1984). *J. Rheumatol.* (In press.)
Fishbein, E., Frajman, M., and Alarcón-Segovia, D. (1985). *Rev. Invest. Clin.* (In press).
Fleishmajer, R., Perlish, J. S., and Reeves, J. R. T. (1977). *Arthritis Rheum.* **20,** 975–984.
Fournier, C., and Charreire, J. (1982). *J. Immunol.* **128,** 2698–2703.
Fraudenberg, J., Holzmann, H., Schneider, S., and Korting, G. W. (1978). *Arch. Dermatol. Res.* **263,** 197–205.
Fritzler, M. J., and Kinsella, T. D. (1980). *Am. J. Med.* **69,** 520–526.
Garza-Elizondo, M. A., Díaz-Jouanen, E., Franco-Casique, J., and Alarcón-Segovia, D. (1983). *J. Rheumatol.* **10,** 797–800.
Gershwin, M. E., Abplanalp, H., Bastles, J. J., Ikeda, R. M., Van der Water, J., Eklund, J., and Haynes, D. (1981). *J. Exp. Med.* **153,** 1640–1659.
Gladman, D. D., Keystone, E. G., Baron, M., Lee, P., Cane, D., and Mervet, H. (1981). *Arthritis Rheum.* **28,** 854–865.
Goetz, R. H. (1945). *Clin. Proc. (Cape Town)* **4,** 337–392.
González-Amaro, R., Alcocer-Varela, J., Díaz de León, L., and Alarcón-Segovia, D. (1984). (Submitted for publication.)
Gupta, S., Malaviya, A. N., Rajagopalan, P., and Good, R. A. (1979). *Clin. Exp. Immunol.* **38,** 342–347.
Harris, E. D., and Krane, S. A. (1971). *Arthritis Rheum.* **14,** 669–684.
Heinzerling, R. H., Weyer, R., Dziuba, D. S., and Burnham, T. K. (1980). *J. Invest. Dermatol.* **75,** 224–227.
Herrera-Acosta, J., Guerrero, J., Erbessd, M. L., Paz-Barahona, M., Chessal, F., Alarcón-Segovia, D., and Peña, J. C. (1978). *Nephron* **22,** 128–137.
Hibbs, M. S., Postlethwaite, A. E., Mainardi, C. L., Seyer, J. M., and Kang, A. H. (1983). *J. Exp. Med.* **157,** 47–59.
Horwitz, D. A., Garrett, M. A., and Craig, A. H. (1977). *Clin. Exp. Immunol.* **27,** 92–99.
Huffstutter, J. E., DeLustro, F., Timpl, R., and LeRoy, E. C. (1984). *Arthritis Rheum. (Suppl.)* **27,** 573. (Abstr.)
Hughes, P., Gelsthorpe, K., Doughty, R. W., Rowell, N. R., Rosenthal, F. D., and Sneedon, I. B. (1978). *Clin. Exp. Immunol.* **31,** 351–356.
Hughes, P., Cunningham, J., Day, M., Fitzgerald, J. C., French, M. A. H., Wright, J. K., and Rowell, N. R. (1983). *J. Clin. Lab. Immunol.* **10,** 133–138.
Inoshita, T., Whiteside, T. L., Rodnan, G. P., and Taylor, F. H. (1981). *J. Lab. Clin. Med.* **97,** 264–277.
Jablonska, S., Blaszcyk, M., and Glinski, W. (1974). *In* "Tenth Symposium on Advanced Medicine" (J. G. G. Ledingham, ed.), pp. 243–261, Pitman, London.
Jaffe, B. D., and Claman, H. N. (1983). *Cell. Immunol.* **77,** 1–12.
Jaffe, I. A. (1982). *Arthritis Rheum.* **25,** 1267–1269.
Jiménez, S. A. (1983). *Semin. Arthritis Rheum.* **13,** 104–113.
Jiménez, S. A., McArthur, W., and Rosenbloom, J. (1979). *J. Exp. Med.* **150,** 1421–1431.
Johnson, R. L., and Ziff, M. (1976). *J. Clin. Invest.* **58,** 240.
Kahaleh, M. B., Sherer, G. K., and LeRoy, E. C. (1979). *J. Exp. Med.* **149,** 1326–1335.
Kahan, A., Bour, B., Cuturier, D., Amor, B., and Menkes, C. J. (1984). *Arthritis Rheum. (Suppl.)* **27,** 573. (Abstr.)
Kallenberg, C. G. M., Van der Boort-Beelen, J. M., D'Amaro, J., and The, T. H. (1981).
Keyser, A. J., Cooper, S. M., and Nimni, M. E. (1980). *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **39,** 874. (Abstr.)
Keystone, E. C., Lan, C., Gladman, D. D., Wilkinson, S., Lee, P., and Shore, A. (1982). *Clin. Exp. Immunol.* **48,** 443–448.

Kleinsmith, D'A. M., Heinzerling, P. H., and Burnham, T. K. (1982). *Arch. Dermatol.* **118,** 882–885.
Klemperer, P. (1950). *Am. J. Pathol.* **26,** 505–519.
Kondo, H., Rabin, B. S., and Rodnan, G. P. (1976). *J. Clin. Invest.* **58,** 1388–1394.
Korn, J. H. (1981) *Cell Immunol.* **63,** 374–378.
Korn, J. H., Halushka, P. V., and LeRoy, E. C. (1980). *J. Clin. Invest.* **65,** 543–554.
Krakauer, R. S., Sundee, J., Sander, D. N., and Scherbel, A. (1981). *Arch. Dermatol.* **117,** 80–82.
Krawitt, E. L., Holdstock, G., Bland, J. H., Chastenay, B. F., and Albertini, R. J. (1982). *J. Rheumatol.* **9,** 263–267.
Kumagai, Y., Shiokawa, Y., Medsger, T. A., and Rodnan, G. P. (1984).
Laffón, A., Alcocer-Varela, J., and Alarcón-Segovia, D. (1983). *Rheum. Int.* **3,** 117–128.
LeRoy, E. C. (1981). *In* "Textbook of Rheumatology" (W. N. Kelley, E. D. Harris, S. Ruddy, and C. B. Sledge, eds.), pp. 1211–1230. Saunders, Philadelphia, Pennsylvania.
López-Ovejero, J. A., Saal, S. D., D'Angelo, W. A., Cheigh, J. S., Stenzel, K. H., and Laragh, J. H. (1979). *N. Engl. J. Med.* **300,** 1417–1419.
Lundbawk, K. (1957). *Acta Med. Scand.* **158,** 447–451.
McCarty, G. A., Rice, J. R., Bembe, M. L., and Barada, F. A., Jr. (1983). *Arthritis Rheum.* **26,** 1–7.
McCoy, R. C., Tisher, C. C., Pepe, P. F., and Clevelend, L. A. (1976). *Lab. Invest.* **35,** 124.
Mackel, A. M., De Lustro, F., Harper, F. E., and LeRoy, E. C. (1982). *Arthritis Rheum.* **25,** 522–531.
Maddison, P. (1983). *Proc. Scleroderma Symp. London, May 12–13,* pp. 35–37.
Maricq, H. R., Johnson, M. N., Whetstone, C. L., and LeRoy, E. C. (1976). *JAMA, J. Am. Med. Assoc.* **236,** 1368.
Melendro, E. I., Saldate, C., Rivero, S. J., and Alarcón-Segovia D. (1983). *Clin. Immunol. Immunopathol.* **27,** 340–347.
Menzel, J., Steffen, C., Kolarz, G., Kojer, M., and Smolen, J. (1978). *Arthritis Rheum* **21,** 243–248.
Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S., and Homma, M. (1981). *J. Clin. Invest.* **68,** 611–620.
Miyawaki, S., and Ritchie, R. F. (1973). *Arthritis Rheum.* **16,** 726–736.
Miyawaki, S., and Ritchie, R. F. (1974). *J. Immunol.* **133,** 1346–1352.
Mizel, S. B., Dayer, J. M., Krane, S. M., and Mergenhagen, S. E. (1981). *Proc. Natl. Acad. Sci. (U.S.A.)* **78,** 2474–2478.
Moretta, L., Webb, S. R., Grossi, C. E., Lydyard, P. M., Cooper, M. D. (1977). *J. Exp. Med.* **146,** 184–200.
Morse, J. H., and Bodi, B. S. (1982). *Arthritis Rheum.* **25,** 390–395.
Nassonova, V. A., and Ivanova, M. M. (1979). *Clin. Rheum. Dis.* **5,** 277–288.
Neilson, E. G., Phillips, S. M., and Jiménez, S. A. (1982). *J. Immunol.* **128,** 1484–1486.
Nimelstein, S. H., Brody, S., McShane, D., and Holman, H. R. (1980). *Medicine (Baltimore)* **59,** 239–248.
Nimni, M., and Bavetta, J. (1965). *Science* **150,** 905–907.
Ocken, S., Reinitz, E., and Strom, J. (1983). *Arthritis Rheum.* **26,** 794–796.
Olmedo-Garzon, F. J., Zea-Mendoza, A. C., Alonso-Ruíz, A., Salazar-Vallinas, J. M., García-Villaneuva, M., Rocamora-Ripoll, A., and Beltrán-Gutiérrez, J. (1982). *Med. Clin. (Barcelona)* **79,** 1–8.
Opelz, G., Kuchi, M., Takasugi, M., and Terasaki, P. I. (1975). *J. Exp. Med.* **142,** 1327–1333.
Palacios, R., Llorente, L., Alarcón-Segovia, D., Ruíz-Arguelles, A., and Díaz-Jouanen, E. (1980). *J. Clin. Invest.* **65,** 1527–1530.

Palacios, R., Alarcón-Segovia, D., Llorente, L., Ruíz-Arguelles, A., and Díaz-Jouanen, E. (1981). *Immunology* **42,** 127–135.

Pisko, E., Gallup, K., Turner, R., Parker, M., Normer, A. M., Box, J., Davis, J., Box, P., and Zothberger, H. (1979). *Arthritis Rheum.* **22,** 518–523.

Postlethwaite, A. E., and Kang, A. H. (1983). *Arthritis Rheum.* **26,** 22–27.

Postlethwaite, A., Snyderman, R., and Kang, A. H. (1976). *J. Exp. Med.* **144,** 1188–1203.

Postlethwaite, A. E., Lachman, L. B., Mainardi, C. L., and Kang, A. H. (1983). *J. Exp. Med.* **157,** 801–806.

Postlethwaite, A. E., Smith, G. N., Mainardi, D. L., Seyer, J. M., and Kang, A. H. (1984). *J. Immunol.* **132,** 2470–2477.

Ramírez-Mata, M., Ibánñez, G., and Alarcón-Segovia, D. (1977). *Arthritis Rheum.* **20,** 30–34.

Ramos-Niembro, F., Alarcón-Segovia, D., and Hernández-Ortíz, J. (1979). *Arthritis Rheum.* **22,** 43–51.

Reynolds, T. B., Denison, E. K., Frankl, H. D., Lieberman, F. L., and Peters, R. L. (1971). *Am. J. Med.* **50,** 302–312.

Ridolfi, R. L., Bulkley, B. H., and Hutchins, G. M. (1976). *Am. J. Med.* **61,** 361–366.

Rodnan, G. P. (1978). *In* "Immunological Diseases" (M. Samter, ed.), 3rd ed., Vol. II, pp. 1109–1141. Little, Brown, Boston, Massachusetts.

Rodnan, G. P., Lipinski, E., Rabin, B. S., and Reichlin, M. (1977). *Arthritis Rheum.* **20,** 133. (Abstr.)

Rosson, R. S., and Yesner, R. (1965). *N. Engl. J. Med.* **272,** 391–394.

Rothfield, N. F., and Rodnan, G. P. (1968). *Arthritis Rheum.* **11,** 607–617.

Sauza, J., Kraus, A., González-Amaro, R., and Alarcón-Segovia, D. (1984). *J. Rheumatol.* **11,** 362–364.

Schmidt, J. A., Mizel, S. B., Cohen, D., and Green, I. (1982). *J. Immunol.* **128,** 2177–2182.

Segond, P., Salliere, D., Galanaud, P., Desmottes, R. M., Massias, P., and Fiessinger, J. N. (1982). *Clin. Exp. Immunol.* **47,** 147–154.

Seibold, J. R., Medsger, T. A., Winkestem, A., Kelly, R. H., and Rodnan, G. P. (1982). *Arthritis Rhem.* **25,** 1167–1173.

Shanahan, W. R., Jr., and Korn, J. H. (1982). *Arthritis Rheum.* **25,** 1391–1395.

Sharp, G. C., Irvin, W. S., Tan, E. M., Gould, R. G., and Holman, H. R. (1972). *Am. J. Med.* **52,** 148–159.

Smith, C. D., and McKendry, R. (1982). *Lancet* **2,** 1299–1301.

Spielvogel, R. L., Goltz, R. W., and Kersey, J. H. (1977). *Arch Dermatol.* **113,** 1424.

Steen, V. D., Medsger, T. A., and Rodnan, G. P. (1982). *Am. Intern. Med.* **97,** 652–659.

Steen, V. D., Medsger, T. A., Osial, T. A., Ziegler, G. L., Shapiro, A. P., and Rodnan, G. P. (1984). *Am. J. Med.* **76,** 779–786.

Steigerwald, J. C. (1979). *Clin. Rheum. Dis.* **5,** 289–294.

Sternberg, E. M., Van Woert, M. H., Young, S. N., Magnissen, I., Baker, H., Gauthier, S., and Osterland, C. K. (1980). *N. Engl. J. Med.* **303,** 782–787.

Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee (1980). *Arthritis Rheum.* **23,** 581–590.

Surwit, R. S., Gilgor, R. S., Allen, L. M., and Duvic, M. (1984). *Arch. Dermatol.* **120,** 329–331.

Takehara, K., Moroi, Y., Nakabayashi, Y., and Ishibashi, Y. (1983). *Arthritis Rheum.* **26,** 612–616.

Tan, E. M., and Kunkel, H. (1966). *J. Immunol.* **96,** 464–471.

Tan, E. M., Rodnan, G. P., García, I., Moroi, Y., Fritzler, M. J., and Peebles, C. (1980). *Arthritis Rheum.* **23,** 617–625.

Thurm, R. H., and Alexander, J. C. (1984). *Arch. Intern. Med.* **144,** 733–735.
Tomonari, K. A., Wakisaka, A., and Sigawa, M. (1980). *J. Immunol.* **125,** 1596–1600.
Tuffanelli, D. L., McKeon, F., Kleinsmith, D'A. M., Burnham, T. K., and Kirschner, M. (1983). *Arch. Dermatol.* **119,** 560–566.
Uitto, J., Helin, P., Rasmussen, O., and Lorenzen, I. (1970). *Ann. Clin. Res.* **2,** 228–234.
Weinstein, P. S., High, K. A., D'Ercole, J., and Jennette, J. C. (1980). *Arthritis Rheum.* **23,** 101–105.
Weksler, M. E., and Kozak, R. (1977). *J. Exp. Med.* **146,** 1833–1838.
Welsh, K. I. (1983). *Proc. Scleroderma Symp., London, May 12–13,* pp. 16–20.
Whal, S. M., and Gately, C. L. (1983). *J. Immunol.* **130,** 1226–1230.
Whiteside, T. L., Medsger, J. R., and Rodnan, G. P. (1983). *J. Rheumatol.* **10,** 128–131.
Whitman, H. H., III, Case, D. B., Laragh, J. H., Christian, C. L., Botstein, G., Maricq, H., and LeRoy, E. C. (1982). *Arthritis Rheum.* 241–248.
Winterbauer, R. H. (1964). *Bull. Johns Hopkins Hosp.* **114,** 361–383.
Zarafonetis, C. J. D., Lorber, S. H., and Hanson, S. M. (1958). *Am. J. Med. Sci.* **236,** 1–14.

CHAPTER 6

Sjögren's Syndrome

NORMAN TALAL
Division of Clinical Immunology
Department of Medicine
The University of Texas Health Science Center at San Antonio
and Clinical Immunology Section
Audie L. Murphy Memorial Veterans Administration Hospital
San Antonio, Texas

I. INTRODUCTION: HISTORICAL BACKGROUND

Sjögren's syndrome (SS) is a chronic inflammatory disease characterized by diminished lacrimal and salivary gland secretion (the sicca complex) resulting in keratoconjunctivitis sicca (KCS) and xerostomia (Talal,

1979; Strand and Talal, 1980; Shearn, 1977). The glandular insufficiency is secondary to lymphocytic and plasma cell infiltration. There is both a primary and secondary form of this disease as described below. Sjögren's syndrome is particularly important among the autoimmune diseases for two reasons. First, it is a common condition with perhaps 2–3 million affected individuals in the United States, the majority undiagnosed. Second, SS is a disorder in which a benign autoimmune process can terminate in a lymphoid malignancy. Thus, it is a crossroads disease that offers potential insight into the mechanisms whereby immunological dysregulation may predispose to a malignant transformation of B cells already involved in an autoimmune process.

The history of the disease goes back some 100 years. Patients with features of dry mouth, dry eyes, and chronic arthritis were described by European clinicians between 1882 and 1925. In 1892, Mikulicz reported a man with bilateral parotid and lacrimal gland enlargement associated with massive round cell infiltration. In 1933, Henrick Sjögren reported detailed clinical and histological findings in 19 women with xerostomia and keratoconjunctivitis sicca, of whom 13 had chronic arthritis. In 1953, Morgan and Castleman concluded that Sjögren's syndrome and Mikulicz's disease were the same entity.

II. GENERAL DESCRIPTION, CLASSIFICATION, AND ANIMAL MODELS

A. General Description

Sjögren's syndrome has been termed an autoimmune exocrinopathy. Ninety percent of the patients are females, probably explainable by the important influence of sex steroid hormones on immunoregulation and autoimmunity (Talal, 1982). The classic lesion is a lymphocytic and plasma cell infiltrate that may involve salivary, lacrimal, and often other exocrine glands, for example, those in the respiratory tract, gastrointestinal tract, and vagina (Greenspan, 1974). The term "benign lymphoepithelial lesion" has been used to describe the characteristic histological appearance in the salivary glands.

The exact mechanism responsible for the glandular destruction in SS is not known. Both T and B lymphocytes are present in the tissue lesions. Large amounts of immunoglobulin and rheumatoid factor are synthesized locally. β-2-microglobulin, a component of lymphocyte membranes that may play an immunological role, is increased in serum and saliva.

B. Classification

Sjögren's syndrome is classified into primary and secondary types. The clinical basis for this classification is that primary SS occurs in the absence of another connective tissue or autoimmune disease, whereas secondary SS accompanies another disease, usually rheumatoid arthritis. Secondary SS with rheumatoid arthritis constitutes ~50% of all patients with SS. In a small percentage of cases of SS there is another connective tissue disease accompanying the sicca complex, or the associated disease may be primary biliary cirrhosis (see Chapter 9). This clinical differentiation can now be validated by other markers; thus, only in primary SS is there an increased frequency of the HLA specificities B8 and DR3, and only in primary SS are the antibodies SS-A(Ro) and SS-B(La) demonstratable. However, rheumatoid factor is present in serum in both types.

C. Animal Models

There have been proposals for animal models of SS in two settings during the past decades. A "spontaneous" model was postulated on the basis of lymphoid infiltrates into salivary tissues in "autoimmune" strains of mice, initially NZB (Kessler, 1968) and more recently the MRL/lpr strain. The other setting has been the attempted induction of an autoimmune sialadenitis by immunization with salivary tissue in Freund's adjuvant, but this met with no consistent success. Neither type of model appears to have been strongly pursued by investigators nor would seem to provide clear insights into pathogenesis.

III. CLINICAL PRESENTATIONS

A. Primary Sjögren's Syndrome

The two most common presentations are, first, the insidious and slowly progressive development of the sicca complex and, second, the more rapid development of a severe oral and ocular dryness, often accompanied by episodic parotitis, in an otherwise well patient. The average patient's age is 50 years. The disease occurs in all races and in both children and adults.

About 50% of patients with the sicca syndrome have additional features of Sjögren's syndrome. The most common ocular complaint is a sensation, described as "gritty" or "sandy," of a foreign body in the eye.

Other symptoms include burning, accumulation of thick ropy strands at the inner canthus particularly on awakening, decreased tear formation, redness, photosensitivity, eye fatigue, itching, and a "filmy" sensation that interferes with vision. Gross inspection of the eyes may reveal nothing abnormal, mucous threads, or conjunctival congestion. An abnormal Schirmer test and rose bengal staining of the conjunctiva or inflammatory keratitis on slip-lamp examination are indications of lacrimal gland involvement.

Patients will complain of eye discomfort and difficulty in reading or watching television. Inability to cry is not a common complaint. Lacrimal gland enlargement occurs infrequently. Ocular complications include corneal ulceration, vascularization, or opacification, followed rarely by perforation.

The distressing manifestations of salivary insufficiency include difficulty with chewing, swallowing, and phonation; adherence of food to buccal surfaces, abnormalities of taste and/or smell; fissures of the tongue, buccal membranes, and lips, particularly at the corners of the mouth, the need for frequent ingestion of liquids, especially at mealtimes; and rampant dental caries. Patients are unable to swallow a dry cracker or toast without ingesting fluids and will manifest displeasure at the suggestion that they do so. They may carry a bottle of water or other lubricants in their purses and may awaken at night for sips of water. The dentist may notice that fillings are loosening. The tongue and mucous membranes are characteristically dry, red, and "parchment like." The lips may be dry and cracked. The tongue depressor adheres to oral surfaces, and the normal pool of saliva in the sublingual vestibule, visible when the tongue is elevated, is not present. Dryness may also involve the nose, the posterior pharynx and larynx, and the tracheobronchial tree, and may lead to epistaxis, hoarseness, recurrent otitis media, bronchitis, or pneumonia.

Half the patients have parotid gland enlargement, often recurrent and symmetrical and sometimes accompanied by fever, tenderness, or erythema. The parotid glands may have any consistency but are usually firm and nontender. Severe bilateral parotid swelling may give rise to the so-called chipmunk facies. Superimposed infection is rare. Rapid fluctuations in gland size are not unusual. A particularly hard or nodular gland may suggest a neoplasm.

Stimulated parotid salivary flow rates can be measured using lemon juice and cups that are placed over the orifices of the parotid ducts and secured by suction. Flow rates are reduced or unobtainable in SS. Salivary scintigraphy employing technetium pertechnetate shows an abnormal salivary gland function. Biopsy of the minor salivary glands in the lower lip makes it possible to obtain histological confirmation and to rule

out other diseases, including sarcoidosis or amyloidosis, entering into the differential diagnosis.

B. Secondary Sjögren's Syndrome and Associated Diseases

1. Rheumatoid Arthritis

The arthritis of Sjögren's syndrome resembles classic rheumatoid arthritis in its clinical, pathological, and roentgenographic features. The sicca syndrome develops in ~10–15% of patients with rheumatoid arthritis. Arthralgias and morning stiffness not progressing to joint deformity may occur in patients with the sicca complex. Fluctuations in the course of the arthritis are not accompanied by parallel alterations in the sicca symptoms. Splenomegaly and leukopenia suggestive of Felty's syndrome and vasculitis with leg ulcers and peripheral neuropathy may appear even in the absence of clinically evident rheumatoid disease.

2. Scleroderma

This is a recognized accompaniment of SS; Raynaud's phenomenon occurs in 20% of patients.

3. Skin

Skin or vaginal dryness and allergic drug eruptions occur frequently. Episodic lower extremity purpura, sometimes preceded by itching or other prodromata, may be the presenting complaint.

4. Renal Disease

Glomerulonephritis develops rarely and should suggest either coexisting systemic lupus erythematosus (SLE) or mixed (IgM-IgG) cryoglobulinemia. Overt or latent abnormalities of the renal tubules including diabetes insipidus, renal tubular acidosis, or Fanconi syndrome occur with some frequency.

5. Myopathy

Severe proximal muscle weakness, and rarely, tenderness may be early symptoms, leading to a diagnosis of polymyositis. Weakness may also be associated with electrolyte imbalance, nephrocalcinosis, and the clinical findings of renal tubular acidosis. Peripheral or cranial neuropathy may cause symptoms of dysesthesia or paresthesia. Facial pain and numbness

can accompany trigeminal neuropathy and contribute to the oral discomfort caused by dryness.

6. Other Immunopathic Diseases

Chronic thyroiditis of the Hashimoto type is present in 5% of patients, but clinical hypothyroidism is rare. Sjögren's syndrome was reported in 52% of patients with primary biliary cirrhosis and in 35% of patients with chronic active hepatitis, although some would regard these frequencies as overly high. Gastric achlorhydria, acute pancreatitis, and adult celiac disease have been reported in Sjögren's syndrome.

7. Lymphadenopathy

Cervical or other lymph node enlargement may the the first indication of malignant lymphoma or pseudolymphoma. It must be emphasized that in the vast majority of cases, lymphoproliferation remains confined to salivary and lacrimal tissue, resulting in a chronic benign course of stable or progressive xerostomia and xerophthalmia. In some, however, evidence of extension of lymphoproliferation to extraglandular sites appears even after 15 or more years of benign disease.

IV. HISTOPATHOLOGY

The description of Morgan and Castleman (1953) remains the standard reference for the histopathology of SS, and further observations are provided by Bloch *et al.* (1965) and Shearn (1971). The histological abnormalities in salivary and lacrimal glands in SS involve parenchyma and ducts. Overall the changes are patchy, initially with the lobular architecture of the gland being preserved and with varying degrees of lobular involvement. The parenchymal changes are marked by a dense lymphocyte-plasma cell infiltration, the degree of which tends to correlate with the extent of acinar damage. Lymph follicles with germinal centers may be evident. In the later stages, increases in the lobular septal fibrous tissue and scarring with acinar depletion become the dominant features, but lobular architecture can be preserved throughout the course of the disease. The ductal changes include hyperplasia of the lining cells of intraglandular ducts, and these proliferative changes in ducts may progress to the development of the characteristic "epimyoepithelial islands." The histological changes are well represented in the minor labial salivary glands, from which a simple biopsy procedure gives valuable diagnostic information.

Extensive lymphoid infiltrates may involve the lung, kidney, or skeletal muscle, resulting in functional abnormalities of these organs or frank lymphoma. In some patients, it may be impossible to decide between a benign or malignant process. The term "pseudolymphoma" has been employed in such situations.

The extraglandular lymphoid infiltrates are of two general types. They may be highly pleomorphic and include small and large lymphocytes, plasma cells, and large reticulum cells. In a lymph node the cells may distort the normal architecture and extend beyond the capsule, making the diagnosis between a benign and malignant lesion very difficult. The term pseudolymphoma has been applied when the lesions show tumor-like aggregates of lymphoid cells but do not meet histological criteria for malignancy. Periodic acid-Schiff–positive intranuclear inclusions and macroglobulins may be present, as in Waldenstrom's macroglobulinemia.

A. Pseudolymphoma

In pseudolymphoma, the site of extraglandular lymphoproliferation determines the clinical presentation. There may be hyperplasia of lymph nodes near salivary glands, and striking regional lymphadenopathy may be the predominant clinical feature. On the other hand, lymphoid infiltration may be selectively excessive in a distant organ such as kidney or lung. The involved organs may become functionally impaired, giving rise to renal abnormalities or pulmonary insufficiency. Renal tubular acidosis may arise through such a mechanism.

Features that should alert the clinician to the possibility of extraglandular lymphoproliferation in a patient with Sjögren's syndrome are regional or generalized lymphadenopathy, hepatosplenomegaly, pulmonary infiltrates, renal insufficiency, purpura, leukopenia, hypergammaglobulinemia, or elevated serum macroglobulin. Vasculitis may be associated, but arthritis is rare in such patients. The entity of pseudolymphoma cannot be clearly defined, but it occupies the middle portion of the spectrum of lymphoproliferation, merging with benign disease (e.g., hypergammaglobulinemic purpura) on the one end and with frankly malignant disease such as Waldenstrom's macroglobulinemia on the other.

B. Lymphoma

The other type of extraglandular lymphoid infiltrate is histologically malignant, enabling the specific diagnosis of a lymphoproliferative neoplasm. These lesions may also appear after several years of apparently benign disease, may or may not be preceded by pseudolymphoma, and

are often highly resistant to therapy. The most common malignant lymphoproliferation is a highly undifferentiated cell, which by immunohistological study clearly belongs to the B-cell lineage, often containing intracellular IgM-κ.

V. IMMUNOLOGY

A. Immunological Features

Abnormalities on laboratory tests include a mild normocytic, normochromic anemia that occurs in ~25% of patients; leukopenia occurs in 30%, eosinophilia (>6% eosinophils) in 25%, and an elevated erythrocyte sedimentation rate (>30 mm/h by the Westergren method) in over 90%. A mild hypoalbuminemia is common. There are increased levels of immune complexes and defects in reticuloendothelial clearance.

1. Hypergammaglobulinemia

Half of the patients have hypergammaglobulinemia, which is generally a diffuse elevation of all immunoglobulin classes. The patients with the greatest degree of hyperglobulinemia are those with primary SS rather than those with rheumatoid arthritis, and accompanying the hyperglobulinemia may be polymyopathy, purpura, or renal tubular acidosis. Monoclonal IgM may be seen. Cryglobulinemia, often of the mixed IgM-IgG type, may be present particularly in patients with glomerulonephritis or pseudolymphoma. Hyperviscosity associated with IgG rheumatoid factor and intermediate complexes has been reported. Some patients with lymphoma have hypogammaglobulinemia.

A low IgM level may herald the presence or development of malignant lymphoproliferation and is a poor prognostic sign. A fall in serum IgM is often accompanied by a reduction in the rheumatoid factor titer and may precede the onset of generalized hypogammaglobulinemia.

2. Autoantibodies

The presence of multiple serum autoantibodies is a characteristic feature. The majority of patients with primary Sjögren's syndrome are seropositive for rheumatoid factor and for antinuclear antibody (ANA), giving a homogenous, or more usually a speckled, pattern of immunofluorescence. An antibody [rheumatoid-associated precipitin (RAP)] to an Epstein–Barr virus-related nuclear antigen (RANA) occurs in secondary

Sjögren's syndrome with RA. It is now recognized that speckled ANA reactions characteristic of primary Sjögren's syndrome are due to an autoantibody to a ribonucleoprotein antigen known as La or SS-B, which occurs in at least 70% of patients with primary Sjögren's syndrome and to a lesser extent in Sjögren's syndrome with features of SLE (Alspaugh *et al.*, 1976; Martinez-Lavin *et al.*, 1979); antibodies to the related ribonucleoprotein Ro, also known as SS-A, also occur but are less specific for SS and are more a feature of a subset of SLE.

3. Historical Developments

Jones (1958) was the first to describe precipitating autoantibodies in the sera of a patient with Sjögren's syndrome using tissue extracts of salivary and lacrimal glands. Then Anderson and colleagues (1961) described two distinct autoantibody specificities, SjD (Sjögren's syndrome, reference serum from patient DON) and SjT (Sjögren's syndrome, reference serum from patient TRA), in 31% of patients with primary Sjögren's syndrome. These autoantibodies reacted with extracts prepared from a variety of human and animal tissues; sera containing SjD antibodies gave mainly cytoplasmic staining with either homogeneous and/or speckled nuclear staining, and SjT antibodies gave speckled nuclear and diffuse cytoplasmic staining by immunofluorescence on a liver substrate. The SjD antigen was shown to be resistant to heat, trypsin, DNAase, and RNAase, while the SjT antigen was sensitive to heat and trypsin but resistant to DNAase and RNAase.

Autoantibodies to a cytoplasmic antigen were described in the sera of patients with SLE by Clark and colleagues (1969). This antigen was called Ro, after the patient in whom it was first described, and its enzyme sensitivity pattern closely resembled that of Anderson's SjD antigen; the major difference was that antibodies to Ro were detected predominantly in patients with SLE (24%). Mattioli and Reichlin (1974) described another autoantibody specificity, anti-La, which was directed against a predominantly cytoplasmic antigen that was resistant to DNAase and RNAase but sensitive to heat and trypsin. Anti-La autoantibodies were detected in the sera of patients with Sjögren's syndrome and SLE. Alspaugh and Tan (1975) later detected two autoantibody specificities, anti-SS-A and anti-SS-B, in the sera of patients with primary Sjögren's syndrome, and a third autoantibody, anti-SS-C, in the sera of patients with secondary Sjögren's syndrome associated with rheumatoid arthritis. Localization experiments using both cells of the WIL-2 cell line and normal human leukocytes demonstrated that these antigens were present in both the nucleus and the cytoplasm. A further autoantibody specificity, anti-

Ha, was described by Akizuki and colleagues (1977) and was shown to be immunologically identical to the anti-SS-B specificity.

In a subsequent interlaboratory collaborative study, Alspaugh and Maddison (1979) demonstrated that Ro and SS-A were immunologically identical, and that La and SS-B also were immunologically identical. The physicochemical properties and clinical associations of SjD and SjT autoantibody specificities suggested that SjD may be related to Ro(SS-A) and SjT to La(SS-B). However, this has not been confirmed, as sera from Anderson's original patients are no longer available for study.

There has been considerable controversy in the literature over the cellular localization of the Ro and La antigens and their susceptibility to enzymatic treatment, since the cellular localization is dependent on the method of fixation used for immunofluorescence and the tissue employed, and because the enzymatic susceptibility varies depending on the tissue or cellular source of the antigen. As the molecular characterization of these antigens becomes more precise, resolution of these uncertainties should follow. Current information on the characterization of anti-La(SS-B) is described below; although anti-Ro(SS-A) has many resemblances to anti-La, it is more characteristic of SLE than of primary Sjögren's syndrome and hence is not specifically discussed in the following section.

4. Characterization of the La(SS-B) Autoantibody

Autoantibodies to La(SS-B) are reactive with a complex ribonucleoprotein (RNP) antigen, which consists of a heterogeneous group of small nuclear (sn) ribonucleic acids (RNAs) of cellular origin that are associated with a protein of MW 45,000–50,000 carrying the La(SS-B) antigenic determinant (McNeilage *et al.*, 1984). There are also four virus-encoded RNAs that can associate with the La(SS-B) protein. The cellular nuclear RNAs that associate with the La(SS-B) protein range in size from 80 to 300 nucleotides and include precursors for 5 S ribosomal RNA and for transfer RNA, plus other snRNAs that are also recognized by Ro(SS-A) autoantibodies when associated with a protein of MW 60,000 that carries the Ro(SS-A) antigenic determinant (Wolin and Steitz, 1984). The four virus-encoded RNAs that can associate with the La(SS-B) protein comprise two small RNAs of ~160 nucleotides, VAI and VAII, which are encoded by adenovirus and two small RNAs of ~170 nucleotides, EBER 1 and EBER 2, which are encoded by the Epstein–Barr virus (Lerner *et al.*, 1981). It is to be noted that neither the cell-derived nor the viral RNAs are antigenic unless bound to the cellular La(SS-B) protein. The fact that both the cellular and viral RNAs of the La(SS-B) antigen are transcribed by RNA polymerase III (Steitz *et al.*, 1982) raises the likelihood that the

La antigenic complex is important for RNA synthesis by RNA polymerase III.

B. Immunological Derangements: Sjögren's Syndrome as a Disorder of Immunoregulation

Sjögren's syndrome is a crossroads disease in which several aspects of immunological dysregulation come into play. I have elsewhere presented a hypothesis that views autoimmune disorders as arising from a background of physiological self-recognition (Talal, 1978). The immune system depends for its internal regulation on the recognition of self in the form of cell-surface Ia (DR) antigens and idiotypic determinants. Superimposed upon this intrinsic system of control is an extrinsic system in which immunity is influenced by classical endocrine and neuroendocrine mechanisms. These are mutual interactions insofar as the immune system itself probably helps regulate endocrine and perhaps even CNS pathways.

The immune system was discovered because of its role in natural defence against infections and its ability to respond to environmental infectious agents. However, over the last decade it has become apparent that the immune system is perhaps even more preoccupied with the organism's internal environment, to the point where many foreign invaders (e.g., viruses) are only recognized as a consequence of the simultaneous recognition of self. Thus, the immune system is both inward looking (surveying self-Ia and receptors on cell surfaces) and outward looking (responding to infectious agents). The former may be more primitive, having evolved over centuries to the system that we study today that is competent in defence as a result of evolutionary pressures and natural selection.

The proper maintenance of immune homeostasis depends on the functioning of these dual aspects of immunity. Immunoregulation is a dynamic process whose equilibrium is best maintained when both recognition systems function simultaneously. It follows, then, that immunoregulatory disturbances may arise when there are inappropriate pertubations or intrinsic defects in these recognition events.

Intercellular communication is at the heart of immunoregulation. Whether through the language of Ia (DR) or the language of idiotype, the cells involved with immunity talk to each other. A breakdown in this communication may result in the immunological chaos that we call autoimmunity.

Although it has been claimed that T cells provide the major means by which the immune system regulates itself, this is certainly not the only way. The balance between help and suppression achieved by immunore-

gulatory T cells is probably the major controlling element in the immune response. Abnormalities in this balance are characteristic of many autoimmune diseases. Yet, in some patients with SS, the numbers of helper and suppressor T cells are normal, and in others there are alterations in T-cell subsets. The autologous mixed lymphocyte response (AMLR) (Miyasaka *et al.*, 1980), which we have suggested may be a common denominator for the autoimmune and lymphoproliferative diseases of immunoregulation (Smith and Talal, 1982), are abnormal in SS due to deficiencies in the stimulating non-T cells as well as in the responding T cells.

More and more, in SS one is struck with defects in immune cells other than T and B cells, notably NK cells and macrophages. Natural killer (NK) cell activity in the blood is diminished as a consequence of immunoregulatory abnormalities. Associated defects may involve adherent cells and soluble factors, with these possibly more important than intrinsic deficits of NK cells (Miyasaka *et al.*, 1983).

C. Mechanisms of Damage

The lymphoid aggregates in the salivary tissues and their apparent relationship to the destruction of acinar tissue are most impressive; however, as yet there are neither adequate functional interpretations of these histopathological appearances nor any explanations for the occurrence of the lesions. B cells appear to be the earliest cells in the salivary gland lesions in SS, and salivary glands may take on the characteristics of organized lymphoid tissue with germinal centers, as seen in the thyroid gland in Hashimoto's disease. Evidence of the intense local B-cell activity in the glands includes polyclonal immunoglobulin production, autoantibody synthesis, and appearance of monoclonal immunoglobulins, and the latter can be present intracellularly as well as synthesitized and secreted.

T-cell infiltrates appear later in the salivary glands as the lesions become more extensive. An immunohistological analysis of the lymphoid infiltrates in primary SS was reported by Adamson *et al.* (1983), who used labial biopsy specimens as a source of tissue, various monoclonal antibodies to lymphoid cells, and a four-step biotin-avidin immunoperoxidase technique on frozen tissue sections. T lymphocytes predominated heavily over other cell types, numbering >75% of all infiltrating cells; a small minority of the cells carried B-cell markers, and a specific subset was present in salivary tissue that was lacking from peripheral blood. The majority of the T lymphocytes present belonged to the T-helper subset, and >50% of the T cells carried activation antigens (Ia-positive, OKT10-positive).

D. IMMUNOGENETICS

Autoimmune diseases are multifactorial, combining genetic, hormonal, immunological, and possibly viral influences. Sjögren's syndome is one of several autoimmune diseases associated with the histocompatibility antigens HLA-B8, -DR3; this genetic predisposition is also seen in celiac disease, dermatitis herpetiformis, myasthenia gravis, Graves' disease, chronic active hepatitis, idiopathic Addison's disease, and insulin-dependent diabetes mellitus. Certain Ia determinants occur more frequently in SS than in the general population.

It is noteworthy, from the standpoint of classification, that the increase in HLA-B8, -DR3 is found only in the primary and not in secondary types of SS. In particular, in SS with rheumatoid arthritis there is an increase in HLA-DR4 but no increase in HLA-B8. As described above, there is a strong association between antibody to the ribonucleoprotein SS-B(La) and primary SS; the association between SS-B(La) and HLA-B8 appears to be as strong if not stronger than with primary SS itself.

There is another strong genetic association with the SS-B(La) antibody, this being with the *Km(1)* immunoglobulin allotype; this was found to be present in 42% of subjects positive for SS-B(La) as opposed to 16% of controls (Whittingham *et al.*, 1984). Thus, taken together, there are strong genetic determinants for production of anti SS-B(La), namely female gender, *HLA-B8*, and *Km(1)*. The implications of these immunogenetic associations are considered in Chapter 3.

VI. TREATMENT AND OUTCOME

Most often, Sjögren's syndrome is a benign disease in which conservative management is the best guide to therapy.

The sicca complex is treated with fluid replacement supplied as often as necessary. There are several readily available ophthalmic preparations (e.g., Tearisol, Liquifilm, and 0.5% methylcellulose) that will adequately replace the deficient tears. In severe situations, patients instill these as often as every 1/2 to 1 h. If corneal ulceration is present, eyepatching and boric acid ointment may be necessary. It is more difficult to compensate for the salivary insufficiency. The frequent ingestion of fluids, particularly with meals, is often the best solution. Patients should see their dentists every 4 months and pay scrupulous attention to proper oral hygiene. The careful use of a Water Pic after eating may reduce the incidence of caries.

A more aggressive lymphocytic disorder develops in a minority of patients with the lymphocytes in extrasalivary sites, kidney, or lung, being primarily B cells which often produce monoclonal IgM: This character-

izes the pseudolymphoma stage. Corticosteroids or immunosuppressive drugs are often indicated in the treatment of pseudolymphoma, particularly when there is renal or pulmonary involvement. Cyclophosphamide at a dosage of 75–100 mg daily may diminish extraglandular lymphoid infiltrates and may even restore salivary gland function in some patients.

The truly malignant disease lymphoma of SS again involves B cells, although now there may be inadequate serum immunoglobulins and humoral immunodeficiency; autoantibodies can disappear from the blood and patients may be unable to resist infections such as pneumococcal pneumonia. Malignant lymphoma should be treated with intensive chemotherapy, surgery, and/or radiotherapy as indicated by the location and extent of disease. These lymphomas can be highly malignant and often fatal tumors that require rapid and skilled intervention; even with this treatment, the prognosis is poor.

VII. CONCLUDING REMARKS: FUTURE PROSPECTS

Sjögren's syndrome represents the expression of autoimmune processes of considerable complexity. Future investigations will be directed towards establishing the specificity of the anti-La(SS-B) antinuclear antibody for diagnosis of primary SS, elucidating the pathogenetic differences between the primary and secondary (rheumatoid-associated) types of the disease, discovering the target of the autoimmune attack, which is not as yet evident, and exploring the possibility that virus infection has an initiating role in primary SS.

VIII. SUMMARY

Sjögren's syndrome is an autoimmune exocrinopathy that predominantly affects females (90%) and has two clearly defined forms, a primary form strongly associated with HLA-B8, -DR3 and a secondary form associated with rheumatoid arthritis and certain other immune-mediated diseases. There exists no convincing animal model.

The disease primarily affects salivary and lacrimal tissues but can have diffuse expressions attributable either to an involvement of other exocrine secretory tissues, associated autoimmune disease, or deposition of immune complexes. Histologically, there is intense lymphocyte infiltration into the affected glands with progressive acinar destriction. Lymphoma occasionally supervenes. The primary type of SS is marked by a characteristic antinuclear autoantibody, anti-La(SS-B), which also occurs in

SLE. Treatment is palliative and is directed to artificial replacement of tears and to oral hygiene.

REFERENCES

Adamson, T. C., Fox, R. I., Frisman, D. M., and Howell, F. V. (1983). *J. Immunol.* **130,** 203.

Akizuki, M., Powers, R., and Holman, H. R. (1977). *J. Clin. Invest.* **59,** 264–272.

Alspaugh, M. A., and Maddison, P. (1979). *Arthritis Rheum.* **22,** 796–798.

Alspaugh, M. A., and Tan, E. M. (1975). *J. Clin. Invest.* **55,** 1067–1073.

Alspaugh, M. A., Talal, N., and Tan, E. M. (1976). *Arthritis Rheum.* **19,** 329–340.

Anderson, J. R., Gray, K. G., and Beck, J. S. (1961). *Lancet* **ii,** 456–460.

Bloch, K. J., Buchanan, W. W., Wohl, M. J., and Bunim, J. J. (1965). *Medicine (Baltimore)* **44,** 187–231.

Clark, G., Reichlin, M., and Tomasi, T. B. (1969). *J. Immunol.* **102,** 117–122.

Greenspan, J. S., Daniels, T. E., Talal, N., and Sylvester, T. A. (1974). *Oral Surg., Oral Med. Oral Pathol.* **37,** 217–229.

Jones, B. R. (1958). *Lancet* **ii,** 773–776.

Lerner, M. A., Andrews, N. C., Miller, G., and Steitz, J. A. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 805–809.

McNeilage, L. J., Whittingham, S. F., and Mackay, I. R. (1984). *J. Clin. Lab. Immunol.* **15,** 1–17.

Martinez-Lavin, M., Vaughan, J. H., and Tan, E. M. (1979). *Ann. Intern. Med.* **91,** 185–190.

Mattioli, M., and Reichlin, M. (1974). *Arthritis Rheum.* **17,** 421–429.

Miyasaka, N., Sauvezie, B., Pierce, D. A., Daniels, T. E., and Talal, N. (1980). *J. Clin. Invest.* **66,** 928–933.

Miyasaka, N., Seaman, W., Bakshi, A., Sauvezie, B., Strand, V., Pope, R., and Talal, N. (1983). *Arthritis Rheum.* **26,** 954–960.

Morgan, W. S., and Castleman, B. (1953). *Am. J. Pathol.* **29,** 471.

Shearn, M. A. (1971). "Sjögren's Syndrome." Saunders, Philadelphia, Pennsylvania. Shearn, M. A. (1977). *Med. Clin. North Am.* **61,** 271–282.

Sjögren, H. (1943). "A New Conception of Keratoconjunctivitis Sicca (Keratitis Filiformis) in Hypofunction of the Lachyrmal Glands" Australasian Medical Publ. Co., Sydney. (Transl. by J. B. Hamilton.)

Smith, J. B., and Talal, N. (1982). *Scand. J. Immunol.* **16,** 269–278.

Steitz, J. A., Wolin, S. L., Rinke, J., Petterson, I., Mount, S. M., Lerner, E. A., Hinterberger, M., and Gottlieb, E. (1982). *Cold Spring Harbor Symp. Quant. Biol.* **47,** 893–900.

Strand, V., and Talal, N. (1980). *Bull. Rheum. Dis.* **30,** 1046–1052.

Talal, N. (1978). *Arthritis Rheum.* **21,** 853–861.

Talal, N. (1979). *In* "Arthritis" (D. J. McCarty, ed.), 9th ed. pp. 810–824. Lea & Febiger, Philadelphia, Pennsylvania.

Talal, N. (1982). *Rheum. Dis.* **8,** 23–28.

Wolin, S. L., and Steitz, J. A. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81,** 1996–2000.

CHAPTER 7

Autoimmune Thyroid Disease

PIERLUIGI E. BIGAZZI
Department of Pathology
University of Connecticut Health Center
Farmington, Connecticut

NOEL R. ROSE*
Department of Immunology and Infectious Diseases, and Department of Medicine
The Johns Hopkins Medical Institutions
Baltimore, Maryland

I. AUTOIMMUNE THYROIDITIS

A. Introduction: Historical Background

Autoimmune thyroiditis is a chronic disorder of the thyroid gland that is characterized by humoral and cell-mediated immune responses to thyroid-specific antigens as well as inflammatory infiltration and tissue damage localized in the thyroid. It was first recognized in 1912 by Hashimoto, a surgeon from the University of Kushu (Japan), who reported the histo-

* Aided by PHS research grants AM 31632 and AG 04342 from the NIH.

pathological findings in the resected thyroids of four middle-aged women with goiter (Hashimoto, 1912). Because of the massive and diffuse lymphocyte infiltration, Hashimoto defined the condition as "struma lymphomatosa." Over the years, other synonyms have been lymphadenoid goiter, lymphoid goiter, chronic lymphocytic throiditis, and more commonly, Hashimoto's thyroiditis. Although both histopathological and clinical aspects of this disorder have been known for a long time, its autoimmune nature has been appreciated only in the last 27 years, following the pioneering studies of Witebsky and his associates on experimentally induced thyroiditis and of Doniach and her associates on Hashimoto's thyroiditis.

Ernest Witebsky initiated his investigations of organ-specific antigens of the brain, adrenals, and thyroid during his stay at the Institute for Experimental Cancer Research in Heidelberg; his studies were directed by Hans Sachs, one of Paul Erlich's students and associates (Witebsky and Steinfeld, 1928; Witebsky, 1929; Witebsky and Klinke, 1933). Witebsky resumed these studies after his move to the United States (Witebsky *et al.*, 1955; Rose and Witebsky, 1955; Shulman *et al.*, 1955). In Witebsky's own words (1968), "thyroglobulin seemed to be the logical candidate for investigating the feasibility of autosensitization." Therefore, a series of experiments was undertaken to determine whether thyroid antibodies could be detected in the sera of rabbits injected with rabbit thyroid extract in complete Freund's adjuvant. Indeed, both autoantibodies to thyroglobulin and inflammatory changes of the thyroid gland were observed in immunized animals (Rose and Witebsky, 1955, 1956; Witebsky and Rose, 1956).

At the same time, Deborah Doniach had begun an investigation of the diagnostic and biochemical aspects of Hashimoto's disease. The raised serum immunoglobulin levels, their delayed return to normal after thyroidectomy, and the infiltration of the thyroid with lymphocytes and plasma cells suggested that an immune response against an antigen of the thyroid gland might be involved in this disease. When the sera of nine patients with Hashimoto's thyroiditis were tested by precipitation against thyroid extract or thyroglobulin, positive reactions were obtained with seven sera (Roitt *et al.*, 1956, 1958; Doniach and Vaughan Hudson, 1957; Doniach and Roitt, 1957; Doniach *et al.*, 1957).

These seminal investigations established the autoimmune nature of Hashimoto's thyroiditis, and were followed in rapid succession by numerous publications from other investigators that confirmed and extended the findings of both Witebsky and Doniach. As a consequence, the term "autoimmune thyroiditis" has gained increasing acceptance over the years, although it has been recently criticized because it covers too many disparate clinical states (Doniach *et al.*, 1979).

B. General Description

The following points were already evident in the clinical description of the first four patients with autoimmune thyroiditis (Hashimoto, 1912):

1. The patients were all middle-aged women between 40 and 61 years old.
2. Their major if not only complaint was the development of a mass in the neck.
3. The goiter affected both lobes of the thyroid and was a moderate size (from a pigeon's to a hen's egg).

These points are still valid, even though later studies showed that the clinical picture of autoimmune thyroiditis can be quite complex and heterogeneous (reviewed in Doniach *et al.*, 1979; Pinchera *et al.*, 1980; DeGroot and Sridama, 1983). Hashimoto's thyroiditis occurs more frequently (5–20 times more) in females than males, with a peak incidence between 30 and 50 years of age. A "juvenile" variant occurs in children and adolescents. The prevalence of thyroiditis in the general population may vary between 0.5 and 1.2%, and its annual rate of incidence is increasing (Furszyfer *et al.*, 1972). Many patients have no symptoms, and a small goiter is detected during a routine checkup; in others, a feeling of pressure in the neck or of general malaise may lead to the discovery of a goiter. Occasionally the presenting symptoms are hoarseness, mild dysphagia, and coughing. The thyroid gland is moderately enlarged, horseshoe shaped, of firm rubbery consistency, and has a finely lobulated surface. Thyroid function is usually normal, but it can be decreased, or less commonly, increased. Spontaneous remissions with decrease in goiter size and thyroid antibody titers have been described, especially in the variant affecting adolescents.

C. Animal Models

Autoimmune thyroiditis can be experimentally induced, but also occurs spontaneously in animals (reviewed by Bigazzi and Rose, 1975; Rose *et al.*, 1977, 1981; Weigle, 1980). Animal models of autoimmune thyroiditis have provided valuable information on the mechanisms underlying this disorder, and their study remains an indispensable adjunct of clinical investigations in humans.

1. Experimental Autoimmune Thyroiditis

Autoimmune responses to autologous or syngeneic thyroglobulin, as well as autoimmune thyroiditis, can be deliberately induced in virtually all

animal species. The literature on experimental autoimmune thyroiditis (EAT) is quite extensive and several reviews have recently appeared (Lewis and Rose, 1985). Therefore we will summarize here the more interesting results obtained from the study of some animal models.

a. Experimental Autoimmune Thyroiditis in Rabbits. Investigations of thyroid organ-specific autoantigens in rabbits (Witebsky and Rose, 1956; Rose and Witebsky, 1956) opened the way to immunological studies of human autoimmune thyroiditis. Rabbits immunized with thyroglobulin in complete Freund's adjuvant develop autoimmune responses to thyroglobulin as well as infiltration of their thyroids with inflammatory cells, with a histopathological picture resembling that of Hashimoto's disease. In rabbits, delayed skin reactions to autologous thyroglobulin occur only several weeks after thyroid lesions and autoantibodies to thyroglobulin have developed, an observation suggesting that in this species the thyroiditis may be more closely associated with humoral autoimmune responses than with delayed hypersensitivity. Interestingly, rabbits with EAT may also develop an immune complex-mediated glomerulonephritis, another finding underlying the role of humoral factors in rabbits.

b. Experimental Autoimmune Thyroiditis in Guinea Pigs. The more interesting data obtained in studies of guinea pigs with EAT have resulted from transfer experiments (Sharp *et al.*, 1974). These experiments have demonstrated that EAT can be transferred to normal inbred guinea pigs by the inoculation of lymphocytes from syngeneic animals with EAT, but more severe thyroid lesions are observed in animals receiving both autoantibodies to thyroglobulin and lymphocytes. Therefore, in guinea pigs the thyroiditis seems to be determined by both humoral and cellular autoimmune responses to thyroglobulin.

c. Experimental Autoimmune Thyroiditis in Mice. Investigations in inbred mice with EAT have provided the most sophisticated evidence available to date on the inheritance of organ-specific autoimmunity. In this species, autoimmune responses to murine thyroglobulin and EAT are controlled by at least two genes mapping within the MHC and one or more non-*H-2* genes (reviewed in Rose *et al.*, 1981; Rose and Kong, 1982). One *Ir* gene (located in the *I-A* subregion) determines the extent of immunological response in terms both of antibody production and pathological lesions. Other genes (located at the *K* and *D* ends) act on effector responses and moderate the severity of histological damage. In mice, serum levels of autoantibodies to thyroglobulin are usually a good predictor of the severity of thyroiditis. However, it is possible to dissociate autoantibody formation from thyroid damage. This suggests a more im-

portant role of effector T cells in the development of thyroid inflammatory lesions. The action of effector T cells was recently demonstrated *in vitro* (Creemers *et al.*, 1983). Investigations in mice have also indicated a major role of T suppressor cells in the maintenance of tolerance to thyroglobulin (Rose *et al.*, 1981).

d. Experimental Autoimmune Thyroiditis in Rats. Studies of the genetic control of EAT in rats have provided results quite different from those obtained in mice (Rose, 1975). Immune responsiveness in terms of autoantibody formation is not linked to the MHC, while thyroid damage may be linked to genes mapping within the MHC. With the increasing availability of congenic strains of rats, it should now be possible to study the effects of background genes that could be obscuring the detection of MHC-linked genes.

Equally important are the results obtained in rats neonatally thymectomized and sublethally irradiated (Penhale *et al.*, 1973, 1975, 1976; Kotani *et al.*, 1981, 1982). These animals develop both autoimmune responses to thyroglobulin and inflammatory changes in their thyroids.

2. Spontaneous Autoimmune Thyroiditis

Autoimmune responses to thyroglobulin and inflammatory infiltration of the thyroid both may occur in animals without any experimental manipulation. The term "spontaneous autoimmune thyroiditis" (SAT) underlines our ignorance of the initiating mechanisms of this condition. SAT can occur in beagles, marmoset monkeys, chickens, and rats. To date, the more interesting observations have been obtained in the latter two species.

a. Chickens. The obese strain (OS) chickens were initially selected from a White Leghorn flock bred at Cornell University on the basis of their appearance, that is, the expression of severe hypothyroidism. It was quickly realized that the affected birds experienced intense autoimmune responses to thyroiditis as well as a severe, destructive thyroiditis. The condition is made worse by neonatal thymectomy and improved by early bursectomy. The spontaneous development of autoimmune thyroiditis is determined by the combination of three independent genetically determined lesions: a strong autoimmune response to thyroglobulin encoded by an immune response gene linked to the MHC of the chicken, a malfunction of thymic effector–suppressor balance, and a defect in thyroid cell function. Genes outside the MHC control SAT in chickens, as shown by studies of F_2 hybrids between the susceptible OS chicken and the normal progenitor CS line.

b. Rats. Autoimmune responses to thyroglobulin and inflammatory infiltration of thyroid occur spontaneously in inbred Buffalo (BUF) rats. The disorder resembles Hashimoto's thyroiditis in that it preferentially affects older females: ~25% of retired female breeders older than 1 year were found to have SAT. Neonatal thymectomy as well as the administration of methylcholanthrene or other chemicals increase the severity of SAT and accelerate its appearance. We have succeeded in obtaining antiidiotypic antibodies directed against autoantibodies to rat thyroglobulin from BUF rats with SAT (Zanetti and Bigazzi, 1981). Passive transfer experiments have shown that autoimmune responses to thyroglobulin could be reduced by the repeated injection of such antiidiotypic antibodies.

A new model of autoimmune endocrinopathy was recently described in the Bio Breeding/Worcester (BB/W) rats that develop spontaneously insulin-dependent diabetes mellitus and lymphocytic thyroiditis (Sternthal *et al.*, 1981).

D. Histopathology

As first described by Hashimoto (1912), the thyroid gland of patients with autoimmune thyroiditis is usually enlarged, with abundant inflammatory infiltration, hyperplasia, and damage of the parenchyma (reviewed in Bigazzi, 1979).

The inflammatory infiltrate is composed of lymphocytes, plasma cells, and macrophages. Lymphocytes are closely associated with thyroid follicular cells and are often observed crossing into the follicular lumen, between epithelial cells, or within damaged cells. Many thyroid follicles contain inflammatory cells (mostly plasma cells, but also lymphocytes and macrophages) within their lumen. Characteristically, lymphoid follicles with distinct germinal centers are present in the thyroid tissue. The basement membrane of the follicular epithelium shows gaps as well as diffuse and irregular thickening, occasionally assuming a multilayered aspect (Kalderon *et al.*, 1973). The most significant ultrastructural finding at this level is the presence of distinct and irregular globoid electron-opaque deposits, almost always associated with plasma cells in close apposition to the follicular basement membrane. The parenchymal damage is focal, with follicles undergoing changes from normal to necrotic. The epithelial cells of the thyroid follicles may be hyperplastic, and especially in more advanced stages, swollen and oxyphilic. Such cells have been termed "oxyphil cells," "oncocytes," "Hurthle cells," or "Askanazy's cells." They are characterized by increased cell size and the accumulation in their cytoplasm of eosiniphilic granules that correspond to mitochondria increased in number and size. It should be noted that

these cells are not typical of thyroiditis alone, and can also be found in other organs. Cell death with destruction of follicles is also observed, and frequently there is irregular focal fibrosis or extensive fibrous replacement of the thyroid parenchyma (fibrous variant of Hashimoto's thyroiditis).

Recent ultrastructural studies have provided additional information on the histopathology of autoimmune thyroiditis. Knecht and Hedinger (1982) noted three major types of thyroid epithelial cells, that is, normal (stimulated) follicular cells, oncocytes, and degenerating follicular cells. The largest areas of infiltration were represented by lymphoid follicles with active germinal centers. The infiltrating cells included lymphocytes, plasma cells, monocytes, macrophages, and giant cells. Plasma cells were the most numerous among the infiltrating cells, and were observed within the lumen of follicles and within macrophages. Two types of macrophages could be distinguished, one with abundant cytoplasmic organelles and another characterized by phagocytic properties. Among the ingested material were lymphocytes, plasma cells, and possibly remnants of epithelial cells. Giant cells, most likely of macrophage origin, were observed within the lumen of thyroid follicles. Finally, "interdigitating" cells, with a bulky electron-lucent cytoplasm, a well-developed tubulovesicular system, and long fingerlike processes were identified within small clusters of lymphocytes and in close contact with thyroid epithelial cells. Shamsuddin and Lane (1981) also reported variable morphological features of thyroid follicular cells and noted that inflammatory cells infiltrating the thyroid follicles were located in between the epithelial cells, with an evident clear space between them. No evidence of migration by inflammatory cells through follicular epithelial cytoplasm (emperipolesis) was observed.

There are numerous histopathological classifications of autoimmune thyroiditis (reviewed by Bigazzi, 1979). One of the most commonly accepted was first introduced by Woolner (Woolner *et al.*, 1959; Woolner, 1964), who distinguished "focal" from "diffuse" thyroiditis and within the latter group identified a "lymphoid" type, an "oxyphilic epithelium" type, and one with "pronounced epithelial destruction." Recently Yagi (1981) performed electron microscopic and immunohistochemical studies on thyroids classified within one of these three groups. He has found that follicular structure was usually maintained in the "lymphoid" (L) and "oxyphilic epithelium" (O) types, as compared to that observed in the "pronounced epithelial destruction" (P) type, but degenerating follicles with focally broken or very stretched basement membranes were often found in both L and O types. Destruction of basement membranes by infiltrating lymphocytes and plasma cells was most often observed in the O type. Cells containing immunoglobulins (mostly IgG) in the stroma and the follicles and immunoglobulin deposits (mostly IgG) in the basement

membranes, colloid, and stroma were observed in the three types of thyroiditis. Similarly, Knecht *et al.* (1981) reported that most of the infiltrating plasma cells contain IgG.

Doniach *et al.* (1979) reviewed the main features of the three histological variants of Hashimoto's thyroiditis and pointed out that the pattern of the disease has changed considerably in the last 27 years. In the period between 1940 and 1950, Hashimoto's disease was considered quite rare and the patients usually had firm, diffuse, and rubbery goiters, They could be hypothyroid, euthyroid, and even hyperthyroid. The first patients studied immunologically had high titers of circulating autoantibodies to thyroglobulin, capable of precipitating in agar when reacted against that autoantigen. Histopathologically, these cases showed severely destructive thyroiditis. On the other hand, in more recent years Hashimoto's disease has been diagnosed at earlier stages because of our increased awareness and understanding of this condition. Therefore, milder cases have become commonplace in thyroid clinics. The majority of patients now have small or moderate-sized goiters and thyroid function tests that are normal or indicative of borderline hypothyroidism. The sera of such patients contain only traces of autoantibodies to thyroglobulin, while autoantibodies to microsomal antigens are invariably present. Histologically, their thyroids belong to the "hypercellular" or "oxyphil" variants of thyroiditis, characterized by a dense lymphocytic infiltrate with germinal centers and numerous Askenazy cells.

Autoimmune thyroiditis may be associated with various other disorders that possibly are of an autoimmune nature (reviewed by Bigazzi, 1979). Thyroiditis and Addison's disease can occur in the same patient, a condition defined as "Schmidt's syndrome." Thyroiditis may also be associated with myasthenia gravis, diabetes mellitus, systemic lupus erythematosus, and rheumatoid arthritis.

Finally, a few cases of immune complex-mediated glomerulonephritis have been reported (O'Reagan *et al.*, 1976; Ploth *et al.*, 1978; Jordan *et al.*, 1978, 1981; Kalderon, 1980; Weetman *et al.*, 1981; Verger *et al.*, 1983). In these cases, the immune deposits were found to contain autoantibodies to thyroglobulin and thyroglobulin, but autoantibodies to thyroid microsomal antigens were also present or their presence could not be excluded. Therefore, it is difficult to determine which of the two types of autoantibodies was responsible for the glomerulonephritis. It is well known that complexes formed between thyroglobulin and autoantibodies to thyroglobulin do not bind complement in significant amounts (see p. 176). This observation may suggest that the autoantibodies to microsomal antigens may have a major role in the few cases of immune complex glomerulonephritis observed in thyroiditis patients. On the other hand,

rabbits injected with denatured thyroglobulin experience a glomerulonephritis due to the deposition of complexes containing thyroglobulin and autoantibodies to thyroglobulin (Nakamura and Weigle, 1969; Germuth *et al.*, 1978) and it is possible that such a situation may occasionally occur in humans.

E. Immunology

1. Immunological Features

As previously mentioned, patients with autoimmune thyroiditis experience both humoral and cell-mediated immune responses to thyroid-specific antigens. Therefore, in this section we will first discuss the autoantigens characteristic of the thyroid gland and then examine the autoimmune responses that they are capable of generating.

a. Thyroid-specific Antigens. The thyroid contains several autoantigens, both in the colloid and the epithelial cells lining the follicles (reviewed by Shulman, 1971; Bigazzi, 1979).

Thyroglobulin, stored in the thyroid colloid, is the main iodoprotein of the thyroid gland and is the storage form of the thyroid hormones. It is also found in small amounts (10–100 ng/ml) in the circulation. It is a large, compact, water-soluble glycoprotein with a molecular weight of ~660,000. Most thyroglobulin molecules have a sedimentation coefficient of 19S and contain two identical subunits, each consisting of a single polypeptide chain with a molecular weight ~300,000. Only certain portions of the thyroglobulin molecule may be capable of causing autoimmune responses: It was estimated that each molecule possesses ~50 antigenic determinants, of which 6 may be autoantigenic. Recently, Nye *et al.* (1980) noted that autoantibodies to thyroglobulin in different human sera are directed against the same epitopes of thyroglobulin, confirming once more that only certain restricted parts of this molecule are capable of inducing autoimmune responses. Their data are consistent with an estimate of two different major epitopic specificities, with occasional sera having antibodies directed against a third site. By a different approach and utilizing monoclonal antibodies to mouse thyroglobulin, Rose and his associates (1982) determined the presence of at least 8 different autoantigenic determinants on the thyroglobulin molecule.

A second autoantigen, also stored in the colloid and termed *colloid antigen second* (CA2), is less well defined than thyroglobulin. The autoantigen CA2 constitutes a small percentage of the proteins in the colloid, is

different from thyroglobulin, appears to contain no iodine, and is distinct from the pH 3.6 thyroid protease. Although it was first described in 1961, there is still no definite information about its structure and antigenic determinants.

The epithelial cells of the thyroid follicles possess an antigen localized in their cytoplasm and termed *microsomal* because it was separated by ultracentrifugation in the microsomal fraction, distinct from the mitochondrial and nuclear components. The antigen is intimately associated with the lipoproteins of microvesicles with smooth profiles, which are particularly abundant at the apical margin of the thyroid cells and represent newly synthesized thyroglobulin droplets originating from the Golgi complex (reviewed by Doniach and Roitt, 1976). Mariotti *et al.* (1979) solubilized thyroid microsomal antigen(s) by several agents, including detergents, hypertonic salts, and proteolytic enzymes. The greatest efficiency in terms of absolute amounts of solubilized antigen was provided by Triton X-100, followed by deoxycholate. These results support the concept that thyroid microsomal antigen(s) are an integral membrane component (Pinchera *et al.*, 1980). Additional evidence has also been provided by Khoury *et al.* (1981), who showed that some cell-surface antigens are identical with the thyroid "microsomal" antigens and that "microvillar/microsomal" surface antigens are normally localized on the apical pole of thyroid follicular cells.

Very little is known about other organ-specific cell-surface autoantigens of the thyroid epithelium. They have been identified by mixed hemadsorption (Kite *et al.*, 1965; Jansson *et al.*, 1968) and indirect immunofluorescence, and have been found to be distributed in discrete areas over the cell surface (Fagraeus and Jansson, 1970). Their relationship to the microvillar/microsomal autoantigens is still unclear and is currently under investigation.

Other autoantigens that may cause autoimmune responses in patients with autoimmune thyroiditis are the thyroid hormones T_3 and T_4, the receptor for TSH (see section on Graves' disease), and tubulin, the fundamental protein of microtubules.

b. Autoimmune Responses to Thyroid Antigens. Patients with autoimmune thyroiditis experience both humoral and cell-mediated immune responses to the various thyroid antigens described in the previous sections.

Humoral Immune Responses. Most patients with autoimmune thyroiditis have circulating autoantibodies capable of reacting *in vitro* with thyroglobulin, CA2, "microsomal" antigen, and other cell-surface au-

toantigens (reviewed in Bigazzi, 1979). Occasionally, serum autoantibodies to T_3, T_4, and the TSH receptor are also detected.

Autoantibodies to Human Thyroglobulin (AHT). AHT can be detected by a variety of procedures, such as precipitation in agar, indirect immunofluorescence, passive hemagglutination, radioimmunoassay, and ELISA. However, most of the data reported in the literature have been obtained using the tanned-cell passive hemagglutination procedure; as shown in Table I, AHT have been detected by this method in the sera of 76–91% of the patients with autoimmune thyroiditis. Sera from histologically proven cases may be negative by the passive hemagglutination procedure; thus, a negative result does not exclude the diagnosis of autoimmune thyroiditis. In such cases, indirect immunofluorescence for CA2 and tests for microsomal antibodies may be useful. Autoantibodies to thyroglobulin are also detected in the sera of 63–82% of patients with primary myxedema, and when this condition has been staged on a clinical basis, AHT have been found in 47% of the patients with subclinical hypothyroidism, 75% of those with mild hypothyroidism, and in 43% of those with overt hypothyroidism. Autoantibodies to human thyroglobulin are also detected in 28–65% of patients with thyroid carcinomas, 28–50% of patients with pernicious anemia, and up to 11% of subjects without overt thyroid disorders (Tables I and II). When the incidence of thyroglobulin antibodies in normal subjects is analyzed in relationship to age and sex, it is found that normal women between the ages of 21 and 70 have a higher incidence of such autoantibodies than men. Titers of AHT detected by tanned-cell passive hemagglutination are extremely high in patients with the fibrotic variant of Hashimoto's thyroiditis, and tend to be lower in most of the other patients with positive reactions.

Recently, both radioimmunoassays and ELISA have been employed to detect serum AHT. Approximately 86–100% of patients with autoimmune thyroiditis, as well as 87–89% of patients with untreated Graves' disease and 69–94% of those with primary myxedema were found to have circulating AHT demonstrable by radioimmunoassay. Results obtained by ELISA have been published by several investigators (Engvall and Perlmann, 1975; Endo *et al.*, 1980; Voller *et al.*, 1980; McCoy *et al.*, 1983). In general, it has been stated that the results by ELISA correlate quite well with those obtained by passive hemagglutination and radioimmunoassay (Endo *et al.*, 1980). However, we tend to agree that, at least in the published protocols, ELISA seems to have a lower sensitivity than radioimmunoassay HX (Voller *et al.*, 1980).

Autoantibodies to human thyroglobulin are present in all immunoglobulin classes, with most of the antibody in the IgG class; up to 20% is IgA

TABLE I

Incidence of Thyroglobulin Autoantibodies in Various Thyroid Diseases[a]

Thyroiditis[b]				
Acute thyroiditis	Subacute thyroiditis	Struma lymphomatosa (Hashimoto)	Chronic nonspecific thyroiditis	Fibrotic thyroiditis (Riedel)
2/8	7/20 (35)	10/32 (31)	72/95 (76)	3/6
—	13/19	96/106 (91)	—	—
7/15	8/15	21/23 (90)	—	—
1/3	—	8/9	—	—
—	—	99/111 (89)	—	—
0/3	1/2	3/5	—	—
2/2	1/1	14/16	—	—
—	—	55/66 (83)	—	—

Other Thyroid Diseases[b]						
Primary (spontaneous) myxedema	Thyroid hyperplasia or thyrotoxicosis	Colloid goiter	Nodular goiter	Adenoma	Carcinoma	Controls (hospital population)
23/32 (72)	88/266 (33)	15/198 (8)	—	—	3/16	72/1432 (5)
66/101 (65)	103/181 (57)	62/198 (31)	38/102 (37)	—	11/39 (28)	10/148 (7)
11/15	53/93 (57)	43/328 (13)	—	17/48 (28)	46/78 (59)	—
24/38 (63)	49/128 (38)	2/78 (3)	8/97 (8)	1/9	5/17	19/178 (11)
18/22 (82)	25/29 (86)	28/36 (78)	—	—	13/20 (65)	—
7/10	34/94 (46)	8/19	8/19	1/8	0/2	8/219 (4)
4/9	4/26 (15)	3/19	3/16	1/9	2/5	—
—	—	—	9/43 (21)	7/33 (21)	5/37 (13)	—

[a] From Rose and Bigazzi (1978).

[b] Numbers in parentheses are percentages

TABLE II

Incidence of Thyroglobulin Antibodies in Individuals without Overt Thyroid Disorders[a]

Number of patients	Positive (%)
146	5.5
52	6.0
219	3.7
387	18.0
1000	6.0
1297	5.7
180	10.5
336	13.1
785	12.0
154	14.2
268	17.5
Total tested 4824	8.7

[a] From Rose and Bigazzi (1978).

and never >1% is IgM. As far as IgG subclasses are concerned, 68% of AHT is IgG1, 19% IgG2, 7% IgG3, and 6% IgG4 (Hay and Torrigiani, 1974). Thus, >70% of AHT are in subclasses that fix complement, and the poor complement fixation observed in this system may be dependent on the nature of the thyroglobulin molecule, endowed with a limited number of autoantigenic determinants, rather than on the autoantibodies.

When AHT have been investigated by isoelectric focusing, all sera but one gave spectrotypes characteristic of polyclonal responses (Nye *et al.*, 1981). These results are similar to those we have obtained by preparative isoelectric focusing of autoantibodies to rat thyroglobulin (Zanetti and Bigazzi, 1981) and indicate that humoral autoimmune responses to thyroglobulin are polyclonal in nature. In spite of this, we have succeeded in obtaining "conventional" (i.e., heterologous and polyclonal) antiidiotypic antibodies against AHT (McCoy *et al.*, 1983). These results have been confirmed by Matsuyama *et al.* (1983), who detected cross-reacting idiotopes of AHT.

Autoantibodies to the "Colloid Antigen Second." Autoantibodies to the "colloid antigen second" are detectable only by indirect immunofluorescence on cryostat sections of thyroid, with a characteristic bright green uniform staining pattern of the entire colloid ("ground-glass" pattern) (Table III). They were observed in 5–8% of the sera found to be negative by other tests for thyroiditis, in 41% of the sera of patients with thyrotoxicosis, in 38% of the sera of those with primary myxedema, in

TABLE III

Incidence of Circulating Antibodies to Thyroid "Microsomal" Antigen, CA2, and Gastric Parietal Cells in Patients with Thyroiditis and Other Conditions[a]

Diagnosis	"Microsomal" antibodies[b]	CA2 antibodies[b]	Gastric parietal cell antibodies[b]
Chronic thyroiditis	89	—	32
	100	—	40
	85	—	—
	—	73	—
Primary hypothyroidism	57	—	29
	18	—	—
	—	38	—
Thyrotoxicosis	37	—	28
	66	—	—
	—	41	—
Pernicious anemia	26	—	76
	67	—	100
SLE	27	—	—
Other "collagen" diseases	7	—	—
Liver cirrhosis	13	—	—
Viral hepatitis	35	—	—
Patients with allergic reactions	12	31	—
Normal controls	4	—	5
	7	—	7
	0.4	—	—
	2	4	—
	4	6	5
	5	—	—

[a] Modified from Rose and Bigazzi (1978).
[b] Data represent percentage of positive.

30% of the sera of relatives of patients with Hashimoto's thyroiditis, and in 4–6% of the sera of control subjects.

Autoantibodies to the Thyroid "Microsomal" Antigen (ATM) ATM can be detected by complement fixation, indirect immunofluorescence, passive hemagglutination, radioimmunoassay, and ELISA. For several years the most commonly used method was indirect immunofluorescence, by which ATM were detected in 85–100% of patients with autoimmune thyroiditis, 57% of those with primary hypothyroidism, 26–67% of those with pernicious anemia, and 2–7% of normal controls (Table

TABLE IV
Incidence of Autoantibodies to Thyroid Microsomal Antigen (ATM) Detected by Passive Hemagglutination in the Sera of 466 Normal Subjects of Various Ages[a]

Age in years	Males	Females
1–10	0	8
11–20	0	12
21–30	0	4
31–40	4	16
41–50	0	8
51–60	12	28
61–70	4	28
71–80	16	20
81–90	8	20
91–100	0	20

[a] Modified from Romanin (1979).

III). Less sensitive is complement fixation, which has given positive results in 55–73% of sera found to be positive by indirect immunofluorescence, and is not in common use. A hemagglutination method for ATM first developed in Japan has proved to be more sensitive than complement fixation and as sensitive as indirect immunofluorescence. Because of its simplicity, this procedure has become the preferred test to demonstrate circulating ATM. By this method, ATM have been detected in the sera of 85–95% of patients with Hashimoto's thyroiditis and 66–86% of those with Graves' disease; ATM are also present in the sera of normal subjects, with a higher incidence in older women (Table IV). Recently, ATM have been measured by ELISA in the sera of 104 patients with thyroid autoantibodies (Schardt *et al.*, 1982) and a highly significant correlation was observed with the results obtained by passive hemagglutination. As for the detection of AHT, it is to be expected that eventually ELISA will be the preferred method for ATM, especially when testing large numbers of sera.

Cytotoxic Autoantibodies. Cytotoxic autoantibodies are present in the sera of most patients with autoimmune thyroiditis. These antibodies are effective in the presence of complement against cells dispersed with trypsin from thyrotoxic and normal thyroids. Khoury *et al.* (1981) demonstrated a complete correlation between complement-mediated antibody-

dependent cytotoxicity and cell-surface immunofluorescence on freshly dispersed thyroid cells. This finding confirmed previous observations suggesting that cytotoxic antibodies are identical with ATM.

Autoantibodies to Thyroid Cell-Surface Antigens. Autoantibodies to thyroid cell-surface antigens have been found in 98% of sera from patients with autoimmune thyroiditis, 76% with thyrotoxicosis, and 34% with nontoxic goiter (Fagraeus and Jansson, 1970). Ongoing investigations should finally determine whether these autoantibodies are directed against the same cell-surface components that are the target of ATM.

Autoantibodies to the TSH Receptor. Circulating autoantibodies to the TSH receptor are found in 10–15% of patients with autoimmune thyroiditis (Smith, 1981; Davies, 1983). In some cases these antibodies can act as TSH agonists and stimulate thyroid function (through stimulation of intracellular production of cyclic AMP), even though hyperthyroidism may not occur because of extensive thyroid destruction. *Thyroid growth-promoting autoantibodies,* a distinct population of thyroid-stimulating antibodies, have also been demonstrated in the circulation of thyroiditis patients (Drexhage *et al.,* 1980; Valente *et al.,* 1983). These autoantibodies might be responsible for the goiter observed in patients with autoimmune thyroiditis.

Autoantibodies against Iodothyronins. Rarely, patients with autoimmune thyroiditis are found to have serum autoantibodies against iodothyronins (Ginsberg *et al.,* 1978; Faerch *et al.,* 1980). It has been shown that such autoantibodies are directed against parts of the thyroglobulin molecule containing thyroxyl residues, even though the thyroxyl may not be the moiety originally involved in initiating the autoimmune response (Himsworth *et al.,* 1983).

Circulating Immune Complexes. Finally, circulating immune complexes have been demonstrated by the Raji cell radioimmunoassay and Clq solid phase radioassay in up to 78% of patients with autoimmune thyroiditis. Mariotti *et al.* (1979) demonstrated that some of these immune complexes contain thyroglobulin; however, they do not bind complement in significant amounts. This may explain why immune complex-mediated renal disease is not a characteristic feature of autoimmune thyroiditis.

Autoantibodies to Other Antigens. Autoantibodies to other antigens, such as those of gastric parietal cells, are also present in the circulation of patients with autoimmune thyroiditis (Table III).

c. Cell-Mediated Immune Responses. Most patients with autoimmune thyroiditis have cell-mediated immune responses to thyroid autoantigens, detectable by various procedures (reviewed by Calder and Irvine, 1975; Bigazzi, 1979).

TABLE V

Delayed Hypersensitivity Reactions to Thyroid Antigens in Patients with Thyroid Disorders[a]

Diagnosis	Percentage of positive reactions	Test
Chronic thyroiditis	40	Leukocyte migration inhibition
	63	Leukocyte migration inhibition
	87	Leukocyte migration inhibition
	67	Leukocyte migration inhibition
	64	Lymphocyte cytotoxicity
Graves' disease	82	Leukocyte migration inhibition
Thyroid controls	15	Leukocyte migration inhibition
Autoimmune controls	12	Leukocyte migration inhibition
Normal controls	0	Leukocyte migration inhibition
	0	Leukocyte migration inhibition
	4	Leukocyte migration inhibition
	0	Lymphocyte cytotoxicity

[a] Modified from Rose and Bagazzi (1978).

In early *in vivo* studies, positive skin reactions were obtained in the majority of patients after the intradermal injection of thyroid extract. However, it is difficult to determine in retrospect whether such reactions were due to delayed-type hypersensitivity or an Arthus phenomenon.

In vitro tests have shown that lymphocytes from patients with autoimmune thyroiditis produce lymphokines after stimulation with thyroid antigens (Table V). Soborg and Halberg (1968) first demonstrated that the addition of crude thyroid extract to culture chambers containing peripheral blood leukocytes in capillary tubes caused the inhibition of migration of leukocytes from 6 of 15 patients with autoimmune thyroiditis. These results were later confirmed by other investigators, who noted inhibition of leukocyte migration in 63–86% of patients with autoimmune thyroiditis. The percentage of positive reactions observed varies according to the antigens used for lymphocyte stimulation. Calder and associates (1972) observed inhibition of migration in 66% of patients when crude thyroid extract was used versus 40% with purified thyroglobulin and 27% with thyroid "microsomal" antigen. In a similar study by Wartenberg and associates (1973a,b), leukocyte migration inhibition was obtained in all patients with autoimmune thyroiditis after stimulation with thyroid "microsomal" antigen but not after incubation with thyroglobulin.

A direct cytotoxic effect of lymphocytes from patients with autoimmune thyroiditis has been reported by several investigators. Podleski

(1972) first showed that mouse mastocytoma cells coated with purified human thyroglobulin were lysed by lymphocytes from 6 of 11 patients with Hashimoto's thyroiditis, while similar target cells coated with "microsomal" antigen were destroyed by lymphocytes from 7 of the same patients. Similar results have also been obtained using chicken red blood cells coated with thyroglobulin and monolayers of dog or human thyroid epithelial cells (Laryea *et al.*, 1973; Seybold *et al.*, 1981; Iwatani *et al.*, 1982). Recently, Chow *et al.* (1983) demonstrated that natural killer (NK) cell activity against thyroid cells was present in Hashimoto's patients, but with no significant difference from normal subjects. On the other hand, thyroiditis patients had depressed NK activity against K562 human tumor cells. These observations do not exclude a possible role of NK cells as a mechanism for thyroid cell destruction in Hashimoto's thyroiditis, but indicate that other mechanisms (perhaps K-cell-mediated lysis) may have a more important role.

Results obtained with a third *in vitro* assay, lymphocyte stimulation, have been controversial. Some investigators have failed to observe blast transformation or increased thymidine incorporation when lymphocytes from patients with autoimmune thyroiditis were incubated with thyroid antigens, while others have obtained positive results in most or all patients with Hashimoto's thyroiditis (reviewed by Calder and Irvine, 1975; Bigazzi, 1979). In view of these contradictory reports, a more recent study has been performed using improved methods in cell separation, cell culture, and cell harvesting (Aoki and DeGroot, 1979). By using increasing levels of thyroglobulin in lymphocyte cultures, blastogenic responses have been observed in 71.4% of Hashimoto's patients versus 9.1% of normal controls.

d. Antibody-Dependent Cellular Cytotoxicity (ADCC). This immunological phenomenon is a combination of both humoral and cellular immunity, since it requires the binding of specific antibodies to their target cells and the attachment of nonsensitized lymphoreticular cells to the Fc region of those antibodies, followed by target cell death through an extracellular nonphagocytic mechanism. Among the cells involved in ADCC are lymphocytes that have been termed K cells, bear Fc receptors, and do not belong to either the mature T- or B-cell population.

Reactions of ADCC have been observed using sera or lymphocytes from patients with autoimmune thyroiditis (reviewed by Calder and Irvine, 1975; Bigazzi, 1979). When sera from these patients were incubated with thyroglobulin-coated chicken erythrocytes and normal human lymphocytes, lysis of the target cells was obtained in 66–74% of cases. One group of investigators observed a correlation with the titer of IgG AHT,

while another group did not notice any correlation. When lymphocytes from patients with Hashimoto's thyroiditis were used in an ADCC assay performed with chicken erythrocytes coated with IgG antibodies against chicken red cells, a significant increase in K-cell cytotoxic activity was found in patients that were either newly diagnosed and untreated or had received treatment for <1 year (Calder *et al.*, 1976). Interestingly, Suzuki *et al.* (1980) reported that 30–45% of patients with Hashimoto's thyroiditis have circulating AHT cytophilic for human monocytes, and that such "armed" monocytes become cytotoxic against thyroglobulin-coated chicken erythrocytes. Finally, a microcytotoxicity assay for thyroid-specific ADCC, developed using human thyroid epithelial cells as targets, confirmed that Hashimoto's sera have significantly more cytotoxic activity than normal sera (Iwatani *et al.*, 1982). When peripheral K cells were measured by a plaque assay, it was noted that patients with euthyroid or hypothyroid Hashimoto's disease have higher numbers of K cells than normal controls, but the difference is not statistically significant. On the other hand, the K-lymphocyte counts in thyrotoxic patients with Hashimoto's disease are significantly higher.

2. Immunological Derangements

At present, we have no understanding of the underlying cause(s) of autoimmune thyroiditis. As in many other conditions of uncertain origin, autoimmune or not, one can suspect that viral infections and hormonal and/or genetic factors contribute to the disease. It has also been hypothesized that one or more abnormalities of the immune system, such as deficiencies of T suppressor cells, excess of T-cell help, polyclonal B-cell activation, and derangements of the idiotype–antiidiotype network, may be the cause of autoimmune thyroiditis as well as of other autoimmune diseases. In addition there may be underlying defects in the thyroid gland itself (reviewed by Rose *et al.*, 1981; Bigazzi, 1983; Rose, 1983).

In 1971, Allison suggested that T lymphocytes play a key role in preventing autoimmunity and that failure of T-cell feedback control is a major factor allowing the development of autoimmune disease (Allison *et al.*, 1971; Allison, 1971). This suggestion has generated an interesting series of investigations on the number and function of T lymphocytes and their subpopulations.

Initial studies on the number of T lymphocytes present in the peripheral blood of patients with autoimmune thyroiditis produced rather contradictory results, since some investigators found higher numbers than in controls, others observed lower numbers, and others did not notice any significant difference (reviewed by Bigazzi, 1979). Similarly conflicting are

TABLE VI

Subpopulations of T Lymphocytes in Peripheral Blood Lymphocytes of Patients with Autoimmune Thyroiditis (AT) and Normal Controls

References	Method	Percentage of cytotoxic/suppressor cells		Ratio of suppressor to helper cells	
		AT	Controls	AT	Controls
Thielemans *et al.* (1981)	IF microscopy, OKT Mabs	20.3 ± 1.1	25.9 ± 0.9		
Canonica *et al.* (1982)	IF microscopy, OKT Mabs	36 ± 4	29 ± 1		
Sridama *et al.* (1982)	IF microscopy, OKT Mabs	18.5 ± 7.6	23.7 ± 5.1	2.7 ± 1.0	1.9 ± 0.6
Bonnyns *et al.* (1983)	IF microscopy and flow cytofluorimetry, OKT Mabs	77 ± 36 (of normal)	100 ± 26		
Iwatani *et al.* (1983)	Flow cytofluorimetry Leu Mabs	22.1 ± 5.7[a] 23.2 ± 8.1[b] 23.5 ± 5.1[c]	23.3 ± 5.2		
Jansson *et al.* (1983)	IF microscopy, OKT Mabs	25 ± 4	28 ± 8	2.1 ± 0.3	2.0 ± 0.7
McLachlan *et al.* (1983b)	IF microscopy, OKT Mabs	30 ± 2	32 ± 2	2.0 ± 0.2	1.7 ± 0.2
Wall *et al.* (1983)	IF microscopy, OKT Mabs			1.62 ± 0.18	1.55 ± 0.18

[a] Euthyroid.
[b] Hypothyroid.
[c] Hyperthyroid.

the results obtained by more recent studies performed with monoclonal antibodies to surface antigens of suppressor/cytotoxic T lymphocytes. As shown in Table VI, some investigators have reported that patients with autoimmune thyroiditis have a significant decrease of OKT8+ (suppressor/cytotoxic) cells in their peripheral blood (Theilemans *et al.*, 1981; Sridama *et al.*, 1982; Misaki *et al.*, 1983), but numerous others have failed to confirm these findings (Canonica *et al.*, 1982; Bonnyns *et al.*, 1983; Jansson *et al.*, 1983; McLachlan *et al.*, 1983b; Wall *et al.*, 1983).

Studies aimed at an evaluation of suppressor T-cell *function* rather than

numbers have also produced contradictory results. Some investigators have reported a deficiency of such function in patients with autoimmune thyroiditis (Okita *et al.*, 1981; Noma *et al.*, 1982; Pacini *et al.*, 1982), while others have found no significant difference from normal controls (Beall and Kruger, 1980b; MacLean *et al.*, 1981; Wall and Chartier, 1981).

Finally, there are only a few published studies of T suppressor cells within the infiltrated thyroid gland of Hashimoto's patients. One group of investigators (Baur *et al.*, 1983; Jansson *et al.*, 1983) reported a relative decrease in OKT8+ cells (13 ± 6%), a higher ratio of OKT4+ to OKT8+ cells (4.7 ± 1.7) and an increase in B cells (27 ± 6%). Another group (MacLachlan *et al.*, 1983b) found no decrease in OKT8+ cells (22%), with normal OKT4+:OKT8+ (1.8), while B cells were increased (28%).

The contrasting results obtained in studies enumerating T lymphocytes and their subsets or evaluating suppressor function may be due to numerous variables, such as the methods employed to count lymphocyte populations (usually rosetting or fluorescence microscopy on a relatively small number of cells), the monoclonal antibodies used to detect T-cell subsets and differences in patients' populations. Investigations of thyroid lymphocytes are also biased by the lack of appropriate controls, such as the use of lymphocytes from lymph nodes adjacent to the thyroid rather than peripheral blood lymphocytes as is commonly done. Obviously, ethical considerations render the use of such controls practically impossible. In any case, at present it seems reasonable to conclude that investigations of T suppressor cells have failed to clarify the role of these cells in autoimmune thyroiditis. However, in spite of the lack of conclusive confirmatory evidence, Allison's attractive hypothesis keeps being resurrected, like an immunological phoenix. As an example, Iwatani *et al.* (1983) recently found, by microfluorimetry, no significant difference in absolute counts and percentages of Leu2a+ (suppressor/cytotoxic) lymphocytes in the peripheral blood of Hashimoto's patients versus normal controls. When they examined the peak position of fluorescence intensity of each T-cell subset, they found that patients with hypothyroid or thyrotoxic Hashimoto's disease had a decrease in the peak of fluorescence intensity of Leu2a+ cells, while patients with euthyroid disease had normal values. Assuming that antigen density of Leu2a (represented by the peak position of fluorescence intensity) is correlated with suppressor T-cell function, these investigators have suggested that generalized T suppressor-cell function is decreased in hypothyroid or hyperthyroid Hashimoto's disease. Obviously, even if their assumption were correct, the normal values observed in euthyroid patients with autoimmune thyroiditis would still raise some doubts about the role of T suppressor cells. In this respect, the observations by Bonnyns *et al.* (1983) may be of some relevance and

suggest alternative possibilities. These investigators found a negative correlation between levels of free thyroxine in the serum and percentage of suppressor/cytotoxic T cells in the peripheral blood, so that hyperthyroid patients or patients treated with thyroxine have lower levels of OKT8+.

Other investigations of lymphocyte function in autoimmune thyroiditis have been less controversial. There is general agreement that antigen-binding lymphocytes reacting with thyroid antigens are increased in patients with Hashimoto's disease (reviewed by Bigazzi, 1979). In addition, it has been shown that both PBM and thyroid lymphocytes are capable of producing autoantibodies to thyroid antigens. Beall and Kruger (1980a,b) first demonstrated that cultures of PBM from patients with autoimmune thyroiditis produce AHT after stimulation with pokeweed mitogen (PWM). In this culture system, no deficiencies of T-cell suppression or excesses of T-helper activity were noted. These observations have been confirmed and expanded by other investigators, who have also reported significant numbers of plaque-forming cells against human thyroglobulin in PWM-stimulated cultures of PBM from patients with autoimmune thyroiditis (McLachlan *et al.*, 1981; Weetman *et al.*, 1982; Weiss *et al.*, 1982; Weiss and Davies, 1982; McLachlan *et al.*, 1983a). Lymphocytes bearing receptors for human thyroglobulin have been purified from PBM of Hashimoto's patients and were found to synthesize AHT of the IgM class after activation with Epstein–Barr virus (McLachlan *et al.*, 1983c). When similar studies have been performed on thyroid lymphocytes from infiltrated thyroids of Hashimoto's patients, it has been noted that such cells synthesize large amounts of AHT and/or ATM without the need of stimulation with mitogens (Weetman *et al.*, 1982; McLachlan *et al.*, 1983c). Thus, thyroid lymphocytes are activated sites of autoantibody synthesis, while peripheral blood lymphocytes are not.

3. Mechanisms of Damage

In spite of 27 years of research, it is not clear whether thyroid damage is caused by autoantibodies, sensitized lymphocytes, or both mechanisms combined.

Circulating autoantibodies to thyroid antigens are present in most patients with autoimmune thyroiditis and may cause cell death and inflammatory infiltration. Autoantibodies to surface antigens of thyroid cells (possibly including "microsomal" antibodies) may directly damage the epithelial cells of the thyroid gland by their cytotoxic activity. However, the presence and titers of circulating ATM and AHT do not seem to correlate well with the severity of thyroiditis. Still, autoantibodies to

thyroid antigens might cause damage indirectly through the *in situ* formation of immune complexes in the thyroid.

Initial studies performed by immunofluorescence on thyroids from patients with Hashimoto's disease showed that immunoglobulins and complement components were bound to colloid material within the thyroid follicles and were also present in the interstitial spaces (reviewed by Bigazzi, 1979). Later investigations demonstrated heavy granular deposits of IgG, IgM, IgA, and C3 as well as thyroglobulin at the level of the follicular basement membrane and in the interfollicular spaces (Werner *et al.*, 1972; Kalderon and Bogaars, 1977; Koh *et al.*, 1983). The major objection against a pathogenic role of such deposits is based on the well-known finding that immune complexes composed of thyroglobulin and autoantibodies to human thyroglobulin do not bind complement in significant amounts, possibly because of the spatial arrangement of the few autoantigenic determinants on the thyroglobulin molecule. However, fragments of thyroglobulin containing the appropriate epitopes might be involved in *in situ* immune complex formation within the thyroid, generating deposits capable of binding complement.

Autoantibodies may also cause damage when associated with lymphocytes, as is the case in antibody-dependent cell-mediated cytotoxicity. This is another attractive possibility, but the evidence available is very scarce. In any case, the inflammatory infiltrates observed in autoimmune thyroiditis contain both T and B lymphocytes, as well as thyroglobulin-binding cells and lymphocytes containing antibodies to thyroglobulin (reviewed by Bigazzi, 1979). Thus, a pathogenetic role of cell-mediated immunity seems quite possible.

Sensitized lymphocytes might mediate thyroid damage directly or through the release of lymphokines. A number of these mediators, from migration inhibitory factor (MIF) or leukocyte inhibitory factor (LIF) to lymphotoxin, have been demonstrated in autoimmune thyroiditis. Either MIF or LIF might cause thyroid infiltration and damage by the stimulation and immobilization of macrophages within the thyroid, while lymphotoxin might have a direct toxic effect on the thyroid epithelial cells themselves. However, as in the case of thyroid autoantibodies, no correlation has been established between lymphokines and the course and severity of the disease.

Over the years, the cell necrosis and the inflammatory infiltration observed in autoimmune thyroiditis have been explained on the basis of either a humoral or a cell-mediated mechanism. However, because of the lack of decisive evidence, it seems more reasonable to assume that both humoral and cellular mechanisms may be responsible for the tissue damage, acting concurrently or in a stepwise sequence.

Until recently, the mechanisms underlying another important manifestation of autoimmune thyroiditis, the goiter, had not been well understood. Doniach *et al.* (1979) suggested that growth-stimulating autoantibodies might be responsible for thyroid cell hyperplasia. These autoantibodies would act on TSH receptors through other mechanisms than AMP stimulation. Evidence in favor of this hypothesis has been recently presented by two research groups (Drexhage *et al.*, 1980; Valente *et al.*, 1983).

4. Immunogenetics

Initial studies of the immunogenetics of autoimmune thyroiditis have given contradictory and inconclusive results (reviewed by Bigazzi, 1979). More recently, an association between *HLA-DR5* and goitrous thyroiditis has been reported (Weissel *et al.*, 1980; Farid *et al.*, 1981; Thomsen *et al.*, 1983). On the other hand, patients with atrophic thyroiditis have a high frequency of *DR3* and a low frequency of *DR5* (Farid *et al.*, 1981). Previous observations had shown that primary atrophic hypothyroidism (but not Hashimoto's goiter) is associated with *HLA-B8*.

On the basis of these results, it seems likely that atrophic and hypertrophic thyroiditis are two different entities and not different stages of the same disease. It is possible that the early immunogenetic studies of autoimmune thyroiditis obtained contradictory results because they did not differentiate between various clinical manifestations. More precise investigations may actually suggest different pathogenetic and etiological mechanisms of thyroiditis.

5. Laboratory Diagnosis

As mentioned previously, the immunological diagnosis of autoimmune thyroiditis is based on the determination of circulating AHT and ATM by passive hemagglutination. Doniach *et al.* (1979) noted that "classical" Hashimoto's goiters presenting with hypothyroidism almost always have high titers of both autoantibodies. However, the majority of today's patients are euthyroid with small goiters and usually have only ATM in their serum. ELISAs for both autoantibodies are now available, and it is likely that in the future they will be routinely employed.

F. Concluding Remarks: Future Prospects

In the 28 years since Witebsky and Rose first reported on the induction of experimental autoimmune thyroiditis, enormous progress has been made in our understanding of autoantigens, humoral and cell-mediated

autoimmune responses, and the pathogenesis and etiology of autoimmune thyroiditis. However, it is only fair to point out that there are still numerous and crucial areas of uncertainty.

As far as autoantigens are concerned, it is hoped that future studies will reveal whether thyroglobulin, microsomal, or perhaps some other still unidentified antigen is the one primarily responsible for the initiation of thyroiditis. Alternatively, the various autoantigens may be found to be involved in different variants of this condition. Further research on the mechanisms of thyroiditis may identify which one of the numerous possibilities is involved in the initial thyroid damage. Conversely, the pathogenesis may vary according to the different forms. It is obvious that well-focused immunogenetic studies will be of great importance in the solution of these issues, and will be expected to shed some light on the heterogeneity of autoimmune thyroiditis.

Another area requiring additional study relates to the control of autoimmune responses to thyroid antigens. One can hope that future research will reveal whether suppressor T cells, idiotype–antiidiotype network, both mechanisms, or some other yet unidentified regulation system is involved in the maintenance of self-tolerance (or nondestructive levels of self-recognition). Recent developments of basic immunology, such as monoclonal antibody production, T- and B-cell cloning, and nucleic acid probes will undoubtedly help in the solution of these problems.

II. GRAVES' DISEASE

A. Introduction: Historical Background

Graves (1835) and von Basedow (1840) independently recognized the disease that came to be called after the former investigator in English-speaking countries and after the latter on the continent of Europe. Its major features are tachycardia, nervousness, tremor, enlargement of the thyroid, and in some cases, exophthalmus. Based on the similarity of this syndrome to the effects of overtreatment of myxoedema with desiccated thyroid, Möbius (1887) suggested that overactivity of the thyroid gland was the basis of the disease. Histological studies showed that the thyroid gland was hyperplastic in Graves' disease. Moreover, extracts of thyroid tissue administered to animals elicited the essential features of Graves' disease. Significantly, extracts of glands from patients with Graves' disease were no more potent in eliciting these symptoms than were extracts of normal thyroids, indicating that the thyroids of patients with Graves' disease did not contain any unique agent absent from normal glands.

The pioneering investigations of Marine and Lenhart (1909) showed that a deficiency in iodine caused thyroid hyperplasia and that administration of iodine corrected this effect. In the face of iodine deficiency, the thyroid stored iodine in an effort to maintain function at a normal rate. In contrast, the hyperplastic thyroid of Graves' disease took up extra iodine and continued to pour out its hormones at an increased rate, an effect that could sometimes be remedied by iodine treatment.

The role of the anterior pituitary in regulating thyroid activity was demonstrated by Rumph and Smith (1926), who described the action of thyroid-stimulating hormone (TSH) or thyrotropin in hypophysectomized rats. These studies led to the logical suggestion that some cases of hyperthyroidism may be attributable to overactivity of the anterior pituitary. Other evidence, however, pointed to primary defects in the thyroid gland itself. One group of cases, first identified by Plummer (1913) as hyperfunctioning adenomas, produced symptoms of hyperthyroidism. However, the most important step in discovering the etiology of Graves' disease was taken by Adams and Purves (1956) when they demonstrated the presence of long-acting thyroid stimulator (LATS) in the serum of a patient with recurrent thyrotoxicosis and exophthalmus.

The discovery of LATS depended on the earlier demonstration that thyroid function can be assayed *in vivo* by measuring uptake of ^{131}I. Injection of thyrotropin into guinea pigs produces as much as a 500-fold increase in ^{131}I uptake by the thyroid within 3 to 6 h. Adams and Purves (1956) noticed that the serum of some patients had only a modest effect at 3 h, but continued to produce elevated ^{131}I uptakes as long as 24 h after injection. They were thereby able to develop a bioassay that distinguished between the effects of thyrotropin at 3 h and LATS at 16 h. By adapting the procedure to the mouse, McKenzie (1958) developed a practical bioassay for LATS.

The nature of LATS remained a mystery for many years. However, Kriss *et al.* (1964) showed the presence of LATS activity in highly purified 7S gamma globulin from Graves' serum. Antiserum to human IgG neutralized the LATS bioassay. Moreover, Kriss *et al.* (1964) demonstrated that LATS activity could be absorbed by the insoluble sediment of human thyroid tissue. LATS is not neutralized by antiserum to TSH, showing that the IgG is not serving simply as a carrier protein for thyrotropin. Meek *et al.* (1964) split immunoglobulin-containing LATS activity into its constituent heavy and light chains by reduction and alkylation, and found that biological activity was present in isolated heavy chains but not in light chains. Activity was localized in the Fab, but not Fc, portions. These studies clearly showed that LATS is an antibody and led to the introduc-

tion of the term "thyroid-stimulating antibody" (TSAb) as a more apt designation for LATS.

The early investigations of LATS depended upon the mouse bioassay. Unfortunately, only about half of the cases of Graves' disease were found to have LATS activity. This inability to demonstrate LATS was first attributed to a lack of sensitivity of the bioassay. In the course of further investigations of this apparent insensitivity of the assay, Adams and Kennedy (1967) absorbed LATS with human thyroid homogenates, repeating the earlier experiments of Kriss *et al.* (1964). They found, unexpectedly, that the sera of certain LATS-negative Graves' patients interfered with the absorbing capacity of thyroid homogenates. This blocking agent was shown to be an immunoglobulin, and was named "LATS protector." When both LATS and LATS protector were measured, positive reactions were found in almost all cases of Graves' disease. LATS seems to be a TSAb that reacts with both mouse and human thyroids, whereas LATS protector is restricted to human thyroid antigen.

B. General Description

Graves' disease, also called toxic diffuse goiter, is a triad of hyperthyroidism, infiltrative ophthalmopathy, and infiltrative dermopathy. These manifestations do not necessarily appear together, and often run courses independent of each other. Although Graves' disease is the most common cause of hyperthyroidism, the same clinical features can be produced by any condition producing an excess of thyroid hormones. Thyroid function tests designed to assess the homeostatic control of thyroid functions, such as the thyroid suppression test, are therefore of little use in the differential diagnosis of Graves' disease. The most specific indicator of Graves' disease is the presence of TSAb, which is found in 93% of Graves' patients and in ~20% of patients with Hashimoto's thyroiditis, especially those with some evidence of Graves' disease such as ophthalmopathy (Rapoport *et al.*, 1984). TSAb are not found in normal subjects or in euthyroid relatives of Graves' patients, even those with antibody to thyroid microsomes or thyroglobulins (Banovac *et al.*, 1981).

In some respects, the clinical manifestations of TSAb mimic the actions of thyrotropin. Thyrotropin reacts with a specific glycoprotein receptor located on the plasma membrane of thyroid cells, and activates receptor-linked adenyl cyclase, which produces cAMP and cAMP-mediated activation of cellular metabolism to produce increased amounts of the thyroid hormones T_4 and T_3. Normally, these thyroid hormones reduce the synthesis of thyrotropin through feedback inhibition of the adenohypophysial

thyrotropin-releasing hormone (TRH). Produced by the hypothalamus, TRH reaches the anterior pituitary through the blood vessels of the hypothalamic pituitary portal system. This complex interplay of hormones results in the normal homeostasis of thyroid function. However, if an abnormal factor acts on the thyroid as if it were TSH and continues to be produced independently of the circulating levels of T_4 and T_3, the hormone-producing thyroid cells will hypertrophy; the patient will become profoundly hyperthyroid, and thyroid suppression tests will be abnormal; that is, TSH concentration in the serum will be low in the face of increased iodide uptake and thyroid production. This situation is characteristic of Graves' disease.

C. Histopathology

In Graves' disease, lymphocytic infiltration with germinal center formation is a regular feature, but is generally less in incidence and intensity than that seen in Hashimoto's thyroiditis. Tötterman (1978) studied the distribution of T and B lymphocytes in needle aspirates from the thyroid glands of seven patients with Graves' disease, including four who were euthyroid due to medication, and found an approximately equal proportion of T and B lymphocytes. In four patients, T lymphocytes were in excess. Aanderud *et al.* (1982) also reported a preponderance of T cells in the thyroid glands, even though the levels of T cells in circulating blood were below normal in Graves' patients. They attributed the reduction in circulating T cells to drug therapy. Sköldstam *et al.* (1978) examined the glands of six Graves' patients and found that the proportions of B and T lymphocytes were comparable to their venous blood levels.

Analysis of T cells using monoclonal antisera demonstrated that there was a significant decrease in total T cells (OKT3+) and helper T cells (OKT4+) in thyroid tissue compared with peripheral blood, but the ratio of helper to suppressor cells (OKT4+:OKT8+) was not different from the values obtained from peripheral blood (McLachlan *et al.*, 1983b). Thyroid lymphocyte suspensions synthesized relatively large amounts of antibody to thyroid microsomes or thyroglobulin in culture without the need for mitogenic stimulation. In contrast, peripheral blood lymphocytes secreted antibody only after addition of mitogen or antigen to the cultures, as described later. These findings emphasize the importance of the thyroid gland as the major site of autoantibody synthesis.

Kalderon and Bogaars (1977) observed electron-dense deposits in the follicular basal laminae of the basement membrane associated with lymphocytic and plasma cell infiltration of thyroids from patients with Graves' disease. By immunofluorescence, these deposits were found to

correspond with the distribution of immunoglobulins and thyroglobulin, suggesting that they represented locally produced immune complexes.

D. Immunology

1. Immunological Features

Although much evidence shows that TSAb is an antibody to the thyrotropin receptor, Fenzi *et al.* (1980) were able to remove thyrotropin receptors by affinity chromatography and find that TSAb, but not TSH, still bound to thyroid membranes. In addition, Solomon and Chopra (1980) found that TSH did not prevent absorption of LATS protector into thyroid membranes. These experiments have thrown doubt on the view that TSAb reacts directly with the TSH receptor. In an attempt to resolve this point, TSAb and TSH were mixed with solubilized thyroid membranes. TSH receptor and TSAb receptor complexes could be found, but there were no TSH–TSAb receptor complexes. These data suggest that TSAb and TSH compete with each other for the neighboring binding sites on the TSH-receptor molecule. Endo *et al.* (1981) showed that TSAb from Graves' disease patients will bind to TSH receptor in tissues other than thyroid; for example, on fat cells. Finally, Wenzel *et al.* (1981) demonstrated that the peripheral blood of Graves' patients contained increased numbers of B lymphocytes that bound partially purified thyrotropin receptor.

Human monoclonal antibodies were produced from heterohybridomas obtained by fusing mouse myeloma cells with peripheral lymphocytes from patients with Graves' disease (Valente *et al.*, 1982). Two of these antibodies stimulated thyroid function in the mouse bioassay. These stimulating antibodies reacted strongly with galactoside preparations from human thyroids. Other monoclonals did not produce stimulation, but they were capable of inhibiting or blocking TSAb from Graves' disease sera. These two antibodies did not react with human thyroid glalactosides, but were strong inhibitors of thyrotropin binding to liposomes, which contained a glycoprotein with high affinity for TSH. Thus, all four monoclonal antibodies reacted with the determinants on or closely neighboring the functional TSH receptor on the thyroid cell.

In patients with hyperthyroidism, it is possible to show a close correlation between the levels of LATS protector and ^{131}I uptake by the thyroid. The data provide strong evidence that LATS protector is the direct cause of the hyperthyroidism of Graves' disease. LATS itself shows no significant correlation with thyroid uptake. On the other hand, LATS has a powerful effect in raising the blood iodide level of rhesus monkeys,

whereas potent sera with LATS protector were inactive. Adams *et al.* (1974) performed experiments using LATS protector sera in human volunteers. They found that an infusion of plasma with LATS protector produced a highly significant and prolonged increase in the recipients' blood iodine levels, giving direct evidence that LATS protector has a stimulatory effect on the human thyroid.

Thyroid-stimulating antibody is detectable by bioassay in most but not all patients who have been diagnosed as having Graves' disease. One problem is that the disease is difficult to diagnose. Graves' disease is not the only cause of hyperthyroidism; for example, toxic adenomas or diffuse multinodular goiters can produce hyperthyroidism not associated with ophthalmopathy or other immunological disturbances. In addition, the bioassay is difficult to carry out and may not be sufficiently sensitive to demonstrate a stimulating factor in the serum of all patients with Graves' disease.

Shishiba *et al.* (1973) were the first to show an *in vitro* effect of LATS protector. Applied to slices of human thyroid tissue, Graves' sera caused an increase in the number of intracellular colloid droplets. Onaya *et al.* (1973) later reported that sera from thyrotoxic patients caused the accumulation of cyclic AMP when applied to human thyroid slices. This method was standardized by Rapoport *et al.* (1984). Tests that measure blocking of TSH binding to thyroid membrane preparations have also been used to measure TSAb (Tao and Kriss, 1982; Kishihara *et al.*, 1978).

Conflicting results have been reported for T-cell transformation when lymphocytes from Graves' patients are cultured with phytohemagglutinin (Calder and Irvine, 1975). Yet, suppressor T-cell function is decreased in peripheral lymphocytes from Graves' disease patients (Pacini *et al.*, 1982). Moreover, it has been shown that B lymphocytes cultured with either phytohemagglutinin or crude human thyroid antigen will secrete TSAb into the medium (Knox *et al.*, 1976a,b). These findings support the view that B cells capable of responding to normal thyroid antigens are permitted to respond by a lowered level of suppressor T cells.

Recently, Drexhage and his colleagues (1980) found that patients with goitrous Graves' disease or goitrous Hashimoto's thyroiditis have an antibody in their serum that promotes thyroid growth, but does not increase iodine uptake. In some respects, these thyroid growth-promoting antibodies mimic the growth-promoting effects of thyrotropin.

Tao and Kriss (1982) described high levels of membrane-binding antibodies in patients with Graves' disease. This binding antibody did not correspond with the presence of TSAb or other thyrotropin-binding activity. Furthermore, the binding activity was present in a variety of other

autoimmune diseases, such as Hashimoto's thyroiditis, systemic lupus erythematosus, mixed connective tissue disease, and rheumatoid arthritis.

Although opthalmopathy is frequently associated with the hyperthyroidism of Graves' disease, it is probably a distinct pathological process; TSAb correlate with hyperthyroidism, but not with eye disease. Winand and Kohn (1979) demonstrated an interaction between eye muscle membranes, thyrotropin subfragments, and an antibody in the sera of Graves' patients. On the other hand, Kriss *et al.* (1967) postulated that there is a primary binding of thyroglobulin to eye muscle membranes followed by the formation of thyroglobulin–antithyroglobulin immune complexes. Finally, Kodama *et al.* (1982) demonstrated circulating organ-specific autoantibody against a soluble eye muscle antigen in Graves' ophthalmopathy.

2. Immunogenetics

A significant association between *HLA-B8* and Graves' disease has been described in several studies of Caucasian populations (Farid *et al.*, 1975; Grumet *et al.*, 1975). An even closer association was recorded between Graves' disease and *HLA-Dw3* (Thorsby *et al.*, 1975). In the Japanese, Graves' disease has been found to be associated with *HLA-Bw35* (Grumet *et al.*, 1975). Farid *et al.* (1976) found that *HLA-B8* homozygosity increased the relative risk of disease severalfold over heterozygosity. This finding was interpreted to mean that the disease-susceptibility gene for Graves' disease is recessive.

More recently, Farid (1981) reported an association of the IgG heavy chain marker *Gmfb* haplotype and Graves' disease. These studies show that two distinct genetic systems have a major effect on susceptibility to Graves' disease, one associated with *HLA* and the other with *Gm*. Since they sort independently, a combination of these two variables may be highly predictive of genetic predisposition to develop Graves' disease.

3. Laboratory Diagnosis

At present, there is no satisfactory method for measuring TSAb for clinical purposes. The McKenzie mouse bioassay (McKenzie, 1958) (together with the test for LATS protector) is performed with concentrates of serum immunoglobulins and seems to be the most sensitive and dependable method, but it is laborious and expensive and requires a relatively large volume of serum. The *in vitro* assays are less tedious, but do

not yet provide dependable diagnostic information in the doubtful cases where they are most needed. The best hope for the future seems to reside in developing better binding immunoassays with the isolated thyrotropin receptor.

Interestingly, Banovac *et al.* (1981) found that euthyroid relatives of patients with Graves' disease did not have LATS in their serum, nor was thyroid-stimulating antibody demonstrable using the cAMP stimulation test on human thyroid slices. Only a few sera were weakly positive in the TSH-binding inhibition assay. On the other hand, many of the sera of asymptomatic relatives of Graves' patients contained antibodies to thyroid microsomes or to thyroglobulin (see Table I). Thus, although euthyroid relatives have evidence of a basic immunological defect leading to thyroid autoimmunity, thyroid-stimulating antibody is not present.

E. Treatment

Graves' disease tends to follow either a relapsing or a remitting course. The natural progression of the disease may be interrupted by ablation of the thyroid through subtotal thyroidectomy or the use of radioiodine, but the most usual treatment is effected by antithyroid drugs. Irvine *et al.* (1977) reported that the rate of relapse of Graves' patients after withdrawal of antithyroid drugs was much higher in patients with the *HLA-B8* haplotype than in *B8*-negative patients. The persistence of thyroid microsomal antibodies after withdrawal of therapy correlated significantly with the presence of *HLA-B8*. The association was more pronounced in patients who remained in remission. In view of these findings, Irvine *et al.* (1977) suggested that Graves' patients be typed for *HLA* and those who are *B8* positive be given longer term treatment with antithyroid drugs.

The influence of thyroidectomy on thyroid-stimulating antibodies in Graves' disease was studied by Werner *et al.* (1967), who found that LATS decreased during the months after operation. Smith *et al.* (1976) found a dramatic fall in TSAb to undetectable levels within 1 to 5 days after subtotal thyroidectomy in most patients with Graves' disease. Within 8 h after surgery, TSAb increased in Graves' patients (Bech *et al.*, 1982), suggesting that removal of the thyroid gland released antibody into the circulation. Interestingly, there was a parallel increase in serum levels of free thyroglobulin and a decline in AHT. These changes probably indicated a shower of thyroglobulin from the gland and subsequent formation of immune complexes. A secondary rise in TSAb correlated with an increase in AHT, probably attributable to stimulation of peripheral lymphocytes by antigens liberated from the thyroid (Knox *et al.*, 1976a).

F. Concluding Remarks

The primary cause of Graves' disease is still unknown, but there is good reason to believe that autoimmunity is responsible for the pathogenesis of this disorder. In addition to their endocrinological symptoms, Graves' patients have a variety of immunological abnormalities, including generalized lymphoid hyperplasia, lymphocytic infiltration of the thyroid and retroorbital tissues, and autoantibodies to thyroid and other endocrine organs. There is a familial association of Graves' disease with other autoimmune endocrinopathies. Furthermore, lymphocytes isolated from the thyroid glands of Graves' patients as well as their peripheral blood lymphocytes can be stimulated to produce TSAb by culturing with either a nonspecific mitogen or a specific antigen. There is also evidence that cell-mediated immunity may play a role in Graves' disease. However, the major pathogenetic agent is certainly antibody to the thyrotropin receptor. These antibodies appear to be heterogeneous and associated with a number of biological activities, including binding, blocking, and stimulating of the receptor. Perhaps in the future the general term "thyrotropin receptor antibody" will replace the other operational terms for this antibody.

Thyrotropin receptor antibodies are not the only autoantibodies to thyroid found in this disease. High titers of thyroglobulin antibodies occur in ~25% of Graves' patients (Table I) and ~50% of Graves' patients have antibodies to thyroid microsomes (Table III). The reason why the thyroid should be subjected to several distinct autoimmune aggressions is unclear, but it is plausible that an initial immunological injury to the thyroid stimulates secondary or even tertiary autoimmune responses in genetically susceptible individuals.

All thyroidologists agree that the demonstration of thyroid-stimulating antibodies has contributed importantly to our understanding of Graves' disease. However, considerable controversy surrounds the practical application of these assays. Most clinicans feel the assays are of little assistance in diagnosis and management of patients and are too time consuming and expensive. We suggest that the technical improvement of these assays will add significantly to the earlier detection, more accurate diagnosis, and more effective treatment of Graves' disease.

REFERENCES

Aanderud, S., Matre, R., and Varhaug, J. E. (1982). *Int. Arch. Allergy Appl Immunol.* **69,** 137–142.

Adams, D. D., and Kennedy, T. H. (1967). *J. Clin. Endocrinol. Metab.* **27,** 173–177.
Adams, D. D., and Purves, H. D. (1956). *Proc. Univ. Otago Med. Sch.* **34,** 11–12.
Adams, D. D., Fastier, F. N., Howie, J. B., Kennedy, T. H., Kilpatrick, J. C., and Stewart, R. D. H. (1974). *J. Clin. Endocrinol. Metab.* **39,** 826–832.
Allison, A. C. (1971). *Lancet* **2,** 1401–1403.
Allison, A. C., Denman, A. M., and Barnes, R. D. (1971). *Lancet* **2,** 135–140.
Aoki, N., and DeGroot, L. J. (1979). *Clin. Exp. Immunol.* **38,** 532–530.
Banovac, K., Zakarija, M., McKenzie, J. M., Witte, A., and Sekso, M. (1981). *J. Clin. Endocrinol. Metab.* **53,** 651–653.
Baur, R. J., Wall, J. R., and Schleusener, H. (1983). *Life Sci.* **32,** 55–62.
Beall, G. N., and Kruger, S. R. (1980a). *Clin. Immunol. Immunopathol.* **16,** 485–497.
Beall, G. N., and Kruger, S. R. (1980b). *Clin. Immunol. Immunopathol.* **16,** 498–503.
Bech, K., Feldt-Rasmussen, U., Bliddal, H., Date, J., and Blichert-Toft, M. (1982). *Clin. Endocrinol.* **16,** 235–242.
Bigazzi, P. E. (1979). *In* "Mechanisms of Immunopathology" (S. Cohen, P. A. Ward, and R. T. McCluskey, eds.), pp. 157–180. Wiley, New York.
Bigazzi, P. E (1983). *In* "Principles and Practice of Surgical Pathology" (S. G. Silverberg, ed.), pp. 57–75. Wiley, New York.
Bigazzi, P. E., and Rose, N. R. (1975). *Prog. Allergy* **19,** 245–274.
Bonnyns, M., Bentin, J., DeVetter, G., and Duchateau, J. (1983). *Clin. Exp. Immunol.* **52,** 629–634.
Calder, E. A., and Irvine, W. J. (1975). *Clin. Endocrinol. Metab.* **4,** 287–318.
Calder, E. A., McLeman, D., Barnes, R. M., and Irvine, W. J. (1972). *Clin. Exp. Immunol.* **12,** 429–438.
Calder, E. A., Irvine, W. J., Davidson, N. McD., and Wu, F. (1976). *Clin. Exp. Immunol.* **25,** 17–22.
Canonica, G. W., Bagnasco, M., Corte, G., Ferrini, S., Ferrini, O., and Giordano, G. (1982). *Clin. Immunol. Immunopathol.* **23,** 616–625.
Chow, A., Baur, R. J., Schleusener, H., and Wall, J. R. (1983). *Life Sci.* **32,** 67–74.
Creemers, P., Rose, N. R., and Kong, Y. M. (1983). *J. Exp. Med.* **157,** 559–571.
Davies, T. F. (1983). *Clin. Endocrinol. Metab.* **12,** 79–100.
DeGroot, L., and Sridama, V. (1983). *Prog. Clin. Biol. Res.* **116,** 1–22.
Doniach, D., and Roitt, I. M. (1957). *J. Clin. Endocrinol. Metab.* **17,** 1293–1304.
Doniach, D., and Roitt, I. M. (1976). *In* "Textbook of Immunopathology" (P. A. Miescher and H. J. Müller-Eberhard, eds.), Vol. I, pp. 22–34. Schwabe, Basel.
Doniach, D., and Vaughan Hudson, R. (1957). *Br. Med. J.* **1,** 672–678.
Doniach, D., Vaughan Hudson, R., and Roitt, I. M. (1957). *Proc. R. Soc. Med.* **50,** 936–939.
Doniach, D., Bottazzo, G. F., and Russell, R. C. G. (1979). *Clin. Endocrinol. Metab.* **8,** 63–80.
Drexhage, H. A., Bottazzo, G. F., Doniach, D., Bitensky, L., and Chayen, J. (1980). *Lancet* **2,** 287–292.
Endo, Y., Nakana, J., Horinouchi, K., Ohtaki, S., Izumi, M., and Ishikawa, E. (1980). *Clin. Chim. Acta* **103,** 67–77.
Endo, K., Amir, S. M., and Ingbar, S. H. (1981). *J. Clin. Endocrinol. Metab.* **52,** 1113–1123.
Engvall, E., and Perlmann, P. (1975). *In* "Automation in Microbiology and Immunology" (C. G. Heden and T. Ileni, eds.), pp. 529–542. Wiley, New York.
Faerch, T., Hojslet, P. E., Mose, C., and Paby, P. (1980). *Scand. J. Clin. Lab. Invest.* **40,** 263–269.

Fagraeus, A., and Jansson, J. (1970). *Immunology* **18,** 413–416.
Farid, N. R. (1981). *In* "HLA in Endocrine and Metabolic Disorders" (N. R. Farid, ed.), pp. 85–143. Academic Press, New York.
Farid, N. R., Barnard, J. M., Kutas, C., Noel, E. P., and Marshall, W. H. (1975). *Int. Arch. Allergy* **49,** 837–842.
Farid, N. R., Barnard, J. M., and Marshall, W. H. (1976). *Tissue Antigens* **8,** 181–189.
Farid, N. R., Sampson, L., Moens, H., and Barnard, J. M. (1981). *Tissue Antigens* **17,** 265–268.
Fenzi, G. F., Pinchera, A., Bartalena, L., Manzani, F., Macchia, E., and Baschieri, L. (1980). *In* "Thyroid Research VIII: Proceedings of the Eighth International Thyroid Congress, Sydney, Australia, 3-8 February 1980" (J. R. Stockigt and S. Natagaki, eds.), pp. 703–706. Pergamon, Oxford.
Furszyfer, J., Kurland, L. T., McConahey, W. M., Coolner, L. B., and Elveback, L. R. (1972). *Metab. Clin. Exp.* **21,** 197–204.
Gurmuth, F. G., Rodriquez, E., Siddiqui, S. Y., Lorelle, C. A., McGee, S., Milano, L. L., and Wise, O. (1978). *Lab. Invest.* **38,** 404–408.
Ginsberg, J., Segal, D., Ehrlich, R. M., and Walfish, P. G. (1978). *Clin. Endocrinol.* **8,** 133–139.
Graves, R. J. (1835). *London Med. Surg. J.* **7,** 516–520.
Grumet, F. C., Payne, R. O., Konishi, J., Mori, T., and Kriss, J. P. (1975). *Tissue Antigens* **6,** 347–352.
Hashimoto, H. (1912). *Arch. Klin. Chir.* **97,** 219–248.
Hay, F. C., and Torrigiani, G. (1974). *Clin. Exp. Immunol.* **16,** 517–521.
Himsworth, R. L., Byfield, P. G. H., and Copping, S. (1983). *Life Sci.* **32,** 119–125.
Irvine, W. J., Gray, R. S., Morris, P. J., and Ting, A. (1977). *Lancet* **2,** 898–900.
Iwatani, Y., Amino, N., Mori, H., Asari, S., Matsuzuka, F., Kuma, K., and Miyai, K. (1982). *J. Immunol. Methods* **48,** 241–250.
Iwatani, Y., Amino, N., Mori, H., Asari, S., Izumiguchi, M., Kumahara, Y., and Miyai, K. (1983). *J. Clin. Endocrinol. Metab.* **56,** 251–254.
Jansson, J., Fagraeus, A., and Biberfeld, G. (1968). *Clin. Exp. Immunol.* **3,** 287–304.
Jansson, J., Tötterman, T. H., Sallstrom, J., and Dahlberg, P. A. (1983). *J. Clin. Endocrinol. Metab.* **56,** 1164–1168.
Jordan, S. C., Johnston, W. H., and Bergstein, J. M. (1978). *Arch. Pathol. Lab. Med.* **102,** 530–533.
Jordan, S. C., Buckingham, B., Sakai, R., and Olson, D. (1981). *N. Engl. J. Med.* **304,** 1212–1215.
Kalderon, A. E. (1980). *Pathol. Annu.* **15,** 23–35.
Kalderon, A. E., and Bogaars, H. A. (1977). *Am. J. Med.* **63,** 729–734.
Kalderon, A. E., Bogaars, H. A., and Diamond, I. (1973). *Am. J. Med.* **55,** 485–491.
Khoury, E. L., Hammond, L., Bottazzo, G. F., and Doniach, D. (1981). *Clin. Exp. Immunol.* **45,** 316–328.
Kishihara, M., Yoshinobu, N., Baba, Y., Ohgo, S., Matsukara, S., Kuma, K., and Imura, H. (1978). *Acta Endocrinol.* **88,** 65–74.
Kite, J. H., Jr., Rose, N. R., Kano, K., and Witebsky, E. (1965). *Ann. N. Y. Acad. Sci.* **124,** 626–643.
Knecht, H., and Hedinger, C. E. (1982). *Histopathology* **6,** 511–538.
Knecht, H., Saremaslani, P., and Hedinger, C. (1981). *Virchows Arch. A: Pathol. Anat.* **393,** 215–231.
Knox, A. J. S., von Westarp, C., Row, V. V., and Volpé, R. (1976a). *Metabolism* **25,** 1217–1223.

Knox, A. J. S., von Westarp, C., Row, V. V., and Volpé, R. (1976b). *J. Clin. Endocrinol. Metab.* **43,** 330–337.

Kodama, K., Bandy-Dafoe, P., Sikorska, H., Bayly, R., and Wall, J. R. (1982). *Lancet* **2,** 1353–1356.

Koh, C. S., Ahn, I. M., Cho, B. Y., and Kim, Y. I. (1983). *In* "Current Problems in Thyroid Research" (N. Ui, K. Torizuka, S. Nagataki, and K. Miyai, eds.), pp. 489–492. Excerpta Medica, Princeton.

Kotani, T., Komuru, K., Yoshiki, T., Itoh, T., and Aizawa, M. (1981). *Clin. Exp. Immunol.* **45,** 329–337.

Kotani, T., Komuru, K., Yoshiki, T., Itoh, T., and Aizawa, M. (1982). *Clin. Immunol. Immunopathol.* **24,** 111–121.

Kriss, J. P., Pleshakov, V., and Chien, J. R. (1964). *J. Clin. Endocrinol. Metab.* **24,** 1005–1028.

Kriss, J. P., Pleshakov, V., Rosenblum, A. L., Holderness, M., Sharp, G., and Utiger, R. J. (1967). *J. Clin. Endocrinol. Metab.* **27,** 582–593.

Laryea, E., Row, V. V., and Volpé, R. (1973). *Clin. Endocrinol. (Oxford)* **2,** 23–25.

Lewis, M., and Rose, N. R. (1985). *In* "Autoimmunity and Its Implications for Diseases of the Endocrine Glands" (R. Volpé, ed.), Dekker, New York. (In press.)

McCoy, J. P., Jr., Michaelson, J. H., and Bigazzi, P. E. (1983). *Life Sci.* **32,** 109–118.

McKenzie, J. M. (1958). *Endocrinology (Baltimore)* **63,** 372–382.

McLachlan, S. M., Bird, A. G., Weetman, A. P., Smith, R., and Hall, R. (1981). *Scand. J. Immunol.* **14,** 233–242.

McLachlan, S. M., Dickinson, A., Baylis, P., Proctor, S., and Smith, B. R. (1983a). *Clin. Exp. Immunol.* **53,** 397–405.

McLachlan, S. M., Dickinson, A. M., Malcolm, A., Farndon, J. R., Young, E., Proctor, S., and Smith, B. R. (1983b). *Clin. Exp. Immunol.* **52,** 45–53.

McLachlan, S. M., Fawcett, J., Atherton, M. C., Thompson, P., Baylis, P., and Smith, B. R. (1983c). *Clin. Exp. Immunol.* **52,** 620–628.

MacLean, D. B., Miller, K. B., Brown, R., and Reichlin, S. (1981). *J. Clin. Endocrinol. Metab.* **53,** 801–805.

Marine, D., and Lenhart, C. H. (1909). *Arch. Int. Med.* **4,** 253–270.

Mariotti, S., DeGroot, L. J., Scarborough, D., and Medof, M. E. (1979). *J. Clin. Endocrinol. Metab.* **49,** 679–686.

Matsuyama, T., Fukumori, J., and Tanaka, H. (1983). *Clin. Exp. Immunol.* **51,** 381–386.

Meek, J. C., Jones, A. E., Lewis, V. J., and Vanderlaan, W. P. (1964). *Proc. Natl. Acad. Sci. U.S.A.* **52,** 342–349.

Misaki, T., Konishi, J., Iida, Y., Endo, K., and Torizuka, K. (1983). *In* "Current Problems in Thyroid Research" (N. Ui, K. Torizuka, S. Nagataki, and K. Miyai, eds.), pp. 429–432. Excerpta Medica, Princeton.

Möbius, P. J. (1887). *Centralbl. Nervenh.* **10,** 225–229.

Nakamura, R. M., and Weigle, W. O. (1969). *J. Exp. Med.* **130,** 263–285.

Noma, T., Yata, J., Shishiba, Y., and Inatsuki, B. (1982). *Clin. Exp. Immunol.* **49,** 565–571.

Nye, L., DeCarvalho, L. P. C., and Roitt, I. M. (1980). *Clin. Exp. Immunol.* **41,** 252–263.

Nye, L., DeCarvalho, L. P., and Roitt, I. M. (1981). *Clin. Exp. Immunol.* **46,** 161–170.

Okita, N., Row, V. V., and Volpé, R. (1981). *J. Clin. Endocrinol. Metab.* **52,** 528–533.

Onaya, T., Kotani, M., Yamada, T., and Ochi, Y. (1973). *J. Clin. Endocrinol. Metab.* **36,** 859–866.

O'Reagan, S., Fong, J. S. C., Kaplan, B. S., de Chadarevian, J.-P., Lapointe, N., and Drummond, K. N. (1976). *Clin. Immunol. Immunopathol.* **6,** 341–346.

Pacini, F., Fragli, P., Mariotti, S., and DeGroot, L. J. (1982). *J. Clin. Immunol.* **2,** 335–342.

Penhale, W. J., Farmer, A., McKenna, R. P., and Irvine, W. J. (1973). *Clin. Exp. Immunol.* **15,** 225–236.

Penhale, W. J., Farmer, A., Urbaniak, S. J., and Irvine, W. J. (1975). *Clin. Exp. Immunol.* **19,** 179–191.

Penhale, W. J., Irvine, W. J., Inglis, J. R., and Farmer, A. (1976). *Clin. Exp. Immunol.* **25,** 6–16.

Pinchera, A., Fenzi, G. F., Bartalena, L., Chiovato, L., and Marcocci, C. (1980). *In* "The Thyroid Gland" (M. De Visscher, ed.), pp. 413–441. Raven, New York.

Ploth, D. W., Fitz, A., Schnetzler, D., Seidenfeld, J., and Wilson, C. B. (1978). *Clin. Immunol. Immunopathol.* **9,** 327–334.

Plummer, H. S. (1913). *Trans. Assoc. Am. Physicians* **28,** 587–594.

Podleski, W. K. (1972). *Clin. Exp. Immunol.* **11,** 543–548.

Rapoport, B., Greenspan, F. S., Filetti, S., and Pepitone, M. (1984). *J. Clin. Endocrinol. Metab.* **58,** 332–338.

Roitt, I. M., Doniach, D., Campbell, P. N., and Vaughan Hudson, R. (1956). *Lancet* **2,** 820–821.

Roitt, I. M., Campbell, P. N., and Doniach, D. (1958). *Biochem. J.* **69,** 248–256.

Romanin, L. (1979). *Can. J. Med. Tech.* **41,** 141–147.

Rose, N. R. (1975). *Cell. Immunol.* **18,** 360–364.

Rose, N. R. (1983). *Immunol. Allergy Pract.* **5,** 36–42.

Rose, N. R., and Bigazzi, P. E. (1978). *In* "Handbook Series in Clinical Laboratory Science, Immunology" (A. Baumgarten and F. F. Richards, eds.), pp. 305–384. CRC Press, Boca Raton, Florida.

Rose, N. R., and Kong, Y. M. (1982). *Life Sci.* **32,** 85–91.

Rose, N. R., and Witebsky, E. (1955). *J. Immunol.* **75,** 291–300.

Rose, N. R., and Witebsky, E. (1956). *J. Immunol.* **76,** 417–427.

Rose, N. R., Bacon, L. D., Sundick, R. S., Kong, Y. M., Esquivel, P., and Bigazzi, P. E. (1977). *In* "Autoimmunity: Genetic, Immunologic, Virologic and Clinical Aspects" (N. Talal, ed.), pp. 63–87. Academic Press, New York.

Rose, N. R., Kong, Y. M., Okayasu, I., Giraldo, A. A., Beisel, K., and Sundick, R. S. (1981). *Immunol. Rev.* **55,** 299–314.

Rose, N. R., Accavitti, M., Pydyn, E. F., Leon, M. A., and Brown, R. K. (1982). *Adv. Exp. Med. Biol.* **150,** 23–35.

Rumph, P., and Smith, P. E. (1926). *Anat. Rec.* **33,** 289–298.

Schardt, C. W., McLachlan, S. M., Matheson, J., and Smith, B. R. (1982). *J. Immunol. Med.* **55,** 155–168.

Seybold, D., Ryan, E. A., and Wall, J. R. (1981). *J. Clin. Lab. Immunol.* **6,** 241–244.

Shamshuddin, A. K. M., and Lane, R. A. (1981). *Hum. Pathol.* **12,** 561–573.

Sharp, G. C., Mullen, H., and Kyriakos, M. (1974). *J. Immunol.* **112,** 478–487.

Shishiba, Y., Shimizu, T., Shizuko, Y., and Shizuma, K. (1973). *J. Clin. Endocrinol. Metab.* **36,** 517–521.

Shulman, S. (1971). *Adv. Immunol.* **14,** 85–185.

Shulman, S., Rose, N. R., and Witebsky, E. (1955). *J. Immunol.* **75,** 291–300.

Sköldstam, L., Anderberg, B., and Norrby, K. (1978). *Clin. Exp. Immunol.* **31,** 524–525.

Smith, B. R. (1981). *In* "Receptors and Recognition, Series B: Receptor Regulation" (R. J. Lefkowitz, ed.), Vol. XIII, pp. 217–244. Chapman & Hall, London.

Smith, B. R., Mukhtar, G. A., Pyle, P., Kendall-Taylor, P., and Hall, R. (1976). *In* "Thyroid Research, Excerpta Medica International Congress Series" (J. Robbins and L. E. Braverman, eds.), pp. 411–413. Excerpta Medica, Amsterdam.

Soborg, M., and Halberg, P. (1968). *Acta Med. Scand.* **183,** 101–105.

Solomon, D. H., and Chopra, I. J. (1980). *In* "Thyroid Research VIII: Proceedings of the Eighth International Thyroid Congress, Sydney, Australia, 3-8 February 1980" (J. R. Stockigt and S. Nagataki, eds.), pp. 717–720. Pergamon, Oxford.

Sridama, V., Pacini, F., and DeGroot, L. J. (1982). *J. Clin. Endocrinol. Metab.* **54,** 316–319.

Sternthal, E., Like, A. A., Sarantis, K., and Braverman, L. E. (1981). *Diabetes* **30,** 1058–1061.

Suzuki, S., Mitsunaga, M., Miyoshi, M., Hirakawa, S., Nakagawa, O., Miura, H., and Ofuji, T. (1980). *J. Clin. Endocrinol. Metab.* **51,** 446–453.

Tao, T. W., and Kriss, J. P. (1982). *J. Clin. Endocrinol. Metab.* **55,** 935–940.

Thielemans, C., Vanhaealst, L., de Waele, M., Jonckheer, M., and Van Camp, B. (1981). *Clin. Endocrinol. (Oxford)* **15,** 259–263.

Thomsen, M., Ryder, L. P., Bech, K., Bliddal, H., Feldt-Rasmussen, U., Molholm, J., Kappelgaard, E., Nielsen, H., and Svejgaard, A. (1983). *Tissue Antigens* **21,** 173–175.

Thorsby, E., Segaard, E., Solem, J. H., and Kornstad, L. (1975). *Tissue Antigens* **6,** 54–55.

Tötterman, T. H. (1978). *Clin. Immunol. Immunopathol.* **10,** 270–277.

Valente, W. A., Vitti, P., Yavin, Z., Yavin, E., Botella, C. M., Grollman, E. F., Toccafondi, R. S., and Kohn, L. D. (1982). *Proc. Natl. Acad. Sci. U.S.A.* **79,** 6680–6684.

Valente, W. A., Vitti, P., Rotella, C. M., Vaughan, M. M., Aloj, S. M., Grollman, E. F., Ambesi-Impiombato, F. S., and Kohn, L. D. (1983). *N. Engl. J. Med.* **309,** 1028–1034.

Verger, M. F., Droz, D., and Vantelon, J. (1983). *Presse Méd.* **12,** 83–86.

Voller, A., Bidwell, D. E., and Burek, C. L. (1980). *Proc. Soc. Exp. Biol. Med.* **163,** 402–405.

von Basedow, C. A. (1840). *Wschr. Gesamte Heilkunde* **13,** 197–204.

Wall, J. R., and Chartier, B. (1981). *J. Clin. Lab. Immunol.* **6,** 115–119.

Wall, J. R., Baur, R., Schleusener, H., and Bandy-Dafoe, P. (1983). *J. Clin. Endocrinol. Metab.* **56,** 164–169.

Wartenberg, J., Doniach, D., Brostoff, J., and Roitt, I. M. (1973a). *Int. Arch. Allergy* **44,** 396–408.

Wartenberg, J., Doniach, D., Brostoff, J., and Roitt, I. M. (1973b). *Clin. Exp. Immunol.* **14,** 203–212.

Weetman, A. P., Pinching, A. J., Pussel, B. A., Evans, D. J., Sweny, P., and Rees, A. J. (1981). *Clin. Nephrol.* **15,** 50–51.

Weetman, A. P., McGregor, A. M., Lazarus, J. H., and Hall, R. (1982). *Clin. Exp. Immunol.* **48,** 196–200.

Weigle, W. O. (1980). *Adv. Immunol.* **30,** 159–273.

Weiss, I., and Davies, T. F. (1982). *J. Clin. Endocrinol. Metab.* **54,** 282–285.

Weiss, I., De Bernardo, E., and Davies, T. F. (1982). *Clin. Immunol. Immunopathol.* **23,** 50–57.

Weissel, M., Höfer, R., Zasmeta, H., and Mayr, W. R. (1980). *Tissue Antigens* **16,** 256–257.

Wenzel, B., Wenzel, K. W., Kotulla, P., and Schleusener, H. (1981). *J. Endocrinol. Invest.* **4,** 161–166.

Werner, S. C., Feind, C. R., and Mitsuyasu, A. (1967). *N. Engl. J. Med.* **276,** 132–138.

Werner, S. C., Wigelius, O., Fierer, J. A., and Hsu, K. C. (1972). *N. Engl. J. Med.* **287,** 421–425.

Winand, R. J., and Kohn, L. D. (1979). *J. Biol. Chem.* **245,** 967–975.

Witebsky, E. (1929). *Naturwissenschaften* **17,** 771–776.

Witebsky, E. (1968). *Am. J. Clin. Pathol.* **49,** 301–311.

Witebsky, E., and Klinke, J. (1933). *Z. Immunitaetsforsch.* **78,** 509–523.

Witebsky, E., and Rose, N. R. (1956). *J. Immunol.* **76,** 408–416.

Witebsky, E., and Steinfeld, J. (1928). *Z. Immunitaetsforsch.* **58,** 271–296.

Witebsky, E., Rose, N. R., and Shulman, S. (1955). *J. Immunol.* **75,** 269–281.

Woolner, L. B. (1964). *In* "The Thyroid" (J. B. Hazard and D. E. Smith, eds.), pp. 123–142. Williams & Wilkins, Baltimore, Maryland.

Woolner, L. B., McConahey, W. M., and Beahrs, O. H. (1959). *J. Clin. Endocrinol. Metab.* **19,** 53–83.

Yagi, Y. (1981). *Acta Pathol. Jpn.* **31,** 611–622.

Zanetti, M., and Bigazzi, P. E. (1981). *Eur. J. Immunol.* **11,** 187–195.

CHAPTER 8

Adrenal Autoimmunity and Autoimmune Polyglandular Syndromes

NOEL K. MACLAREN
Department of Clinical Pathology
College of Medicine
University of Florida
Gainesville, Florida

ROBERT M. BLIZZARD
Department of Pediatrics
University of Virginia Medical Center
Charlottesville, Virginia

I. INTRODUCTION

Autoimmune endocrinopathies coexist in single patients or in in individual pedigrees, and thus are often component entities of an autoimmune polyglandular syndrome. This fact has been appreciated most readily among patients with autoimmune Addison's disease, accounting for the considerable interest shown in this relatively rare disease. Whereas the frequency of autoimmune involvement of many endocrine tissues at the same time suggests an underlying defect in the host's immunoregulatory network responsible for maintaining tolerance to self antigens, it also is clear that other putative autoimmune diseases such as the collagenoses or multiple sclerosis are not more common in patients with endocrine autoimmunities. These findings suggest that endocrine glands may share common antigenic determinants and that autoimmunity against one of them frequently is associated with involvement of other glands.

In common with all diseases considered to be of autoimmune origin, the autoimmune endocrinopathies are thought to be polygenic, with one or more genes linked to the *B-D* region of the *HLA* gene complex on the short arm of chromosome VI. Incomplete penetrance is often seen, and the role of environmental factors is believed to be of importance in triggering expression of the disease. Viral agents often are suspected in the latter regard.

In this chapter, we will present the knowledge that is currently available pertaining to autoimmunity of the adrenal gland, including the historical and etiological considerations. Emphasis will be placed upon the associations of Addison's disease with other autoimmune diseases. Such associations comprise the polyglandular autoimmune syndromes.

II. HISTORICAL ASPECTS

In 1849, Thomas Addison, a physician at Guy's Hospital, London, described a group of patients with severe anemia. In three patients upon whom he performed autopsies, disease of the suprarenal capsules was identified, suggesting to Addison a state of cause and effect. In this initial paper (Addison, 1849), there is no mention of increased dermal pigmentation but rather descriptions of waxy pallor of the skin. It is plausible that Addison described pernicious anemia that coincidentally occurred in patients with coexistent adrenalitis and/or adrenal atrophy. Six years later in 1855, Addison published his classical observations of clinical adrenocortical failure. Of the 11 cases reported, 7 were men and 4 were women. Four had tuberculosis, 4 had malignancies, and the origin of disease of the

adrenal glands was unexplained in 3. Interestingly because of the association, one of the latter had vitiligo (Addison, 1855).

In 1862, Wilks proposed to name the suprarenal syndrome "Addison's disease" (Wilks, 1862). Sixty years later, Brenner noted that Addison's disease was clinically apparent only when most of the adrenal cortex had been destroyed, and symptoms were present even if the medullary portions of the gland remained intact (Brenner, 1928). Shortly afterward, Swingle and Pfiffner demonstrated that suprarenal cortical extracts could maintain the lives of adrenalectomized cats (Swingle and Pfiffner, 1930). These extracts eventually proved to be beneficial when used on patients by Rountree at the Mayo Clinic (Rountree and Snell, 1931). The cycle of cause (i.e., adrenal atrophy or destruction) and effect (i.e., clinical Addison's disease; which was reversible by adrenocortical hormones) had been closed.

Over the decades of the twentieth century, tuberculosis as a significant cause of adrenal failure was to fade in Western countries. However, patients with Addison's disease continued to be seen. In 1926, Schmidt described the simultaneous lymphocytic infiltration of both the thyroid gland and the adrenal cortices in two patients with Addison's disease (Schmidt, 1926). An autoimmune polyglandular syndrome had thus been described. The observation of Schmidt provided a foundation for an immunological explanation for both Hashimoto's and Addison's diseases. In 1964, Carpenter was to add significantly to Schmidt's syndrome by confirming that insulin-dependent diabetes was more common than normal in patients with "Schmidt's syndrome" (Carpenter *et al.*, 1964), a situation reported earlier (Wehrmacher, 1961).

Evidence for an autoimmune pathogenesis of any of these entities was first provided by Doniach and Roitt in 1957, in their initial descriptions of the first organ-specific autoantibody. The disease was Hashimoto's disease and the autoantibody was a circulating precipitin reacting to thyroglobulin (Doniach and Roitt, 1957). Shortly afterward, autoantibodies reacting to adrenocortical cells were discovered (Anderson *et al.*, 1957; Blizzard and Kyle, 1963) in the sera of patients with nontuberculous, "idiopathic" Addison's disease. Many studies confirming the immunological nature of "idiopathic" Addison's disease followed, and the complexity of the polyglandular syndromes associated with adrenal autoimmunity grew rapidly. In 1981, Neufeld, Blizzard, and Maclaren proposed that these syndrome complexes were divisible into two entities, provisionally named types I and II autoimmune polyglandular syndromes (Neufeld *et al.*, 1981). This division has enhanced our understanding of polyglandular autoimmune disease.

III. ADDISON'S DISEASE AND AUTOIMMUNE POLYGLANDULAR SYNDROMES (APS)

As outlined above, Addison's original cases included several that may have had combined idiopathic adrenal atrophy with pernicious anemia and at least one with vitiligo (Addison, 1849, 1855). In addition to the reports of Schmidt (1926) and Carpenter *et al.* (1964), there have been multiple reports of polyglandular failures (Dunlap, 1963; Turkington and Lebovitz, 1967; Appel and Holub, 1976; Irvine, 1978; Irvine and Barnes, 1975; Neufeld *et al.*, 1981) (see Table I). Beaven expanded on the association between Addison's disease and diabetes mellitus (Beaven *et al.*, 1959), as did others (Nerup and Binder, 1973; Bottazzo *et al.*, 1974; Nerup 1974c; Riley *et al.*, 1980). Other investigators noted increased frequencies of Graves' disease (Irvine, 1978; Males *et al.*, 1971), of pernicious anemia/atrophic gastritis (Addison, 1849; Irvine, 1978; Doniach *et al.*, 1963; Strickland, 1969; Forcier *et al.*, 1972; Meecham and Jones, 1967), of goiter/thyroiditis (Schmidt, 1926; Carpenter *et al.*, 1964; Neufeld *et al.*, 1981; Irvine and Barnes, 1975; Beaven *et al.*, 1959; Nerup and Biner, 1973; Nerup, 1974c; Riley *et al.*, 1980; Cunliffe *et al.*, 1968b; McHardy-Young *et al.*, 1972; Kiaer and Nørgaard, 1969; Christy *et al.*, 1962; Edmonds *et al.*, 1973; Gastineau *et al.*, 1964; Gharib *et al.*, 1972; Karlish and McGregor, 1970; Nerup, 1974a; Solomon *et al.*, 1965), of hypoparathyroidism and/or chronic mucocutaneous moniliasis (Neufeld *et al.*, 1981; Thorpe and Handley, 1929; Blizzard *et al.*, 1966; Blizzard and Gibbs, 1968; Block *et al.*, 1971; Castells *et al.*, 1971; Craig *et al.*, 1955; Hung *et al.*, 1963; Irvine and Barnes, 1975; Kenny and Holliday, 1964; Morse *et al.*, 1961; Myllarniemi and Perheentupa, 1978; Perheentupa and Hickkala, 1973; Pearlman, 1961; Spinner *et al.*, 1968; Stankler and Bensher, 1972; Tomar *et al.*, 1979; Vazquez and Kenny, 1973; Weinberg *et al.*, 1976; Wells *et al.*, 1972; Whitaker *et al.*, 1956; Wuepper and Fudenberg, 1967; Bronsky *et al.*, 1958; Drury *et al.*, 1970; Golonka and Goodman, 1968; Spinner *et al.*, 1969; Arulanantham *et al.*, 1979; Neufeld *et al.*, 1981, as reviewed in Eisenbarth *et al.*, 1979), vitiligo and/or alopecia (Addison, 1855; Neufeld *et al.*, 1981; Irvine and Barnes, 1975; Cunliffe *et al.*, 1968b; Blizzard and Gibbs, 1968; Perheentupa and Hickkala, 1973; Eisenbarth *et al.*, 1979; Betterle *et al.*, 1979; Bor *et al.*, 1969; Brostoff *et al.*, 1969; Collen *et al.*, 1979; Cunliffe *et al.*, 1968a; Dawber, 1970; Frey *et al.*, 1973; Grunnet *et al.*, 1970; Harmans *et al.*, 1969; Hertz *et al.*, 1977; Howitz and Schwartz, 1971; Kern *et al.*, 1973; McGregor *et al.*, 1972; Muller and Winkleman, 1963), of myasthenia gravis (Kane and Weed, 1950; Eisenbarth *et al.*, 1978; Fritze *et al.*, 1976), of hypogonadism/infertility/amenorrhea (Neufeld *et al.*, 1981; Blizzard and Kyle, 1963; Turkington and Le-

TABLE I

The Autoimmune Entities That Occurred in 295 Patients (183 Females and 112 Males) Who Had Addison's Disease as Part of a PGA Syndrome[a]

Associated entity	Number	%	Females	Males	Female to male Ratio
Autoimmune thyroid disease	162	55	112	50	2.2
Insulin-requiring diabetes	118	40	66	52	1.3
Hypoparathyroidism	54	18	33	21	1.6
Mucocutaneous candidiasis	52	15	29	23	1.3
Alopecia	24	8	14	10	1.4
Gonadal failure	20	7	15	5	3.0
Vitiligo	16	5	8	8	1.0
Malabsorption	5	5	6	10	0.6
Pernicious anemia	10	3	5	6	0.8
Chronic active hepatitis	9	3	6	3	2.0

[a] From Neufeld *et al.* (1981). Reproduced with permission of the editor of *Medicine (Baltimore)*.

bovitz, 1967; Elder *et al.*, 1981; Christy *et al.*, 1962; Edmonds *et al.*, 1973; Irvine and Barnes, 1975; Perheentupa and Hickkala, 1973; Vazquez and Kenny, 1973; Weinberg *et al.*, 1976; Golanka and Goodman, 1968; Collen *et al.*, 1979; Crispell and Parson, 1952; Irvine and Barnes, 1974; Irvine *et al.*, 1968; Matz and Tucker, 1969; Ruehsen *et al.*, 1972; Ayala *et al.*, 1979; Irvine *et al.*, 1967), of coeliac disease (Neufeld *et al.*, 1980a; Cooper *et al.*, 1978), of hypopituitarism/hypophysitis (Kiaer and Nørgaard, 1969; Castells *et al.*, 1971; Bottazzo and Doniach, 1978; Gleason *et al.*, 1978; Ludwig and Schernthaner, 1978; Bottazzo *et al.*, 1975; Goudie *et al.*, 1968; Rupp *et al.*, 1953), and of chronic active hepatitis and malabsorption syndromes (Neufeld *et al.*, 1981; Perheentupa and Hickkala, 1973; Van Thiel *et al.*, 1977) among patients with nontuberculous ("autoimmune") Addison's disease and other patients with autoimmune polyglandular syndromes. From these reports in the literature and from a series of patients that we were able to assemble ourselves with the help of colleagues from the Lawson Wilkins Pediatric Endocrine Society, we were able to divide patients with autoimmune Addison's disease into two types as based principally on disease associations and ages of onsets (Neufeld *et al.*, 1979, 1980a, 1981) as adapted by Eisenbarth and Jackson (1981) (see Table II).

TABLE II

Addison's Disease in Polyglandular Autoimmune Disease: Type I and Type II[a]

Entity	Type I (%)		Type II (%)	
Addison's disease	71/71	(100)	224/224	(100)
Hypoparathyroidism	54/71	(76)	— —	
Chronic mucocutaneous candidiasis	52/71	(73)	— —	
Autoimmune thyroid disease	8/71	(11)	154/224	(69)
Insulin-requiring diabetes	3/71	(4)	117/224	(52)
Chronic active hepatitis	9/71	(13)	— —	
Malabsorption syndromes	16/71	(22)	— —	
Alopecia	23/71	(32)	1/224	(0.5)
Pernicious anemia	9/71	(13)	1/224	(0.5)
Gonadal failure	12/71	(17)	8/224	(3.6)
Vitiligo	6/71	(8)	10/224	(4.5)
Female/male ratio	43/28 1.5		145/79 1/8	
HLA Associations	No constant findings		*B8* (*A1*)	

[a] From Neufeld *et al.* (1981). Reproduced with permission of the editor of *Medicine* (*Baltimore*).

A. Type I Autoimmune Polyglandular Syndrome (Blizzard's Syndrome)

This complicated and relatively rare syndrome complex is seen almost exclusively in children. It is defined by the occurrence of at least two of the three diseases—chronic mucocutaneous moniliasis, acquired hypoparathyroidism, and autoimmune (autoantibody-positive) Addison's disease.

These principal diseases usually make their onset in a uniform order; that is, moniliasis first, hypoparathyroidism second, and Addison's disease last, albeit one of these components may not appear (Neufeld *et al.*, 1981). Further, other complications such as pernicious anemia and hypogonadism tend to occur last, characteristically during the teenage years or later. At times, the hypocalcemia of hypoparathyroidism may be masked in the presence of concomitant Addison's disease, only to declare itself after steroid replacement therapy has been initiated. Alopecia and vitiligo may be seen, and often the degree of these skin lesions is striking; alopecia totalis with virtual absence of body hair, and vitiligo that extends to the head causing large areas of completely depigmented white hair.

Difficulties with malabsorption are common, and these are due to a

variety of causes. Achlorhydria due to gastric atrophy, which is associated with autoantibodies against gastric parietal cells, is common. Deficiency of IgA immunoglobulin has been reported. Perhaps because of the latter two problems, overgrowth of bacteria in the upper intestine (jejunal-duodenum) is not uncommon, as is invasive giardiasis (J. Andres, N. Maclaren, and D. Barrett, pers. obs.). Further, some of the patients may be heavily afflicted with oral–eosophageal–eustachial moniliasis, often with concomitant enlargement of the parotid glands resembling that seen with the sicca syndrome, which may indeed coexist. Chronic active hepatitis (HA.Ag-negative type) is distressingly common among such patients, and all patients should be screened routinely for this problem whenever type I APS has been diagnosed. Pernicious anemia is also seen, especially among older patients. The pernicious anemia is of the classical type although it presents at an unusually young age, while bacterial overgrowth of the intestine and chronic active hepatic disease are compounding variables that contribute to vitamin B_{12} deficiency seen. Hypogonadism in association with autoantibodies against steroid hormone-producing cells is also fairly common, as discussed below.

In many instances, elevations of pituitary gonadotrophins may be less than expected for primary hypogonadism, suggesting a more complex explanation for the hypogonadism manifesting as lack of sexual development or amenorrhea than solely that of destructive lymphocytic gonadal inflammation. In addition, many patients are malnourished and chronically ill from chronic hepatitis and malabsorption, etc., and may have a secondary hypogonadotrophic hypogonadism on this basis. Others may have involvement of the anterior pituitary gland as mentioned above. In contrast to the associated features of type I APS mentioned above, Hashimoto's disease and insulin-requiring diabetes are unusual. Notwithstanding, the latter diseases are described in patients with type I APS, which suggests to us that such instances may represent interactions of separate autoimmune genotypes.

We have been following a 15-year-old boy with full expression of the type I APS phenotype for >5 years. In all that time, he has exhibited strongly positive pancreatic islet-cell autoantibodies, including the complement-fixing type, without insulin-dependent diabetes becoming apparent. Other patients with type I APS and no diabetes whom we have studied have lacked islet-cell autoantibodies. One additional patient, however, had autoantibodies to pancreatic α cells, although there was a normal reserve of glucagon.

Several recent reports suggest that patients with type I APS may have an underlying defect in T lymphocytes. To our knowledge, the numbers of circulating lymphocytes have not been reported to be deficient; however,

the inability to eradicate monilia, which is characteristic of this disorder, strongly suggests the presence of a T-cell defect (Ehligen, 1969). This may be partially correctable with the use of transfer factor (Kirkpatrick, 1979). One group has suggested that such patients may have defective T-cell suppressor functions (Arulanantham *et al.*, 1979), which seems a reasonable explanation; however, in several patients whom we studied using the OK series of monoclonal antibodies, we found an increase rather than a deficiency of the OKT8+ cytotoxic/suppressor T-cell subset. Noting the aggregation of these several immunological defects, including possible polyclonal B-cell expansion with proliferation of multiple autoantibodies, defective functioning T-cell immunosuppression, and defective cell-mediated immunity, we speculated that defective T- to B-cell suppression in these patients may be present without similar T- to T-cell suppressive defects. Depressed T-cell immunity could therefore result. Obviously further studies are required on this group of patients to validate this speculation.

There are some similarities between type I APS and the BB rat model (Naknooda *et al.*, 1977), in that widespread autoimmunities are associated with severe defects in T lymphocytes. In this strain of rat, these defects appear to result from defective T-lymphocyte maturation, which results in T lymphopenia and markedly defective T-cell functions (Elder and Maclaren, 1983; Maclaren and Elder, 1983; Elder *et al.*, 1982).

B. Type II Autoimmune Polyglandular Syndrome (Schmidt–Carpenter's Syndrome)

This considerably more common form of APS occurs at all times of life, especially in middle-aged Caucasoid women. It is defined by the occurrence of Addison's disease in association with thyroid autoimmune disease and/or insulin-dependent diabetes. If a sibling of a patient with type II polyglandular autoimmune disease has only autoimmune thyroid disease and diabetes mellitus, that patient also can be classified as having type II polyglandular autoimmune disease. Pernicious anemia occurs more frequently in mid-life or later; pernicious anemia of early onset as in type I APS is not found. Although a few individuals with type II APS have been reported to have gonadal failure, which is associated with high levels of circulating pituitary gonadotrophins (Neufeld *et al.*, 1981; Irvine and Barnes, 1974, 1975; Christy *et al.*, 1962; Edmonds *et al.*, 1973; Crispell and Parson, 1952; Irvine *et al.*, 1968; Matz and Tucker, 1969; Ruehsen *et al.*, 1972; Neufeld *et al.*, 1980a), its occurrence is at least considerably less frequent than in type I APS. Indeed, Elder and Maclaren could not find any of >500 insulin-dependent diabetic patients to have clinical hy-

pergonadotrophic hypogonadism or associated gonadal autoantibodies (Elder *et al.*, 1981).

The occurrence of chronic active hepatitis in type II APS is rare in relationship to its high frequency in type I APS. Gonadal failure and/or chronic active hepatitis are seen, however, in type II APS, which may also reflect the positive interaction of separate autoimmune genotypes.

IV. AUTOANTIBODIES AND AUTOIMMUNE ADDISON'S DISEASE

A. Adrenocortical-Specific Autoantibodies

In 1957, Anderson reported the presence of complement-fixing autoantibodies to saline extracts of human adrenal and thyroid glands in 2 of 10 patients with Addison's disease (Anderson *et al.*, 1957), and later several groups using similar methods or indirect immunofluorescence reported >50% of Addisonian patients to be positive for adrenocortical autoantibodies (Blizzard and Kyle, 1963; Irvine, 1978; Irvine and Barnes, 1974, 1975; Nerup and Binder, 1973; Nerup, 1974c; Riley *et al.*, 1980; Blizzard *et al.*, 1966; Spinner *et al.*, 1969; Irvine *et al.*, 1968; Ruehsen *et al.*, 1972; Elder *et al.*, 1981; Goudie *et al.*, 1966, 1968; Satsion *et al.*, 1980). The antigen or antigens responsible are present in human, monkey, bovine, rabbit, and guinea pig adrenal glands, but not in rodent adrenals or in (at least some) human adrenocortical neoplasms (Blizzard and Kyle, 1963; Nerup, 1974b; Anderson *et al.*, 1968; Andrada *et al.*, 1968). Normal human glands or those with adrenal hyperplasia are preferred as the source of antigen when testing for adrenal autoantibodies (Elder *et al.*, 1981). Fixation of the tissue prior to use renders it unacceptable to demonstrate adrenocortical autoantibodies (Elder *et al.*, 1981), which coincides with the suggestions of Blizzard and of Goudie and colleagues that the antigen is probably a lipoprotein predominantly of microsomal origin (Blizzard and Kyle, 1963; Goudie *et al.*, 1968; Anderson *et al.*, 1968).

Adrenocortical autoantibodies usually react with all layers of the adrenal cortex; however, instances of weak reactions, best demonstrated against the zona glomerulosa layer, have been noted in patients with insulin-dependent diabetes (Riley *et al.*, 1980). Adrenocortical autoantibodies have not been identified in tuberculous Addison's disease, suggesting that they do not arise secondarily as a consequence of destruction of the adrenal cortex (Irvine and Barnes, 1975; Neufeld *et al.*, 1981). The high specificity of adrenocortical autoantibodies for "autoimmune" Addison's disease can also be surmised because of the low frequency of these

autoantibodies in the general population. Nerup found only 0.1% positive, while Goudie in his study of 702 normal control patients found only 4 to be positive, a rate of 0.6%. Riley *et al.* (1980) had identical findings, but in a more recent study of >2000 normal control individuals, we found that the frequencies of adrenal autoantibodies increase somewhat with age, while the autoantibody was more frequent in women than in men at all ages (N. Maclaren and W. Riley, 1985).

Adrenal autoantibodies generally occur at low titers in the serum similar to those seen for pancreatic islet cell autoantibodies in insulin-dependent diabetes (Irvine and Barnes, 1975). However, the natural history of the autoantibody in respect to the appearance of adrenocortical insufficiency is not yet clear. By analogy to insulin-dependent diabetes (Gorsuch *et al.*, 1981), they would be expected to appear long before Addison's disease becomes apparent and to disappear coincident with the loss of the immunizing adrenal autoantigen in adrenal atrophy. C. Ketchum, W. Riley, and N. Maclaren (1984) recently investigated 15 patients with positive adrenocortical autoantibodies (half had insulin-dependent diabetes) but without overt Addison's disease for evidence of adrenocortical insufficiency. The results indicated that more than half of the patients with adrenal autoantibodies had chemical evidence of adrenal hormone deficiencies, which was made most obvious by finding elevations of ACTH levels in the mid- and late afternoon and renin levels in the early morning when the patients were recumbent. These adrenotrophic hormones were elevated in more than half of all patients with adrenocortical autoantibodies, in contrast to the normal levels, which were found in both diabetic and normal controls without adrenal antibodies (see Section VII). This situation is similar to the elevations of plasma TSH levels seen with incipient hypothyroidism in Hashimoto's disease (Riley *et al.*, 1981).

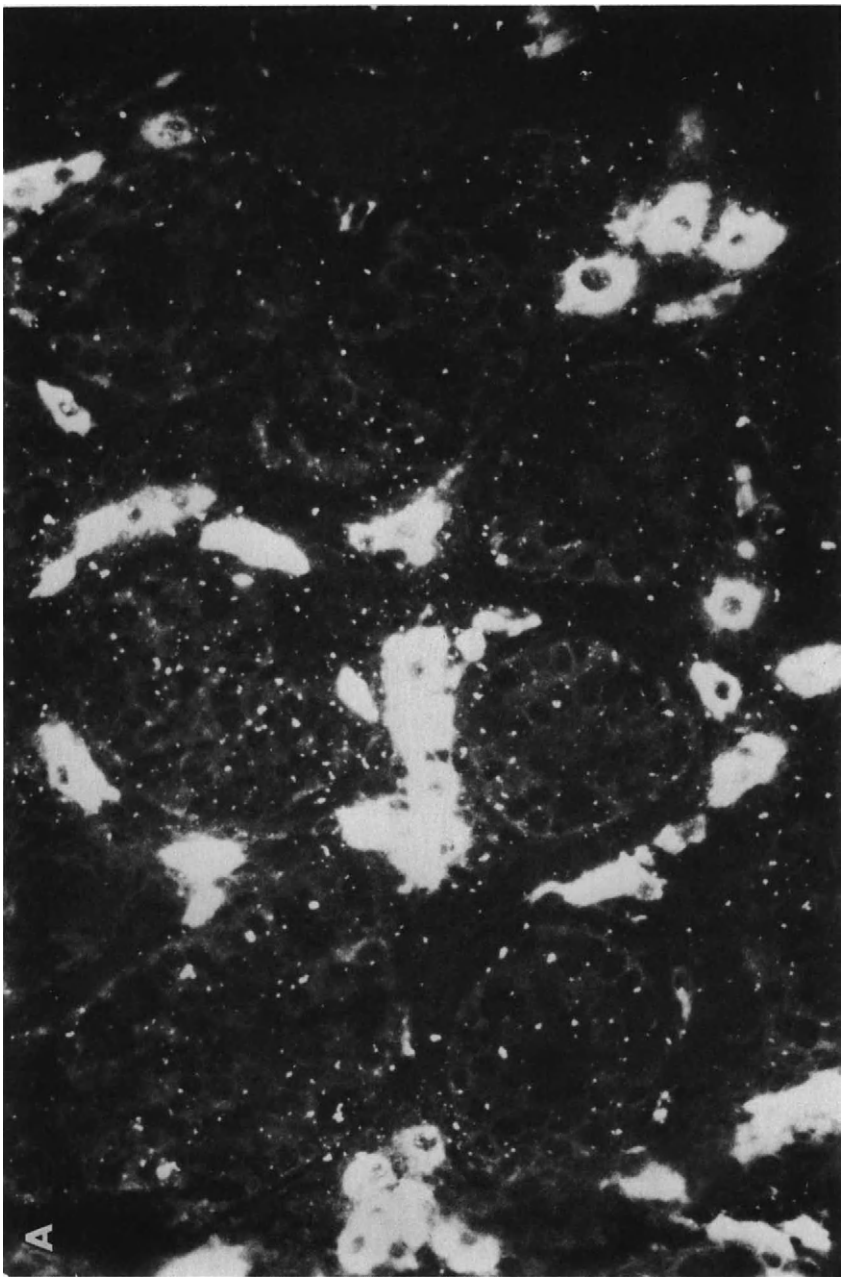
B. Autoantibodies to Steroid Hormone-Producing Cells

In 1968, Anderson described autoantibodies reacting with theca interna and corpus luteal cells of the ovary, interstitial (Leydig) cells of testis, placental trophoblasts, and cells of the adrenal cortex in two patients with gonadal insufficiency and Addison's disease (Anderson *et al.*, 1968). Irvine and colleagues found strong correlations between such autoantibodies and the presence of hypogonadism in patients with Addison's disease (Irvine and Barnes, 1975; Andrada *et al.*, 1968). The immediate problem to be resolved was the nature of the relationships between autoantibodies reacting to adrenal and/or other steroid hormone-producing cells.

Irvine argued that there were a variety of different autoantibodies reacting to several constituent antigens of steroid hormone producing cells in the gonads and placenta, but that not all of these autoantibodies cross-reacted with adrenocortical antigens (Irvine and Barnes, 1975; Andrada *et al.*, 1968; Irvine *et al.*, 1969). Ruehsen and colleagues agreed (1972), but wondered whether staining of the corpus luteum reported by Irvine was not merely a function of titer. In the more recent studies of Elder *et al.* (1981), a number of potential problems were identified. Using ovaries from rabbits as substrate, a sizable number of normal sera reacted in a clumped pattern by indirect immunofluorescence that could be considered nonspecific with respect to presence of adrenal autoimmunity. Similarly, nonspecific staining of a connective tissue antigen in the theca interna-granulosa layer of Graafian follicles of rabbit ovary, and around Leydig cells of the human testes, was fairly frequent and could presumably be removed by prior absorptions with rabbit tissue powders and erythrocytes. Twelve sera that unequivocally contained autoantibodies against gonadal tissues were found among 325 controls, 21 patients with Turner's syndrome, 505 insulin-dependent diabetics, and 37 patients with Addison's disease. All 12 sera also reacted to adrenal cortex and to other steroid-producing cells in varying degrees.

The best nonadrenal tissue in this study to identify steroid-producing cell autoantibodies was human testes from a patient with testicular feminization (Fig. 1). Adrenal-specific autoantibodies could not be absorbed out by repeated contact with placenta, ovary, or testes. However, titers of autoantibodies to steroid-producing cells, as determined against adrenal slices, were clearly diminished by a similar procedure, indicating the latter sera had a mixture of both adrenal-specific and adrenal-nonspecific (or steroid-producing cell) autoantibodies. Further, the frequencies of autoantibodies to steroid-producing cells were much greater and at higher titers in type I APS patients than in type II APS patients. These findings coincided with those described by Satsion and colleagues (1980). More recently, we found unusual sera that had autoantibodies that were positive only for steroid-producing cells of the gonad and not for those of the adrenal cortex, but this specificity has not yet been confirmed by absorption studies. In one pedigree in which three individuals in two generations had premature menopause, we also identified a common *A1-*, *B8-*, *DR3-*bearing *HLA* haplotype, steroidal autoantibodies, and absence of follicles on ovarian biopsies. Further studies are indicated for similar groups of patients.

In summary, adrenocortical autoantibodies usually react only to adrenal cortical cells, whereas steroid cell autoantibodies react ubiquitously


A

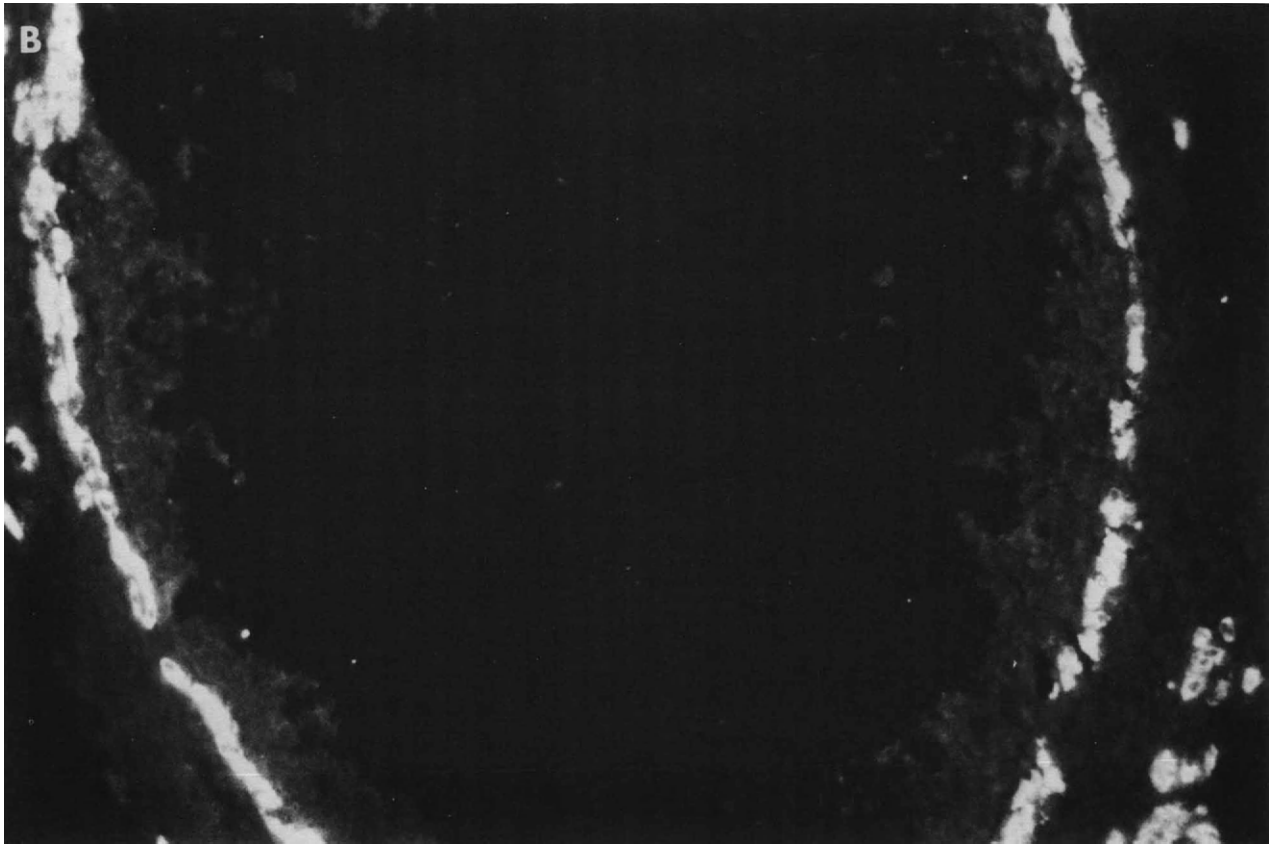


FIG. 1. Positive indirect immunofluorescent tests for steroid hormone cell autoantibodies using various cryocut unfixed tissue sections are shown. (A) Reactive Leydig cells of the testes on an infant with testicular feminization syndrome. (B) Reactive theca interna and granulosa layers of a Graafian follicle in pregnant rabbit ovary. (C) Reactive theca interna-granulosa layer of a Graafian follicle in human ovary. (D) Reactive syncytiotrophoblasts of human (blood group O) placenta. (E) Reactive human adrenal cortex. (Continued on following pages.)

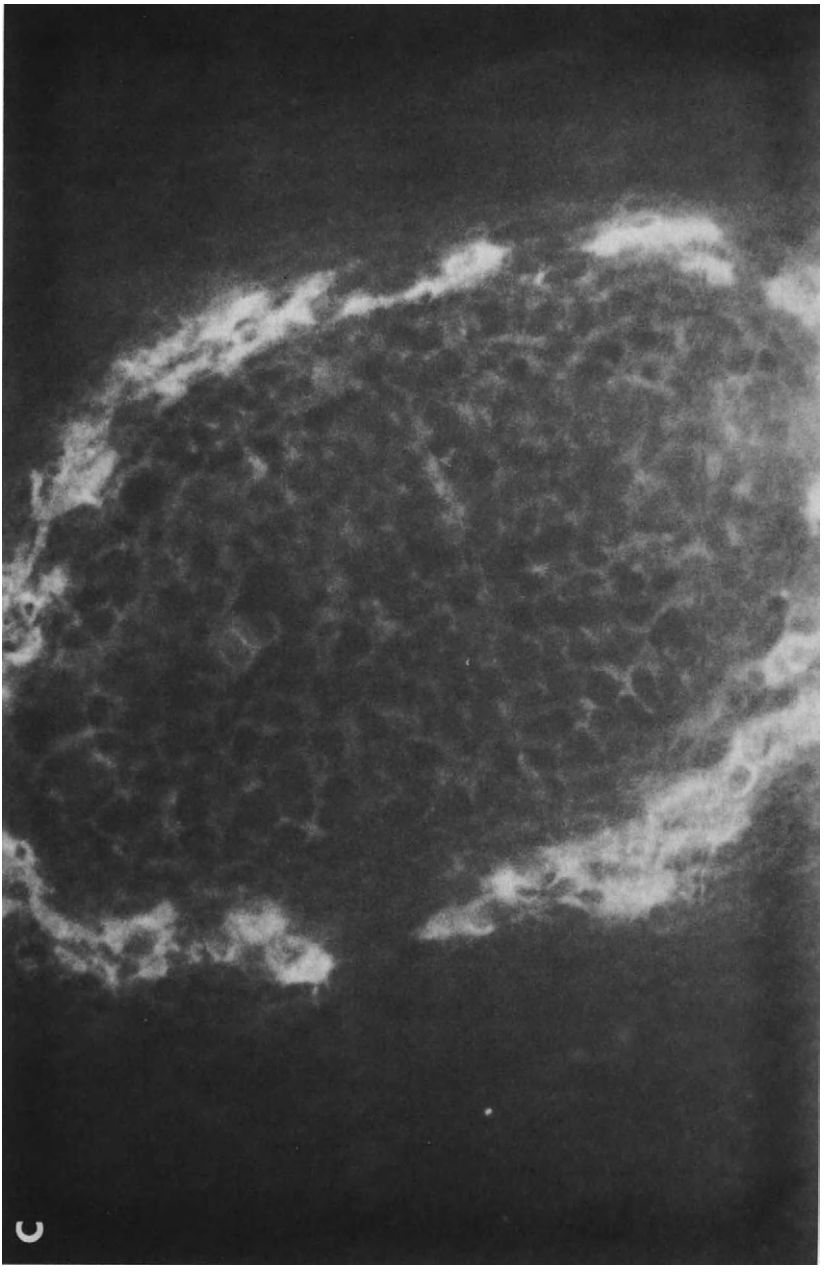


FIG. 1. (Continued)

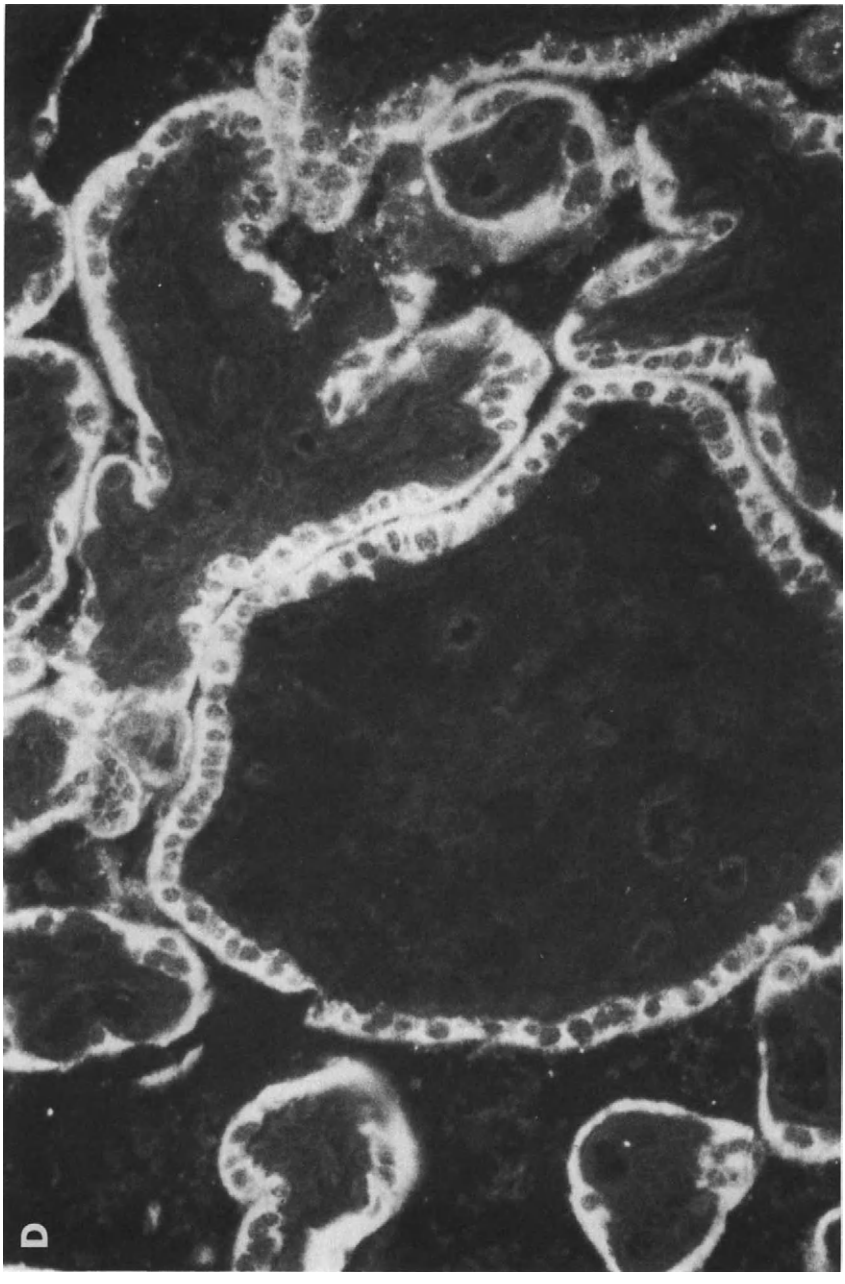
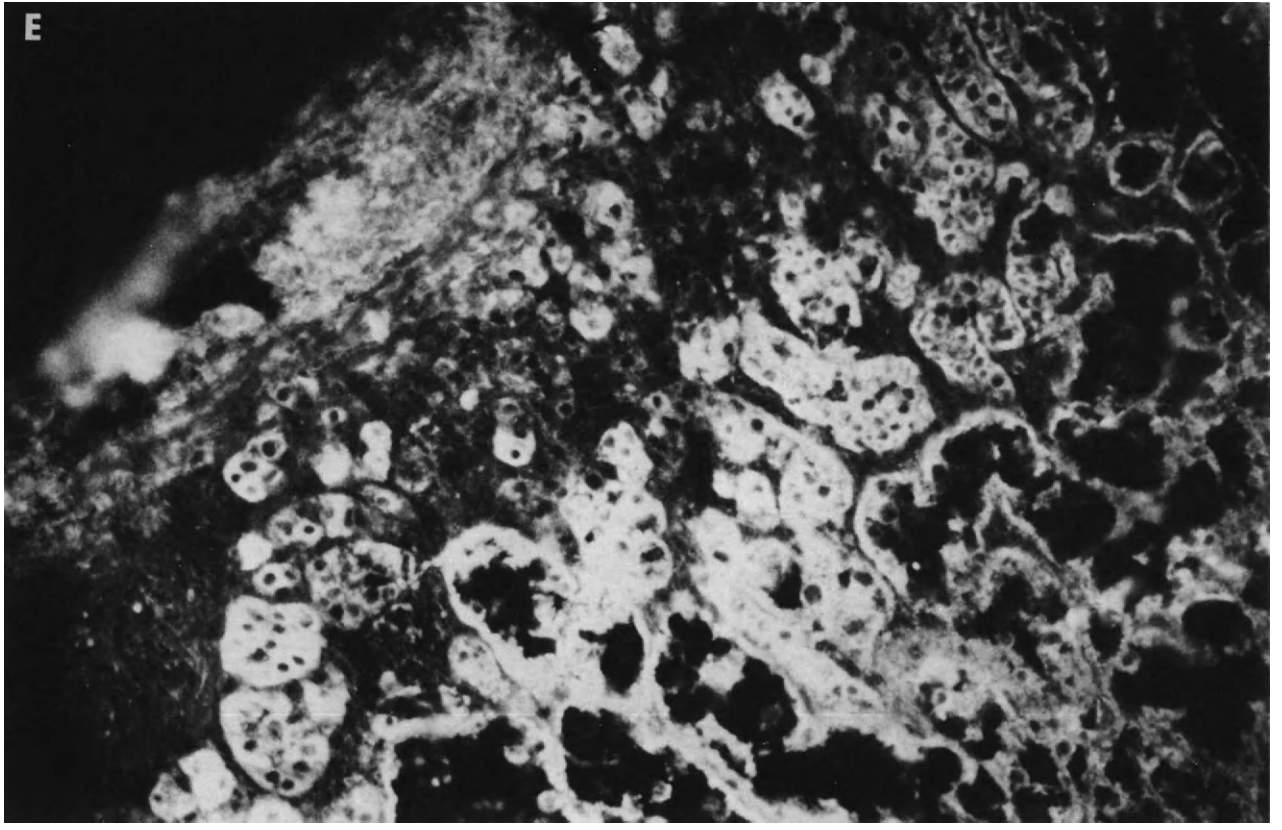


FIG. 1. (Continued)


E

with all such cell types in placenta, gonads, and adrenals, albeit such sera usually contain adrenal-specific autoantibodies as well.

As indicated by Satsion *et al.* (1980), human or monkey tissues are preferred substrates to minimize false-positive reactions when using the indirect immunofluorescent test to detect steroid cell autoantibodies, and blood group O human tissues may be the safest to use routinely if available. These investigators also indicated that autoantibodies reacting with ova may coexist with steroid hormone cell autoantibodies, although they do not cross-react to related autoantigens.

V. CELL-MEDIATED IMMUNITY TO ADRENAL GLANDS IN ADDISON'S DISEASE

Adrenal autoantibodies are immunoglobulins that react to cytoplasmic antigens of adrenocortical cells. They cannot, therefore, be involved in the pathogenic process unless the same antigens also are exposed on the surface of adrenal cells. Whereas such antibodies are of the IgG immunoglobulin class and readily cross the placenta, no confirmed instance of "congenital" autoimmune adrenal failure in a newborn who was exposed to maternal adrenocortical autoantibodies *in utero* has been described. Notwithstanding, McNatty has demonstrated complement-fixing steroid cell autoantibodies that proved to be cytotoxic to human granulosa cells growing in tissue culture *in vitro*. Sera showing clumped staining within the corpus luteal cells appeared to be the most toxic (Irvine *et al.*, 1969; McNatty *et al.*, 1975).

In 1969, Nerup and Bendixen found evidence for delayed hypersensitivity to adrenal cortex in patients with "idiopathic" Addison's disease. These workers used in their methods human fetal adrenal extracts as the exciting antigen with *in vitro* leucocyte migration inhibition and in intracutaneous skin testing (Nerup and Bendixen, 1969). Blast transformation with patients' leukocytes in culture in the presence of adrenal extracts could not be identified. However, mitochondria from human fetal adrenal, but not from human fetal liver, elicited leukocyte migration inhibition while some patients had positive skin tests to this antigen (Nerup *et al.*, 1970).

Like other organ-specific autoimmunities, definite studies to prove a cell-mediated mechanism for loss of adrenocortical cells in patients with autoimmune Addison's disease are difficult to perform because of limited availability of suitable fresh tissues and perhaps requirements such as *HLA* compatibilities between T cells (patients') and target adrenal cells in culture. The T-cell defect of type I APS has been discussed in Section III.

VI. GENETICS OF AUTOIMMUNE ADDISON'S DISEASE

A. Type I APS (Blizzard's Syndrome)

This syndrome is relatively rare, and most reports in the literature concern small numbers of patients. In assembling available data, however, Neufeld *et al.* (1981) could identify no clear *HLA* associations with type I APS. This lack of association with *HLA-B8* in particular was also found by Eisenbarth, being present in only 3 of 21 patients with candidiasis and hypoparathyroidism (Arulanantham *et al.*, 1979; Eisenbarth *et al.*, 1979; Eisenbarth and Jackson, 1981). Furthermore, in the few informative pedigrees studied to date, there were no associations between affected family members and their *HLA* haplotypes (Arulanantham *et al.*, 1979; Eisenbarth *et al.*, 1979; Brueton *et al.*, 1980). Spinner *et al.* (1969) previously had shown the disease to be inherited as a recessive characteristic; however, single sporadic cases are often encountered.

Thus, autoimmune Addison's disease in type I APS seems distinctive in its disease and serological associations, as well as in its lack of associations with *HLA* genes. These findings are in contrast to those in Addison's disease associated with type II APS. We conclude that two separate genetic systems involving different pathogenic mechanism probably are responsible for autoimmune adrenocortical insufficiency.

B. Type II APS (Schmidt–Carpenter's Syndrome)

Most of the component diseases of this syndrome complex are strongly and primarily associated with *HLA-DR3/Dw3* and secondarily with *HLA-B8* through linkage disequilibrium. These diseases are Addison's disease (Riley *et al.*, 1980; Thomsen *et al.*, 1975), insulin-dependent diabetes (Riley *et al.*, 1980; Thomsen *et al.*, 1975; also review by Cudworth and Wolff, 1982), Grave's disease (Farid, 1981), myasthenia gravis (Naeim *et al.*, 1978), and atrophic thyroiditis/hypothyroidism (Moens *et al.*, 1979). In contrast, pernicious anemia and goitrous thyroiditis/hypothyroidism are only weakly associated with *HLA-DR3*, and weakly if at all with *HLA-B8*, and only then in the presence of other associated autoimmunities that are themselves associated with *HLA-B8* and *-DR3* (Eisenbarth and Jackson, 1981). In an important study, Gorsuch and colleagues found that thyrogastric autoantibodies and thyrogastric autoimmune diseases that occurred in pedigrees with insulin-dependent diabetes did not segregate with *HLA* haplotypes as was the case for diabetes itself, suggesting that

the diabetes and the thyrogastric autoimmunity have different genetic determinants (Gorsuch *et al.*, 1980). Riley *et al.* (1982a) reported similar findings. Assuming then that thyroid and gastric autoimmunity (in association) is inherited by non-*HLA*-associated genes, which are separate from those for insulin-dependent diabetes, we need to explain their increased frequencies in patients with insulin-dependent diabetes or autoimmune Addison's disease. We hypothesize that there are positive and separate genetic interactions taking place between these separate autoimmune systems. It also seems probably that there are multiple genetic influences impacting on the disease spectra and the clinical outcomes seen. Thus, to speculate that the pattern of inheritance in type II APS is dominant or recessive appears to be too simplistic an exercise. The occurrence of multiple autoantibodies within pedigrees with type II APS, however, is seen in each vertical generation as a dominant trait, while in many multiplex pedigrees, *HLA*-associated genes occurring singly or in double doses may be operative (Maclaren *et al.*, 1981) in the outcome of insulin-dependent diabetes.

VII. ADRENAL AND OTHER EXTRAPANCREATIC AUTOIMMUNITIES IN INSULIN-DEPENDENT DIABETES

Insulin-dependent diabetes often occurs as part of an autoimmune polyglandular syndrome. Thyroid microsomal, gastric parietal cell, and adrenocortical autoantibodies occur at increased frequencies with similar increases in the diseases associated with these autoantibodies (Neufeld *et al.*, 1980b) at some 4 to 8 times the normal background rates (see Table II). These autoantibodies occur more commonly in females, and with respect to pancreatic islet cell and thyroid microsomal autoantibodies, more commonly in the Caucasian population than among Blacks. The frequencies of autoantibodies seen in patients with insulin-dependent diabetic patients, in fact, resemble those seen in the geriatric general population (N. Maclaren and W. Riley, 1985). Furthermore, there are increased frequencies of thyrogastric autoantibodies among the family members of a diabetic proband who also has such antibodies (Riley *et al.*, 1980; Neufeld *et al.*, 1980b; Nissley *et al.*, 1973).

These observations, coupled with the probable disassociation of *HLA*-related diabetetogenic genes from those for thyrogastric autoimmunities (Gorsuch *et al.*, 1980; Riley *et al.*, 1982a) imply that these genetic systems interact when they coincide in individual patients. Riley was able to demonstrate, in cross-sectional and longitudinal studies, that thyrogastric au-

TABLE III

Adrenocortical Autoantibodies and Addison's Disease in Children and Young Adults with Insulin-Dependent Diabetes (IDD)[a,b]

	n	Adrenocortical autoantibodies (%)		Addison's disease (%)	
Controls	200	(1)	0.5	—	
IDD + ICA	466	(7)	1.5	(2)	0.4
IDD + PCA	38	(1)	2.6	(1)	2.6
IDD + TMA	77	(5)	6.5	(2)	2.6

[a] After Riley *et al.* (1980). The table shows the value of screening patients with thyroid antibodies for adrenal antibodies.

[b] Abbreviations: ICA, islet cell autoantibody; PCA, gastric (parietal) cell autoantibody; TMA, thyroid microsomal autoantibody.

toantibodies appeared in patients with insulin-dependent diabetes in the great majority of instances by the time of onset of clinical diabetes, albeit the resulting clinical diseases occurred with different natural histories (Riley *et al.*, in press, 1983). This situation is also seen in the BB rat (Elder and Maclaren, 1983; Maclaren and Elder, 1983; Elder *et al.*, 1981, 1982).

Thus Graves' disease tended to appear near the time of onset of diabetes, while hypothyroidism occurred progressively after diagnosis of diabetes (Riley *et al.*, 1981), as did gastric achlorhydria in those with gastric parietal autoantibodies. Pernicious anemia, however, only appeared in mid- to late life (Riley *et al.*, 1982c), often years after gastric atrophy had occurred. We have strongly argued, on the bases of these results, that all patients with insulin-dependent diabetes should be screened at least once for thyroid microsomal autoantibodies. For those individuals with positive tests, thyroid monitoring should follow, with yearly plasma TSH and perhaps free T_4 (Riley *et al.*, 1981, 1982b).

In our previous studies, we demonstrated that the incidence of adrenocortical autoantibodies was ~2% in Caucasian children with insulin-dependent diabetes and that overt Addison's disease occurred in ~1 of 250 patients, albeit many patients with adrenal antibodies who do not have overt adrenal insufficiency will have biochemical evidence for hypoadrenocorticalism, that is, raised basal levels of plasma ACTH and renin (C. Ketchum, W. Riley, and N. Maclaren, 1984). Although the occurrence of Addison's disease in a patient with insulin-dependent diabetes is

a life-threatening situation and one that easily can be clinically misdiagnosed, the cost effectiveness of screening all insulin-dependent diabetic patients for adrenal antibodies is less convincing than is the case for screening patients with coincident thyroid autoimmune disease. In our experience, most (80%) diabetic patients with adrenal autoimmunities will have coincident thyroid autoimmunity (positive microsomal autoantibodies). We therefore recommend testing the patients with both insulin-dependent diabetes and thyroid autoimmune disease for adrenal autoantibodies. We also have found remarkably high frequencies of *HLA-B8/DR3* among diabetic patients with adrenal autoantibodies (Riley *et al.*, 1980). This was substantiated by Eisenbarth (Eisenbarth and Jackson, 1981).

VIII. CONCLUSION

The importance of autoimmune Addison's disease to the physician may not be in its occurrence as a common disease, but in the dramatic and yet treatable nature of the clinical illness and in its frequent associations with other organ-specific autoimmunities. We proposed that the autoimmune polyendocrine syndromes that include Addison's disease are divisible into two types, with different genetics, unique serological findings, and distinct associated diseases. Recognition of the latter two polyglandular autoimmunities, which include Addison's disease, has proved useful in anticipating additional problems that affected patients may incur. The early identification of "idiopathic" Addison's disease as a multisystem disorder has prompted widespread recognition that all autoimmune endocrinopathies are potential components of autoimmune polyglandular syndromes. This recognition has allowed for associated entities to be identified and treated in both the patient and his family by use of autoantibody screening. This is true despite the basic ignorance about the pathogenic mechanisms of these disorders. Increased awareness of the possibility of autoimmune adrenal insufficiency in association with other diseases is important to obviate needless consequences of these processes.

Dr. Addison's statement in 1855, regarding adrenocortical insufficiency alone, is equally important regarding polyglandular disease in 1983. Addison states:

> My experience, though necessarily limited, leads to a belief that the disease is by no means of rare occurrence, and that were we better acquainted with its symptoms and progress, we should probably succeed in detecting many cases, which in the present state of our knowledge, may be entirely overlooked or misunderstood; and, I think, I may with some confidence affirm, that although partial disease of the capsules may give rise to symptoms, and to a condition of the general system, extremely equivocal and inconclusive, yet that a more extensive

lesion will be found to produce a state, which may not only create a suspicion, but be pronounced with some confidence to arise from the lesion in question. When the lesion is acute and rapid, I believe the anemia, prostration, and peculiar condition of the skin will present a corresponding character, and that whether acute or chronic, provided the lesion involve the entire structure of both organs, death will inevitably be the consequence [Addison, 1855].

REFERENCES

Addison, T. (1849). *London Med. Gaz.* **43,** 517–518.

Addison, T. (1855). VIII. *11* pl. tol., 43. Samuel Highley, London.

Anderson, J., Goudie, R., Gray, K., and Timbury, G. (1957). *Lancet* **1,** 1123–1125.

Anderson, J., Goudie, R., Gray, K., and Stuart-Smith, D. (1968). *Clin. Exp. Immunol.* **3,** 107–112.

Andrada, J., Bigazzi, P., Andrada, E., Milgrom, F., and Witebsky, E. (1968). *JAMA, J. Am. Med. Assoc.* **206,** 1535–1541.

Andres, J., Maclaren, N., and Barrett, D. (1983). Personal observation.

Appel, G., and Holub, D. (1976). *Am. J. Med.* **61,** 129–133.

Arulanantham, K., Dwyer, M., and Genel, M. (1979). *N. Engl. J. Med.* **300,** 164–168.

Ayala, A., Canales, E., Karchmer, S., Alarcon, D., and Zarate, A. (1979). *Obstet. Gynecol. (N. Y.)* **53,** 98–101(S).

Beaven, D., Nelson, D., Renold, A., and Thorn, G. W. (1959). *N. Engl. J. Med.* **261,** 443–454.

Betterle, C., Peserico, A., and Bersani, G. (1979). *Arch. Dermatol.* **155,** 364–368.

Blizzard, R., and Gibbs, J. (1968). *Pediatrics* **42,** 231–237.

Blizzard, R., and Kyle, M. (1963). *J. Clin. Invest.* **42,** 1653–1657.

Blizzard, R., Chee, D., and Davis, W. (1966). *Clin. Exp. Immunol.* **1,** 119–128.

Block, M., Pachman, L., Windhorst, D., and Goldfine, I. (1971). *Am. J. Med. Sci.* **261,** 213–218.

Bor, S., Feiwel, M., and Chanarin, I. (1969). *Br. J. Dermatol.* **81,** 83–88.

Bottazzo, G., and Doniach, D. (1978). *J. R. Soc. Med.* **71,** 433–436.

Bottazzo, G., Florin-Christensen, A., and Doniach, D. (1974). *Lancet* **2,** 1279–1283.

Bottazzo, G., Pouplard, A., Florin-Christensen, A., and Doniach, D. (1975). *Lancet* **2,** 97–101.

Brenner, O. (1928). *Qt. J. Med.* **22,** 121.

Bronsky, D., Kushner, D., Dubin, A., and Snapper, M. (1958). *Medicine (Baltimore)* **37,** 317–321.

Brostoff, J., Bor, S., and Feiwel, M. (1969). *Lancet* **2,** 177–178.

Brueton, M., Chapel, H., and MacKintosh, L. (1980). *Tissue Antigens* **15,** 101–103.

Carpenter, C., Solomon, N., Silverberg, S., Bledsoe, T., Northcutt, R., Klinenberg, J., Bennett, Z., and Harvey, A. (1964). *Medicine (Baltimore)* **43,** 153–180.

Castells, S., Fikrig, S., Inamdar, S., and Orti, E. (1971). *J. Pediatr. (St. Louis)* **79,** 72–79.

Christy, N., Holub, D., and Tomasi, T. (1962). *J. Clin. Endocrinol.* **22,** 155–159.

Collen, R., Lippe, B., and Kaplan, S. (1979). *Am. J. Dis. Child.* **133,** 598–600.

Cooper, B., Holmes, G., and Cooke, W. (1978). *Br. Med. J.* **1,** 537–549.

Craig, J., Schiff, L., and Boone, J. (1955). *Am. J. Dis. Child.* **89,** 669–684.

Crispell, K., and Parson W. (1952). *J. Clin. Endocrinol.* **12,** 881.

Cudworth, A., and Wolff, E. (1982). *Clin. Endocrinol. Metab.* **11**(2), 389–408.

Cunliffe, W., Hall, R., Newell, D., and Stevenson, C. (1968a). *Br. J. Dermatol.* **80,** 135–139.
Cunliffe, W., Hall, R., Stevenson, C., and Weightman, D. (1968b). *Br. J. Dermatol.* **81,** 877–881.
Dawber, R. (1970). *Postgrad. Med.* **46,** 276–277.
Doniach, D., and Roitt, I. (1957). *J. Clin. Endocrinol. Metab.* **17,** 1293–1304.
Doniach, D., Roitt, I., and Taylor, K. (1963). *Br. Med. J.* **1,** 1374–1379.
Drury, M., Keelan, D., Timoney, F., and Irvine, W. (1970). *Clin. Exp. Immunol.* **7,** 125–132.
Dunlap, D. (1963). *Br. Med. J.* **2,** 887–891.
Edmonds, M., Lamki, L., Killinger, D., and Volpé, E. (1973). *Am. J. Med.* **54,** 782–787.
Ehligen, R. (1969). *Lancet* **2,** 1286–1288.
Eisenbarth, G., and Jackson, R. (1981). *In* "HLA in Endocrine and Metabolic Disorders" (N. R. Farid, ed.), pp. 235–264. Academic Press, New York.
Eisenbarth, G., Wilson, P., Ward, F., and Lebovitz, H. (1978). *N. Engl. J. Med.* **298,** 92–94.
Eisenbarth, G., Wilson, P., Ward, F., Buckley, C., and Lebovitz, H. (1979). *Ann. Intern. Med.* **91,** 528–533.
Elder, M., and Maclaren, N. (1983). *J. Immunol.* **130**(4), 1723–1731.
Elder, M., Maclaren, N., and Riley, W. (1981). *J. Clin. Endocrinol. Metab.* **52,** 1137–1142.
Elder, M., Maclaren, N., Riley, W., and McConnell, T. (1982). *Diabetes* **31,** 313–318.
Farid, N. R. (1981). *In* "HLA in Endocrine and Metabolic Disorders" (N. R. Farid, ed.), pp. 85–143. Academic Press, New York.
Forcier, R., McIntyre, O., Frey, W., Andrada, J., and Strieff, R. (1972). *Arch. Intern. Med.* **129,** 638–641.
Frey, H., Vogt, J., and Nerup, J. (1973). *Acta Endocrinol. (Copenhagen)* **72,** 401–416.
Fritze, D., Hermann, C., Jr., Naeim, F., Smith, G., Zeller, E., and Walford, R. (1976). *Ann. N. Y. Acad. Sci.* **274,** 440–450.
Gastineau, C., Meyers, W., Arnold, J., and McConahey, W. (1964). *Mayo Clin. Proc.* **39,** 939.
Gharib, H., Hodgson, S., Gastineau, C., Sholz, D., and Smith, L. (1972). *Lancet* **II,** 734–736.
Gleason, T., Stebbins, P., and Shanahan, M. (1978). *Arch. Pathol. Lab. Med.* **102,** 46–48.
Golonka, J., and Goodman, A. (1968). *J. Clin. Endocrinol.* **28,** 79–83.
Goudie, R., and Pinkerton, P. (1962). *J. Pathol. Bacteriol.* **83,** 584–585.
Goudie, R., Gray, K., and Whyte, W. (1966). *Lancet* **1,** 1173–1174.
Goudie, R., McDonald, E., Anderson, J., and Gray, K. (1968). *Clin. Exp. Immunol.* **3,** 119–123.
Gorsuch, A., Dean, B., Bottazzo, G., Lister, J., and Cudworth, A. (1980). *Br. Med. J.* **1,** 145–147.
Gorsuch, A., Spencer, K., Lister, J., McNally, J., Dean, B., Bottazzo, G., and Cudworth, A. (1981). *Lancet* **2,** 1363–1365.
Grunnet, I., Howitz, J., Reymann, F., and Schwartz, M. (1970). *Arch. Dermatol.* **101,** 82–85.
Harmans, P., Ulrich, J., and Markowitz, H. (1969). *Clin. Stnd.* **47,** 503–519.
Hertz, K., Gazze, L., Kirkpatrick, C., and Katz, S. (1977). *N. Engl. J. Med.* **297,** 634–637.
Howitz, J., and Schwartz, M. (1971). *Lancet* **1,** 1331–1335.
Hung, W., Migeon, C., and Parrott, R. (1963). *N. Engl. J. Med.* **269,** 658–663.
Irvine, W. (1978). *In* "Genetic Control of Autoimmune Disease" (N. Rose, P. Bigazzi, and N. Warner, eds.), pp. 77–110. Elsevier, Amsterdam.
Irvine, W., and Barnes, E. (1974). *J. Reprod. Fertil.* **21,** 1–31.
Irvine, W., and Barnes E. (1975). *Clin. Endocrinol. Metab.* **4**(2), 379–434.

Irvine, W., Stewart, A., and Scarth, L. (1967). *Clin. Exp. Immunol.* **2,** 31–39.
Irvine, W., Chan, M., Scarth, L., Kolb, F., Hartog, M., Bayliss, R., and Drury, M. (1968). *Lancet* **2,** 883–887.
Irvine, W., Chan, M., and Scarth, L. (1969). *Clin. Exp. Immunol.* **4,** 489–493.
Kane, C. and Weed, L. (1950). *N. Engl. J. Med.* **243,** 939–944.
Karlish, A., and McGregor, G. (1970). *Lancet* **II,** 330–332.
Kenny, F., and Holliday, M. (1964). *N. Engl. J. Med.* **271,** 708–713.
Kern, F., Hoffman, W., Hambrick, G., and Blizzard, R. (1973). *Arch. Dermatol.* **107,** 407–412.
Ketchum, C., Riley, W., and Maclaren, N. (1984). *J. Clin. Endocrinol. Metab.* **58,** 1–5.
Kier, W., and Nørgaard, R. (1969). *Acta Pathol. Microbiol. Scand.* **76,** 229–233.
Kirkpatrick, C. (1979). *In* "Immune Regulators in Transfer Factor" (A. Khan, C. Kirkpatrick, and N. Hill, eds.), pp. 547–562. Academic Press, New York.
Ludwig, H., and Schernthaner, G. (1978). *Wein. Klin. Wochenschr.* **90,** 736–741.
McGregor, B., Katz, H., and Doe, R. (1972). *JAMA, J. Am. Med. Assoc.* **218,** 724–725.
McHardy-Young, S., Lessof, M., and Maisey, M. (1972). *Clin. Endocrinol.* **1,** 45–46.
Maclaren, N., and Elder, M. (1983). *Metabolism.* 32 (Suppl. 1), 92–96, 1983.
Maclaren, N., and Riley, W. (1985). *Diabetes Care* (In press).
Maclaren, N., Rosenbloom, A., McLaughlin, J., Bruck, E., and Lezotte, D. (1981). *IRCS Med. Sci: Libr. Compend.* **9,** 631–632.
McNatty, K., Short, R., Barnes, E., and Irvine, W. (1975). *Clin. Exp. Immunol.* **22,** 378–384.
Males, J., Spitler, A., and Townsend, J. (1971). *J. Okla. State Med. Assoc.* **64,** 298–302.
Matz, R., and Tucker, S. (1969). *Metabolism* **16,** 522–525.
Meecham, J., and Jones, E. (1967). *Lancet* **1,** 535.
Moens, H., Barnard, J., Bear, J., and Farid, N. (1979). *Tissue Antigens* **13,** 342–348.
Morse, W., Cochrane, W., and Landrigan, P. (1961). *N. Engl. J. Med.* **264,** 1021–1026.
Muller, S., and Winkleman, R. (1963). *Arch. Dermatol.* **88,** 290–297.
Myllarniemi, S., and Perheentupa, J. (1978). *Oral Surg., Oral Med., Oral Pathol.* **45,** 721–729.
Naeim, F., Keesey, J., Herrmann, C., Jr., Lindstrom, J., Zeller, E., and Walford, R. (1978). *Tissue Antigens* **12,** 381–386.
Nakhooda, A., Like, A., Chappel, C., Murray, F., and Marliss, E. (1977). *Diabetes* **26,** 100–112.
Nerup, J. (1974a). *Acta Endocrinol. (Copenhagen)* **76,** 127–131.
Nerup, J. (1974b). *Acta Endocrinol. (Copenhagen)* **76,** 142–149.
Nerup, J. (1974c). *In* "Immunity and Autoimmunity in Diabetes Mellitus" (P. Bastenie and W. Gepts, eds.), pp. 149–152. Excerpta Medica, Amsterdam.
Nerup, J., and Bendixen, G. (1969). *Clin. Exp. Immunol.* **5,** 355–359.
Nerup, J., and Biner, C. (1973). *Acta Endocrinol. (Copenhagen)* **72,** 279–286.
Nerup, J., Anderson, V., and Bendixen, G. (1970). *Clin. Exp. Immunol.* **6,** 733–738.
Neufeld, M., Blizzard, R., and Maclaren, N. (1979). *Clin. Res.* **27,** 812A.
Neufeld, M., Maclaren, N., and Blizzard, R. (1980a). *Pediatr. Ann.* **9,** 43–53.
Neufeld, M., Maclaren, N., Riley, W., McLaughlin, J., Silverstein, J., and Rosenbloom, A. (1980b). *Diabetes* **29,** 589–593.
Neufeld, M., Maclaren, N., and Blizzard, R. (1981). *Medicine (Baltimore)* **60,** 355–362.
Nissley, S., Drash, A., Blizzard, R., Sperling, M., and Childs, B. (1973). *Diabetes* **22,** 63–65.
Pearlman, L. (1961). *Ann. Intern. Med.* **55,** 796–799.
Perheentupa, J., and Hickkala, H. (1973). *Acta Paediatr. Scand.* **62,** 110–111.

Riley, W., Maclaren, N., and Neufeld, M. (1980). *J. Pediatr. (St. Louis)* **97,** 191–195.
Riley, W., Maclaren, N., Lezotte, D., Spillar, R., and Rosenbloom, A. (1981). *J. Pediatr. (St. Louis)* **98,** 350–354.
Riley, W., Glancey, L., and Phinney, R. (1982a). *Diabetes* **31,** (Suppl. 2), 45A.
Riley, W., Maclaren, N., and Rosenbloom, A. (1982b). *Lancet* **2,** 489–490.
Riley, W., Toskes, P., and Maclaren, N. (1982c). *Diabetes* **31,** 1051–1055.
Riley, W., Winer, A., and Goldstein, D. (1983). *Diabetologia* **24,** 418–421.
Rountree, L., and Snell, A. (1931). *JAMA, J. Am. Med. Assoc.* **96,** 231; **97,** 1446.
Ruehsen, M., de M., Blizzard, R., Garcia-Bunnuel, R., and Jones, G. (1972). *Am. J. Obstetr. Gynecol.* **112,** 693–703.
Rupp, J., and Paschkis, K. (1953). *Ann. Intern. Med.* **39,** 1103–1106.
Satsion, F., Bottozzo, G., and Doniach, D. (1980). *Clin. Exp. Immunol.* **39,** 97–111.
Schmidt, M. B. (1926). *Verh. Dtsch. Ges. Pathol.* **21,** 212–221.
Solomon, N., Carpenter, C., Bennett, I., and Harvey, A. (1965). *Diabetes* **14,** 300–304.
Spinner, M., Blizzard, R., and Childs, B. (1968). *J. Clin. Endocrinol. Metab.* **28,** 795–804.
Spinner, M., Blizzard, R., Gibbs, J., Abbey, H., and Childs, B. (1969). *Clin. Exp. Immunol.* **5,** 461–465.
Stankler, L., and Bensher, P. (1972). *Br. J. Dermatol.* **86,** 238–245.
Strickland, R. (1969). *Ann. Intern. Med.* **70,** 1001–1005.
Swingle, W., and Pfiffner, J. (1930). *Science* **71,** 321–322.
Thomsen, M., Platz, P., Anderson, O., Christy, M., Lyngsoe, J., Nerup, J., Rasmussen, K., Ryder, L., Staub, J. H., Nielson, J., and Svejgaard, A. (1975). *Transplant. Rev.* **22,** 125–147.
Thorpe, E., and Handley, H. (1929). *Am. J. Dis. Child.* **38,** 328–338.
Tomar, R., Rao, R., Lawrence, A., and Moses, A. (1979). *Ann. Allergy* **42,** 241–245.
Turkington, R., and Lebovitz, H. (1967). *Am. J. Med.* **43,** 499–507.
Van Thiel, D., Smith, W., Rabin, B., Fisher, S., and Lester, R. (1977). *Ann. Intern. Med.* **86,** 10–19.
Vazquez, A., and Kenny, M. (1973). *Obstet. Gynecol. (N. Y.)* **41,** 414–418.
Wehrmacher, W. (1961). *Arch. Intern. Med.* **108,** 114.
Weinberg, U., Kraemer, F., and Kammerman, S. (1976). *Am. J. Med. Sci.* **272,** 200–215.
Wells, R., Higgs, J., McDonald, A., Valdimarsson, H., and Holt, P. (1972). *J. Med. Genet* **9,** 302–310.
Whitaker, J., Landing, B., Esselborn, V., and Williams, R. (1956). *J. Clin. Endocrinol. Metab.* **16,** 1374–1387.
Wilks, S. (1862). *Addisonii Guy's Hosp. Rep.* **8,** 1.
Wuepper, K., and Fudenberg, H. (1967). *Clin. Exp. Immunol.* **2,** 71–82.
Zinkernagal, R. (1977). *In* "Autoimmunity" (N. Talal, ed.), pp. 363–381. Academic Press, New York.

CHAPTER 9

Immunology of Insulin-Dependent Diabetes

DEBORAH DONIACH
Department of Immunology
The Middlesex Hospital Medical School
London, England

KATE M. SPENCER
Department of Diabetes and Immunology
St. Batholomew's Hospital
London, England

GIAN FRANCO BOTTAZZO
Department of Immunology
The Middlesex Hospital Medical School
London, England

I. INTRODUCTION

Ideas on the role of autoimmunity in the causation of diabetes are in a state of flux, mainly because of prospective immunological studies and new cloning techniques. Investigators are also becoming aware of the

possible participation of other endocrine glands in the final metabolic status, and new animal models resembling the human disease are helping to define pathogenetic mechanisms. A few years ago the division into type I or insulin-dependent (IDDM) and type II or non-insulin-dependent diabetes (NIDDM) helped to overcome the problem of age at onset, when it was realized that IDDM could occur at any age and that at least two genetic subspecies of NIDDM existed in young people (maturity-onset diabetes of the young, MODY).

In considering IDDM itself, it was thought for some time that polyendocrine or autoimmune IDDM could be clearly distinguished from the more common juvenile onset cases (Bottazzo *et al.*, 1981). The prospective studies in unselected families of diabetic children revealed that many of the characteristics of endocrine patients with IDDM, especially the prolonged latency period, also applied to the relatives of these children. Observations in these predisposed individuals and in adult endocrine patients found to have islet-cell antibodies by systematic autoimmune screening now seem to indicate that within these groups with serological markers of autoimmune insulitis, a proportion of individuals might not necessarily become insulin-dependent within a lifetime. It is now realized that even when the insulin-secreting cells are destroyed by chemical poisons (Karam *et al.*, 1980; Helgason and Jonasson, 1981) similar in their action to streptozotocin used for experimental diabetes or by well-defined virus infections (Rayfield and Yoon, 1981), autoimmunity still plays an important accessory role. Finally, new concepts related to idiotype networks may affect the treatment of IDDM patients in future generations.

II. HUMORAL IMMUNITY IN DIABETICS AND PREDISPOSED INDIVIDUALS

Islet-cell antibodies are organ specific for endocrine pancreas; they cross-react with other species, and they are of IgG class. Several antigens including some that are only present in β cells are now envisaged in what is basically a polyclonal autoimmune response with subclass restrictions (Dean *et al.*, 1983). The antibodies may be complement fixing (CF-ICA), and this is a separate variant (Bottazzo *et al.*, 1980). The more common islet-cell antibodies are not able to fix complement. As the islet-cell antigens have not yet been isolated and are at present only partially characterized (Baekkeskov *et al.*, 1982), cytoplasmic ICA are normally detected by standard immunofluorescence (IFL). Unfixed group O human pancreas is still the substrate of choice (Bottazzo *et al.*, 1974). Fixed pancreas substrates give controversial readings, especially in complement-fixation im-

munofluorescence tests with anti-C3 conjugates used to detect CF-ICA (Dean *et al.*, 1982).

In the screening test with anti-IgG, ICA reacts with all four endocrine cells (Bottazzo and Doniach, 1978). This shared autoantigen is not represented on the cell surface and therefore cannot be in direct contact with sensitized lymphocytes in the living gland. Since circulating antibodies exert cytotoxic effects through complement-dependent mechanisms, extensive parallel tests for ICA and CF-ICA revealed that in all diabetic groups tested, only half the islet-cell antibodies fixed complement (Bottazzo *et al.*, 1980). Most significantly, this was also found in genetically predisposed first-degree relatives and in endocrine patients who later became diabetic. This led to the conclusion that CF-ICA reactions include the β-cell-specific autoantibodies that are the most relevant markers for ongoing insulitis. We now know that some CF-ICA selectively stain β cells (Bottazzo *et al.*, 1982). We also found some sera containing complement-fixing variants specific for glucagon or somatostatin cells. These single-cell antibodies were known to exist in some diabetic sera (Bottazzo and Lendrum, 1976).

Islet-cell surface antibodies (ICSA) are detected on viable cultured human fetal cells (Pujol-Borrell *et al.*, 1982) or adult animal islet cells (Papadopoulos and Lernmark, 1983); here also separate specificities exist for α and β cells, respectively (Van de Winkel *et al.*, 1982). These authors used the cell sorter to separate ICSA-positive cells, and identified a small number of sera that contained surface-reacting antibodies to pancreatic polypeptide (PP) cells. Antibodies to these cells have not so far been detected on cryostat sections, probably because it is necessary to employ PP lobe, that is, the dorsal portion of the head of the pancreas known to be rich in PP cells (Orci, 1982). By analogy with the surface expression of other organ-specific microsomal systems such as the thyroid, gastric, or adrenal, in which the cytoplasmic autoantigens are also expressed on the cell surface (Doniach *et al.*, 1982), it is probable that some selective CF-ICA represent the cytoplasmic portion of those ICSAs that react with β cells and are therefore cytotoxic to cultured animal islets, or interfere with glucose-stimulated insulin secretion in these cultures (Sai *et al.*, 1981). However, surface islet-cell staining is obtained in ~30% of diabetic sera that give negative results on sections (Freedman *et al.*, 1980). This strongly suggests the involvement of an additional antigen that is expressed entirely on the plasma membrane of β cells. These data support the existence of multiple antibodies specific to different islet antigens, and it is only when the antigens are isolated and characterized that we will be in a position to study their effects separately. Monoclonal antibodies make it possible to analyze human polyclonal antibody mixtures in recep-

tor diseases such as myasthenia gravis and thyrotoxicosis (Kohn *et al.*, 1983). Attempts to produce monoclonal ICA have not yet yielded specificities found in diabetic patients.

Humoral Immunity in Predisposed Individuals

Because IDDM often starts as an abrupt illness, it was a surprise to discover, by regular testing of unaffected relatives of diabetic children, that autoimmune phenomena exist for years in those who inherit the diabetogenic genes, as identified by selecting sibs who are *HLA* identical or haploidentical with the proband and who possess CF-ICA during the latency period (Gorsuch *et al.*, 1981). These sibs proved to be the most vulnerable for future diabetes.

Data from two British prospective family studies (Bart; Windsor and Oxford) have been pooled. The cohort contains 283 probands and 1040 first-degree relatives. During the follow-up, IDDM developed in 13 siblings (2.7%) and 4 parents (0.7%). Ten were followed up for ICA; all were positive for 6 to 76 months (median 48 months) before diagnosis. CF-ICA were present in 8 of these 10 cases. A panel of viral antibodies measured in 5 of them showed no significant rising titers. HLA typing was done on 12 siblings, 7 of whom were haploidentical and 5 of whom were HLA identical with the proband. All carried DR3 and/or DR4.

There are still a number of sibs who have the same features and are euglycemic. In five unaffected sibs CF-ICA became negative. A further 4 sibs had these antibodies on and off for many months, while one unaffected brother has shown this reactivity for over 56 months without developing evidence of glucose intolerance. One sib and 1 parent who were constantly negative for conventional ICA as well as for CF-ICA became positive after 24 and 39 months of follow-up, respectively, and the antibodies have so far persisted in their sera.

In the diabetic families, fluctuations of ICA reactions are often seen, and sometimes these fluctuations appear simultaneously in more than one member, including the IDDM proband. Overall, ICA reappeared in 27% of diabetic children who had given many negative results for up to 5 years and had been diabetic for up to 20 years. This was somewhat unexpected in view of previous studies on large groups of juvenile diabetics showing disappearance of ICA between 6 months and 3 years after onset of the disease in the great majority of cases (Lendrum et al., 1976). The fluctuations observed in the islet-cell antibodies perhaps reflect silent intermittent infections with insulotropic viruses. We are looking at viral antibody patterns in these individuals. A retrospective study on a set of identical diabetic triplets revealed the first appearance of ICA-IgG in a discordant

case 5 years before he became diabetic, and the total period of discordance was 39 years (Srikanta *et al.*, 1983). Identical twins are known to remain discordant for IDDM in at least 50% of cases (Barnet *et al.*, 1981). Prospective tests for islet-cell antibodies over a period of 7 years in 46 discordant co-twins showed that only 2 had CF-ICA, and 1 of these became diabetic 5 years later (Colin Johnston, personal communication).

III. CELLULAR IMMUNITY IN DIABETICS

Cell-mediated immune (CMI) mechanisms are as important for the initiation of insulitis as are the humoral responses. Leukocyte migration inhibition and lymphocyte transformation tests have been shown to be positive using islet-enriched pancreatic antigens (Nerup *et al.*, 1974). Lymphocyte adherence tests have also been employed with islet cells in culture (Huang and MacLaren, 1976). Peripheral lymphocytes from IDDM patients produced cytotoxic effects on insulinoma cells in suspension. This was thought to represent both antibody-dependent cell-mediated (ADCC) and direct T-cell cytotoxicity. Cell separations with monoclonal reagents will no doubt identify more precisely the T-cell subsets responsible for β-cell recognition in the islets of Langerhans. Functional inhibition of insulin secretion in suspensions of viable mouse islets by lymphocytes from IDDM cases has also been reported (Boitard *et al.*, 1981). It is felt that this work requires confirmation to determine to what extent the immune attack produced functional β-cell impairment without damaging these cells irreversibly as in the other cytotoxicity studies.

In patients with newly diagnosed diabetes, several abnormalities of lymphocyte subsets have been reported, including raised K-cell levels (Pozzilli *et al.*, 1979), impaired suppressor-cell function (Buschard *et al.*, 1982; Fairchild *et al.*, 1982), and increased natural killer activity over xenogeneic islet cells (Charles *et al.*, 1983). The occurrence of circulating activated T cells has also been described (Jackson *et al.*, 1982; Pozzilli *et al.*, 1983). Evidence of an organ-specific suppressor T-cell defect in diabetes and other endocrine autoimmune diseases has been obtained using the direct leukocyte migration inhibition test (LMI) (Topliss *et al.*, 1983). Leukocytes from patients with type I diabetes, previously shown to respond to an insulinoma extract, lost their MIF specificity when co-cultured with normal lymphocytes, and more significantly, with lymphocytes from patients with autoimmune thyroid disease. Conversely, T lymphocytes from diabetic patients sensitized to islet-cell antigen, but not to thyroid antigen, ameliorated the migration inhibition of T lymphocytes from autoimmune thyroid disease in response to thyroid antigens. Inter-

pretation of these results relied on the assumption that the addition of a small number of normally functioning T-suppressor cells is sufficient to compensate for a selective suppressor T-cell defect in response to islet and other endocrine cell antigens. Even though these experiments are technically complicated and interpretation is controversial, it is of interest that similar results have been obtained in autoimmune liver disease (Vento *et al.*, 1984).

IV. CELL-MEDIATED IMMUNITY IN PREDISPOSED INDIVIDUALS

Cell-mediated immune (CMI) phenomena are also of interest in the unaffected members of diabetic families. We previously showed a raised K-cell number and increased K-cell function in a small proportion of ICA-positive healthy sibs (Pozzilli *et al.*, 1979). These also showed abnormal leukocyte migration inhibition tests in the presence of human pancreatic extracts (Pozzilli *et al.*, 1980). All these individuals shared at least one HLA haplotype with the diabetic proband, but CMI usually did not parallel the circulating antibody results. Suppressor-cell function was found abnormal in some IDDM probands and their healthy relatives, and this defect was more often linked to HLA-DR3 and appeared to run in families irrespective of the age of onset or the duration and severity of the diabetes (Jarowski *et al.*, 1983).

Prospective investigations of T-cell subsets in our family study indicated that helper cells are increased when CF-ICA are present in the serum. However, the ratio between T-helper cells and cytotoxic/suppressor T cells did not differ from that observed in first-degree relatives who did not have ICA in their serum. Only two relatives showed increased activated T cells. These two siblings constantly had positive CF-ICA, and they shared two *HLA* haplotypes with the respective diabetic proband. In addition, they had high levels of killer/*normal*killer (K/NK) cells (Pozzilli *et al.*, 1984).

It appears, therefore, that abnormalities in the cell-mediated immune response can be present in some genetically susceptible individuals who are still disease free. In particular, this finding further supports the concept that expression of regulatory cell surface molecules on activated T cells could be related to an early stage of β-cell damage. The follow-up of CMI in susceptible individuals can therefore give important information on the ultimate cellular events associated with and/or leading to the development of type 1 diabetes.

V. REEXAMINATION OF THE INSULITIS PROCESS

The lesions characteristic of autoimmune insulitis, especially the mononuclear cell infiltrate, can be seen only in pancreatic tissues of newly diagnosed diabetics. Lymphocytes surrounding and infiltrating individual islets have been described mainly in young children, but occasional reports have indicated similar appearances in elderly female patients with polyendocrine autoimmune disorders (Gepts, 1981). Previous histopathology studies on insulitis were carried out on Formalin-fixed pancreases, which do not permit evaluation of lymphocyte subsets by the presently available monoclonal reagents that specifically define the phenotypes of these cells. Furthermore, there are no previous histological reports attempting to demonstrate immunoglobulin deposition in relationship to islets of Langerhans.

A direct demonstration and a better definition of various immune phenomena occurring *in vivo* have recently been achieved by examining fresh frozen blocks of pancreas obtained from a newly diagnosed diabetic child who accidentally died 24 h after diagnosis (Bottazzo *et al.*, 1985). Active insulitis was seen in this pancreas as anticipated. Immunohistological studies confirmed the presence of pseudoatrophic islets. These were interspersed irregularly with islets showing milder lesions and a fairly well-organized architecture. The majority of the islets were devoid of insulin cells, whereas those secreting glucagon or somatostatin were well-preserved. The inflammatory process was scattered in foci throughout the gland and was more obvious in the tail and body of the pancreas. In these areas, multiple blocks had to be sectioned to find the cellular infiltrates of the insulitis. Nevertheless, when such inflammatory foci were identified, the mononuclear cells were abundant and tended to surround or embrace individual islets or small groups of these. When various monoclonal reagents were applied by immunofluorescence with single or double fluorochrome techniques, the majority of mononuclear cells were found to be T lymphocytes. Cytotoxic/suppressor lymphocytes appeared to constitute the main bulk of the infiltrate, but the other known lymphocyte subpopulations were also represented. Ninety percent of autoreactive T lymphocytes expressed *HLA-DR* antigens. This indicated that they were activated T cells, suggesting a specific immune response directed against islet antigens.

When anti-IgG reagents were applied, several interesting phenomena could be observed. Preplasma cells synthesizing IgG, as shown by their bright intracytoplasmic fluorescence, were seen emerging from pancreatic vessels. They appeared to be streaming in the interstitium between exo-

crine lobules, ultimately to converge toward islets. In other fields, these IgG-positive immunocytes were seen stationed around individual islets. It is of interest that mature plasma cells are apparently rarely seen in diabetic insulitis by conventional histology. Perhaps the most intriguing finding was the presence of IgG within some of the islets. In certain fields, endocrine cells varying from a few isolated cells to large groups constituting up to two-thirds of the islet appeared to be coated with a smooth layer of IgG on their outer membrane. Granular Ig deposits indicative of immune complexes were not seen in the pancreas. The most striking picture, however, was the presence of IgG inside the cytoplasm of islet cells. This suggests penetration of antibodies following injury to the cell membrane. Similar appearances have been observed affecting thyrocytes in autoimmune thyroiditis and in renal tubular cells in lupus nephritis.

VI. COULD ABERRANT *HLA-DR* EXPRESSION ON B CELLS BE RELEVANT TO AUTOIMMUNE INSULITIS?

The class II major histocompatibility antigens encoded by genes in the *HLA-D* region play a key role in the presentation of antigens and regulation of the immune response. The expression of these cell-surface glycoprotein molecules is normally restricted to B-lymphocytes, macrophages, dendritic and other antigen-presenting cells, and capillary endothelium (Barclay and Mason, 1983). Modulation of *HLA-DR* expression has been observed only exceptionally outside the immune system; guinea pig mammary gland/duct/epithelium becomes *Ia* positive during pregnancy and lactation (Klareskog *et al.*, 1980). *Ia* antigen expression can be induced during graft-versus-host disease (Barclay and Mason, 1982) and in some cancers (Lloyd *et al.*, 1981).

HLA-DR molecules are not expressed on resting endocrine cells under physiological conditions, although the class I MHC products *HLA-A*, *-B*, and *-C*, can be demonstrated on the membrane and in the cytoplasm of some endocrine organs, including the thyroid gland (Pujol-Borrell *et al.*, 1983). Normal thyroid cells in culture can be stimulated by mitogens, and under these circumstances they actively synthesize and express *HLA-DR* molecules on their plasma membranes (Pujol-Borrell *et al.*, 1983). Perhaps more relevant to human autoimmune disease is the finding that thyrocytes isolated from glands affected by Graves' disease do not require any mitogenic stimulus and spontaneously showed *DR* molecules when cultured in monolayers. This was demonstrated by immunofluorescence with monoclonal antibodies to the nonpolymorphic region of these sur-

face markers. On cryostat sections from glands removed at partial thyroidectomy, the same phenomenon was more clearly seen and could be analyzed in relation to the other histological and immunological features of the goiters. The brightest *DR* staining was seen in Hashimoto's thyroiditis, and in Graves' disease it appeared to be a very early index of focal lymphocytic thyroiditis (Hanafusa *et al.*, 1983).

In the human diabetic pancreas examined in detail by immunofluorescence (Bottazzo *et al.*, 1985), experiments suggest that some endocrine cells in diseased islets could be stained with the monoclonal anti-*DR* antibodies used in the thyroid work. After many sections from different portions of the gland were scanned, it was found that in some islets the few remaining insulin-producing cells were also expressing HLA-DR molecules. Glucagon and somatostatin cells, known to be spared in the insulitis process, were consistently negative for DR staining. The precise identification of DR-positive insulin cells and differential staining of other pancreatic endocrine cells was made possible by using several combinations of monoclonal reagents, including those to pancreatic hormones, and observing these with different fluorochromes. In this context, it is also worth mentioning that diseased bile duct cells in liver biopsies from patients with primary biliary cirrhosis also expressed HLA-DR molecules inappropriately, so this may be a general phenomenon in autoimmunity (Ballardini *et al.*, 1984).

Immune responses are initiated by *HLA-DR*-positive cells presenting antigen to T lymphocytes. Based on the thyroid experiments and on the finding that *DR* molecules are found on thyrocytes in autoimmune thyroid diseases and possibly also in the islet cells in diabetic insulitis, a new hypothesis was put forward for the mechanism of induction of autoimmunity in endocrine organs (Bottazzo *et al.*, 1983). It is known that interferon is the best inducer of *DR* antigen expression. It was postulated that its production could be stimulated by viruses or other environmental agents that might be present in the relevant endocrine gland of genetically predisposed individuals without causing signs of infection. The series of events that would follow is that of aberrant *DR* expression, presentation of surface autoantigens, and subsequent induction of autoreactive T cells. These T cells would in turn activate effector B and T lymphocytes.

Whether the initial induction of autoimmune T cells led to autoimmune disease would depend on a variety of other factors such as abnormalities of the suppressor T-cell pathway. This mechanism of autoimmune disease induction explains vague associations with viral infections and long latency periods before disease becomes manifest, and gives a simple explanation for the well-documented HLA-DR association of autoimmune diseases in man.

VII. HOW DO OTHER ENDOCRINE GLANDS CONTRIBUTE TO DIABETES?

Diabetologists have noticed for some years that prepubertal boys were too tall for their age when they presented with diabetes (Edelsten *et al.*, 1981). This prompted a search for evidence of pituitary involvement. Previous studies showed anterior pituitary antibodies in ~7% of type I adult diabetics who also suffered from polyendocrine autoimmunity, and most of these appeared to react with prolactin cells. In the family studies, predisposed relatives having CF-ICA proved to have in addition antibodies reacting with several pituitary cell types (Mirakian *et al.*, 1982). The highest prevalence was in those who became diabetic during the follow-up, 4 of 7. In the prediabetic latency period, 36% of ICA-positive relatives had these antibodies; in newly diagnosed DM, 16% were positive by IFL on pituitary, and in long-standing DM, only 2% were positive.

This suggests either that similar viruses may infect pituicytes and β cells (Onodera *et al.*, 1981) or that the antibodies detected are markers of some as yet unknown stimulating immunoglobulins, akin to those found in thyrotoxicosis but acting on the pituitary receptors either directly or through increased secretion of as yet uncharacterized hypothalamic (Bobbioni and Jeanrenaud, 1982) or pituitary (Beloff-Chain *et al.*, 1983) hormones that affect insulin secretion. In this context it is of interest that some *HLA*-identical relatives in other prospective studies have increased insulin responses to arginine (Hollander *et al.*, 1982) suggesting that something stimulates the pancreas to regenerate its β cells, which, as we know from the presence of CF-ICA, are gradually being damaged. In polyendocrine patients, who have similar long latency periods, there is usually a decreased insulin production (Irvine *et al.*, 1980).

Because some diabetics showed diminished responses of gastric inhibitory peptide (GIP) to protein test meals, the IFL tests were carried out on human duodenum. A small proportion of NIDDM proved to have antibodies to the GIP-secreting cells, and, when IDDM sera were applied to gut sections, ~20% of the sera were positive on GIP and/or secretin cells (Mirakian *et al.*, 1981). The significance of these unexpected organ-specific reactions either indicates further heterogeneity in this complex syndrome or suggests that in some cases the enteroinsular axis is also involved in the final development of the diabetic syndromes.

Another interesting development is the report of human hybridomas raised with lymphocytes taken from young diabetics whose sera were positive by immunofluorescence on pancreas, thyroid, stomach, or pituitary gland (Satoh *et al.*, 1983). The monoclonal antibodies secreted by some of these clones reacted with several endocrine glands, suggesting

the existence of cross-reactive antigens in these organs. In human sera, each of the microsomal autoantibodies is strictly organ specific; it is possible, for instance, to absorb out all the thyroid cytoplasmic IFL without affecting the titer of gastric parietal cell antibodies coexisting in the same serum. Also, the human organ-specific antibodies are mainly of IgG class, whereas the multiple organ monoclonals were all IgM. Another important difference was that some clones cross-reacted with cytoskeletal elements, which was not seen in human organ-specific autoimmunity.

VIII. ARE WE READY TO TREAT DIABETICS WITH IMMUNOSUPPRESSIVE DRUGS?

Partial success was obtained with massive immunosuppressive therapy in a few newly diagnosed diabetics (Leslie and Pyke, 1980; Spencer *et al.*, 1982), and more selective results were obtained with cyclosporin-A (Stiller *et al.*, 1983; Assan *et al.*, 1985). In the BB rat model, it proved possible to prevent the onset of diabetes by a variety of conventional immunosuppressive agents (Marliss *et al.*, 1982). Whole blood transfusion from nonpredisposed control rats to young BB animals also prevented clinical expression of the disease and diminished the frequency of insulitis seen on histology (Rossini *et al.*, 1983). Now that we can identify predisposed individuals in human diabetic families, the question remains: Is it justifiable to immunosuppress them in order to slow down the insulitis damage and possibly to prevent IDDM? We fully agree with Rossini (1983) that the available immunosuppressive regimens do not select the right subsets of islet-reactive immunocytes and are therefore too drastic. At the same time it is probable that blockage of the autoimmune response to insulin-secreting cells would effectively prevent the disease, irrespective of whether the insulitis is a primary autoimmune phenomenon or a secondary event initiated by environmental agents. The fact that ~50% of CF-ICA-positive relatives in our study escaped the clinical expression of the disease shows that we are not yet ready.

It is impossible to envisage giving any of the available forms of nonselective immunosuppression to symptom-free individuals. There is an urgent need to dissect further the autoreactive immune networks in order to identify future diabetics with a 100% certainty. We hope to approach this goal by cloning activated T cells isolated from the blood of newly diagnosed diabetics. Some of the circulating *DR*-positive T cells are likely to be specifically involved in recognizing relevant autoantigens in the islets. It is now possible to establish self-replicating clones of organ-specific activated T cells as shown in thyroiditis (Londei *et al.*, 1985). Specific

monoclonal antibodies to surface markers on these cells might then be beneficially used to inhibit a selected subset of helper or suppressor regulatory T cells and to prevent triggering or expansion of organ-specific effector/cytotoxic lymphocytes.

IX. FUTURE PROSPECTS FOR DIABETES RESEARCH

A prolonged autoimmune process underlies certain endocrine diseases such as myxedema and Addison's disease, and the thyroid or adrenal antibodies are usually present for many years before organ failure supervenes (Doniach and Bottazzo, 1983). A similar slow time course has been demonstrated in nondiabetic adults in whom islet-cell antibodies are found when they come to hospital for thyrotoxicosis or one of the other autoimmune endocrine conditions. We have found the same prolonged prediabetic stage in selected sibs of diabetic children. The question now remains as to whether or not the pathogenic mechanisms operating in the polyendocrine cases are similar to those in the early-onset disease. Although the most vulnerable sibs of diabetic children (those 2% of all relatives who had CF-ICA in their sera) belong to autoimmune families, the diabetic probands themselves showed only a slightly increased prevalence of thyroid antibodies in the absence of other endocrine insufficiencies. There was the usual excess of males, and the probands were very young patients.

It appears that several sets of genes predispose certain families to type I IDDM. In our family studies, the distribution of thyroid, gastric, pituitary, and ICA-IgG reactions did not correspond to that of the diabetogenic *HLA* haplotypes inherited by the sibs from both parents. Only the CF-ICA seemed to occur more often in *HLA*-identical sibs. Epidemiological studies tell us that 10–20% of such sibs can be expected to get diabetes. However, when two markers were employed, half of those who had CF-ICA became diabetic; hence the genes connected with autoimmunity are heavily skewed toward the pancreas, just as in thyroiditis families the autoimmunity is biased towards the thyroid gland and pernicious anemia relatives have an excess of gastric autoimmunity.

In type I diabetes, only the β cells are destroyed. Some of the CF-ICA detected on unfixed pancreatic sections and some of the cell surface-reactive antibodies (ICSA) are now known to be β-cell specific.

A detailed analysis of *HLA-DR* molecules by DNA probing techniques (Owerbach *et al.*, 1983) underlies the polymorphism of the DR region and reinforces the concept of heterogeneity in the family makeup of diabetic patients. New monoclonal reagents to the other molecules of the *HLA-D*

region will soon be more generally available (Guy and Van Heyningen, 1983), and when applied to population studies they are certain to disclose further differences between diabetics and nondiabetics (Cudworth and Wolf, 1982).

Another interesting development in the study of IDDM was the unexpected finding of insulin-receptor antibodies of the IgM class in untreated juvenile diabetics (Maron *et al.*, 1983). There are also spontaneous antibodies to insulin that appear before any insulin injections (Palmer *et al.*, 1983). Findings in recent juvenile diabetics have been confirmed and extended to unaffected identical twins of IDDM cases (Wilkin *et al.*, 1985). The insulin antibodies in the twins were not related to the presence of ICA or to HLA-DR 3 or 4 haplotypes. Since the unaffected twins of long-standing diabetics are very unlikely to develop the disease, these observations suggest that insulin autoantibodies do not presage future diabetes but rather reflect an inherited diathesis for pancreatic autoimmunity.

The presence of these two types of autoimmune reactions may suggest that insulin or proinsulin acts as powerful immunogens, especially if an inappropriate secretion of immature hormone molecules is envisaged during the slow process of the autoimmune β-cell damage, as suggested by Rimoin and Rotter (1983). *HLA-DR4* individuals are more liable to produce insulin antibodies to both animal and human preparations (Schernthaner *et al.*, 1983), and these patients may be the right candidates for spontaneous autoimmunization to their own insulin or proinsulin (Rotter and Rimoin, 1983). The operation of antiidiotype networks can then explain the apparent insulin-receptor antibodies as internal image antiidiotypes made in response to the original antiinsulin idiotypic epitope (Cohen *et al.*, 1983; Roitt, 1984).

ACKNOWLEDGMENTS

We would like to thank our research fellows, Drs. Rita Mirakian, Betty Dean, Ricardo Pujol-Borrell, and Toshiaki Hanafusa for helpful discussion and for allowing us to quote their works in progress. Work in the laboratory was supported by the Wellcome Trust, The Juvenile Diabetes Foundation (United States), and Novo Research Laboratories.

REFERENCES

Assan, R., Feutren, G., Debray-Sachs, M., Quiniou-Debrie, M. C., Laborie, C., Thomas, G., Chatenoud, L., and Bach, J. F. (1985). *Lancet* **1**, 67–71.

Baekkeskov, S., Nielsen, J. H., Marner, B., Bilde, T., Ludvigsson, J., and Lernmark, A. (1982). *Nature (London)* **298**, 167–169.

Ballardini, G., Mirakian, R., Bianchi, F. B., Pisi, E., Doniach, D., and Bottazzo, G. F. (1984). *Lancet* **2,** 1009–1013.

Barclay, A. N., and Mason, D. W. (1982). *J. Exp. Med.* **156,** 1665–1676.

Barclay, A. N., and Mason, D. W., (1983). *Nature (London)* **303,** 382–383.

Barnet, A. H., Eff, C., Leslie, R. D. G., and Pyke, D. A. (1981). *Diabetologia* **20,** 87–93.

Beloff-Chain, A., Morton, J., Dunmore, S., Taylor, G. W., and Morris, M. R. (1983). *Nature (London)* **301,** 255–258.

Bobbioni, E., and Jeanrenaud, B. (1982). *Endocrinology (Baltimore)* **110,** 631–636.

Boitard, C., Debray-Sachs, M., Pouplard, A., Assan, R., and Hamburger, J. L. (1981). *Diabetologia* **21,** 41–46.

Bottazzo, G. F., and Doniach, D. (1978). *Clin. Lab.* **8,** 29–38.

Bottazzo, G. F., and Lendrum, R. (1976). *Lancet* **2,** 873–876.

Bottazzo, G. F., Florin-Christensen, A., and Doniach, D. (1974). *Lancet* **2,** 1279–1283.

Bottazzo, G. F., Dean, B. M., Gorsuch, A. N., Cudworth, A. G., and Doniach, D. (1980). *Lancet* **1,** 668–672.

Bottazzo, G. F., Pujol-Borrel, R., and Doniach, D. (1981). *Clin. Immunol. Allergy* **1,** 63–80.

Bottazzo, G. F., Mirakian, R., Dean, B. M., McNally, J. M., and Doniach, D. (1982). *In* "Genetics of Diabetes Mellitus" (R. B. Tattersall and J. K. Koberling, eds.), pp. 79–90. Academic Press, London.

Bottazzo, G. F., Pujol-Borrell, R., Hanafusa, T., and Feldman, M. (1983b). *Lancet* **2,** 1115–1119.

Bottazzo, G. F., Dean, B. M., McNally, J. M., McKay, E. H., Swift, P. G. F., and Gamble, D. R. (1985). *N. Engl. J. Med.* (in press).

Buschard, K., Madsbad, S., and Rygaard, J. (1982). *J. Clin. Lab. Immunol.* **8,** 19–24.

Charles, M. A., Auzuki, M., Waldeck, N., Dodson, L. E., Slater, L., Ong, K., Kershnar, A., Buckinngham, B., and Golder, M. (1983). *J. Immunol.* **130,** 1189–1196.

Cohen, I. R., Elias, D., Maron, R., and Schechter, Y. (1983). *In* "Idiotypic Manipulations in Biological Systems" (H. Kholer, P. A. Cazenave, and J. Urbain, eds.), pp. 1–22. Academic Press, New York.

Cudworth, A. G., and Wolf, E. (1982). *Clin. Endocrinol. Metab.* **11,** 389–406.

Dean, B., Pujol-Borrell, R., and Bottazzo, G. F. (1982). *Lancet* **II,** 1343–1344.

Dean, B. M., Bottazzo, G. F., and Cudworth, A. G. (1983). *Clin. Exp. Immunol.* **52,** 61–66.

Doniach, D., and Bottazzo, G. F. (1983). *Hosp. Up-Date* **9,** 1145–1159.

Doniach, D., Cudworth, A. G., Khoury, E. L., and Bottazzo, G. F. (1982). *Recent Prog. Endocrinol.* **2,** 99–132.

Edelsten, A. D., Hughes, I. A., Oakes, S., Gordon, I. R. S., and Savage, D. C. L. (1981). *Arch. Dis. Child.* **56,** 40–44.

Fairchild, R. S., Kyner, J. L., and Abdou, N. I. (1982). *J. Lab. Clin. Med.* **99,** 175–186.

Freedman, Z. R., Irvine, W. J., Lernmark, A., Huen, A. H. J., Steiner, O. F., and Rubenstein, A. H. (1980). *In* "Immunology of Diabetes" (W. J. Irvine, ed.), pp. 169–175. Teviot Sci. Publ., Edinburgh.

Gepts, W. (1981). *In* "The Islets of Langerhans" (S. J. Cooperstein, and D. Watkins, eds.), pp. 321–356. Academic Press, New York.

Gorsuch, A. N., Spencer, K. M., Lister, J., McNally, J. M., Dean, B. M., Bottazzo, G. F., and Cudworth, A. G. (1981). *Lancet* **2,** 1363–1365.

Guy, K. V., and Van Heyningen, V. (1983). *Immunol. Today* **4,** 186–189.

Hanafusa, T., Pujol-Borrell, R., Chiovato, R., Russell, R. C. G., Doniach, D., and Bottazzo, G. F. (1983). *Lancet* **2,** 1111–1115.

Helgason, T., and Jonasson, M. R. (1981). *Lancet* **2,** 716–720.

Hollander, P. H., Asplin, C. M., Kniaz, D., Hansen, J. A., and Palmer, J. P. (1982). *Diabetes* **31,** 149–153.

Huang, S. W., and MacLaren, N. K. (1976). *Science* **192,** 64–66.

Irvine, W. J., Gray, R. S., and Steel, J. M. (1980). *In* "Immunology of Diabetes" (W. J. Irvine, ed.), pp. 117–154. Teviot Sci. Publ., Edinburgh.

Jackson, R. A., Morris, M. A., Haynes, B. F., and Eisenbarth, G. S. (1982). *N. Engl. J. Med.* **306,** 785–788.

Jarowski, M. A., Colle, E., and Guttmann, R. D. (1983). *Hum. Immunol.* **7,** 25–34.

Karam, J. H., Lewitt, P. A., Young, C. W., Nowlain, R. E., Frankek, B. J., Fujiya, H., Freedman, Z. R., and Grodsky, G. M. (1980). *Diabetes* **29,** 971–978.

Klareskog, L., Forsum, U., and Peterson, P. A. (1980). *Eur. J. Immunol.* **10,** 958–963.

Knight, A., Knight, J., Laing, P., and Adams, D. (1984). *J. Clin. Lab. Immunol.* **14,** 141–144.

Kohn, L. D., Yavin, E., Yavin, Z., Laccetti, P., Vitti, P., Grollman, E. F., and Valente, W. A. (1983). *In* "Monoclonal Antibodies: Probes for Study of Autoimmunity and Immunodeficiency" (G. Eisenbarth and R. Haynes, eds.), pp. 221–258. Academic Press, New York.

Lendrum, R., Walker, J. G., Cudworth, A. G., Theophanides, C., Pyke, D. A., Bloom, A., and Gamble, D. R. (1976). *Lancet,* **2,** 1273–1276.

Leslie, R. D. G., and Pyke, D. A. (1980). *In* "Immunology of Diabetes" (W. J. Irvine, ed.), pp. 345–347. Teviot Sci. Publ., Edinburgh.

Lloyd, K. O., Ng, J., and Diffold, W. G. (1981). *J. Immunol.* **126,** 2408–2413.

Londei, M., Bottazzo, G. F., and Feldmann, M. (1985). *Science* **228,** 85–87.

Marliss, E. B., Nakhooda, A. F., Poussier, P., and Sima, A. A. F. (1982). *Diabetologia* **22,** 225–232.

Maron, R., Elias, D., De Jongh, B. M., Bruining, G. J., Van Rood, J. J., Schechter, Y., and Cohen, I. R. (1983). *Nature* (*London*) **303,** 817–818.

Mirakian, R., Richardson, C. A., Bottazzo, G. F., and Doniach, D. (1981). *Clin. Immunol. Newslett.* **2,** 161–167.

Mirakian, R., Cudworth, A. G., Bottazzo, G. F., Richardson, C. A., and Doniach, D. (1982). *Lancet* **1,** 755–759.

Nerup, J., Anderson, O. O., Bendixen, G., Egeberg, J., Gunnarsson, R., Kromann, G., and Poulsen, J. E. (1974). *Proc. R. Soc. Med.* **67,** 506–513.

Onodera, T., Toniolo, A., Ray, U. R., Jenson, A. B., Knazek, A., and Notkins, A. L. (1981). *J. Exp. Med.* **153,** 1473–1477.

Orci, L. (1982). *Diabetes* **31,** 538–565.

Owerbach, D., Lernmark, A., Platz, P., Ryder, L. P., Rask, L., Peterson, P. A., and Ludvigsson, J. (1983). *Nature* (*London*) **303,** 815–817.

Palmer, J. P., Asplin, C. M., Clemons, P., Lyen, K., Tatpali, O., Raghu, P. K., Paquette, T. L. (1983). *Science* **222,** 1337–1339.

Papadopoulos, G. K., and Lernmark, A. (1983). *In* "Autoimmune Endocrine Disease" (T. F. Davies, ed.), pp. 167–180. Wiley, New York.

Pozzilli, P., Sensi, M., Gorsuch, A. N., Bottazzo, G. F., and Cudworth, A. G. (1979). *Lancet* **2,** 173–175.

Pozzilli, P., Sensi, M., and Cudworth, A. G. (1980). *Diabetologia Croat., IX Suppl.* **1,** pp. 132–135.

Pozzilli, P., Zuccarini, O., Iavicoli, M., Andreani, D., Sensi, M., Spencer, K. M., Bottazzo, G. F., Beverley, P. C. L., Kyner, J. L., and Cudworth, A. G. (1983). *Diabetes* **32,** 91–94.

Pozzilli, P., Sensi, M., Al-Sakkaf, L., Tarn, A., Zuccarini, O., and Bottazzo, G. F. (1984). *Diabetologia* **27,** 132–135.

Pujol-Borrell, R., Khoury, E. L., and Bottazzo, G. F. (1982). *Diabetologia* **22,** 89–95.
Pujol-Borrell, R., Hanafusa, T., Chiovato, L., and Bottazzo, G. F. (1983). *Nature (London)* **304,** 71–73.
Rayfield, E. J., and Yoon, J. W. (1981). *In* "The Islet of Langerhans" (S. J. Cooperstein and D. Watkins, eds.), pp. 427–445. Academic Press, New York.
Roitt, I. M. (1984). *Triangle* **23,** 67–76.
Rossini, A. A. (1983). *N. Engl. J. Med.* **308,** 333–335.
Rossini, A. A., Mordes, J. P., Pelletier, A. M., and Like, A. A. (1983). *Science* **219,** 975–976.
Rotter, J. I., and Rimoin, D. L. (1983). *Acta Endocrinol. Suppl. 256* **103,** 26.
Sai, P., Boitard, C., Debray-Sachs, M., Pouplard, A., Assan, R., and Hamburger, J. (1981). *Diabetes* **30,** 1051–1057.
Satoh, J., Prabhakar, B. S., Haspel, M. V., Ginsberg-Fellner, F., and Notkins, A. L. (1983). *N. Engl. J. Med.* **309,** 217–220.
Schernthaner, G., Borkenstein, M., Fink, M., Mayr, W. R., Menzel, J., and Schuber, E. (1983). *Diabetes Care, Suppl. 1* **6,** 43–48.
Spencer, K. M., Dean, B. M., Bottazzo, G. F., Medbak, S., and Cudworth, A. G. (1982). *Diabetologia* **70,** 474A.
Srikanta, S., Ganda, O. P., Eisenbarth, G. S., and Soeldner, J. S. (1983). *N. Engl. J. Med.* **308,** 322–325.
Stiller, C. R., Laupacis, A., Dupre, J., Jenner, M. R., Keown, P. A., Rodger, W., and Wolfe, B. M. J. (1983). *N. Engl. J. Med.* **308,** 1226–1227.
Topliss, D., How, T., Lewis, M., Row, V., and Volpe, R. (1983). *J. Clin. Endocrinol. Metab.* **57,** 700–705.
Van de Winkel, M., Smets, G., Gepts, W., and Pipeleers, D. (1982). *J. Clin. Invest.* **70,** 41–49.
Vento, S., Hegarty, J. E., Bottazzo, G. F., Macchia, E., Williams, R., and Eddleston, A. L. W. F. (1984). *Lancet* **1,** 1200–1204.
Wilkin, T., Hoskins, P. J., Armitage, M., Rodier, M., Casey, C., Diaz, J.-L., Pyke, D. A., and Leslie, R. D. G. (1985). *Lancet* **1,** 480–482.

CHAPTER 10

Pernicious Anemia and Gastric Atrophy

SENGA WHITTINGHAM
IAN R. MACKAY
Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital
Melbourne, Victoria, Australia

I. INTRODUCTION: HISTORICAL BACKGROUND

Thomas Addison, at a meeting of the South London Medical Society in 1849, described the clinical features of a disease identifiable as pernicious anemia although he did not give it that name. However, he recognized that this "very remarkable form of general anemia" was a disease that was different from other forms of anemia. According to Castle (1953), it was Austin Flint who intuitively remarked,

> I suspect that in these cases there exists degenerative disease of the glandular tubuli of the stomach . . . . Fatal anemia must follow an amount of degenerative disease reducing the amount of gastric juice so far that assimilation of food is rendered wholly inadequate to the wants of the body. I shall be ready to claim the

> merit of this idea when the difficult and laborious researches of someone have shown it to be correct.

This requisite histological evidence for gastric atrophy was provided by Fenwick (1870), and a year later the anemia was called *pernicious anemia* by Biermer (1872).

Treatment also played a role in the eventual understanding of the development of pernicious anemia (PA). At first the causal connection between the anemia and gastritis was incomprehensible, although a reticulocyte response was observed after patients had been fed large meals of cooked liver (Minot and Murphy, 1926). The reason for the response was explained by the experiments of Castle (1953), which showed that a combination of "extrinsic factor" subsequently identified as vitamin B_{12} (Lester-Smith, 1948; Rickes *et al.*, 1948), and "intrinsic factor" in gastric juice (Highley *et al.*, 1967) led to a reticulocyte response.

Oral treatment with crude extracts of hog stomach had first been given in the late 1920s (Sharp, 1929; Sturgis and Isaacs, 1929), and with such treatment remission was achieved for several years (Wilkinson, 1949). However, in the course of time, it was observed that relapses tended to occur and some cases became refractory to increasing amounts of the extract (Berlin *et al.*, 1958). The reason for this state was shown to be the development of an inhibitory factor in serum; serum from such patients mixed with intrinsic factor inhibited its effectiveness (Schwartz, 1958). Moreover, it was observed that the serum of rabbits injected with intrinsic factor developed similar inhibitory properties that were attributed to antibodies (Taylor and Morton, 1958). Next, sera from patients with PA who had not been treated with intrinsic factor were shown to contain antibodies (Schwartz, 1960; Taylor, 1959), and various studies were developed to demonstrate and examine these antibodies (Abels *et al.*, 1963; Ardeman and Chanarin, 1963; Gottlieb *et al.*, 1965; Jeffries *et al.*, 1962; Ungar, 1967). In addition to antibodies to the binding site for vitamin B_{12}, other antibodies were shown to restrict mobility of the intrinsic factor–vitamin B_{12} complex on gel electrophoresis. Thus the notion emerged that there were two antibodies, one reactive with the binding site for vitamin B_{12} (type 1 antibody) and a second reactive with another site on intrinsic factor (type 2 antibody) (Bardhan *et al.*, 1968; Jacob and Schilling, 1966; Jeffries *et al.*, 1962).

Besides these studies on antibodies to gastric intrinsic factor, other studies reported the finding of complement-fixing antibodies to gastric parietal cells (Irvine *et al.*, 1962; Markson and Moore, 1962; Taylor *et al.*, 1962), which are now universally demonstrated by immunofluorescence. Little progress was made in the understanding of the gastric lesion until a

gastric biopsy tube was designed that would permit the study of the histological appearance of the mucosa (Wood *et al.*, 1949).

Pernicious anemia was observed to cluster in families (McIntyre *et al.*, 1959) and coexist with autoimmune thyroid diseases (Tudhope and Wilson, 1960), suggesting there was a genetic component to the disorder. The recognition that PA fulfilled the markers of autoimmune disease promulgated by Mackay and Burnet (1963) led to its acceptance as an autoimmune disease of the stomach and to the question that is still plaguing immunologists today: How does the immune system recognize self?

II. GENERAL DESCRIPTION: ANIMAL MODELS

Pernicious anemia (PA) is the end stage of an autoimmune disease that results in the destruction of gastric mucosa. It is the gastric equivalent of myxedema, which follows functional failure of the thyroid gland due to autoimmune thyroiditis of the Hashimoto type. The autoimmune process is limited to the body of the stomach with sparing of the antrum, and it culminates in gastric atrophy associated with which is a varying degree of infiltration of mononuclear cells including plasma cells.

The striking histological lesion, often unsuspected clinically, appears to begin early in adult life and may span a lifetime before the onset of the functional failure, which is expressed hematologically as a deficiency of vitamin B_{12} associated with megaloblastic anemia. Although "silent" until the end stage, the gastric lesion is readily diagnosed and can be predicted years before the clinical presentation by immunological and genetic markers specific for gastric autoimmunity.

Until recently there was no informative animal model of naturally occurring autoimmune gastritis, which precluded precise analysis of the contribution of immunogenetics to the development of gastric atrophy. An experimental animal model that was developed in the rhesus monkey by immunization with extracts of gastric mucosa in Freund's complete adjuvant (Andrada *et al.*, 1969) resulted in the development of an autoimmune gastritis, gastric parietal cell antibodies, and positive cutaneous delayed-type hypersensitivity reactions to gastric antigens. Experimental immunization of dogs with gastric juice and extracts of gastric mucosa also resulted in the development of gastritis and equivalent humoral and cellular reactions to gastric antigens, but in this model, unlike the human disease, the inflammatory response in the gastric mucosa was minimal although atrophy was severe (Hennes *et al.*, 1962; Fixa *et al.*, 1964, 1972; Krohn and Findlayson, 1973). The cellular response is thought to be the more important in the production of gastritis in dogs (Krohn and Findlay-

son, 1973), although a combined humoral and cellular response to immunization has been demonstrated. However, in rats, hypochlorhydria and atrophy of gastric parietal cells were induced by infusion of an immunoglobulin G (IgG) fraction of human serum containing gastric parietal cell antibodies (Tanaka and Glass, 1970); this suggested that antibody is predominantly or even exclusively involved in the induction of the gastric lesion.

There is also evidence that clustering of autoimmunity occurs in the animal models of human diseases associated with autoimmune gastritis. Gastric parietal cell antibodies reactive with rat stomach were demonstrated in 68% of BB/W rats with insulin-dependent diabetes mellitus of spontaneous onset (Elder *et al.*, 1982). These antibodies were associated with histological evidence of mild to moderate gastritis and loss of specialized cells, fibrosis, and squamous metaplasia, although no significant reduction in acid secretion or fall in the serum level of vitamin B_{12} was detected in the 90- to 300-day-old rats studied. Also, antibodies to gastric parietal cells were shown to coexist serologically with thyroid antibodies in obese strain chickens with the avian counterpart of human Hashimoto's throiditis (Khoury *et al.*, 1982). Sera from 9 of 47 of these chickens gave a pattern of reactivity comparable with that of human serum containing gastric parietal cell antibodies when tested against the proventriculus of the chicken.

The animal models suggested that the gastric atrophy of PA is initiated and sustained by an autoimmune process. However, there have been no meaningful studies on how genetic effects influence the lesion, what effector mechanisms act on the gastric mucosa, and what effects minor extrinsic damage may have on the initiation and persistence of the lesion. If the contributory factors are comparable with those involved in the "sister" autoimmune disease, autoimmune thyroiditis, then two "susceptibility" genetic effects would require to be expressed concurrently, one expressing a primary defect in the gastric mucosal cell and another expressing a defect in the thymus resulting in immunoregulatory dysfunction (Maron and Cohen, 1980; Penhale *et al.*, 1975; Tomazic *et al.*, 1974).

III. CLINICAL PRESENTATION

Classically, the patient with PA due to autoimmune gastritis is a gray-haired, middle-aged woman of northern European origin with signs and symptoms of anemia and possibly of associated neuropathy. There may be relatives with pernicious anemia and/or a coexisting or family history of other autoimmune diseases in the thyrogastric cluster (Table I).

TABLE I
Diseases Associated with PA That Cluster in "PA Families"

Diseases	Citations
Thyrotoxicosis, Hashimoto's thyroiditis	Ardeman *et al.* (1966b)
Primary hypothyroidism	Irvine *et al.* (1965) Irvine (1975) Schiller *et al.* (1968) Tudhope and Wilson (1960)
Insulin-dependent diabetes mellitus	Irvine *et al.* (1970) Whittingham *et al.* (1971)
Primary Addison's disease	Blizzard *et al.* (1967) Irvine (1978)
Primary ovarian failure	Irvine and Barnes (1974)
Primary hypoparathyroidism	Blizzard *et al.* (1966)
Premature graying of hair	Whittingham *et al.* (1969)
Vitiligo	Bor *et al.* (1969)
Myasthenia gravis	Mackay (1971)
Lambert–Eaton syndrome	Guttmann *et al.* (1972)

The clinical features of PA are well known. Usually the patient presents in late middle age, is female more often than male, and is usually of northern European origin. The patient will appear pale, physically tired, mentally depressed, and may complain of abdominal discomfort. The anemia is megaloblastic, due to deficiency of vitamin B_{12}, and may in later stages be accompanied by peripheral neuropathy.

The laboratory tests are diagnostic. The blood film and bone marrow show features of megaloblastic anemia, the gastric aspirate shows achlorhydria, and the gastric mucosal biopsy shows gastritis and varying degrees of gastric atrophy (Fig. 1). Almost invariably (>90%) there are serum autoantibodies to gastric parietal cells (Fig. 2), and frequently (60%) there are antibodies to gastric intrinsic factor. There is a low level (<200 pg/ml) of serum vitamin B_{12} in serum, and a low level (<20 ng/ml) of serum pepsinogen I and a high (>100 pmol/liter) level of serum gastrin. The Schilling test result is <5% (Chanarin, 1979), suggesting that there is gross impairment of absorption of vitamin B_{12}.

It is important to appreciate that the immunological processes that finally result in vitamin B_{12} deficiency are operative for many years before there is any functional evidence of exhaustion of the gastric parietal cells. Two points can be made to which we will return in later sections. One is that the chief cells of the stomach, which secrete pepsin, are also affected, although there is no evidence that these are a target for immune-mediated damage. The other is that, for historical reasons, we are burdened by

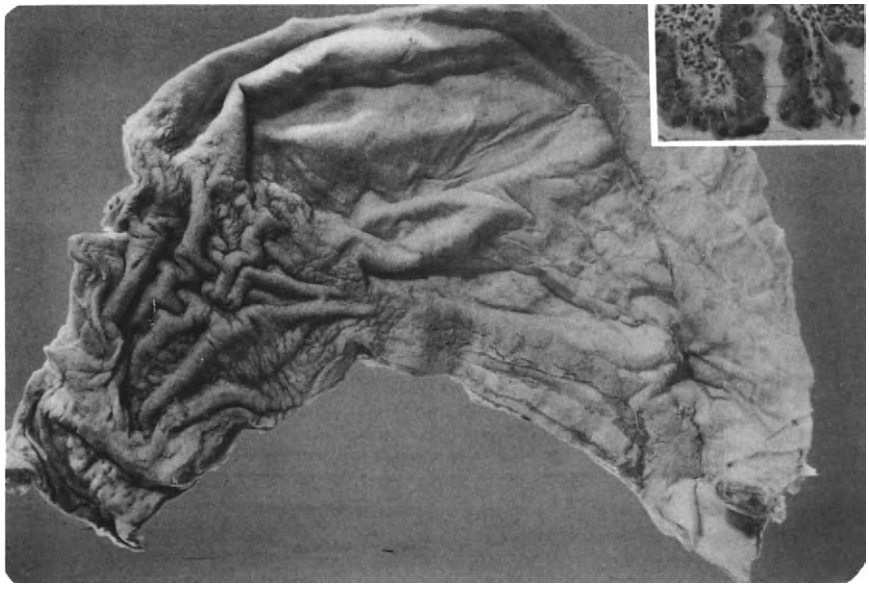


FIG. 1. Macroscopic appearance of the gastric mucosa in pernicious anemia showing the extreme thinning of the body (right) because of gastric atrophy with loss of specialized cells and replacement by mucous glands in contrast to the healthy mucosal folds of the antrum (left). Inset is the microscopic appearance when a mucicarmine stain is used to detect mucus-secreting cells.

problems in terminology. Since the eventual clinical expression of the disease is a blood disorder, our chapter heading includes "pernicious anemia," but the disease process with which we are concerned pathogenetically is a *chronic autoimmune gastritis,* and this is the process that terminates in gastric atrophy and is the cause of the megaloblastic or "pernicious" anemia.

Pernicious anemia associates with a number of diseases, but predominantly with endocrinopathies and diseases in which autoantibodies to important cell receptors account for the immunopathogenesis of the disorder. The associated autoimmune diseases include thyrotoxicosis, Hashimoto's thyroiditis, insulin-dependent diabetes mellitus, primary Addison's disease of the adrenal gland, primary ovarian failure, primary hypoparathyroidism, premature greying of the hair, vitiligo, myasthenia gravis, and the Lambert–Eaton syndrome which is a more recently observed association (Table I). These "thyrogastric" autoimmune diseases not only occur together in the same patient with PA but aggregate in "PA

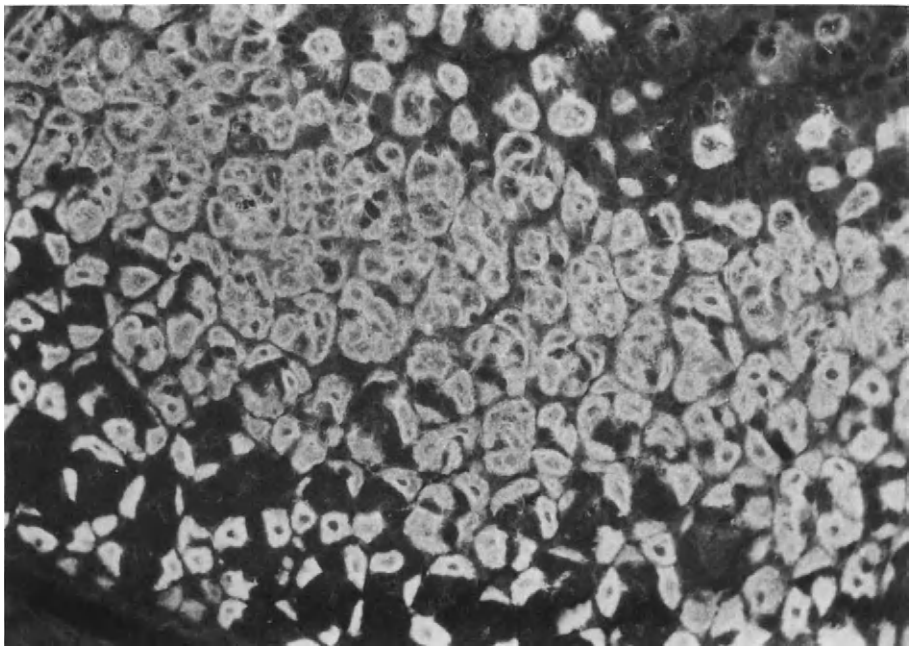


FIG. 2. Indirect immunofluorescence preparation showing gastric parietal cell antibody detected using a serum of patient with pernicious anemia applied to a frozen air-dried section of mouse stomach. There is reactivity with the parietal cells of the tubular glands (×50).

families" (Ardeman *et al.*, 1966a; Wangel *et al.*, 1968a,b; Whittingham *et al.*, 1969). One common link between these disorders is autoimmunity.

IV. HISTOPATHOLOGY

A. Evolution of Gastric Atrophy

The evolution of gastric atrophy (Fig. 3, A–C) in most cases of PA probably spans 20 to 30 years, but this is difficult to assess in individual cases. Nearly all patients with gastric parietal cell antibody whose gastric mucosae have been examined histologically have shown some evidence of gastritis (Serafini *et al.*, 1970). Thus, the presence of gastric parietal cell antibody in the serum is predictive of autoimmune type gastritis (Irvine *et al.*, 1965). Conversely, gastric parietal cell antibody is not observed when gastritis is due to diseases affecting the body of the stomach which are not

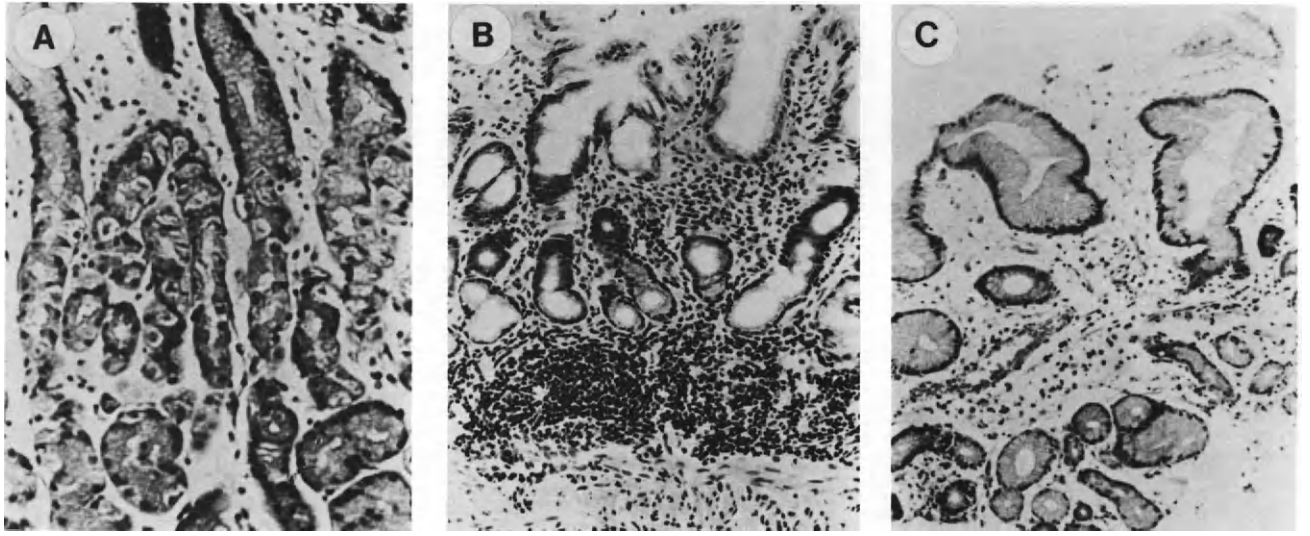


FIG. 3. The figures illustrate the evolution of autoimmune gastritis to gastric atrophy. (A) The mucosa of the body of the stomach of a normal person showing intact glands and absence of mononuclear cell infiltration (HE × 30). (B) An early stage of autoimmune gastritis showing loss of specialized cells and many lymphocytes and plasma cells in the lamina propria (HE × 30). (C) The mucosa of the body of a stomach from a patient with pernicious anemia illustrating gastric atrophy with extreme thinning of the mucosa, loss of gastric parietal and chief cells, and extensive intestinal metaplasia of superficial glands. This is the terminal stage of autoimmune gastritis (HE × 30).

TABLE II

Frequency of Gastric Autoantibodies in Diseases in the Autoimmune Cluster Compared with Control Populations

	Frequency (%) of autoantibodies to		
	Gastric parietal cells	Gastric intrinsic factor	Citation
Disease group (number tested)			
Thyrotoxicosis (302)	24	3.0	Irvine (1965)
Hashimoto's thyroiditis (120)	NT[a]	5.0	Irvine (1975)
Primary hypothyroidism (297)	NT	6.7	Irvine (1975)
Insulin-dependent diabetes mellitus			
Aged <30 years (771)	9.0	NT	Riley *et al.* (1982)
Aged 10–90 years (200)	28.0	4.0	Ungar *et al.* (1968)
Primary Addison's disease (261)	31.0	8.4	Irvine and Barnes (1975)
Primary ovarian failure (5)	40.0	40.0	Irvine *et al.* (1968)
Primary hypoparathyroidism (68)	22.0	NT	Blizzard *et al.* (1966)
Vitiligo (80)	21.0	NT	Brostoff *et al.* (1969)
Lambert–Eaton syndrome (46)	26.0	11.4	Lennon *et al.* (1982)
Control (number tested)			
Chronic atrophic gastritis, Type B (non-PA) (13)	0	0	Whittingham *et al.* (1969)
Duodenal ulcer (200)	5.0	0	Ungar *et al.* (1976)
Australian population (3492)	4.8	0	Hooper et al. (1972)
Aged 21–30 years (551)	2.2	NT	
Aged 61–65 years (317)	6.3	NT	
Australian blood donors (500)	NT	0	Ungar *et al.* (1968)
Scottish blood donors, females aged 40–60 years (141)	9.0	0	Irvine (1965)

[a] NT, not tested.

autoimmune (Table II). Why some patients with autoimmune gastritis progress to PA while others maintain sufficient vitamin B_{12} absorption for long periods is not known. However this point is unimportant, because PA is merely the terminal stage of the process. The youngest person we have studied with gastric parietal cell antibody was aged 13 years, and her gastric mucosa biopsied at the age of 17 years showed mild gastritis. Whether and when she will develop PA cannot be predicted.

B. The Type A–Type B Classification

Gastric atrophy is readily recognized macroscopically and microscopically. The wall of the body of the stomach becomes paper thin because the gastric glands are markedly reduced or absent (Fig. 1). There may be intestinal metaplasia and residues of an earlier and denser infiltration of mononuclear cells (Fig. 3,B).

Strickland and Mackay (1973) proposed a classification of gastritis on the basis of histological findings of the mucosa of the body of the stomach and antrum, the presence of gastric parietal cell antibody, and serum levels of gastrin. Type A gastritis, the "PA type," involves only the body of the stomach, is associated with antibodies to gastric parietal cells, achlorhydria, and high levels of serum gastrin secreted by the intact antral glands. Type B gastritis, the non-PA or simple type, involves both the body and antral mucosa, is not associated with gastric parietal cell antibody, and shows incomplete failure of acid secretion and low levels of serum gastrin because the antrum is affected by the gastritic process. Of 30 patients with type A gastritis and able to absorb vitamin B_{12} normally, Strickland and Mackay observed that 5 (16%) developed overt or latent PA during a follow-up period of 3–24 years. Type A gastritis has also been shown to be the gastritis characteristic of families in which PA predominates (Kekki *et al.*, 1983; Varis, 1981). This classification suggests that since the histological findings in PA are restricted to the body of the stomach and autoantibodies are directed against components of the gastric parietal cell, this cell must be the target of the autoimmune process. However, gastric atrophy with PA is characterized by complete loss of both parietal cells and chief cells, and it is difficult to explain why chief cells are involved unless they are environmentally implicated in the destructive process affecting the parietal cells.

It is of interest that PA in patients with the common variable type of immunodeficiency associated with low levels of serum immunoglobulins can be distinguished from classical PA on this classification. This former type of PA usually occurs in a younger age group, is histologically type B, is associated with a negative test for antibodies to gastric parietal cells and intrinsic factor, and shows a low level of serum gastrin (Cowling *et al.*, 1974; Hughes *et al.*, 1972; Twomey *et al.*, 1969, 1970).

C. Reversibility of the Lesion

There are many reports of regeneration of gastric parietal cells, improvement in gastric function, and hematological remission after corticosteroid drugs (Ardeman and Chanarin, 1965; Baggett and Welsh, 1970; Doig *et al.*, 1957; Gordin, 1959; Jeffries *et al.*, 1966; Kristensen and Friis,

1962; Rødbro *et al.*, 1967; Strickland *et al.*, 1969; Wall *et al.*, 1968) or azathioprine (Jorge and Sanchez, 1973). This suggests that the gastric mucosa is the direct target of an autoimmune process that can be checked by the immunomodulating effect of these drugs.

V. IMMUNOLOGY

A. Immunological Features

The cell implicated in the autoimmune process in PA is the gastric parietal cell. Autoantibodies develop against various components of this cell; the membrane, the cytoplasm, and a secreted product, intrinsic factor, which normally binds avidly to dietary vitamin B_{12} (Glass, 1963) and promotes its transport to the ileum for absorption (Donaldson *et al.*, 1967; Kapadia *et al.*, 1983). Like most autoantibodies, gastric autoantibodies are polyvalent but are predominantly of the immunoglobulin G (IgG) isotype (Serafini *et al.*, 1970). IgA antibodies to gastric intrinsic factor have been demonstrated in gastric juice (Goldberg and Bluestone, 1970).

Antibodies to gastric parietal cells are reactive with a cell-surface membrane antigen (De Aizpurua *et al.*, 1983a; Masala *et al.*, 1980) and a lipoprotein on membranes of the canalicular microvilli in the cytoplasm (Hoedemaeker and Ito, 1970; Ward and Nairn, 1972). There are opposing views as to whether these antibodies are the same or distinct (De Aizapurua *et al.*, 1983a; Masala *et al.*, 1980). Both are demonstrable by immunofluorescence, but as the test for antibody to surface membranes requires living cells, the test for antibody to cytoplasm using frozen sections of mucosa is simpler to perform (Fig. 2). The substrate recommended for the latter is an unfixed, air-dried, frozen section of human stomach, but if this is not available, rodent stomach may be used. Mouse stomach is preferable to rat stomach because with mouse stomach there is a lower frequency of heterophile reactions (Muller *et al.*, 1971) that could be misinterpreted as antibody specific for parietal cells.

Since complement binds the antibody to the cell cytoplasm, and autoantibodies to gastric parietal cells have been shown to be cytotoxic to these cells *in vitro* (De Aizpurua *et al.*, 1983b), both antibodies are potentially cytolytic *in vivo*. However, the rapid turnover of gastric parietal cells normally provides a reserve of cells far in excess of the number required for adequate function. Gastric parietal cell antibodies may also have an inhibitory effect on the secretion of acid by binding to the cell receptors for gastrin (Loveridge *et al.*, 1980). The cited frequency for which gastric antibodies in PA are detected serologically by immuno-

fluorescence is ~90%. Explanations for the residual 10% of seronegative cases could include faulty diagnosis, a "burning-out" of the autoimmune response as the autoantigen becomes exhausted, or dependence exclusively of the damaging process on cellular rather than humoral immunity. The problem of a frequency of positivity <100% for the relevant autoantibody is not confined to PA.

Human intrinsic factor is a glycoprotein with a molecular weight of ~44,000. Each molecule of intrinsic factor has the capacity to bind to one molecule of vitamin B_{12} (Chanarin, 1979). Two distinct antibodies are detected by radioimmunoassay; one reacts with the binding site for vitamin B_{12} and blocks subsequent binding of intrinsic factor with the free vitamin, and the other reacts with an antigenic determinant remote from this site (Rothenberg *et al.*, 1971; Samloff *et al.*, 1968; Samloff and Turner, 1968). Antibodies to the vitamin B_{12} binding site are demonstrable in serum of ~70% of patients with PA and to the other site in ~50%. These frequencies are greater if gastric juice is assayed.

Cell-mediated immune reactions in PA have been demonstrated *in vitro* by transformation of peripheral blood lymphocytes in the presence of human intrinsic factor, gastric juice, or an homogenate of gastric mucosa (Tai and McGuigan, 1969), and by inhibition of the migration of peripheral blood leucocytes in the presence of human gastric juice or intrinsic factor (Finlayson *et al.*, 1972; Fixa *et al.*, 1972; Goldstone *et al.*, 1973; Rose *et al.*, 1970). Glass (1977) also reported positive cutaneous delayed-type hypersensitivity reactions to gastric intrinsic factor and extracts of gastric mucosa. Some of the results are difficult to interpret, since the number of reactors was low (38%) in tests for lymphocyte transformation, and inhibition of leucocyte migration was as effective with liver mitochondria as it was with the gastric antigens (Goldstone *et al.*, 1973). However, Irvine *et al.* (1965) have shown very clearly and elegantly by electron microscopy that lymphocytes line up against the membranes of gastric parietal cells and chief cells, suggesting that these cells are actually responsible for the death of gastric mucosal cells.

B. Immunological Derangements

The detection in patients with PA of a high frequency of thyroid autoantibodies and a lower but clearly increased frequency of autoantibodies to pancreatic islet β cells, adrenal cortical cells, and ovary (Table III) is in keeping with the clinical association PA has with the autoimmune diseases which these autoantibodies specify. Also in accord with this special clustering of diseases is the increased frequency of gastric parietal cell and intrinsic factor antibodies observed in these diseases (Table II). Of

TABLE III
Frequency of Coexisting Autoantibodies in 90 Patients with Pernicious Anemia

Autoantibodies to	Number positive (%)
Thyroid microsomes	41 (46)
Thyroglobulin	9 (10)
Pancreatic islet cells	5 (6)
Adrenal cortical cells	4 (4)
Ovarian cells	2 (2)

interest is the most recent member to join the cluster, the Lambert–Eaton syndrome (Lang *et al.*, 1981; Lennon *et al.*, 1982). The high frequency of thyroid antibodies present in this syndrome, myasthenia gravis, and PA is illustrated in Fig. 4.

C. Mechanisms of Damage

There are two processes, both immunologically mediated, which lead to malabsorption of vitamin B_{12}, the determinant of pernicious anemia. The first is depletion of gastric parietal cells, which secrete gastric intrinsic factor, and the second is the blockade of intrinsic factor by autoantibody to it. Normally, intrinsic factor is secreted in volumes far in excess of that required to maintain absorption of vitamin B_{12}, but as immunological destruction of gastric parietal cells proceeds to gastric atrophy, the level of secretion falls below that required for complexing with dietary vitamin B_{12} and subsequent absorption of the complex. Over and above this effect, there are autoantibodies reactive with the receptor site for vitamin B_{12} on intrinsic factor. These are demonstrable in serum and gastric juice (Fisher *et al.*, 1966; Schade *et al.*, 1966), and interfere with the formation of the stable complex required for the transport of vitamin B_{12} from the stomach to the absorption site on epithelial cells in the distal ileum (Donaldson *et al.*, 1967; Kapadia *et al.*, 1983). This complex may be present exclusively with no free intrinsic factor detectable (Goldberg and Bluestone, 1970; Rose and Chanarin, 1969). Megaloblastic anemia and neuropathy are the direct effects of the ensuing deficiency of vitamin B_{12} (Chanarin *et al.*, 1981).

Following the exciting findings of earlier times on the role of vitamin B_{12} in development of anemia and the effect of autoimmunity on the gastric parietal cell and intrinsic factor, research on the pathogenesis of gastric atrophy has slowed in recent years. Explanations for the initiation and

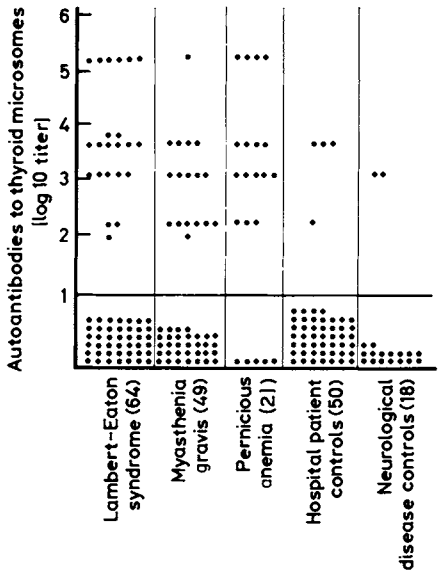


FIG. 4. There is increased frequency of autoantibodies to thyroid antigens in several diseases that cluster with pernicious anemia (PA). Results of tests for autoantibodies to thyroid microsomes in patients with the neuromuscular disorders Lambert–Eaton syndrome (LES) and myasthenia gravis (MG) are shown to illustrate the clustering of these diseases with pernicious anemia as compared with patients with other neurological diseases (predominantly amyotrophic lateral sclerosis in this study) and control patients with various nonimmunologically mediated diseases. The study illustrates that LES and MG overlap serologically with PA.

progression of the gastric lesion appear to lie in the understanding of two major effects, immunological and genetic.

To explain the lesion, we put forward two alternative (although not mutually exclusive hypotheses) that are based on the immunological mechanisms believed to be involved in the two autoimmune diseases most closely associated with PA; Graves' disease, in which an autoantibody to a cell-surface receptor (TSH receptor) is operative, and Hashimoto's thyroiditis, in which an immunoinflammatory and tissue-destructive process is operative (Fig. 5).

The first hypothesis, *the antireceptor antibody concept* (see Chapter 23), proposes that autoantibody to a receptor on gastric parietal cells is a major determinant of gastric atrophy. The idea that antibody to a cell receptor may account for the clinical expression of disease was first suggested by Simpson (1960) as an explanation for the features of myasthenia gravis. This hypothesis lay dormant for a decade until it was reintroduced

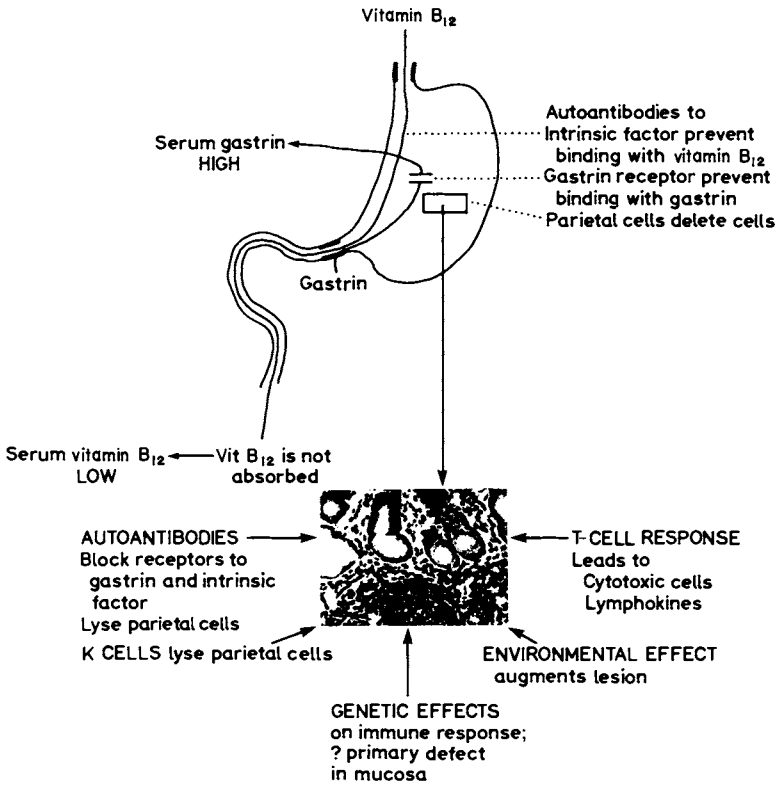


FIG. 5. A portrayal of the evolution of autoimmune gastritis and the mechanisms that may be operative in the cases in which there is evolution to the end stage of gastric atrophy.

by Lennon and Carnegie (1971) as a concept for explaining "immunopharmacological diseases." Thereafter, various clinical observations and experimental studies established that immunological reactions with cell receptors were implicated in a number of disease states (Carnegie and Mackay, 1975). Such processes could account for a number of findings in gastric parietal cell dysfunction.

In an experimental preparation it was shown that gastric parietal cell antibodies inhibited the response of the parietal cell to gastrin, pointing to blockade by antibody reactive with the cell-surface receptor for gastrin (Loveridge *et al.*, 1980). This inhibitory effect on the receptor may prevent proliferation and maturation of gastric parietal cells so leading to mucosal atrophy. One reason for the greater susceptibility of females to gastric atrophy could be the lower number in females of gastrin receptors on parietal cells (Johnson *et al.*, 1982) and therefore fewer binding sites

for gastrin. Since the inflammatory process in PA spares the antrum and the antrum is the site of production of gastrin (McGuigan and Trudeau, 1972), high levels of gastrin should be demonstrable in the serum of patients PA. In fact, raised levels of gastrin in serum were found in 75–80% of cases, even in patients in whom there is a considerable reserve of parietal cells (Ganguli *et al.*, 1971).

The increased level of serum gastrin in PA is a matter of much interest. One reason for this is the lack of hydrochloric acid in the stomach, since acid inhibits the secretion of gastrin. However, in a more general sense, the situation is analogous to certain other diseases marked either by interference with binding of the ligand at the receptor site, or by disappearance of cells bearing the receptor (Carnegie and Mackay, 1982): There is a compensatory increase in serum levels of the natural ligand, or otherwise stated, failure of "feedback inhibition." The mechanism is similar to that which occurs in the syndrome of insulin resistance associated with acanthosis nigricans, wherein the main effect of the antibody on the receptor for insulin is inhibitory (Kahn *et al.*, 1976), leading to profound resistance to insulin.

The alternative hypothesis, *the multiple gene concept,* proposes that gastric atrophy is the result of genetic defects involving the body mucosa and the immune system. This would be similar to that proposed for an avian model of spontaneous human autoimmune thyroiditis that occurs in a strain of chicken known as the obese strain (OS). Studies on this OS chicken suggest that three main gene effects are operative in the development of a disease that is a close analog of human Hashimoto's thyroiditis and may bear some resemblances to human thyrogastric autoimmune disease (see Chapter 1). The first genetic effect is operative on the thyroid gland; the second genetic effect is mediated through the major histocompatibility complex (MCH) (the *B* locus of the chicken), and the third promotes immunological hyperreactivity (Wick *et al.*, 1982).

Considering such effects in humans, there is no definite evidence for a primary defect in the parietal cell system in the stomach of patients predisposed to developing PA, although we can cite biopsy studies of the gastric mucosa of relatives of probands with PA showing an increased frequency of type A gastritis (Varis *et al.*, 1979) that has been attributed to an inherited defect (Varis, 1981). The apparent absence of PA in patients with duodenal ulcer with gastric parietal cell antibodies may be due to a structurally and functionally sound gastric mucosa (Ungar *et al.*, 1976). Considering the second gene effect, there is evidence in PA for increased frequencies of particular *HLA* antigens (Ungar *et al.*, 1981; Van den Berg-Loonen *et al.*, 1982). With respect to the third gene effect, there is height-

ened reactivity of the immune response to particular autoantigens. Presuming that the two main gene effects promoting the development of gastric atrophy of the PA type are a defect in the mucosa and a defect in the immune system, then both would be required to act in concert for expression of PA and gastric atrophy.

While antireceptor antibodies are possible participants in immunopathogensis, other effector mechanisms may contribute singly or in concert to the progressive inflammatory lesion. These are complement-dependent cytotoxic antibodies, cytotoxic damage by T lymphocytes, and antibody-dependent cellular cytotoxicity (ADCC) by killer (K) cells.

Whether gastric atrophy occurs alone or in association with other autoimmune diseases in the cluster would appear to be genetically determined. At least this is suggested by the reported high frequency of *HLA-DR2* in PA not associated with other autoimmune diseases, and of *HLA-DR3* in PA associated with other autoimmune diseases (Ungar *et al.*, 1981).

D. Immunogenetics

Heredity is thought to affect the development of PA and gastric atrophy, although the mode of inheritance is not understood. Evidence for genetic factors influencing the expression of PA includes clustering of PA in families, clustering of PA with other autoimmune diseases (Table I), racial predilection for subjects of northern European origin, and an increased frequency of particular genetic markers in patients with PA.

A number of Caucasian families have been reported in which there has been a high frequency of PA over several generations (Ardeman *et al.*, 1966b; Callender and Denborough, 1957; Doniach *et al.*, 1965; McIntyre *et al.*, 1959; Te Velde *et al.*, 1964; Wangel *et al.*, 1968a,b; Whittingham *et al.*, 1969), and there are also reports of concordance of PA in monozygotic twins (Delva *et al.*, 1965; Irvine *et al.*, 1965). However, discordance of PA has also been observed in monozygotic twins (Balcerzak *et al.*, 1968), suggesting that some unknown environmental factor has influenced the development of the disease.

Pernicious anemia is rare among southern Europeans and almost nonexistent among blacks and Asians (Irvine *et al.*, 1969; Jayaratnam *et al.*, 1967), and in keeping with racial differences, PA is claimed to be associated with phenotypic markers that are absent or occur with low frequency in the above racial groups. These markers include blue eyes, fair skin, blood group A (Callender and Denborough, 1957), and certain *HLA* markers. *HLA-B7* and B_{12} are increased in frequency in patients with severely

impaired vitamin B_{12} absorption, and *HLA-B8*, *-B18*, and *-Bw15* are increased in frequency in patients with autoimmune endocrinopathies (Ungar *et al.*, 1977). *HLA-DR4* is increased in frequency in all patients, but interestingly this increase is associated with increased frequencies of *DR2* in patients with PA alone and of *DR3* in patients with PA and autoimmune endocrinopathies. The increased frequency of *HLA-DR2* in PA has been independently corroborated (Van den Berg-Loonen *et al.*, 1982). The data of Ungar *et al.* (1981) suggested that *DR2* may protect against the development of autoimmune endocrinopathies in patients with PA.

E. Laboratory Diagnosis

The anemia can be demonstrated to be of the megaloblastic type by examination of blood and bone marrow (Chanarin, 1979). Of patients, 90% will have a positive immunofluorescence test for gastric parietal cell antibody and ~60% will have serum antibodies to gastric intrinsic factor. There may be coexisting autoantibodies specific for the various other autoimmune diseases in the thyrogastric cluster (Table II). At the stage of pernicious anemia, all patients will have a serum vitamin B_{12} level of <200 pg/ml (Ungar *et al.*, 1968); >95% will have a level of serum pepsinogen I <20 ng/ml (Samloff *et al.*, 1982), and 75–80% will have a serum gastrin level >100 pmol/liter (Varis *et al.*, 1979). In keeping with the sparing of the antrum from the inflammatory process, the level of pepsinogen II is normal and the low ratio (<1.0; normal 6.2) of pepsinogen I to pepsinogen II is predictive of the histological status of the gastric mucosa (Samloff *et al.*, 1982). At earlier stages of autoimmune chronic gastritis when the destructive lesion is "silent," there may be alterations in levels of these various laboratory indices suggestive of the disease process that can eventually lead to PA (Irvine *et al.*, 1974).

A biopsy of the body of the stomach stained by hematoxylin and eosin shows marked reduction of the secretory glands of the stomach or frank gastric atrophy (Figs. 1 and 3C), varying degrees of intestinalization (metaplasia) of the gastric glands (Strickland and Mackay, 1973) (Fig. 3C), and in the earlier lesions (Fig. 3B), prominence of lymphocytes and plasma cells in the inflammatory infiltrate in the lamina propria.

Acidity of gastric juice is lost due to functional impairment of gastric parietal cells. Absence of HCl from the gastric juice is an essential prerequisite for the diagnosis of PA (Chanarin, 1979). Confirmation of the diagnosis is by the Schilling test, in which it can be shown after a loading dose of vitamin B_{12} given intravenously that the excretion of orally administered vitamin B_{12} into the urine is <5% of the dose given (it is usually

<1%), and that following the giving of intrinsic factor with vitamin B_{12} the absorption of vitamin B_{12} is restored, as is shown by a urinary excretion 10–25% of the dose given.

VI. TREATMENT AND OUTCOME

The standard treatment is that of correction of deficiency of vitamin B_{12} and maintenance of an adequate level by parenteral administration of vitamin B_{12}. This corrects the anemia and may, particularly if given in high doses, correct the neuropathy. There is no necessity to give corticosteroid or immunosuppressive drugs, although in research studies these have been shown to ameliorate the gastric lesion.

There is historical evidence for a link between type A atrophic gastritis and its sequel, gastric atrophy, and adenocarcinoma of the stomach, but the association may not be strong. Only 2% of all gastric cancers were associated with PA in a large autopsy survey in Malmö, Sweden (Eriksson *et al.*, 1981), even when PA had been present >15 years. In a study by Walker *et al.* (1971) of cases of type A and type B gastritis, gastric cancer supervened only in patients with type B gastritis. None was observed among patients with type A gastritis. These findings were corroborated by Irvine *et al.* (1974). The studies of earlier years, which reported an increased risk (Kaplan and Riglar, 1945; Mosbech and Videbach, 1950; Elsborg and Mosbech, 1979), would not have taken account of the heterogeneity that exists among patients with atrophic gastritis.

VII. CONCLUDING REMARKS

Pernicious anemia and gastric atrophy are terminal events in a protracted chronic autoimmune atrophic gastritis that affects the body of the stomach and is usually expressed clinically in late middle age. The disease expression is the result of functional failure at the terminal phase of an organ-specific autoimmune disease rather than the disease process itself, and in that sense the names used are misleading. Moreover, these names *pernicious anemia* and *gastric atrophy* do not capture all the patients with autoimmune gastritis.

Immunologically, autoimmune atrophic gastritis is well defined. It fulfills the criteria for an organ-specific autoimmune disease: predominance in women, autoantibodies to gastric antigens, infiltration of mononuclear cells into the target organ with evidence of destruction, a regenerative

response of the affected tissue to corticosteroid and immunosuppressive drugs, familial predisposition, and association with other autoimmune diseases, which in this setting are mostly the autoimmune endocrinopathies. Questions as to whether a primary defect in the body mucosa as well as autoantibody blockade of the receptors for gastrin and intrinsic factor contributes to the development of the autoimmune process, and about the role of genes of the major histocompatibility complex on the inheritance of autoimmune gastritis, still remain to be answered.

VIII. SUMMARY

Pernicious anemia (PA) was first described in 1849 and named as such in 1872; it was later defined as a disease resulting from gastric secretory failure, then attributed to a deficiency state resulting in bone marrow failure, and finally shown to be due to an autoimmune gastritis. Two autoantibodies are demonstrable, one present in almost all cases reactive with gastric parietal cells, and the other present in some 60% of cases reactive with gastric intrinsic factor. The parietal cell autoantibody detected by immunofluorescence, or a parietal cell surface-reactive antibody, may interfere with the binding of gastrin to the receptor for gastrin on the parietal cell, or may result in cell lysis. Why chief cells also disappear from the gastric mucosa in the disease process is puzzling, since there is no demonstrable autoantibody to these cells. Cell-mediated immune processes in PA have not been convincingly delineated as a cause of cell damage. The BB/W rat and the obese chicken represent spontaneous animal models of autoimmune gastritis; induced animal models have been scarce and neither type of model has contributed substantially to the understanding of the disease.

The autoimmune type of gastritis, type A, has the following features which distinguish it from the other type, simple or type B gastritis, which presumably results from extrinsically induced damage; it clusters with other thyrogastric autoimmune diseases, affects the body of the stomach but not the antrum, and progresses to total glandular atrophy with achylia. This is accompanied by high levels of serum gastrin due to sparing of the antrum, failure of secretion of intrinsic factor, and gastric parietal cell autoantibodies. Questions to be answered include the modus operandi of the strong genetic influence, whether there is antibody blockade of the gastrin receptor, what causes the fault in maintenance of self-tolerance (a recurring question in autoimmunity), and how the autoimmune reaction(s) damage the target tissue.

REFERENCES

Abels, J., Bouma, W., and Nieweg, H. O. (1963). *Biochim. Biophys. Acta* **71,** 227–229.

Andrada, J. A., Rose, N. R., and Andrada, E. C. (1969). *Clin. Exp. Immunol.* **4,** 293–310.

Ardeman, S., and Chanarin, I. (1963). *Lancet* **2,** 1350–1354.

Ardeman, S., and Chanarin, I. (1965). *New Engl. J. Med.* **273,** 1352–1353.

Ardeman, S., Chanarin, I., Jacobs, A., and Griffiths, L. (1966a). *Blood* **27,** 599–610.

Ardeman, S., Chanarin, I., Krafchik, B., and Singer, W. (1966b). *Q. J. Med.* **35,** 421–431.

Baggett, R. T., and Welsh, J. D. (1970). *Am. J. Dig. Dis.* **15,** 871–880.

Balcerzak, S. P., Westerman, M. P., and Heinle, E. W. (1968). *Blood* **32,** 701–710.

Bardhan, K. D., Hall, J. R., Spray, G. H., and Callender, S. T. E. (1968). *Lancet* **1,** 62–64.

Berlin, R., Berlin, H., Brante, G., and Sjöberg, S.-G. (1958). *Acta Med. Scand.* **161,** 143–150.

Biermer, A. (1872). *Schweiz. Ärzteztg.* **2,** 15.

Blizzard, R. M., Clee, D., and Davis, W. (1966). *Clin. Exp. Immunol.* **1,** 119–128.

Blizzard, R. M., Clee, D., and Davis, W. (1967). *Clin. Exp. Immunol.* **2,** 19–30.

Bor, S., Feiwel, M., and Chanarin, I. (1969). *Br. J. Dermatol.* **81,** 83–88.

Brostoff, J., Bor, S., and Feiwel, M. (1969). *Lancet* **2,** 177–178.

Callender, S. T., and Denborough, M. A. (1957). *Br. J. Haematol.* **3,** 88–106.

Carnegie, P. R., and Mackay, I. R. (1975). *Lancet* **2,** 684–687.

Carnegie, P. R., and Mackay, I. R. (1982). *Springer Semin. Immunopathol.* **5,** 379–388.

Castle, W. B. (1953). *N. Engl. J. Med.* **249,** 603–614.

Chanarin, I. (1979). "The Megaloblastic Anaemias," 2nd ed. Blackwell, Oxford.

Chanarin, I., Deacon, R., Perry, J., and Lumb, M. (1981). *Br. J. Haematol.* **47,** 487–491.

Cowling, D. C., Strickland, R. G., Ungar, B., Whittingham, S., and Rose, W. McI. (1974). *Med. J. Aust.* **1,** 15–17.

De Aizapurua, H. J., Toh, B. H., and Ungar, B. (1983a). *Clin. Exp. Immunol.* **52,** 341–349.

De Aizapurua, H. J., Cosgrove, L. J., Ungar, B., and Toh, B. H. (1983b). *N. Engl. J. Med.* **309,** 625–629.

Delva, P. L., Macdonald, J. E., and Macintosh, D. C. (1965). *Can. Med. Assoc. J.* **92,** 1129–1131.

Doig, A., Girdwood, R. H., Duthie, J. J. R., and Knox, J. D. E. (1957). *Lancet* **2,** 966–972.

Donaldson, R. M., Mackenzie, I. L., and Trier, J. S. (1967). *J. Clin. Invest.* **46,** 1215–1228.

Doniach, D., Roitt, I. M., and Taylor, K. B. (1965). *Ann. N. Y. Acad. Sci.* **124,** 605–625.

Elder, M., Maclaren, N., Riley, W., and McConnell, T. (1982). *Diabetes* **31,** 313–318.

Elsborg, L., and Mosbech, J. (1979). *Acta Med. Scand.* **206,** 315–318.

Eriksson, S., Clase, L., and Moquist-Olsson, I. (1981). *Acta Med. Scand.* **210,** 481–484.

Fenwick, S. (1870). *Lancet* **2,** 78–80.

Finlayson, N. D. C., Fauconnet, M. H., and Krohn, K. (1972). *Am. J. Dig. Dis.* **17,** 631–638.

Fisher, J. M., Rees, C., and Taylor, K. B. (1966). *Lancet* **2,** 88–89.

Fixa, B., Vejbora, O., Komarkova, O., Lanr, F., and Parizek, J. (1964). *Gastroenterologia* **102,** 331–338.

Fixa, B., Thiele, H. G., Komarkovs, O., and Nozick, A. Z. (1972). *Scand. J. Gastroenterol.* **7,** 237–240.

Ganguli, P. C., Cullen, D. R., and Irvine, W. J. (1971). *Lancet* **1,** 155–158.

Glass, G. B. J. (1963). *Physiol. Rev.* **43,** 529–849.

Glass, G. B. J. (1977). *N. Y. State J. Med.* **77,** 1697.

Goldberg, L. S., and Bluestone, R. (1970). *J. Lab. Clin. Med.* **75,** 449–456.
Goldstone, A. H., Calder, A. E., Barnes, E. W., and Irvine, W. J. (1973). *Clin. Exp. Immunol.* **14,** 501–508.
Gordin, R. (1959). *Acta Med. Scand.* **164,** 159–165.
Gottlieb, C., Lau, K.-S., Wasserman, L. R., and Herbert, V. (1965). *Blood* **25,** 875–885.
Guttmann, L., Crosby, T. W., Takamori, M., and Martin, J. D. (1972). *Am. J. Med.* **53,** 354–356.
Hennes, A. R., Sevelius, H., Lewellen, T., Joel, W., Woods, A. H., and Wolf, S. (1962). *Arch. Pathol.* **73,** 281–287.
Highley, D. R., Davies, M. C., and Ellenbogen, K. (1967). *J. Biol. Chem.* **242,** 1010–1015.
Hoedemaeker, P. J., and Ito, S. (1970). *Lab. Invest.* **22,** 184–188.
Hooper, B., Whittingham, S., Mathews, J. D., Mackay, I. R., and Curnow, D. H. (1972). *Clin. Exp. Immunol.* **12,** 79–87.
Hughes, W. S., Brooks, F. P., and Conn, H. O. (1972). *Ann. Intern. Med.* **77,** 746–750.
Irvine, W. J. (1965). *N. Engl. J. Med.* **273,** 432–438.
Irvine, W. J. (1975). *Clin. Endocrinol. Metab.* **4,** 351–377.
Irvine, W. J. (1978). *In* "Immunological Diseases" (M. Samter, ed.), 3rd ed. Vol. 2, p. 1284. Little, Brown, Boston, Massachusetts.
Irvine, W. J., and Barnes, E. W. (1974). *J. Reprod. Fertil., Suppl.* **21,** 1–31.
Irvine, W. J., and Barnes, E. W. (1975). *Clin. Endocrinol. Metabol.* **4,** 379–434.
Irvine, W. J., Davies, S. H., Delamore, I. W., and Williams, A. W. (1962). *Br. Med. J.* **2,** 454–456.
Irvine, W. J., Davies, S. H., Teitelbaum, S., Delamore, I. W., and Williams, A. W. (1965). *Ann. N. Y. Acad. Sci.* **124,** 657–691.
Irvine, W. J., Chan, M. M. W., Scarth, L., Kolb, F. O., Hartog, M., Bayliss, R. I. S., and Doury, M. I. (1968). *Lancet* **2,** 883–887.
Irvine, W. J., McFadzean, A. J. S., Todd, D., Tso, C. S., and Yeung, R. T. T. (1969). *Clin. Exp. Immunol.* **4,** 375–386.
Irvine, W. J., Clarke, B. F., Scarth, L., Cullen, D. R., and Duncan, L. J. P. (1970). *Lancet* **2,** 163–168.
Irvine, W. J., Cullen, D. R., and Mawhinney, H. (1974). *Lancet* **1,** 482–485.
Jacob, E., and Schilling, R. G. (1966). *J. Lab. Clin. Med.* **67,** 510–515.
Jayaratnam, F. J., Seah, C. S., Da Costa, J. L., Tan, K. K., and O'Brien, W. (1967). *Br. Med. J.* **3,** 18–20.
Jeffries, G. J., Hoskins, D. W., and Sleisenger, M. H. (1962). *J. Clin. Invest.* **41,** 1106–1115.
Jeffries, G. J., Todd, J. E., and Sleisenger, M. H. (1966). *J. Clin. Invest.* **45,** 803–812.
Johnson, L. R., Peitsch, W., and Takeuchi, K. (1982). *Am. J. Physiol.* **243,** 469–474.
Jorge, A. D., and Sanchez, D. (1973). *Gut* **14,** 104–106.
Kahn, C. R., Flier, J. S., Bar, R. S., Archer, J. A., Gorden, P., Martin, M. M. and Ruth, J. (1976). *N. Engl. J. Med.* **294,** 739–745.
Kapadia, C. R., Serfilippi, D., Voloshin, K., and Donaldson, R. M. (1983). *J. Clin. Invest.* **71,** 440–448.
Kaplan, H. S., and Riglar, L. G. (1945). *Am. J. Med. Sci.* **209,** 339–348.
Kekki, M., Varis, K., Pohjanpalo, H., Isokoski, M., Ihamäki, T., and Siurala, M. (1983). *Dig. Dis. Sci.* **28,** 698–704.
Khoury, E. L., Bottazzo, G. F., Pontes de Carvalho, L. C., Wick, G., and Roitt, I. M. (1982). *Clin. Exp. Immunol.* **49,** 273–282.
Kristensen, H. P., and Friis, T. (1962). *Acta Med. Scand.* **168,** 457–459.
Krohn, K. J. E., and Findlayson, N. D. C. (1973). *Clin. Exp. Immunol.* **14,** 237–245.

Lang, B., Newsom-Davis, J., Wray, D., Vincent, A., and Murray, N. (1981). *Lancet* **2,** 224–226.

Lennon, V. A., and Carnegie, P. R. (1971). *Lancet* **1,** 630–633.

Lennon, V. A., Lambert, E. H., Whittingham, S., and Fairbanks, V. (1982). *Muscle Nerve, Suppl.* pp. 21–25.

Lester-Smith, E. (1948). *Nature (London)* **161,** 638–639.

Loveridge, N., Bitensky, L., Chayen, J., Hausamen, T. U., Fisher, J. M., Taylor, K. B., Gardner, J. D., Bottazzo, G. F., and Doniach, D. (1980). *Clin. Exp. Immunol.* **41,** 264–270.

McGuigan, J. E., and Trudeau, W. L. (1972). *N. Engl. J. Med.* **286,** 184–188.

McIntyre, P. A., Hahn, Conley, C. L., and Glass, G. B. J. (1959). *Johns Hopkins Hosp. Bull.* **104,** 309–342.

Mackay, I. R. (1971). *In* "Immunological Diseases" (M. Samter, ed.), 2nd ed., Vol. 2, pp. 1323–1345. Little, Brown, Boston, Massachusetts.

Mackay, I. R., and Burnet, E. M. (1963). "Autoimmune Disease: Pathogenesis, Chemistry and Therapy." Thomas, Springfield, Illinois.

Markson, J. L., and Moore, J. M. (1962). *Lancet* **2,** 1240–1243.

Maron, R., and Cohen, I. R. (1980). *J. Exp. Med.* **152,** 1115–1120.

Masala, C., Smurra, G., De Prima, M. A., Amendolea, M. A., Celestino, D., and Salsano, F. (1980). *Clin. Exp. Immunol.* **41,** 271–280.

Minot, G. R., and Murphy, W. P. (1926). *J. Am. Med. Assoc.* **87,** 470–476.

Mosbech, J., and Videbach, A. (1950). *Br. Med. J.* **2,** 390–394.

Muller, H. K., McGiven, A. R., and Nairn, R. C. (1971). *J. Clin. Pathol.* **24,** 13–14.

Penhale, W. J., Farmer, A., and Irvine, W. J. (1975). *Clin. Exp. Immunol.* **21,** 362–375.

Rickes, E. L., Brink, N. G., Konivsky, F. R., Wood, T. R., and Folkers, K. (1948). *Science* **107,** 396–397.

Riley, W. J., Toskes, P. P., Maclaren, N. K., and Silverstein, J. H. (1982). *Diabetes* **31,** 1051–1055.

Rødbro, P., Dige-Petersen, H., Schwartz, M., and Dalgaard, O. Z. (1967). *Acta Med. Scand.* **181,** 445–452.

Rose, M. S., and Chanarin, I. (1969). *Br. Med. J.* **1,** 468–470.

Rose, M. S., Chanarin, I., Doniach, D., Brostoff, J., and Ardeman, S. (1970). *Lancet* **2,** 9–13.

Rothenberg, S. P., Kajani Kantha, K. R., and Ficarra, A. (1971). *J. Lab. Clin. Med.* **77,** 476–484.

Samloff, I. M., and Turner, M. D. (1968). *J. Immunol.* **101,** 578–586.

Samloff, I. M., Kleinman, M. S., Turner, M. D., Sohel, M. V., and Jeffries, G. H. (1968). *Gastroenterology* **55,** 575–583.

Samloff, I. M., Varis, K., Ihamaki, T., Siurala, M., and Rotter, J. I. (1982). *Gastroenterology* **83,** 204–209.

Schade, S. G., Feick, P. L., Muckerheide, M., and Schilling, R. F. (1966). *N. Engl. J. Med.* **275,** 528–531.

Schiller, K. F. R., Spray, G. H., Wangel, A. G., and Wright, R. (1968). *Q. J. Med.* **37,** 451–462.

Schwartz, M. (1958). *Lancet* **2,** 61–62.

Schwartz, M. (1960). *Lancet* **2,** 1263–1267.

Serafini, U., Masala, C., and Pala, A. M. (1970). *In* "Proceedings of the Seventh International Congress of Allergology" (U. Serafini, A. W. Frankland, C. Masala, and J. M. Jamar, eds.) pp. 338–349. Excerpta Medica, Amsterdam.

Sharp, E. A. (1929). *J. Am. Med. Assoc.* **93,** 749–751.

Simpson, J. A. (1960). *Scott. Med. J.* **5,** 419–436.
Strickland, R. G., and Mackay, I. R. (1973). *Am. J. Dig. Dis.* **18,** 426–440.
Strickland, R. G., Fisher, J. M., and Taylor, K. B. (1969). *Gastroenterology* **56,** 675–686.
Sturgis, C. C., and Isaacs, R. (1929). *J. Am. Med. Assoc.* **93,** 747–749.
Tai, C., and McGuigan, J. E. (1969). *Blood* **34,** 63–71.
Tanaka, N., and Glass, G. B. J. (1970). *Gastroenterology* **58,** 482–493.
Taylor, K. B. (1959). *Lancet* **2,** 106–108.
Taylor, K. B., and Morton, J. A. (1958). *Lancet* **1,** 29–30.
Taylor, K. B., Roitt, I. M., Doniach, D., Couchman, K. G., and Shapland, C. (1962). *Br. Med. J.* **2,** 1347–1354.
Te Velde, K., Abels, J., Anders, G. J. P. A., Arends, A., Hoedemaeker, P. J., and Nieweg, H. O. (1964). *J. Lab. Clin. Med.* **64,** 177–187.
Tomazic, V., Rose, N. R., and Shreffler, D. C. (1974). *J. Immunol.* **112,** 965–969.
Tudhope, G. R., and Wilson, G. M. (1960). *Q. J. Med.* **29,** 513–537.
Twomey, J. J., Jordon, P. H., Jarrold, T., Trubowitz, S., Ritz., N. D., and Conn, H. O. (1969). *Am. J. Med.* **47,** 340–350.
Twomey, J. J., Jordon, P. H., Laughter, A. H., Menwissen, H. J., and Good, R. A. (1970). *Ann. Intern. Med.* **72,** 499–504.
Ungar, B. (1967). *Aust. J. Exp. Biol. Med. Sci.* **45,** 317–321.
Ungar, B. (1968). *Aust. Ann. Med.* **17,** 107–109.
Ungar, B., Stocks, A. E., Martin, F. I. R., Whittingham, S., and Mackay, I. R. (1968). *Lancet* **2,** 415–418.
Ungar, B., Francis, C. M., and Cowling, D. C. (1976). *Med. J. Aust.* **2,** 900–902.
Ungar, B., Mathews, J. D., Tait, B. D., and Cowling, D. C. (1977). *Br. Med. J.* **1,** 798–800.
Ungar, B., Mathews, J. D., Tait, B. D., and Cowling, D. C. (1981). *Br. Med. J.* **1,** 768–770.
Van den Berg-Loonen, E. M., Hitterman, T. C. M., Bins, M., Nijenhuis, L. E., and Engelfriet, P. (1982). *Tissue Antigens* **19,** 158–160.
Varis, K. (1981). *Ann. Clin. Res.* **13,** 123–129.
Varis, K., Ihamäki, T., Härkönen, M., Samloff, I. M., and Suirala, M. (1979). *Scand. J. Gastroenterol.* **14,** 129–139.
Walker, I. R., Strickland, R. G., Ungar, B., and Mackay, I. R. (1971). *Gut* **12,** 906–911.
Wall, A. J., Whittingham, S., Mackay, I. R., and Ungar, B. (1968). *Clin. Exp. Immunol.* **3,** 359–366.
Wangel, A. G., Callender, S. T., Spray, G. H., and Wright, R. (1968a). *Br. J. Haematol.* **14,** 161–182.
Wangel, A. G., Callenger, S. T., Spray, G. H., and Wright, R. (1968b). *Br. J. Haematol.* **14,** 183–204.
Ward, H. A., and Nairn, R. C. (1972). *Clin. Exp. Immunol.* **10,** 435–451.
Whittingham, S., Ungar, B., Mackay, I. R., and Mathews, J. D. (1969). *Lancet* **1,** 951–954.
Whittingham, S., Mathews, J. D., Mackay, I. R., Stocks, A. E., Ungar, B., and Martin, F. I. R. (1971). *Lancet* **1,** 763–767.
Wick, G., Boyd, R., Hála, K., Thunold, S., and Kofler, H. (1982). *Clin. Exp. Immunol.* **47,** 1–18.
Wilkinson, J. F. (1949). *Lancet* **1,** 249–255.
Wood, I. J., Doig, R. K., Motteram, R., and Hughes, A. (1949). *Lancet* **1,** 18–21.

CHAPTER 11

Inflammatory Bowel Diseases

STEPHEN B. HANAUER
SUMNER C. KRAFT
Department of Medicine
The University of Chicago
Chicago, Illinois

I. INTRODUCTION

Ulcerative colitis and Crohn's disease are acute and chronic inflammatory bowel disorders of unknown etiology that must be excluded from enterocolitides that are due to specific infections, ischemia, irradiation, or toxins. Whereas ulcerative colitis is limited to the colon and usually occurs in a continuous distribution extending proximally from the rectum, Crohn's disease may involve any portion of the alimentary canal from the lips to the anus, often in a discontinuous or segmental distribution (Kirsner

and Shorter, 1982). Both disorders often are accompanied by similar extraintestinal lesions and respond favorably to identical medical therapies. Many of the clinical manifestations of these inflammatory bowel diseases suggest the participation of primary or secondary immunological mechanisms (Kraft and Kirsner, 1971; Greenstein *et al.*, 1976).

Ulcerative colitis involves the superficial colonic mucosa in a diffuse fashion. When acute, the patient may be seriously ill with 20 or more loose and often bloody stools per day; frequent complications include dehydration, anemia, hypoproteinemia, and fever. The illness often follows a chronic course with intermittent exacerbations, but there is great individual variability as regards the duration of remissions. Proctosigmoidoscopic abnormalities include edema, hyperemia, granularity, petechial hemorrhages, pinpoint ulcerations that may become confluent, friability on cotton wiping, and, in more severe cases, a mucopurulent, sanguineous exudate. The disease may be limited to the distal rectum (proctitis) or involve the entire colon (pancolitis). Except for occasional radiographic evidence of a few centimeters of clinically insignificant "backwash ileitis," easily differentiable from Crohn's disease of the ileum, ulcerative colitis never leads to small intestinal disease. It is thus surgically curable by total proctocolectomy. Medical management includes such agents as sedatives, antidiarrheal drugs, sulfasalazine, and topical or systemic corticosteroids (Kirsner and Shorter, 1982).

Crohn's disease is a persistent or relapsing, transmural, granulomatous inflammatory process resulting in strictures, fissured ulcers, sinuses, and fistulas. The inflammatory process occurs in a discontinuous distribution, the focal nature of the process discriminating Crohn's colitis from ulcerative colitis in ~85% of cases. Clinical symptoms and signs include abdominal pain, diarrhea, malnutrition, inflammatory masses, intestinal obstruction, and severe anal disease. Although the inflammatory process may be confined to the small intestine and most frequently involves the terminal ileum, any combination of colonic and more proximal involvement may occur. The rectum frequently is spared but may demonstrate patchy, fine, or coarse granularity or shallow aphthoid ulcers. Proctoscopic or colonoscopic biopsies may reveal a focal inflammatory infiltrate, whereas in ulcerative colitis the microscopic changes are diffuse. Noncaseating, acid fast-negative, sarcoid-type granulomas are detectable in up to 50% of cases, representing pathognomonic evidence of Crohn's disease (Schachter and Kirsner, 1980; Kirsner and Shorter, 1982).

Hypotheses attempting to explain the etiology of inflammatory bowel disease have implicated infectious agents, toxins, destructive enzymes, deficiency states, motility changes, vascular disturbances, autonomic nervous system imbalance, psychogenic mechanisms, metabolic defects,

connective tissue disorders, abnormal mucosal regeneration, and immunological phenomena (Kraft, 1979; Kirsner and Shorter, 1982). The length of this list correlates well with the frustration that investigators have had in attempting to dissect the nature of these conditions.

Indirect evidence implicating immunological mechanisms in the etiology of inflammatory bowel disease includes the frequency of personal and family histories of known allergic disorders, the increasing family clustering of cases, the occasional association with other disorders often linked with immunological mechanisms, the frequency of the disease among young people, and the favorable response to antiinflammatory or immunosuppressive agents (Kraft, 1979). Antigenic responses to intestinal products and contents certainly may generate some of the enteric and extraintestinal features of inflammatory bowel disease, as has been implicated following jejunoileal bypass surgery (Thayer and Kirsner, 1980). Whereas the granulomatous inflammation in Crohn's disease suggests the possibility of tissue hyperreactivity to luminal antigens, hyposplenism and other causes of a decreased ability to handle immune stimuli may limit host defenses and increase septic complications in some cases (Palmer *et al.*, 1981).

While impressive progress has been made over the past two decades in describing the immunological features of inflammatory bowel disease, there is no unifying hypothesis that incorporates all of the available information, and the broad overlap of immunological features does not exclude the possibility that ulcerative colitis and Crohn's disease each may consist of etiologically heterogeneous subgroups (Kirsner and Shorter, 1982). This chapter highlights the major studies of humoral and cellular immune phenomena in these inflammatory bowel diseases.

II. HUMORAL IMMUNE PHENOMENA

The presence of lymphocytes, plasma cells, eosinophils, and mast cells in colonic tissues from patients with ulcerative colitis is consistent with the participation of humoral immunological mechanisms (Kraft, 1979). The additional presence of macrophages, which are known to play an important role in the sequence of events leading to humoral antibody production, sets the stage for the potential induction, perpetuation, and termination of local antigen–antibody reactions (Sommers and Korelitz, 1975). In Crohn's disease, the early tissue changes also include increased proportions of lymphocytes and plasma cells in the lamina propria, in close proximity to macrophages and often degranulated tissue mast cells (Hanauer and Kraft, 1983).

Although ulcerative colitislike syndromes and Crohn's disease have been described in rare patients with major defects of the humoral immune system, which perhaps argues against an important pathogenic role for humoral-type immune processes in the production of these intestinal lesions, neither ulcerative colitis nor Crohn's disease is a frequent component of the recognized immunodeficiency syndromes. Indeed, the integrity of the humoral immune system in most patients with inflammatory bowel disease would appear evidenced by normal or elevated serum gamma globulin concentrations and apparently normal antibody responses to standard bacterial antigens, common upper-respiratory viral agents, and enteric pathogens (Kraft, 1979).

It is well known that the intestine plays important homeostatic roles in regulating levels of albumin and other plasma proteins, and that depletions of some of these components occur in patients with inflammatory bowel disease. Unfortunately, the fact that these changes often are merely secondary to the intestinal lesion has not always been given adequate consideration. Depletion of the "building blocks" of both the humoral and cell-mediated immune systems occurs in a variety of exudative enteropathies, resulting in immunological observations that are essentially epiphenomena; in addition, some are related to the associated malnutrition (Hanauer and Kraft, 1983). In reviewing the results of such studies of the immune system, one also must consider that heterogeneous populations of patients have been reported, highly variable laboratory methods have been used, and that immunological concepts are constantly being updated.

A. Serum Immunoglobulins

Serum concentrations of the major immunoglobulin isotypes follow no predictable pattern in ulcerative colitis, bearing no consistent relationship to the state of activity, extent, or severity of the disease (Kraft, 1979). Serial studies in individual patients may be more revealing than an assessment at a single point in time. While the serum IgA concentration may be elevated when symptoms of ulcerative colitis have been present for >10 years, this value may increase further in some patients following total proctocolectomy (Hardy Smith and Macphee, 1971). Although serum IgG and IgM levels also may be increased, a few patients with active ulcerative colitis have had elevated fractional catabolic rates for these two immunoglobulins on the basis of isotope turnover studies (Bendixen *et al.*, 1970).

Aside from rare instances of associated selective IgA deficiency, serum IgA levels in patients with Crohn's disease usually are normal; they tend

to increase with extensive involvement of the small bowel, the colon, or both (Hodgson and Jewell, 1978). The levels of IgG in the serum often are normal (Hodgson and Jewell, 1978), but may be increased (e.g., with extensive colonic disease or liver involvement) or decreased (perhaps as a result of increased catabolism, due in part to intestinal protein loss) (Bendixen *et al.*, 1968). Circulating concentrations of IgM also may be elevated in patients with marked colonic involvement, with corticosteroid usage, or in the presence of extraintestinal manifestations (Hodgson and Jewell, 1978), while normal or low levels of IgM have been associated with disease chronicity (Weeke and Jarnum, 1971). The demonstration that patients with clinically inactive Crohn's disease may possess "covert" suppressor T cells in the peripheral blood raises the possibility that additional instances of incipient or actual humoral immune deficiency will surface when new methods become available for characterizing the regulatory T-cell system (Elson *et al.*, 1981; MacDermott *et al.*, 1981).

B. Circulating B Cells

Patients with ulcerative colitis have been reported to have increased proportions and absolute numbers of peripheral-blood B lymphocytes, especially those positive for surface IgA and IgM (Strickland *et al.*, 1974). Of course, such B-cell data may be quite unrelated to circulating and secretory immunoglobulin measurements in the same individuals.

Initial studies in Crohn's disease demonstrated increased peripheral-blood B cells bearing surface-membrane IgA and IgM; these observations were unrelated to the activity, site, and duration of disease and to the mode of treatment. In more recent work (Auer *et al.*, 1978b), the proportion of circulating B cells was shown to be significantly above normal in both acute Crohn's disease without prior therapy and chronic drug-treated cases. In the first group, the absolute numbers of B cells also were considerably increased, diminishing with disease chronicity and treatment. These findings were extended to demonstrate that the proportion of B cells bearing IgM was significantly higher in short-duration, untreated Crohn's disease than in chronically treated patients (Auer *et al.*, 1979). Obviously, concomitant studies of B cells in the circulation and in the affected tissues of carefully defined subgroups of patients with inflammatory bowel disease are indicated. It also must be kept in mind that techniques for determining B-cell numbers by staining for surface-membrane immunoglobulins may give falsely elevated values due to the *in situ* nonspecific binding of cytophilic antibody or the attachment of immune complexes to surface Fc receptors.

C. Tissue Immunoglobulins

In analyzing the humoral immune components of intestinal tissues, consideration again must be given to the techniques used to quantitate immunoglobulin-producing cells (Brandtzaeg and Baklien, 1976), assess B-cell surface-membrane characteristics (Rosekrans *et al.*, 1980), or isolate and culture viable lymphoid cells for measurement of immunoglobulin synthesis (Bland *et al.*, 1979). It also is important to be able to correlate such data with the anatomical location and degree of tissue inflammation (Brandtzaeg and Baklien, 1979).

In comparison to control subjects without overt inflammation, colonic tissue specimens from patients with active ulcerative colitis have shown increased numbers of IgA-, IgM-, and IgG-containing plasma cells. The IgG cells were especially prominent in the submucosa, and serial follow-up studies suggested that these cells tend to persist during the inactive stage of the disease (Baklien and Brandtzaeg, 1975). Yet, others have found no significant differences between the numbers of IgA, IgG, and IgM cells in colonic tissues from patients with inactive ulcerative colitis and control subjects (Skinner and Whitehead, 1974).

Early studies of patients with terminal ileal Crohn's disease demonstrated increased numbers of immunoglobulin-containing cells and larger local concentrations of the corresponding immunoglobulins compared to control ileal tissues (Persson and Danielsson, 1973; Persson, 1974). In later work, the total numbers of lymphoid cells containing IgA, IgM, and IgG in Crohn's disease tissues were found to be increased by 2- to 60-fold the numbers detected per mucosal tissue unit in normal specimens. The greater proportions of IgG cells were most striking at sites of pronounced inflammation, in areas of fissure formation, and especially surrounding lymphoid nodules and granulomas (Brandtzaeg and Baklien, 1976, 1979). Others also have described increased numbers of IgM cells in nonulcerated but affected areas of bowel from patients with Crohn's disease (Rosekrans *et al.*, 1980).

The distribution of IgE cells in the colonic mucosa in ulcerative colitis requires additional correlation with disease activity, since findings of large numbers of such cells in some patients have not been noted by other workers (O'Donoghue and Kumar, 1979). Similar controversies revolve around the presence of IgE cells in the intestinal mucosa in Crohn's disease (Baklien and Brandtzaeg, 1975). While concentrations of both IgE and IgA have been normal in the intestinal fluids obtained from patients with Crohn's disease (Jones *et al.*, 1976), quantitative differences may be less important than are the structural and functional characteristics of these secretory immunoglobulins.

In vitro kinetic studies of tissue immunoglobulin secretion have demonstrated increased IgA and IgG synthesis in rectal mucosal biopsies from untreated patients with ulcerative colitis, in comparison with normal controls and patients with bacterial enterocolitis (Fiorilli *et al.*, 1975). Mononuclear cells separated from resected Crohn's disease specimens synthesized IgG at a rate increased 10-fold over that of intestinal mucosal lymphoid cells isolated from control tissues (Bookman and Bull, 1979); however, others found a decreased secretion of IgG and IgM by Crohn's disease and ulcerative colitis intestinal mononuclear cells (MacDermott *et al.*, 1981). Synthesis of IgA was normal in both studies.

D. Circulating and Tissue Antibodies

1. Anticolon Antibodies

Numerous antibodies that react against host or foreign antigens have been detected in sera from patients with inflammatory bowel disease. The specificity, isotype, prevalence, and other characteristics of these antibodies clearly are of interest and potential importance. Yet, any discussion of the immunology of inflammatory bowel disease is handicapped by uncertainty regarding the nature of the antigen(s) involved. In all disorders involving an immunological pathogenesis, antibodies are necessary but not alone sufficient to bring about the untoward physiological response and tissue reaction; a proper understanding of any type of immunological tissue damage requires adequate antigenic identification (Farr, 1963).

The original serological finding of an increased frequency of hemagglutinating antibodies against colon mucosa in patients with ulcerative colitis stimulated much effort to implicate specific colon antigens and antibodies in this disease. It is not surprising that there has been much heterogeneity among the various anticolon antibody populations described to date; the many colonic antigens that have been identified possess varying degrees of digestive tract, organ, and species specificity (Kraft, 1979). Early studies suggested that the colonic antigenic determinants were carbohydrate in nature, probably involving a gastrointestinal mucopolysaccharide as found in mucin (Broberger and Perlmann, 1959). The search for a more available source of antigen led to the finding that extracts of germ-free rat colon and feces also reacted with sera from certain patients with ulcerative colitis (Perlmann *et al.*, 1965). The involved antigen appeared to be immunologically related to the human colonic antigen and additionally shared determinants with a lipopolysaccharide extractable from *Escheri-*

chia coli 014, a heterogenetic constituent of most Enterobacteriaceae strains that is called the common antigen of Kunin (Kraft, 1979). Although immunofluorescent staining initially had not detected circulating anticolon antibodies in patients with Crohn's disease, subsequent work demonstrated serum antibodies to germ-free rat colon/feces and *E. coli,* 014 as frequently as in patients with ulcerative colitis (Lagercrantz *et al.,* 1966; Thayer *et al.,* 1969).

The significance of anticolon antibodies in the pathogenesis of inflammatory bowel disease remains unclear. Although capable of inhibiting DNA synthesis by colonic mucosal epithelial cells in guinea pigs and rats, sera containing anticolon antibodies are not cytotoxic in cultures of human fetal colon cells. Furthermore, specific anticolon antibodies have not been demonstrated in the affected tissue in inflammatory bowel disease, and their presence in the circulation has not correlated with the extent, severity, duration, and course of disease, the presence of extraintestinal complications, or with corticosteroid treatment (Kraft, 1979). The demonstration of antitissue antibodies does not necessarily indicate an immunological basis for any disease, but may reflect an epiphenomenon secondary to tissue damage—similar to the detection of circulating antibodies to heart antigens after an uncomplicated myocardial infarction. The possibility also exists that antibodies to tissue elements may assist in healing by facilitating the removal of debris, perhaps involving interactions with complement and leukocytes. By using an antibody-dependent cell-mediated cytotoxicity assay, sera from some patients with ulcerative colitis were shown to contain an antibody with lytic activity against a colonic cancer cell line (Nagai and Das, 1981).

2. Antibacterial Antibodies

There appears to be no paucity of serum antibodies against microbial antigens in patients with inflammatory bowel disease. In addition to the possible consequences of their sharing antigenic determinants, bacteria could affect the intestinal epithelium by directly damaging mucosal cells, by altering the configuration of the indigenous colonic antigens, or by attaching as in a hapten combination. Circulating antibodies to various *E. coli* O antigens, to cell wall-deficient *Pseudomonas* variants, and to other microorganisms have been demonstrated in both ulcerative colitis and Crohn's disease (Kraft, 1979; Hanauer and Kraft, 1983). Additionally, serum agglutinins to *Eubacterium* and *Peptostreptococcus* species in patients with Crohn's disease have been correlated with localization of the disease to the colon, the presence of fistulas, and serum levels of im-

munoglobulins, but not with the degree of disease activity (van de Merwe *et al.*, 1982).

3. Antiviral Antibodies

The possibility of a viral etiology for inflammatory bowel disease has been under intermittent investigation for >30 years and continues to be considered. A serological survey for antibodies against a spectrum of viral antigens demonstrated that only antibodies to cytomegalovirus were present more frequently and in significantly higher titers in patients with ulcerative colitis than in matched control subjects (Farmer *et al.*, 1973). Yet, subsequent studies have failed to support a role for this virus in the pathogenesis of the disorder, and circulating antibodies to Epstein–Barr virus, rotavirus, or Norwalk agent have not been more prevalent in patients with inflammatory bowel disease (Greenberg *et al.*, 1979). Nevertheless, indirect evidence for viral involvement comes from studies showing that sera obtained from patients with ulcerative colitis and Crohn's disease, as well as from their relatives, contain antibodies that bind to single- and double-strand RNA, but not to DNA, more often than do sera from control subjects and their families. Furthermore, patients with either ulcerative colitis or Crohn's disease and their spouses appear to possess circulating cold-reactive lymphocytotoxic antibody (similar to findings in several liver disorders and in diseases of known viral etiology) more frequently than do matched control groups (Strickland *et al.*, 1977; De-Horatius *et al.*, 1978). Although this heterogeneous antibody is predominantly of the IgM isotype, is not thought to be directed against conventional HLA antigenic determinants, and is especially reactive with B-cell antigens, it seems unlikely that it causes *in vivo* tissue destruction or even represents a marker of a viral agent (Kuiper *et al.*, 1981).

4. Miscellaneous Antibodies

Dietary antigens remain of theoretical etiological significance in inflammatory bowel disease, but immunization to ingested proteins must be recognized as a naturally occurring phenomenon. Therefore, it has been difficult to define and separate the possible contributory roles of immunological and nonimmunological factors in a wide variety of alleged food-induced conditions. Certainly, nonimmunological mechanisms involving chemical, metabolic, or mechanical phenomena are capable of producing symptoms after the ingestion of milk and other foods. Although both groups of patients may show an enhanced humoral antibody response to certain dietary antigens, it is not possible to differentiate patients with

Crohn's disease from those with ulcerative colitis on the basis of the incidence or amounts of antibody detectable against skimmed milk antigens, bovine serum albumin, casein, β lactoglobulin, α lactalbumin, or maize (Hanauer and Kraft, 1983). In attempting to demonstrate reaginic mechanisms in inflammatory bowel disease, intradermal tests and measurements of IgE-specific antibodies to a battery of food antigens failed to differentiate patients with inflammatory bowel disease from healthy subjects (Mee *et al.*, 1979a; Jones *et al.*, 1981). The list of other circulating reactants that have been inconsistently or occasionally detected in sera from patients with ulcerative colitis and Crohn's disease is extensive but as yet of no pathogenetic significance (Kraft, 1979).

E. Antigen–Antibody Complexes

Reports of circulating immune complexes in about one-third of patients with ulcerative colitis or Crohn's disease have been based on methods demonstrating anticomplementary activity in the serum, precipitin reactions with Clq inhibition of antibody-dependent cell-mediated cytotoxicity, demonstration of complement-dependent rosette formation, inhibition of latex agglutination by Clq or rheumatoid factor, or [^{125}I]Clq binding (Hodgson *et al.*, 1977; Kemler and Alpert, 1980a; Richens *et al.*, 1982). Such data must be interpreted with caution, since heat inactivation or storage at 4°C may create aggregates of immunoglobulins that are indistinguishable from true antigen–antibody complexes using many of the above assays. Furthermore, antilymphocyte antibodies and altered albumin/globulin ratios facilitate *in vivo* IgG aggregation and may also interfere with some of these assays (Soltis *et al.*, 1979). When such phenomena have been controlled for, sera from patients with inflammatory bowel disease have not demonstrated immune complexes nearly as frequently (Einstein *et al.*, 1979).

Clearly, the extraintestinal manifestations such as skin lesions, arthritis, pericholangitis, and iritis are suggestive of immune complex-mediated processes. Complement stimulation/activation induced by antigen–antibody complexes could potentiate a localized immune response. The local production of antibodies may contribute to the chronicity and recurrences seen in inflammatory bowel disease directly or in cooperation with lymphocytes and modulation of lymphocyte responses may provide a feedback control loop between cellular and humoral immunity. Nevertheless, the deposition of immune complexes in the affected tissues has not been demonstrated, perhaps because they are degraded after participating in the early stages of the pathologic process (Kirsner and Shorter, 1982).

F. Complement System

In a broader context, the complement system merits continued study in patients with inflammatory bowel disease. For example, Clq has been observed to undergo both increased synthesis and catabolism during active phases of Crohn's disease with normal synthesis in remission, suggesting the possibility of complement activation via the classical pathway (Potter *et al.*, 1979; Ross *et al.*, 1979). Both C3 and factor B appear to behave as acute phase reactants and are elevated during the active disease (Jewell *et al.*, 1981); however, serum levels of C4 do not rise and are actually reduced in the presence of hepatic disease. A report of decreased concentrations of serum properdin and properdin convertase in association with a variety of such extraintestinal complications of inflammatory bowel disease suggested activation of the alternate complement pathway (Lake *et al.*, 1979). Complement inhibitors also may be elevated in the serum in patients with inflammatory bowel disease, again consistent with an acute-phase phenomenon (Potter *et al.*, 1980).

III. CELLULAR IMMUNE PHENOMENA

The apparently limited role of the above humoral-type mechanisms in inflammatory bowel disease has intensified efforts to demonstrate cell-mediated immunological reactions, especially the possible involvement of microorganisms. Furthermore, the rare association of inflammatory bowel disease with severely impaired immunoglobulin synthesis but intact cellular immunity (Söltoft *et al.*, 1972), coupled with reports of altered cellular immune function in certain patients, has heightened interest in the possible role of the cell-mediated immune system in the bowel process.

A. Skin Responsiveness

Conventional skin reactivity has been variously reported to be normal or depressed in patients with inflammatory bowel disease in comparison to healthy controls (Kraft, 1979). Studies of skin sensitization to 2,4-dinitrochlorobenzene in patients with inflammatory bowel disease have shown normal initial responsiveness with failure in certain patients to react on rechallenge (Thayer *et al.*, 1978). This suggested defect in recall function could not be correlated with disease activity, the absolute numbers of T cells, or *in vitro* lymphocyte responses to phythohemagglutinin (PHA), but conceivably may relate to differences in the extent of disease,

inclusion of postoperative patients, the effects of antiinflammatory/immunosuppressive drugs, and malnutrition.

B. Circulating T Cells

Although initial reports of the proportions and absolute numbers of circulating T cells in patients with inflammatory bowel disease were conflicting, additional evidence has pointed to a reduction of these parameters—perhaps directly related to disease duration and activity or to the use of immunosuppressive agents (Auer *et al.*, 1978a, 1979; Lyanga *et al.*, 1979). Preliminary studies to evaluate subpopulations of peripheral-blood T cells that bear Fc receptors for IgG or IgM suggested reduced numbers of both (Victorino and Hodgson, 1980). However, enumeration of circulating T-lymphocyte subsets using monoclonal antibodies reactive against surface phenotypic markers has not elicited consistent differences between patients with inflammatory bowel disease and control groups (Hanauer *et al.*, 1982; Selby and Jewell, 1983; Yuan *et al.*, 1983b).

C. T Cells In Intestinal Tissues

Techniques developed to assess the characteristics of separated human intestinal lymphoid cells are now being applied to inflammatory bowel disease tissues, but the isolation method may influence at least the functional properties of these cells. One also must consider how a particular method may enrich for intraepithelial (mainly T suppressor cells) versus lamina proprial lymphocytes (proportionately more T helper cells). Resected intestinal tissues from patients with Crohn's disease contain abundant T cells, especially in the deeper layer of the inflamed areas where they may surround granulomas and be near larger numbers of B cells (Strickland *et al.*, 1975; Meuwissen *et al.*, 1976; Meijer *et al.*, 1979). Preliminary monoclonal antibody studies of T cells enzymatically separated from the intestinal lamina propria point to increased proportions and absolute numbers of the suppressor/cytotoxic phenotype in Crohn's disease, but not ulcerative colitis, in contrast to histologically normal control specimens (Yuan *et al.*, 1983a).

D. Reactivity of Lymphocytes

1. Lymphocyte Transformation

Tritiated thymidine incorporation into unstimulated peripheral-blood lymphocytes has been normal or decreased in patients with inflammatory

bowel disease in comparison to healthy controls; however, it may not differ from subjects having other diseases (Auer *et al.*, 1978a). Responses to allogeneic lymphocytes have been depressed, often in direct relation to the activity of the bowel disease. Furthermore, one group detected a defect in the autologous mixed-lymphocyte reaction in patients with inflammatory bowel disease that coincided with an abnormality in the generation of suppressor T cells and did not appear to relate to corticosteroid treatment (Ginsburg and Falchuk, 1982). The mean *in vitro* responses of circulating lymphocytes to PHA and concanavalin A in inflammatory bowel disease have been impaired in some studies and normal in many others (Kraft, 1979). While the significance of such data remains uncertain, T-cell mitogenic responses are impaired in some patients, perhaps related to disease activity and possibly affected by helper/suppressor cell imbalances. Whether these abnormalities are of primary significance or secondarily relate to therapy, enteric lymphocyte losses, malnutrition, or other aspects of the bowel lesion needs further elucidation (Hanauer and Kraft, 1983).

Initial studies of gut mucosal lymphocytes isolated from surgical specimens in inflammatory bowel disease have shown no evidence of defective proliferation to bacterial antigens (Fiocchi and Parent, 1981).

2. Cytotoxic Action of Lymphocytes

Preliminary studies of lymphocyte activities in inflammatory bowel disease demonstrate diminished antibody-dependent cell-mediated cytotoxicity that is comparable to disease controls (Auer and Ziemer, 1981). Peripheral-blood mononuclear cells from patients with either ulcerative colitis or Crohn's disease can be cytotoxic to human colonic epithelial cells in tissue culture. Null cells appear to account for this phenomenon, and the cytotoxic potential seems to be transmissible to normal control lymphocytes by preexposure to serum from patients with either ulcerative colitis or Crohn's disease (Watson *et al.*, 1980).

Isolation techniques for extracting tissue mononuclear cells may influence some of their *in vitro* cytotoxic activities (Bland *et al.*, 1979; Chiba *et al.*, 1981). For example, resected mesenteric lymph nodes in Crohn's disease initially were reported to show depressed responses to PHA. Using improved cell-separation methods, however, subsequent studies have shown the normal or even increased responsiveness of Crohn's disease lymph node lymphocytes, perhaps related to the degree of the tissue abnormality (Bird and Britton, 1974; Machado and Falchuk, 1978; Fiocchi *et al.*, 1979a). Since evidence is accumulating that both mesenteric lymph node and intestinal mucosal lymphocytes are capable of ex-

erting cytotoxic reactions by ADCC and SCMC mechanisms, these types of studies of isolated gut-associated lymphoid cells merit continued application in patients with inflammatory bowel disease (Bland *et al.*, 1979; MacDermott *et al.*, 1980; Chiba *et al.*, 1981).

E. Suppressor T-Cell Activity

Concanavalin A induction of suppressor T-cell activity has been found to be reduced in some patients with ulcerative colitis or Crohn's disease, and inhibition of suppressor-cell activity has correlated with the Crohn's disease activity (Hodgson *et al.*, 1978b; Kemler and Alpert, 1980b). Although patients with clinically mild or inactive Crohn's disease have shown no deficiency of the capacity of suppressor T cells to regulate antibody synthesis, many of the patients possessed circulating "covert" suppressor T cells that could be activated *in vitro* (Elson *et al.*, 1981). The occasional patient with inflammatory bowel disease and overt hypogammaglobulinemia may represent the loss or failure of development of opposing "contrasuppressor" T cells.

Suppressor-cell activity in lamina proprial lymphocytes from patients with inflammatory bowel disease has been said to be normal or reduced (Bookman and Bull, 1979; Fiocchi *et al.*, 1979b). On the other hand, many of the intraepithelial lymphocytes appear to be suppressor T cells (Bartnik *et al.*, 1980). While the exact genealogy of a population of granulated epithelial lymphocytes remains uncertain, they may represent either a subset of T cells or a unique type of mast cell that helps determine to what extent the suppressor T cells react to various tissue stimuli (Schrader *et al.*, 1983). Additionally, prostanoid synthesis in relation to mononuclear cells in inflammatory bowel disease is currently being studied in several laboratories (Rachmilewitz *et al.*, 1982).

IV. MACROPHAGES, MONOCYTES, NEUTROPHILS, EOSINOPHILS, AND MAST CELLS

Current concepts of the regulation of the immune response include a central role for macrophages, the exudate variety of which are derived from circulating monocytes. Activated monocytes have been identified in the blood of patients with inflammatory bowel disease, and circulating monocytes appear to be increased in number, apparently related to enhanced bone marrow monocytopoietic activity (Auer *et al.*, 1978b; Mee *et al.*, 1980). Furthermore, concentrations of lysozyme, transcobalamin II, lysosomal enzymes, and other monocyte/macrophage products are in-

creased both in the serum and in circulating monocytes from patients with inflammatory bowel disease, potentially participating in mucosal defense and tissue damage (Hanauer and Kraft, 1983). Monocyte–lymphocyte interactions are involved in the *in vitro* migration-inhibition phenomenon that occurs when inflammatory bowel disease lymphoid cells are exposed to sterile fetal colon homogenates, crude extracts of feces, preparations of diseased intestinal mucosa, and the enterobacterial common antigen (Kraft, 1979). Recently, blood monocytes have been implicated in *in vitro* SCMC reactions in patients with untreated Crohn's disease (Beeken *et al.*, 1980).

The potential role of macrophages in the pathological process in Crohn's disease is especially germane since the epithelioid cell granuloma is a prominent feature of the inflammatory response. Under normal conditions, lamina proprial macrophages may ingest foreign material and present processed antigen(s) to lymphocytes as an early event in the mucosal immune response. In Crohn's disease, mucosal macrophages are plentiful and prominently participate in granuloma formation (Sommers and Korelitz, 1975; Ward, 1979). Indeed, Ward (1977) postulated that the macrophages in patients with Crohn's disease may be defective, having a limited capacity to degrade ingested material. The secretion of plasminogen activator by macrophages suggests still another contributing factor to tissue injury in ulcerative colitis and Crohn's disease (Doe and Dorsman, 1982).

Circulating neutrophils often are increased in number in ulcerative colitis and Crohn's disease (Auer *et al.*, 1978b; Sategna-Guidetti and Rutigliano, 1979), having increased metabolic activity that correlates with the bowel activity (Wandall and Binder, 1982a,b). Abnormal neutrophil chemotaxis in both ulcerative colitis and Crohn's disease appears to be augmented by corticosteroids and is possibly related to the inhibiting effects of such serum factors as immune complexes or aggregated immunoglobulins (Rhodes *et al.*, 1982; Kirk *et al.*, 1983). The activity of leukocyte endogenous mediator, a small polypeptide hormone released by granulocytes, may be increased in patients with active Crohn's disease; this substance also has been implicated in the abnormalities of zinc distribution observed in such patients (Solomons *et al.*, 1978). In addition, neutrophils are capable of migrating into the intestinal lumen in response to specific antigens, even in the absence of mucosal damage (Bellamy and Nielsen, 1974).

Mucosal mast cells may have different morphological and functional properties than other tissue mast cells. These cells are recognized as a component of the tissue reaction in all layers of the intestinal wall in

patients with ulcerative colitis (Hiatt and Katz, 1962); in Crohn's disease, intestinal mast cells have been associated with the presence of granulomas (Rao, 1973). Several early studies reported increased mast cell counts in inflammatory bowel disease tissues; but when mast cells have not been detected or seemed reduced in numbers on routine microscopic sections, the presence of many IgE-containing cells in the same tissues has suggested the occurrence of mast cell degranulation (Dvorak *et al.*, 1978; Matsueda *et al.*, 1981). In addition to the indigenous chemical mediators, the release of which may foster tissue inflammation, mast cell degranulation also may trigger the release of eosinophilic basic protein (Dvorak, 1980). Eosinophils have been well documented in the affected tissues in both ulcerative colitis and Crohn's disease, although the precise role of these cells has yet to be discovered (Willoughby *et al.*, 1979; Dvorak *et al.*, 1980).

V. NUTRITION AND THE IMMUNE RESPONSE

Both macro- and micronutrient deficiencies may be associated with altered humoral and cellular immune responses and are frequent complications of inflammatory bowel disease. This malnutrition-induced immune dysfunction in inflammatory bowel disease is difficult to separate from the immunological abnormalities previously described in such patients, prior to the onset of detectable protein and calorie depletion (Murphy *et al.*, 1980). Lymphocytopenia is frequent in patients with severe malnutrition, and circulating immunoglobulin concentrations and humoral antibody responses often are decreased. Complement components may be diminished, while neutrophil and monocyte functions also may be adversely affected by nutritional deficiencies (Beisel *et al.*, 1981).

In patients with severe protein and calorie malnutrition, lymphatic structures such as the thymus and tonsils are reduced in size, and intradermal responses are diminished, proportionately to decreased circulating concentrations of albumin, hemoglobin, carotene, and vitamin C (Neumann *et al.*, 1975). Cutaneous testing may be a more sensitive indicator of protein-calorie malnutrition than are *in vitro* lymphocyte responses to nonspecific mitogens, although both parameters often are reduced (Chandra and Scrimshaw, 1980). Reduced numbers of T cells and increased proportions of null cells occur in the peripheral blood of marasmic children, perhaps due to an impaired differentiation and maturation of T cells (Chandra, 1980). Similar phenomena may well explain some of the functional immunological abnormalities noted in malnourished patients with inflammatory bowel disease.

VI. ANIMAL MODELS

Reports of immune-mediated animal models of human inflammatory bowel disease have been less than convincing (Cave and Kirsner, 1979). Rabbits intravenously injected with complexes of human serum albumin and its specific antibody after receiving dilute formalin per rectum [a modification of the original Auer reaction (Kraft *et al.*, 1963)] developed severe inflammation in the area of the nonspecific colonic irritation (Hodgson *et al.*, 1978a). A persistent colonic inflammation was produced in previously immunized rabbits by a similar injection with the enterobacterial common antigen of Kunin (Mee *et al.*, 1979b). When 2, 4-dinitrochlorobenzene (DNCB) was used as a sensitizing agent to elicit cell-mediated immune responses in the rabbit colon, animals that had previously received cutaneous applications of the chemical developed a "colitis" within 3 days after its rectal administration (Rabin and Rogers, 1978). However, the lesions that followed repeated DNCB-containing enemas healed at the same rate as those occurring after a single rectal challenge. In the guinea pig, multiple instillations of DNCB incorporated in an adherent paste circumvented the necessity for an initial skin sensitization and produced colonic inflammation (Askenase *et al.*, 1978). Furthermore, T cells isolated from the colonic lamina propria of guinea pigs treated with either DNCB or carrageenan had enhanced *in vitro* mitogen-induced cytotoxic activity (Glick and Falchuk, 1981; Braden *et al.*, 1980). This finding supported the possibility that changes in the surface epithelium of the colon due to exposure to chemical or microbiological agents may induce the production of cytotoxic lymphocytes.

Such observations merit continued study in order to provide testable hypotheses of the possible role of effector T cells in the local tissue reactions in inflammatory bowel disease. However, the absence of a convincing animal model remains a major impediment in achieving a better understanding of the pathogenesis of the acute and chronic tissue lesions in ulcerative colitis and Crohn's disease.

VII. IMMUNE HYPOTHESES: CONCLUDING REMARKS

The many clinical, epidemiological, familial, and immunological similarities between ulcerative colitis and Crohn's disease make a clear-cut separation between the two conditions difficult. However, disparate features in Crohn's disease, including the anatomical distribution of involvement, the histological appearance, a high recurrence rate after resection, and a less predictable response to therapy, suggest the possibility of dif-

ferent underlying pathogenetic mechanisms. The frequent clustering of both types of inflammatory bowel disease within families implies an underlying genetic susceptibility to variable immunopathological events (Kirsner and Shorter, 1982).

Although recognition of a linkage between histocompatibility antigens and immune response genes has led to studies of patients and families with ulcerative colitis and Crohn's disease, common *HLA* haplotypes seem to account for only a small portion of affected individuals. The only consistent correlations have been the presence of *HLA-B27* in patients also having ankylosing spondylitis or iritis. Nevertheless, speculation persists that the presence of specific cell-surface antigens could genetically predispose a person to an intestinal disease by virtue of sharing antigenic determinants with microbial agents, foods, or other exogenous factors. This could then enhance the mucosal penetration of these environmental antigens by providing receptor sites or preventing protective responses, perhaps as a result of a form of immunological tolerance. The possibility also exists that *HLA* antigens could lead to a disorder by interfering with interactions between nonimmunological ligands such as hormones and their cell-surface receptors (Svejgaard and Ryder, 1976).

One working hypothesis regarding the pathogenesis of inflammatory bowel disease is based on a concept of hypersensitivity to microbial antigens that normally are present in the lumen of the lower gastrointestinal tract. The demonstration of determinants that are shared by the heterogenetic enterobacterial antigen of Kunin and human colonic mucosal epithelial cells has suggested a possible mechanism for the tissue destruction in inflammatory bowel disease. It is speculated that ulcerative colitis and Crohn's disease represent parts of a spectrum of a pathogenetic process involving a mucosal "block" and prevention of the free penetration of enteric bacteria that normally colonize the intestine. This block is theorized to develop after bacterial antigens come into contact with gut-associated lymphoid tissues during neonatal life, causing the host to become sensitized to these commensal gut inhabitants. Subsequent disruption of the mucosal barrier by trauma, viral enteritis, bacillary dysentery, amebiasis, or intestinal ischemia, for example, is then theorized to set the stage for bacterial rechallenge and the induction of humoral and cell-mediated immunological reactions involving the determinants shared with bowel wall antigens. The interactions of various immunoregulatory factors could conceivably modulate the exacerbations and remissions of inflammatory bowel disease and account for the recurring extraintestinal manifestations (Watson *et al.*, 1980).

A better understanding of the possible *in vivo* relationships between bacteria and colon antigens is needed, and further efforts also should be

directed towards the identification of specific antigens and antibodies within the involved tissues of patients with inflammatory bowel disease. If immune complexes are indeed present in the serum, they provide a valuable opportunity for antigen isolation and characterization. Although the normal or enhanced general activity of the humoral immune system makes it appear unlikely that a B-cell dysfunction is the primary event in Crohn's disease, local deficiencies of IgA that are induced by altered suppressor/helper T-cell ratios in the intestinal wall could account in part for the local granulomatous tissue reaction (de Sousa, 1980). Evidence of viral-induced cytopathic changes in various cultured cells after exposure to extracts of inflammatory bowel disease tissues also could relate to the presence of interferon, a macrophage product that is capable of activating natural killer cells (McLaren and Gitnick, 1982).

Knowledge of the mechanisms of action of useful therapeutic agents on the various cellular components of the immunological apparatus of the gut offers future promise for more rational approaches to immune modulation in inflammatory bowel disease. The effect of corticosteroids on different lymphocyte subsets and the resultant change in the delicate balance of suppressor and effector activities needs to be elucidated. Also, the precise mechanism of action of sulfasalazine has yet to be defined and may include antilymphocyte effects, perhaps involving an antifolate mechanism (Klotz *et al.*, 1980; Baum *et al.*, 1981).

The differential susceptibilities of subpopulations of immunocompetent cells may account for some of the unexplained actions of azathioprine on cell-mediated immunity (Gyte and Willoughby, 1977). Since cimetidine recently has been reported to reverse anergy in some patients with Crohn's disease (Bicks and Rosenberg, 1980), the known presence of histamine$_2$ receptors on suppressor T cells suggests another reasonable area of investigation in this condition. Most of the therapeutic modalities utilized in inflammatory bowel disease have multiple mechanisms of action (corticosteroids, antibacterials, nutrients); therefore, these illnesses are best viewed as a blend of both immunological and nonimmunological phenomena. It remains to be established whether the immunological observations that have been reviewed are of etiopathogenetic importance, and if some of these phenomena are merely secondary or epiphenomena.

REFERENCES

Askenase, P. W., Boone, W. T., and Binder, H. J. (1978). *J. Immunol.* **120,** 198–201.

Auer, I. O., and Ziemer, E. (1981). *In* "Recent Advances in Crohn's Disease" (A. S. Pena, I. T. Weterman, C. C. Booth, and W. Strober, eds.), pp. 424–430. Nijhoff, The Hague.

Auer, I. O., Buschmann, C. H., and Ziemer, E. (1978a). *Gut* **19,** 618–626.
Auer, I. O., Wechsler, W., Ziemer, E., Malchow, H., and Sommer, H. (1978b). *Scand. J. Gastroenterol.* **13,** 561–571.
Auer, I. O., Götz, S., Ziemer, E., Malchow, H., and Ehms, H. (1979). *Gut* **20,** 261–268.
Baklien, K., and Brandtzaeg, P. (1975). *Clin. Exp. Immunol.* **22,** 197–209.
Bartnik, W., ReMine, S. G., Chiba, M., Thayer, W. R., and Shorter, R. G. (1980). *Gastroenterology* **78,** 976–985.
Baum, C. L., Selhub, J., and Rosenberg, I. H. (1981). *J. Lab. Clin. Med.* **97,** 779–784.
Beeken, W., St. Andre-Ukena S., and Gundel, R. M. (1980). *Gastroenterology* **78,** 1138. (Abstr.)
Beisel, W. R., Edelman, R., Nauss, K., and Suskind, R. M. (1981). *JAMA, J. Am. Med. Assoc.* **245,** 53–58.
Bellamy, J. E. C., and Nielsen N. O. (1974). *Infect. Immun.* **9,** 615–619.
Bendixen, G., Jarnum S., Söltoft, J., Westergaard, H., Weeke, B., and Yssing, M. (1968). *Scand. J. Gastroenterol.* **3,** 481–489.
Bendixen, G., Goltermann, N., Jarnum, S., Jensen, K. B., Weeke, B., and Westergaard, H. (1970). *Scand. J. Gastroenterol.* **5,** 433–441.
Bicks, R. O., and Rosenberg, E. W. (1980). *Lancet* **1,** 552–553.
Bird, A. G., and Britton, S. (1974). *Gastroenterology* **67,** 926–932.
Bland, P. W., Richens, E. R., Britton, D. C., and Lloyd, J. V. (1979). *Gut* **20,** 1037–1046.
Bookman, M. A., and Bull, D. M. (1979). *Gastroenterology* **77,** 503–510.
Braden, G. L., Masters, J. T., Onderdonk, A. B., and Falchuk, Z. M. (1980). *Gastroenterology* **78,** 1144. (Abstr.)
Brandtzaeg, P., and Baklien, K. (1976). *Scand. J. Gastroenterol., Suppl. 36* **11,** 1–45.
Brandtzaeg, P., and Baklien, K. (1979). *Z. Gastroenterol., Suppl.* **17,** 77–82.
Broberger, O., and Perlmann, P. (1959). *J. Exp. Med.* **110,** 657–674.
Cave, D. R., and Kirsner, J. B. (1979). *Z. Gastroenterol., Suppl.* **17,** 125–135.
Chandra, R. K. (1980). "Immunology of Nutritional Disorders." Arnold, London.
Chandra, R. K., and Scrimshaw, N. S. (1980). *Am. J. Clin. Nutr.* **33,** 2694–2697.
Chiba, M., Bartnik, W., ReMine, S. G., Thayer, W. R., and Shorter, R. G. (1981). *Gut* **22,** 177–186.
DeHoratius, R. J., Strickland, R. G., Miller, W. C., Volpicelli, N. A., Gaeke, R. F., Kirsner, J. B., and Williams, R. C., Jr. (1978). *Lancet* **1,** 1116–1119.
de Sousa, M. (1980). *Hosp. Pract.* **15,** 71–87.
Doe, W. F., and Dorsman, B. (1982). *Clin. Exp. Immunol.* **48,** 256–260.
Dvorak, A. M. (1980). *J. Immunol.* **125,** 460–462.
Dvorak, A. M., Monahan, R. A., Osage, J. E., and Dickersin, G. R. (1978). *Lancet* **1,** 498.
Dvorak, A. M., Monahan, R. A., Osage, J. E., and Dickersin, G. R. (1980). *Hum. Pathol.* **11,** 606–619.
Einstein, E., Charland, C., and Thayer, W. R., Jr. (1979). *Digestion* **19,** 65–69.
Elson, C. O., Graeff, A. S., James S. P., and Strober, W. (1981). *Gastroenterology* **80,** 1513–1521.
Farmer, G. W., Vincent, M. M., Fuccillo, D. A., Horta-Barbosa, L., Ritman, S., Sever, J. L., and Gitnick, G. L. (1973). *Gastroenterology* **65,** 8–18.
Farr, R. S. (1963). *Arch. Environ. Health* **6,** 92–98.
Fiocchi, C., and Parent, K. (1981). *In* "Recent Advances in Crohn's Disease" (A. S. Pena, I. T. Weterman, C. C. Booth, and W. Strober, eds.), pp. 433–438. Nijhoff, The Hague.
Fiocchi, C., Battisto, J. R., and Farmer, R. G. (1979a). *Gastroenterology* **76,** 1133. (Abstr.)
Fiocchi, C., Battisto, J. R., and Farmer, R. G. (1979b). *Dig. Dis. Sci.* **24,** 705–717.
Fiorilli, M., Luzi, G., and Aiuti, F. (1975). *Rend. Gastroenterol.* **7,** 1–4.

Ginsburg, C. H., and Falchuk, Z. M. (1982). *Gastroenterology* **83,** 1–9.
Glick, M. E., and Falchuk, Z. M. (1981). *Gut* **22,** 120–125.
Greenberg, H. B., Gebhard, R. L., McClain, C. J., Soltis, R. D., and Kapikian, A. Z. (1979). *Gastroenterology* **76,** 349–350.
Greenstein, A. J., Janowitz, H. D., and Sachar, D. B. (1976). *Medicine (Baltimore)* **55,** 401–412.
Gyte, G. M. L., and Willoughby, J. M. T. (1977). *Clin. Exp. Immunol.* **30,** 242–251.
Hanauer, S. B., Kluskens, L. F., Yuan, S.-Z., and Kraft, S. C. (1982). *Gastroenterology* **82,** 1079. (Abstr.)
Hanauer, S. B., and Kraft, S. C. (1983). *In* "Inflammatory Bowel Diseases" (R. N. Allan, M. R. B. Keighley, J. Alexander-Williams, and C. Hawkin, eds.), pp. 356–371. Churchill, Edinburgh.
Hardy Smith, A., and Macphee, I. W. (1971). *Gut* **12,** 20–26.
Hiatt, R. B., and Katz, L. (1962). *Am. J. Gastroenterol.* **37,** 541–545.
Hodgson, H. J. F., and Jewell, D. P. (1978). *Dig. Dis. Sci.* **23,** 123–128.
Hodgson, H. J. F., Potter, B. J., and Jewell, D. P. (1977). *Clin. Exp. Immunol.* **29,** 187–196.
Hodgson, H. J. F., Potter, B. J., Skinner, J., and Jewell, D. P. (1978a). *Gut* **19,** 225–232.
Hodgson, H. J. F., Wands, W. R., and Isselbacher, K. J. (1978b). *Clin. Exp. Immunol.* **32,** 451–458.
Jewell, D. P., Potter, B. J., Brown, D. J. C., Hodgson, H. J. F., and Mee, A. S. (1981). *In* "Recent Advances in Crohn's Disease" (A. S. Pena, I. T. Weterman, C. C. Booth, and W. Strober, eds.), pp. 333–335. Nijhoff, The Hague.
Jones, D. B., Kerr, G. D., Parker, J. H., and Wilson, R. S. E. (1981). *J. R. Soc. Med.* **74,** 292–293.
Jones, E. G., Beeken, W. L., Roessner, K. D., and Brown, W. R. (1976). *Digestion* **14,** 12–19.
Kemler, B. J., and Alpert, E. (1980a). *Gut* **21,** 195–201.
Kemler, B. J., and Alpert, E. (1980b). *Clin. Exp. Immunol.* **42,** 280–284.
Kirk, A. P., Cason, J., Fordham, J. N., Brown, K. A., Goddard, D. H., Holborow, E. J., and Lennard-Jones, J. E. (1983). *Dig. Dis. Sci.* **28,** 236–248.
Kirsner, J. B., and Shorter, R. G. (1982). *N. Engl. J. Med.* **306,** 775–785, 837–848.
Klotz, U., Maier, K., Fischer, C., and Heinkel, K. (1980). *N. Engl. J. Med.* **303,** 1499–1502.
Kraft, S. C. (1979). *In* "Immunology of the Gastrointestinal Tract" (P. Asquith, ed.), pp. 95–128. Churchill, Edinburgh.
Kraft, S. C., and Kirsner, J. B. (1971). *Gastroenterology* **60,** 922–951.
Kraft, S. C., Fitch, F. W., and Kirsner, J. B. (1963). *Am. J. Pathol.* **43,** 913–927.
Kuiper, I., Weterman, I. T., Biemond, I., Van Rood, J. I., and Pena, A. S. (1981). *In* "Recent Advances in Crohn's Disease" (A. S. Pena, I. T. Weterman, C. C. Booth, and W. Strober, eds.), pp. 341–347. Nijhoff, The Hague.
Lagercrantz, R., Hammarström, S. Perlmann, P., and Gustafsson, B. E. (1966). *Clin. Exp. Immunol.* **1,** 263–276.
Lake, A. M., Stitzel, A. E., Urmson, J. R., Walker, W. A., and Spitzer, R. E. (1979). *Gastroenterology* **76,** 1374–1379.
Lyanga, J. J., Davis, P., and Thomson, A. B. R. (1979). *Clin. Exp. Immunol.* **37,** 120–125.
MacDermott, R. P., Franklin, G. O., Jenkins, K. M., Kodner, I. J., Nash, G. S., and Weinrieb, I. J. (1980). *Gastroenterology* **78,** 47–56.
MacDermott, R. P., Nash, G. S., Bertovich, M. J., Seiden, M. V., Bragdon, M. J., and Beale, M. G. (1981). *Gastroenterology* **81,** 844–852.
Machado, I., and Falchuk, Z. M. (1978). *Gastroenterology* **74,** 1059. (Abstr.)
McLaren, L. C., and Gitnick, G. (1982). *Gastroenterology* **82,** 1381–1388.

Matsueda, K., Rimpila, J. J., Ford, H. E., Levin, B., and Kraft, S. C. (1981). *In* "Recent Advances in Crohn's Disease" (A. S. Pena, I. T. Weterman, C. C. Booth, and W. Strober, eds.), pp. 103–107. Nijhoff, The Hague.

Mee, A. S., Brown, D., and Jewell, D. P. (1979a). *Scand. J. Gastroenterol.* **14,** 743–746.

Mee, A. S., McLaughlin, J. E., Hodgson, H. J. F., and Jewell, D. P. (1979b). *Gut* **20,** 1–5.

Mee, A. S., Berney, J., and Jewell, D. P. (1980). *J. Clin. Pathol.* **33,** 917–920.

Meijer, C. J. L. M., Bosman, F. T., and Lindeman, J. (1979). *Scand. J. Gastroenterol.* **14,** 21–32.

Meuwissen, S. G. M., Feltkamp-Vroom, T. M., Brutel de la Rivièr, A., von dem Borne, A. E. G. K., and Tytgat, G. N. (1976). *Gut* **17,** 770–780.

Murphy, M. J., Charland, C., and Thayer, W. (1980). *Gastroenterology* **78,** 1227. (Abstr.)

Nagai, T., and Das, K. M. (1981). *Gastroenterology* **80,** 1507–1512.

Neumann, C. G., Lawlor, G. J., Jr., Stiehm, E. R., Swendseid, M. E., Newton, C., Herbert, J., Ammann, A. J., and Jacob, M. (1975). *Am. J. Clin. Nutr.* **28,** 89–104.

O'Donoghue, D. P., and Kumar, P. (1979). *Gut* **20,** 149–153.

Palmer, K. R., Sherriff, S. B., Holdsworth, C. D., and Ryan, F. P. (1981). *Q. J. Med.* **50,** 463–471.

Perlmann, P., Hammarström, S., Lagercrantz, R., and Gustafsson, B. E. (1965). *Ann. N. Y. Acad. Sci.* **124,** 377–394.

Persson, S. (1974). *Acta Chir. Scand.* **140,** 64–67.

Persson, S., and Danielsson, D. (1973). *Acta Chir. Scand.* **139,** 735–738.

Potter, B. J., Hodgson, H. J. F., Mee, A. S., and Jewell, D. P. (1979). *Gut* **20,** 1012–1019.

Potter, B. J., Brown, D. J. C., Watson, A., and Jewell, D. P. (1980). *Gut* **21,** 1030–1034.

Rabin, B. S., and Rogers, S. J. (1978). *Gastroenterology* **75,** 29–33.

Rachmilewitz, D., Ligumsky, M., Haimovitz, A., and Treves, A. J. (1982). *Gastroenterology* **82,** 673–679.

Rao, S. N. (1973). *J. Pathol.* **109,** 79–82.

Rhodes, J. M., Potter, B. J., Brown, D. J. C., and Jewell, D. P. (1982). *Gastroenterology* **82,** 1327–1334.

Richens, E. R., Thorp, C. M., Bland, P. W., and Hall, N. D. (1982). *Dig. Dis. Sci.* **27,** 129–138.

Rosekrans, P. C. M., Meijer, C. J. L. M., van der Wal A. M., Cornelisse, C. J., and Lindeman, S. (1980). *Gut* **21,** 941–947.

Ross, I. N., Thompson, R. A., Montgomery, R. D., and Asquith, P. (1979). *J. Clin. Pathol.* **32,** 798–801.

Sategna-Guidetti, C., and Rutigliano, A. R. (1979). *Acta Gastroenterol. Belg.* **42,** 247–256.

Schachter, H., and Kirsner, J. B. (1980). "Crohn's Disease of the Gastrointestinal Tract." Wiley, New York.

Schrader, J. W., Scollay, R., and Battye, F. (1983). *J. Immunol.* **130,** 558–564.

Selby, W. S., and Jewell, D. P. (1983). *Gut* **24,** 99–105.

Skinner, J. M., and Whitehead, R. (1974). *J. Clin. Pathol.* **27,** 643–646.

Solomons, N. W., Elson, C. O., Pekarek, R. S., Jacob, R. A., Sandstead, H. H., and Rosenberg, I. H. (1978). *Infect. Immun.* **22,** 637–639.

Soltis, R. D., Hasz, D., Morris, M. J., and Wilson, I. D. (1979). *Gastroenterology* **76,** 1380–1385.

Söltoft, J., Petersen, L., and Kruse, P. (1972). *Scand. J. Gastroenterol.* **7,** 233–236.

Sommers, S. C., and Korelitz, B. I. (1975). *Am. J. Clin. Pathol.* **63,** 359–365.

Strickland, R. G., Korsmeyer, S., Soltis, R. D., Wilson, I. D., and Williams, R. C., Jr. (1974). *Gastroenterology* **67,** 569–577.

Strickland, R. G., Husby, G., Black, W. C., and Williams, R. C., Jr. (1975). *Gut* **16,** 847–853.

Strickland, R. G., Miller, W. C., Volpicelli, N. A., Gaeke, R. F., Wilson, I. D., Kirsner, J. B., and Williams, R. C., Jr. (1977). *Clin. Exp. Immunol.* **30,** 188–192.

Svejgaard, A., and Ryder, L. P. (1976). *Lancet* **2,** 547–549.

Thayer, W. R., Jr., and Kirsner, J. B. (1980). *Gastroenterology* **78,** 1097–1100.

Thayer, W. R., Jr., Brown, M., Sangree, M. H., Katz, J., and Hersh, T. (1969). *Gastroenterology* **57,** 311–318.

Thayer, W. R., Charland, C., and Fitzgibbons, A. K. (1978). *Gastroenterology* **74,** 1103. (Abstr.)

van de Merwe, J. P., Mol, G. J. J., Wensinck, F., Dees, J., van Blankenstein, M., and Westbroek, D. L. (1982). *Digestion* **23,** 104–109.

Victorino, R. M. N., and Hodgson, H. J. F. (1980). *Clin. Exp. Immunol.* **41,** 156–165.

Wandall, J. H., and Binder, V. (1982a). *Gut* **23,** 173–180.

Wandall, J. H., and Binder, V. (1982b). *Gut* **23,** 758–765.

Ward, M. (1977). *Lancet* **2,** 903–905.

Ward, M. (1979). *Z. Gastroenterol., Suppl.,* **17,** 116–124.

Watson, D. W., Bartnik, W., and Shorter, R. G. (1980). *In* "Inflammatory Bowel Disease" (J. B. Kirsner and R. G. Shorter, eds.), 2nd ed., pp. 121–137. Lea & Febiger, Philadelphia, Pennsylvania.

Weeke, B., and Jarnum, S. (1971). *Gut* **12,** 297–302.

Willoughby, C. P., Piris, J., and Truelove, S. C. (1979). *Scand. J. Gastroenterol.* **14,** 395–399.

Yuan, S.-Z., Foster, L. P., Wang. N.-S., and Kraft, S. C. (1983a). *Gastroenterology* **84,** 1358. (Abstr.)

Yuan, S.-Z., Hanauer, S. B., Kluskens, L. F., and Kraft, S. C. (1983b). *Gastroenterology,* **85,** 1313–1318.

CHAPTER 12

Autoimmune Diseases of the Liver: Chronic Active Hepatitis and Primary Biliary Cirrhosis

IAN R. MACKAY*

The Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital
Melbourne, Victoria, Australia

This chapter on autoimmunity and liver will cover two diseases, chronic active hepatitis (CAH) and primary biliary cirrhosis (PBC). For both diseases, recognition and clinical interest developed from around

* Aided by a grant from the National Health and Medical Research Council of Australia.

1950, and features of autoimmunity were first recognized several years later. Although there are striking expressions of immunological derangement for both diseases, yet for neither has there been discovered organ-specific immune reactivities that could account for the tissue damage observed to hepatic parenchymal cells (in CAH) and to intrahepatic cholangioles (in PBC). Despite occasional examples of overlap, the two diseases are sufficiently distinct to justify separate consideration in this chapter.

I. CHRONIC ACTIVE HEPATITIS

A. Introduction: Historical Background

The term chronic active hepatitis (CAH) describes a form of chronic inflammatory liver disease that precedes, coexists with, and finally terminates in a cirrhosis of the liver. The historical background (see Mackay, 1975) includes descriptions in the 1940s of "subacute hepatitis" and "chronic viral hepatitis," with the recognition in the 1950s of the prototype disease in young females and its designation as "active chronic hepatitis." Chronic active hepatitis became better characterized in the 1960s with the clinical application of serum electrophoresis, which defined hypergammaglobulinemia; liver biopsy, which showed the persisting inflammatory activity and necrosis; measurement of transaminase enzymes in serum, which allowed monitoring of the activity of the disease and response to treatment; and tests for serum autoantibodies, which pointed to immunopathogenetic mechanisms.

The concept of autoimmunity as the cause of CAH was introduced when Joske and King (1955) described a positive test for lupus erythematosus (LE) cells in two patients with "active chronic viral hepatitis"; with the recognition of five further cases, the disease was called "lupoid hepatitis" (Mackay *et al.*, 1956). Later, the serum of patients with lupoid hepatitis, primary biliary cirrhosis (PBC), and systemic lupus erythematosus (SLE) was found to react by complement fixation (CF) with saline extracts of human tissues, particularly liver and kidney (Mackay and Gajdusek, 1958). After 1960, immunofluorescence procedures for antinuclear autoantibodies gave further impetus to the concept of autoimmunity, and subsequently the characteristic autoantibody to smooth muscle was recognized.

During 1955–1965, much attention was given to the histopathology of CAH. Early on, the term "plasma-cell hepatitis" described the heavy

accumulation of plasma cells in portal and periportal areas (Page and Good, 1960), and the term "piecemeal necrosis" described the characteristic periportal destruction of hepatocytes and the inflammatory response at the limiting plate of liver lobules (Paronetto *et al.*, 1962). In 1968, pathologists agreed on a histological definition of chronic hepatitis and defined two distinct groups: chronic persistent hepatitis, which was clinically and histologically benign and nonprogressive; and chronic aggressive hepatitis, which was the histological counterpart of clinically recognized CAH (De Groote *et al.*, 1968).

In 1967, a serum particle called Australia (Au) antigen was shown to be associated with cases of acute hepatitis (Blumberg *et al.*, 1967), and sera containing this particle were found to be infectious. This particle was found to be a marker for infection with the hepatitis B virus, and it eventually became known as hepatitis B surface antigen (HBsAg). After 1970, cases of CAH were found to be associated with the presence in serum of HBsAg, indicative of persistence in the liver of HBV and of a viral causation of the disease. However, the type of chronic hepatitis associated with HBsAg differed from the classic type: It predominantly affected males and was more indolent in its course but in most respects was indistinguishable histologically. Also, serologists segregated the autoantibody-positive type of CAH and the HBsAg-positive type. Hypotheses were framed to provide for an etiological role of the hepatitis B virus in all cases of CAH, whether HBsAg positive or HBsAg negative, but these failed to become substantiated. By 1972 it was evident (Mackay, 1972) that CAH could be due to several causes, each with an independent pathogenesis and outcome.

B. General Description, Classification, and Animal Models

1. General Description

Chronic active hepatitis has been defined by the IASL nomenclature group of the International Association for the Study of the Liver (IASL) (Leevy *et al.*, 1976) and separated histologically from chronic persistent hepatitis (Fig. 1). The IASL definition of CAH described the classic type of disease with immunoserological abnormalities, but did not present a subclassification or differentiate the autoimmune from other types. A disease-specific autoantibody in serum is a desirable marker for specifying any condition including CAH as autoimmune, and, ideally, such autoantibody should be organ- or tissue specific; unfortunately, these criteria

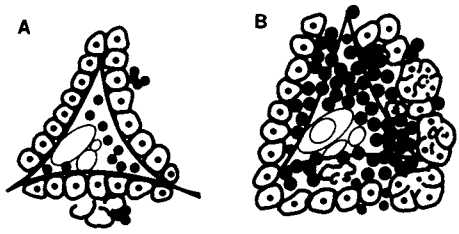


FIG. 1. Diagrammatic representations of histological appearances in (A) chronic persistent hepatitis showing portal cellular infiltrate and (B) chronic active (aggressive) hepatitis showing periportal lymphocytic infiltration with necrosis of liver cells.

are not fully met by the autoantibody markers, antinuclear and/or anti-smooth-muscle, that are presently used to identify autoimmune CAH; the identification of a specific antiliver autoantibody and the development of a clinically applicable assay for this are awaited. Hence, autoimmune CAH is recognized by a profile of features that include female sex, characteristic liver histology, hypergammaglobulinemia, non-tissue-specific autoantibodies, the HLA phenotype B8, DR3, and responsiveness to treatment with prednisolone (see Section I,F). The prevalence of autoimmune CAH among populations of European origin would be of the order of 5 to 8 per 100,000 (Mackay, 1981; Bjarnason *et al.*, 1982).

2. Classification

The major types of CAH include the autoimmune type, the HBsAg-positive type, and the cryptogenic type in which autoantibody markers, HBsAg, and other evidence for a specific etiology are lacking. In addition, there are miscellaneous types in which there may be expression of the histological features of CAH but without resemblance to CAH clinically; these include hepatitic reactions to medicinal drugs, ethanol abuse, non-A, non-B virus infection, Wilson's disease, and α_1 anti-trypsin deficiency. As the differentiation of autoimmune from other types of CAH depends mostly on autoantibody markers, which may vary in titer with disease activity, a "gray edge" is inevitable; however, such markers are lacking in the miscellaneous types of CAH, particularly CAH associated with ethanol abuse (Crapper *et al.*, 1983), non-A, non-B viral infection (Mackay *et al.*, 1985) and α_1 antitrypsin deficiency (Hodges *et al.*, 1981). The one exception is a type of CAH usually seen in middle-aged women and associated with an adverse reaction to drugs, notably oxyphenisatin (Reynolds *et al.*, 1971) and α-methyl dopa (Eliastim and Holmes, 1971);

this resembles spontaneous autoimmune CAH in the autoantibody markers expressed, but there is regression when the offending drug is withdrawn.

The relative frequencies of the entities in the spectrum of CAH were surveyed for southern England by Hodges *et al.* (1982), who recognized that data would differ substantially according to geographic area and referral patterns, as cited by Mackay (1984). Differing frequences of autoimmune CAH (autoantibody-positive, HbsAg-negative), HBV-associated CAH (autoantibody-negative, HBsAg-positive), and cryptogenic CAH for three different geographic areas (Australia, Germany, and Hong Kong) are shown in Fig. 2, based on data of the author (unpublished), Meyer zum Büschenfelde and Hütteroth (1979), and Lam *et al.* (1980). Cryptogenic CAH accounted for about one-third of all cases of CAH in each area; this is often aligned with autoimmune CAH into a general HBsAg-negative category, but there are reasons to implicate antecedent HBV infection (Pedersen *et al.*, 1982). It is also of note that the 106 cases from West Germany included 6% positive for autoantibody markers and HBsAg.

3. Animal Models

The existence of a spontaneous or induced model in animals is an important attribute for ascribing any given human disease to autoimmunity, but this attribute has proven difficult to fulfil in the case of chronic hepatitis. The various attempts to induce a model of autoimmune CAH have included some that are quite artificial (see Mackay, 1978). A frequently cited model is that described by Meyer zum Büschenfelde and Hopf (1974), who repetitively immunized rabbits with proteins extracted

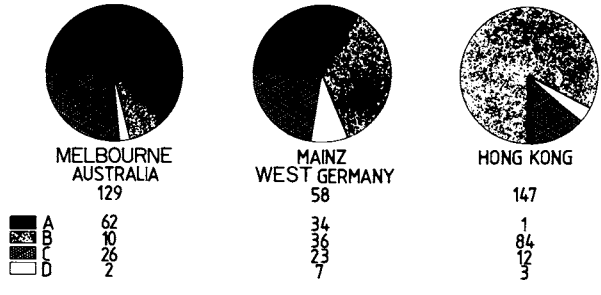


FIG. 2. Differences in frequency of major types of CAH (percent distribution by region), exemplified by data (see discussion in text) from Australia, Europe, and Hong Kong. A, Autoimmune; B, HBV-associated; C, cryptogenic; D, miscellaneous.

from human liver in Freund's complete adjuvant for periods >1 year; the eventual hepatitis resembled human chronic hepatitis histologically, showing portal lymphocyte infiltrates and periportal necrosis, but it was not described as self-perpetuating. The injected rabbits developed autoantibodies to liver-specific protein (LSP; see Section I,E,1,g), and also to a membrane determinant of liver cells. This model has not been adequately confirmed in rabbits by other workers, and rats and mice are quite resistant, although mice readily develop antibody to LSP after immunization (Bartholomaeus *et al.*, 1981).

The liver-specific F antigen (FAg) is a cytoplasmic component of liver that is of interest by reason of its existence in mice as two allotypic variants, type I and type II. Immunization of responder strains of mice with the allotypic form of FAg induces an antiliver antibody that is autoreactive, indicating that the allotypic variant can break tolerance to self (Fravi and Lindenmann, 1968); this is explained by the allotypic variant being part of the carrier portion of the molecule, the antibody-reactive determinant being the hapten. This immune response is the best example of the T-cell bypass theory of induction of autoimmunity (Allison *et al.*, 1971; Weigle, 1977), but autoantibody to FAg is never associated with any evidence of liver disease.

C. Clinical Presentations

Autoimmune CAH is a persisting and progressive disease, which develops insidiously without any definable onset in some 75% of cases acutely, and in some 25%, resembling acute viral hepatitis. Whether the initiating event is a hepatitis of known cause, viral or toxic, or whether autoimmune hepatitis begins *de novo,* is unresolved (Crapper *et al.*, 1984). Autoimmune CAH usually presents in an acute relapse, but it may remain silent until manifested clinically as an advanced cirrhosis with hepatic decompensation or variceal bleeding. A duration of hepatitis for at least 3 months is required to define chronicity. Classically, the disease is one of young women in whom the symptoms are either those of (a) liver inflammation, expressed as jaundice, anorexia, and right upper abdominal pain, (b) multisystem disease expressed as fever and arthralgia, or (c) endocrine disorder expressed as amenorrhea, acne, striae, and spider nevi. Liver enlargement is present only in the early stages, since the liver progressively shrinks as macronodular cirrhosis develops. Splenomegaly is frequent, and the size of the spleen fluctuates with activity of the disease; in the early stages at least, splenomegaly is attributable more to lymphoid hyperplasia than to portal hypertension. In the later stages the

physical findings in autoimmune CAH come to include those associated with liver cirrhosis from any cause.

The biochemical laboratory indices include raised levels of serum aspartate and alanine transaminases (AST, ALT) (100–2000 units), extreme hypergammaglobulinemia (30–80 g/liter), hyperbilirubinemia, and impaired prothrombin efficiency during relapses. Moderately raised levels of serum alkaline phosphatase may be observed, particularly in later stages of the disease when these specify bile duct damage that is observable histologically.

D. Histopathology

The following definition of CAH, provided by an International Group of Pathologists (1977), specifies several features common to all untreated examples of the disease and includes the following. (a) Piecemeal necrosis, which is defined as the destruction of liver cells at the interface between parenchyma and connective tissue with a predominantly lymphocytic infiltrate (Fig. 3). (b) At the advancing edge of piecemeal necrosis, hepatocytes show hydropic swelling and may be arranged in abnormal ductular configurations known as rosettes. (c) Bands of piecemeal necrosis link portal tracts to give the appearance of bridging hepatic necrosis, formerly regarded as a forerunner and determinant of progression of hepatitis to cirrhosis but now less emphasized; other patterns include confluent and submassive necrosis. (d) In acute relapses, there may be foci of intralobular necrosis in association with infiltrating lymphocytes and plasmacytoid cells. (e) A distinct feature is the appearance of eosinophilic necrotic liver cells or apoptotic bodies; apoptosis (falling away) refers to noninflammatory cellular necrosis, perhaps a reflection of injury by cytokines released by activated T cells. (f) A long-accepted feature of the cellular infiltrate seen in portal tracts and areas of piecemeal necrosis in the liver is the predominance of plasma cells, giving rise to the term "plasma cell hepatitis"; surprisingly, application of monoclonal antibody markers to frozen sections of liver in cases of autoimmune hepatitis (see below) reveals a marked preponderance of T cells.

The acute phase of autoimmune hepatitis evolves, seemingly inevitably and within periods of several months to a few years, to a macronodular cirrhosis (Fig. 4). Although the cellular infiltrates persist to some degree indefinitely, the intensity of cellular necrosis and lymphocytic infiltration is greatly modified by regimens of treatment with corticosteroid drugs; in early cases, restitution to almost normal histological appearances, or reversion to changes of chronic persistent hepatitis, can be anticipated.

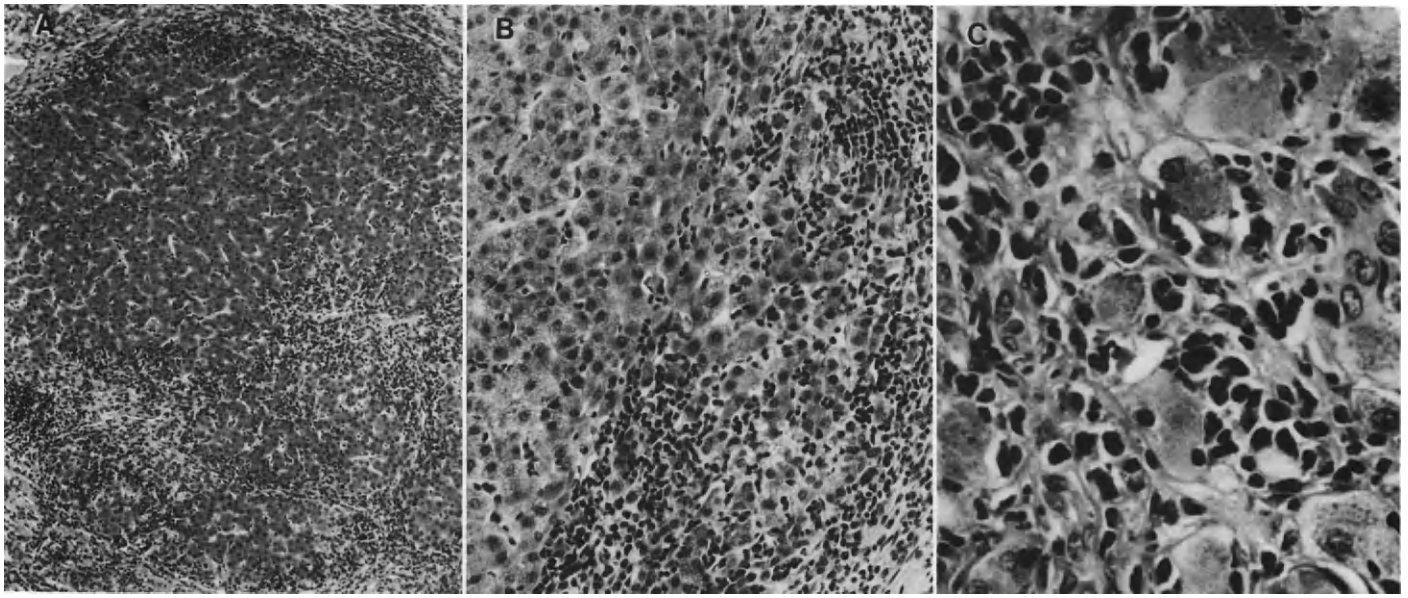


FIG. 3. The histological lesions of autoimmune CAH: from left to right are low (150×), medium (330×), and high (550×) magnifications: (A) perilobular distribution of the cellular infiltrate; (B) invasion of lobule by lymphoid cells and necrotic hepatocytes (piecemeal necrosis); (C) degenerate hepatocytes with lymphoid cells in close apposition.

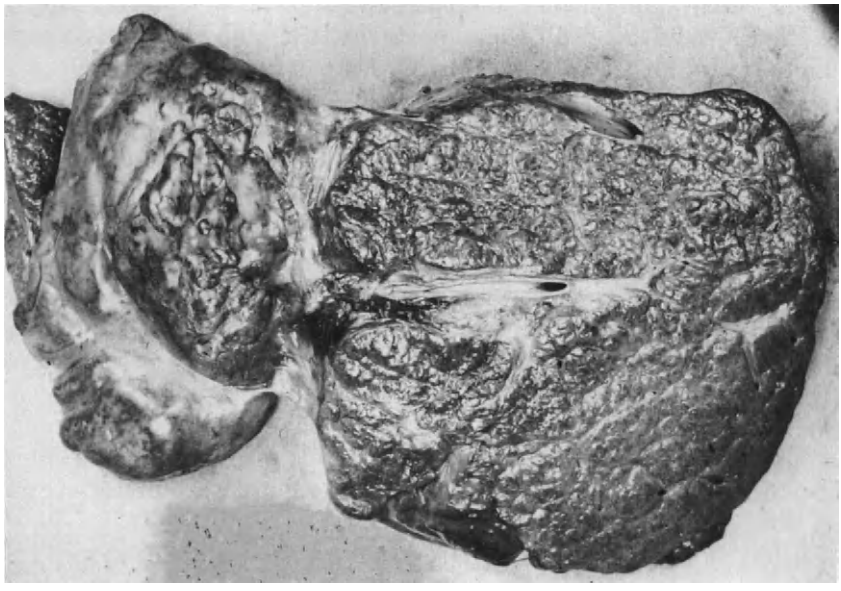


FIG. 4. Macroscopic appearance of liver from patient with autoimmune CAH showing coarse macronodular cirrhosis.

E. IMMUNOLOGY

1. Immunological Features

a. Hypergammaglobulinemia. This was first recognized in the 1940s when electrophoresis was applied to the study of human sera, and it featured prominently in early descriptions of CAH. The increase is polyclonal and is more marked in the autoimmune than in other types of CAH. Different isotypes of immunoglobulin (Ig) are expressed predominantly in different chronic liver diseases, IgG in CAH, IgM in PBC, and IgA in alcoholic cirrhosis; studies on subclasses of IgG, the isotype in autoimmune CAH, are lacking. The degree of hypergammaglobulinemia decreases markedly after treatment with prednisolone (Mackay, 1968). Explanations for the hypergammaglobulinemia include the contribution made by autoantibodies, the bypassing of the Kupffer cell system by enteric antigens, which thus gain access to systemic sites of immunoglobulin synthesis, and the formation of antiimmunoglobulins, either to Fc sites as rheumatoid factors or to F(ab) sites as antiidiotypes.

b. Antinuclear Antibodies (ANA). These were first recognized in CAH by positivity to the lupus erythematosus (LE) cell test and specified "lu-

poid hepatitis," although in comparison with systemic lupus erythematosus (SLE), the degree of LE-cell positivity is less and positive reactions tend to occur only during exacerbations. Reactions of ANA in CAH are now recognized by immunofluorescence. Conventional nuclear substrates for detection of ANA are those provided by frozen sections of rat liver or cells of the HEp2 cell line, which provide for better assessment of patterns of antinuclear reactivity. Human blood smears provide nuclei of granulocytes and allow recognition of granulocyte-specific ANA (gsANA) characteristic of autoimmune CAH (Smalley *et al.*, 1968); this gsANA is a late-appearing differentiation antigen of the granulocyte nucleus (Whittingham *et al.*, 1981b). The pattern of ANA reactivity is usually *homogeneous,* signifying reactivity with DNA–histone; a rim or peripheral pattern is a variant of this, and much less frequently the pattern is *speckled,* denoting reactivity with saline-extractable ribonuclear antigens, or *nucleolar* (Kenneally *et al.*, 1984). It may well be that the nucleoprotein antigen recognized in CAH differs from that recognized in SLE, but this question has not yet been engaged.

Cut-off titers for specifying ANA positively have not been agreed upon but appear to vary from 1 : 10 to 1 : 80 according to individual laboratories. Thus, there will be a degree of variability in the reported frequency of ANA positivity from laboratory to laboratory in CAH, as in other autoimmune diseases, attributable to (a) the particular titer taken to represent a positive reaction, (b) whether single or repeated tests are done during the course of the disease and the phase of the disease at which the serum is tested, (c) the source and nature of the substrate nuclear antigen chosen, and (d) the chosen criteria for diagnosis and subgrouping of cases.

Antibodies to native double-strand DNA detected by radioimmunoassay occur in autoimmune CAH (Davis and Read, 1975; Jain *et al.*, 1976; Smeenk *et al.*, 1982); the reported frequencies vary, presumably because the strength of anti-DNA reactions varies with the activity of the disease. There is a high incidence of reactivity to denatured single-strand DNA (Koffler *et al.*, 1969).

c. Anticytoplasmic Antibodies. Anticytoplasmic antibodies in liver disease detected by complement fixation (CF) were described as early as 1908, and have been reported sporadically since then (Mackay, 1975). However, it was the detection in 1957–1958, in cases of CAH and PBC, of high-titer CF reactions with various tissue extracts that initially promoted the concept of autoimmunity in these diseases; the antigenic constituents for this autoimmune complement fixation (AICF) reaction were found to consist of mitochondria, microsomes, and a soluble cytoplasmic component (Asherson, 1959). The frequency of such reactivity in CAH (titer >8)

was of the order of 30%, according to reports from various centers (Mackay and Larkin, 1958).

d. Anti-Smooth-Muscle Antibody (ASMA). This was first demonstrated in 1965 in CAH by immunofluorescence (Johnson *et al.*, 1965). Anti-smooth-muscle antibody is not entirely specific for CAH, being given (to low titer) by sera from patients with various other forms of liver disease, including diseases such as alcoholic hepatitis, rheumatoid arthritis, multiple sclerosis, cancer, and particularly acute viral diseases. The reaction for ASMA is lacking in SLE, perhaps because major hepatic involvement does not seem to occur in this disease, although there are opinions to the contrary (Runyon *et al.*, 1980). A positive serological reaction for ASMA will be given by some 70% of cases of autoimmune CAH, particularly if tests are performed in phases of activity, in contrast to the reported low incidence for normal populations (3 to 14%). Smooth muscle from any source in the body and from all species down to Amphibia is reactive with ASMA, and positive sera react with smooth muscle of autologous origin. The reaction is not readily demonstrable by complement fixation or by immunodiffusion when smooth muscle extracts are used as antigen.

Actin was recognized as one important determinant for reactivity when ASMA was reacted with cells from fetal liver and lung that were grown as monolayers in culture; the filamentous immunofluorescence reaction suggested that the reactant was a ubiquitous contractile protein resembling actomyosin in the membrane of various cells (Holborow, 1972). Also, rabbit antisera raised to human platelets that are rich in actin reacted with smooth muscle in a variety of sites (Becker and Nachman, 1973); the immunogenic component of platelets is thrombasthenin, which contains actin (Gabbiani *et al.*, 1973). Another site where membrane contractile protein is demonstrable, using ASMA as reactant, is the lymphocyte surface (Fagraeus *et al.*, 1974). Sera with ASMA activity were shown to react with renal glomeruli (Whittingham *et al.*, 1966), attributable to contractile elements containing actin in glomerular mesangial cells (Andrews and Coffey, 1983). Since the liver cell is rich in actin, its release in states of liver damage may serve as a source for the stimulation of ASMA.

Immunofluorescent reactions have been reported in CAH with bile canaliculi (Diederichsen, 1969) or with a structure at the perimeter of liver cells giving a polyglonal pattern of staining; certain of these patterns can be attributed to ASMA that is reactive with cell-surface actomyosin (Kurki *et al.*, 1980). The antiactin reactivity of the ASMA that occurs in autoimmune CAH has been confirmed by several studies (reviewed by Toh, 1979) and can be summarized as follows: ASMA is reactive with actin-rich I-bands of skeletal muscle myofibrils and can be completely

absorbed by F-actin from smooth or skeletal muscle; eluates derived from ASMA passed through actin-immunoabsorbant columns give the same staining reactions as the original sera. Finally, specific antibodies raised in rabbits to purified skeletal muscle actin give the same staining reactions as the human autoantibody.

Recognition that ASMA comprises a family of autoantibodies with a spectrum of tissue reactivity followed the description of three different immunofluorescent staining patterns for ASMA on frozen sections of rat kidney, namely SMA-T (tubules), SMA-G (glomeruli), and SMA-V (vessels) (Botazzo *et al.*, 1976). The autoantibody SMA-T reacted with renal tubules and glomeruli as well as with actin-containing structures in other tissues, SMA-G reacted with renal glomeruli, and SMA-V reacted with walls of blood vessels; the reactivity of most SMA-T and some SMA-G was neutralized by actin, but that of SMA-V was not, pointing to ASMA having actin or nonactin specificity. Further definition of the specificity of ASMA came from studies on acetone-fixed cultured fibroblasts, which indicated that ASMA could be reactive with different components of the cellular cytoskeleton, including three types of ubiquitous cytofilaments recognized ultrastructurally on the basis of filament diameter as microfilaments (6 nm), intermediate filaments (10 nm), and microtubules (25 nm); the main protein subunits for these are, respectively, actin, vimentin (in the case of fibroblasts), and tubulin. Immunofluorescence studies using antibodies raised in rabbits to pure proteins suggested the presence of various proteins among each of these cytoplasmic filaments (Toh, 1979).

Actin has been identified as the cytoskeletal reactant in autoimmune CAH (Lidman *et al.*, 1976). Actin is present in the cytoplasm in three distinct forms: monomeric G-actin, polymeric F-actin filaments, and higher orders of filament complexes including longitudinal stress fibers; the dynamic relationship between these is dependent on a complex group of actin-associated proteins (Weeds, 1982). The antiactin reactivity of sera from cases of autoimmune CAH is demonstrable by immunofluorescence using acetone-fixed cultured rat or human fibroblasts, which allow detection of reactivity with stress fibers (F-actin) (Pedersen *et al.*, 1982) (Fig. 5), or by an ELISA using monomeric G-actin as substrate (Bretherton *et al.*, 1983). Reactivity with F-actin appears to have the greater specificity for autoimmune CAH (Cunningham *et al.*, 1985)

Anticytoskeletal antibodies occur in hepatic diseases other than CAH, as well as in certain other nonhepatic diseases (Kurki and Virtanen, 1984). Some of these reactions are attributable to antibodies to intermediate filaments, demonstrable by the use of cultured fibroblasts, which allow specification of the reactivity as antivimentin. Antibody to vimentin appears to occur as a response to viral infection, including hepatitis B virus,

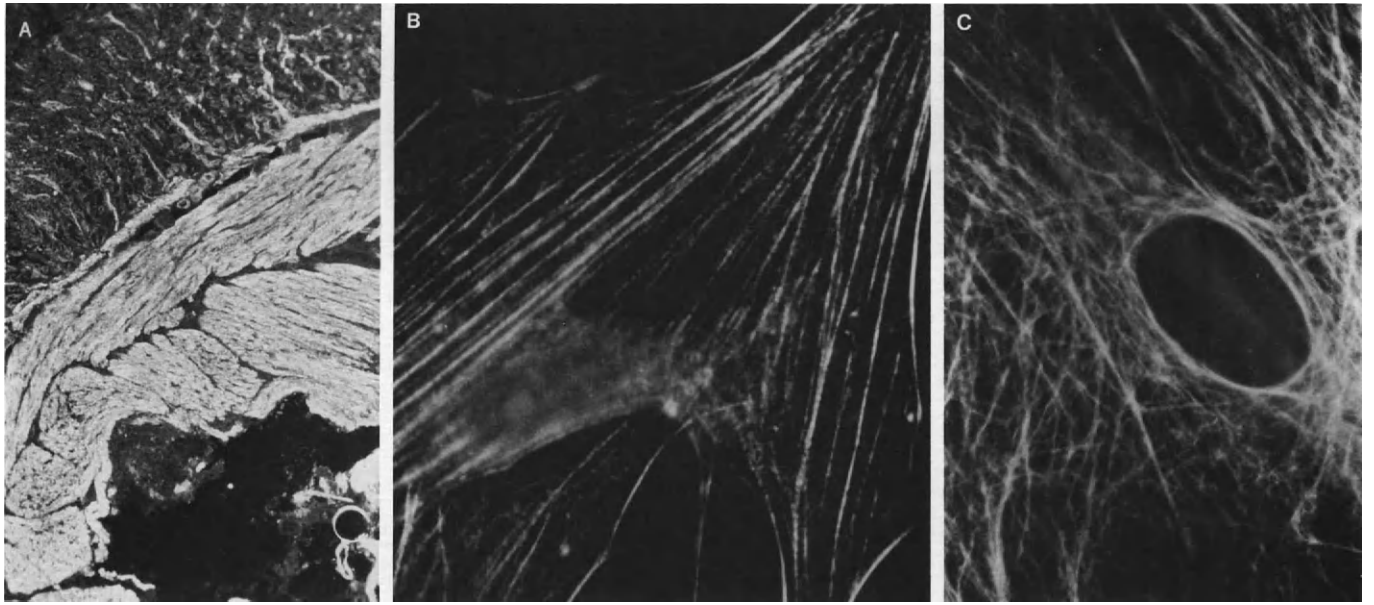


FIG. 5. The smooth muscle antibody (SMA) reactivity of sera in chronic hepatitis by indirect immunofluorescence. (A) Reactivity with smooth muscle of muscularis and muscularis mucosae in section of mouse stomach; (B) reactivity with "stress cables" of cultured fibroblast, illustrative of antiactin component of SMA reactivity, seen particularly in autoimmune hepatitis; (C) reactivity with perinuclear intermediate filaments of cultured fibroblast illustrative of antivimentin component of SMA reactivity, seen particularly in infection with viruses including HBV.

possibly because of shared determinants between viruses and intermediate filaments, as shown at least for measles virus (Fijihami *et al.*, 1983).

e. Antimitochondrial Antibody (AMA). The several types of AMA are discussed in Section II. One type, known as M4, occurs in a variant form of CAH in which there are overlapping features of CAH and PBC.

f. Antimicrosomal Antibody. Antibody to cell microsomes with a limited tissue distribution occurs in association with a rare subset of autoimmune CAH, as yet not well characterized, occurring in young patients (Odièvre *et al.*, 1983). The antibody is demonstrable by immunofluorescence and is complement fixing; it reacts with cells in the third portion of renal proximal tubules and with hepatocytes (Rizzetto *et al.*, 1973) and hence is called liver–kidney microsomal (LKM) antibody, but it may react also with cells in bronchial, esophageal, and duodenal epithelium. Interestingly, LKM antibody is not associated with the presence of other autoantibodies, ANA and ASMA, which are characteristic of the usual type of autoimmune CAH. The antigen, which can be absorbed with a microsomal preparation, is a lipoprotein.

g. Liver-Specific Protein. The expectation that liver tissue would contain organ-specific antigens, presumably associated with the cell membrane, led to a search (as yet unsuccessful) for a membrane antigen for liver-specific antibodies; this could be revealed by immunizing animals or could occur spontaneously in liver disease. Such a reactant could provide the immunogenic stimulus for initiation and/or perpetuation of CAH and present target structures for an immune attack.

Historically, studies extend back some 20 years, when Meyer zum Büschenfelde and Schrank (1966) extracted soluble proteins from liver; heterologous antisera were prepared, and, after appropriate absorptions, these antisera reacted specifically with the liver protein preparation. Subsequently liver-specific reactions were obtained with two chromatographically derived fractions; LP-1, present on the surface of hepatocytes and LP-2, present in the cytoplasm (Meyer zum Büschenfelde and Miescher, 1972). Attention was directed particularly to LP-1, and in earlier studies (1972–1974), cellular immune reactions measured by leucocyte migration-inhibition assays were reported in cases of CAH (Meyer zum Büschenfelde *et al.*, 1974): however, interest in such assays waned, perhaps because of their tedious nature and the capricious results obtained. McFarlane *et al.* (1977) further purified the membrane preparation described as LP-1 and found by using Tris buffer containing EDTA that the material remained stable; it was then named liver-specific protein (LSP).

Liver-specific protein is now recognized as a crude high molecular weight ($4–20 \times 10^6$) preparation with non-organ-specific and non-species-specific moieties. A point recently made by Meyer zum Büschenfelde and Manns (1984) is that "LSP—the preparation" should refer to the material derived by Sepharose 6-B chromatography, whereas "LSP—the molecule" should refer to the organ-specific non-species-specific determinant(s) originally identified by heterologous antisera. However, it may not be legitimate to refer to LSP as a molecule in view of the demonstration that an LSP preparation contained material reactive with antisera to various liver cell constituents, mitochondria, smooth muscle, reticular tissue, and so on (Riisom and Diederichsen, 1983).

In laboratory animals (mice and rabbits), cellular and humoral immune responsiveness to determinants of LSP is demonstrable after immunization with LSP in complete Freund's adjuvant (CFA); humoral antibody to xenogeneic LSP was readily measurable by various techniques and cell-mediated immunity by blast transformation of lymphocytes (see Frazer and Mackay, 1984a). The humoral immune response to LSP in animals is primarily directed, at least initially, against species-nonspecific antigens in LSP, although, in both mice and rabbits, species-specific autoantibody can be produced by repeated immunization. In the mouse, the induction of autoantibody is not accompanied by the development of liver disease, whereas in the rabbit, reports are conflicting (Mackay and Frazer, 1984).

In man, serum antibody to allogeneic human LSP can be measured by radioimmunoassays based on precipitation of complexes of LSP and antibody to it by radioiodine-labeled staphylococcal protein A (Jensen *et al.*, 1978) or antiimmunoglobulin antiserum (Kakumu *et al.*, 1979). Sera from patients with the autoimmune and HBV-associated types of CAH give comparable degrees of reactivity, and lesser degrees of reactivity are demonstrable in liver diseases other than CAH; in Jensen's study, serum titers correlated best with the degree of periportal piecemeal necrosis assessed histologically. Overall, serum titers of antibody to allogeneic LSP in CAH fall well short of titers to xenogeneic LSP developed after deliberate immunization of animals. The occurrence of antibody to LSP as an apparent consequence of liver damage in various diseases leaves the question open as to whether antibody to LSP is any indication that an autoimmune process is operative in pathogenesis.

h. Liver Membrane Antigen. A liver membrane antigen that differs from LSP became recognized when it was found that sera from rabbits injected with preparations of liver proteins reacted by immunofluorescence with the membranes of isolated hepatocytes (Hopf and Meyer zum Büschen-

felde, 1974). Thereafter it was shown by immunofluorescence that sera from certain types of chronic hepatitis would bind to rabbit hepatocytes, and that hepatocytes from patients had bound antibody on their surface: The pattern of binding differed according to whether the hepatocytes were from autoimmune CAH (linear binding) or HBV-associated CAH (granular binding) (Hopf, *et al.*, 1976). The reactive liver membrane antigen (LMAg) is species nonspecific and has been partly separated and purified by affinity chromatography using reactive sera from patients with CAH. Liver membrane antigen differs from LSP as judged by precipitation patterns and counterimmunoelectrophoresis (Meyer zum Büschenfelde *et al.*, 1979).

Since immunofluorescence reactions are difficult to quantify, we developed an immunoassay for antibody to LMAg based on a two-step procedure in which serum was exposed to isolated hepatocytes; binding was revealed by the use of ^{125}I-labeled staphylococcal protein A (Frazer *et al.*, 1983). Monkeys were selected for the source of hepatocytes by reason of the closeness of this species to man and availability of liver tissue. Hepatocytes were treated with glutaraldehyde to fix antigens on the surface and to preserve cells for repetitive assays. Serum binding to hepatocyte membrane had some degree of specificity for autoimmune CAH, although binding could be demonstrated with most normal sera up to dilutions of 1:2048, and there was a clear gradation in the degree of serum binding activity from normal, miscellaneous liver diseases including acute viral hepatitis, HBV-associated CAH, and cryptogenic CAH, to autoimmune CAH in which the highest degrees of binding activity were present. The LMAg that reacted in this immunoassay was found not to be inhibitable by LSP or F-actin, but reactivity decreased after treatment of hepatocytes with agents that degraded protein or glycoproteins; of interest, $F(ab)_2$ fragments from reactive sera used in competitive inhibition experiments indicated that the reactive determinant(s) of LMAg for sera from cases of autoimmune and HBV-associated CAH were the same (Frazer and Mackay, 1984b). In further studies, using the Western immunoblotting procedure on liver membrane components separated by polyacrylamide gel electrophoresis, reactivity of sera with several or more bands on the nitrocellulose transfer was detectable with ^{125}I-protein A; however, none of these reactivities seemed specific for autoimmune CAH (Mackay *et al.*, 1984).

2. Immunological Derangements

The immunological derangement(s) that determines the development and progression of autoimmune CAH is unknown, and many of the ob-

served derangements could be a result rather than a cause of the disease.

a. B-Cell Hyperfunction. This is clearly exemplified by polyclonal hypergammaglobulinemia (Section I,E,1,a) and non-tissue-specific autoantibodies. It is likely that B cells are in a state of release from nonspecific suppressor influences. There have been consistent reports of abnormally high antibody titers to viral antigens, rubella, and measles, but not to bacterial antigens. There are increased numbers of immunoglobulin-producing cells in blood, as judged by hemolytic plaque assays, and this may reflect decreased suppressor activity by either T cells or monocytes, which can influence the behavior of B cells (Holdstock *et al.*, 1982; Burns *et al.*, 1982).

b. Serum Complement. Levels of serum complement in the precirrhotic stage of autoimmune CAH are not abnormal, but low levels may occur in advanced stages due to decreased synthesis (Finlayson *et al.*, 1972).

c. T-Lymphocyte Hypofunction. There is a depressed level of reactivity of T lymphocytes in various assays for T-cell activity (Mackay, 1982). For example, there is decreased responsiveness to challenge with panels of antigens that elicit delayed-type hypersensitivity, decreased inducibility of cutaneous reactivity to DNCB, and a decreased mitogenic response of blood mononuclear cells to phytohemagglutinin. Such changes are seen in various chronic diseases and may be secondary.

d. Suppressor T-Lymphocyte (T_s) Hypofunction. A decreased function of suppressor T cells in blood is demonstrable in CAH, according to functional assays *in vitro*. A below-normal recruitment of T_s cells by concanavalin A (Con A) was shown in two studies, as judged by an impaired capacity of Con A-recruited cells to inhibit the proliferation of lymphocytes in response to this mitogen (Hodgson *et al.*, 1978; Coovadia *et al.*, 1981). The relevance of these non-antigen-specific assays is uncertain, and, moreover, attention can be drawn here to the poor correlation between numerical and functional assays of T suppressor cells (Alexander *et al.*, 1983).

e. Lymphocyte Counts in Blood. Counts in blood of surface immunoglobulin-positive B lymphocytes are normal. Counts of E-rosette-forming T cells have been reported to be decreased (Mackay, 1982), but this has not been substantiated by studies using monoclonal antibodies. Counts of cells with Fc receptors for IgM (T_μ cells, helper T cells) or for IgG (T_γ) are not substantially altered (Coovadia *et al.*, 1981). Results are

now available from several studies on T-lymphocyte subsets in blood using monoclonal antibodies, and as pertains in many autoimmune diseases, the proportions of T4+ (helper) cells and T8+ (suppressor) cells are not markedly altered. Analyses in CAH of blood lymphocyte populations using monoclonal markers for helper (OKT4, Leu 3) T cells and suppressor (OKT8, Leu 2) T cells give added confirmation of regulatory disorder, since the helper subset is overrepresented in HBsAg-negative CAH and the suppressor subset is overrepresented in HBsAg-positive CAH (Mackay, 1983). Exposure of blood lymphocytes from cases of autoimmune CAH to prednisolone *in vitro* was shown to increase depressed T_s function, explained by prednisolone facilitating a differentiation step for the T_s subset (Nouri-Aria *et al.*, 1982).

f. Lymphocyte Counts in Liver. Application to frozen sections of liver of a panel of monoclonal antibodies identifies infiltrating lymphoid cells after appropriate staining procedures, either immunofluorescence or immunoperoxidase. In the past, routine histology indicated that the predominant cell type among the portal and periportal infiltrates in CAH was the plasma cell, but this now comes into question. Several of the recent publications based on application of monoclonal antibodies to sections of liver in different types of CAH have been summarized by Alexander and Williams (1984). In all types of CAH there is a predominance of cells bearing T cell markers, T3, T4, and T8, and relatively few cells bearing markers for monocytes, K cells, or NK cells. In general, in all types of CAH, cells of T_s phenotype predominate in areas of piecemeal necrosis and those of Th phenotype in portal tracts. The majority of the T cells visualized are "activated," as judged by *HLA-DR* positivity; a proportion (some 30%) of cells bear the T10 marker (Frazer *et al.*, 1985).

3. Mechanisms of Damage

The characteristic periportal lymphocyte-plasma cell infiltrate with its provocatively close association with necrotic or "ballooned" hepatocytes in chronic hepatitis (piecemeal necrosis) brings up the as yet unanswered questions of what evokes this response and how the infiltrating cells might be harming the liver: Almost all possible mechanisms known to immunopathology have been invoked, as is illustrated in Fig. 6.

In the liver, there is a clear predominance in areas of periportal piecemeal necrosis of cells with features of activated T lymphocytes, and in ultrastructural studies, a close apposition between infiltrating lympho-

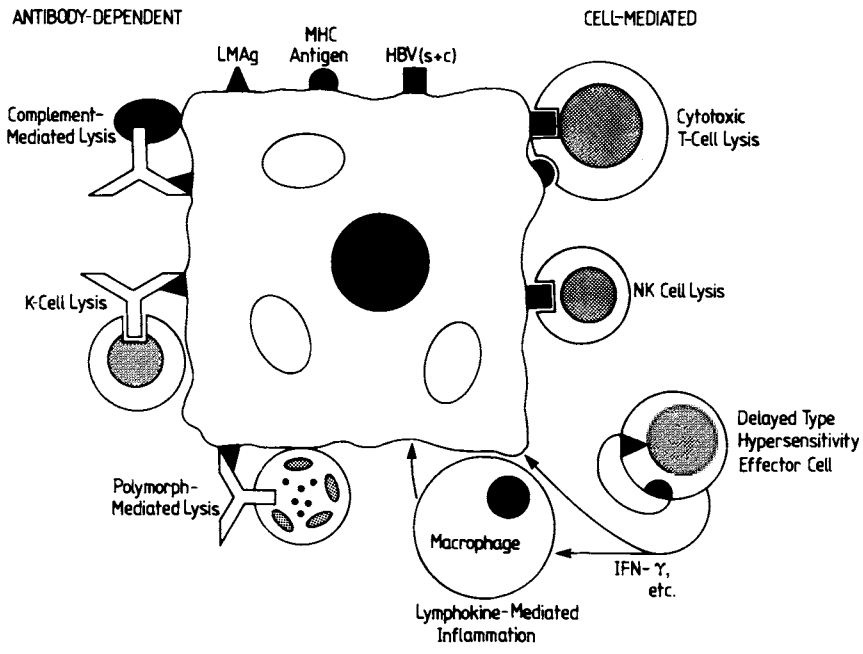


FIG. 6. Postulated cytotoxic mechanisms in CAH. Surface structures relevant to cytotoxicity in autoimmune or viral CAH include liver membrane autoantigens (▲), MHC antigens (HLA-B8, DR3) (●), and HBV antigens (■) (s, surface; c, core). Possible effector agents are shown around the margin of the cell. The Th effector cell is a cell of the T-helper lineage which may release lymphokines that activate other cytotoxic cells such as macrophages (Mϕ). The requirement for T cells to engage both an MHC antigen and a cell-surface antigen is indicated. None of these postulated mechanisms is yet established.

cytes and hepatocytes is evident, consistent with a direct cytotoxic effect of these lymphocytes (Bernuau *et al.*, 1982). Monoclonal markers suggest that the cytotoxic population consists of T8+ lymphocytes.

It is difficult to derive information on mediators of cytotoxicity in the liver, because of nonrepresentation of relevant cell population in peripheral blood and the problem of deriving autologous hepatocytes, which are rather fragile, as targets in reliable cytotoxicity assays. Earlier studies were reviewed by Mackay (1982). Work from Kings College Hospital, reviewed by Mondelli *et al.* (1984), suggested that cytotoxicity is antibody dependent in autoimmune CAH and is mediated by T8+ lymphocytes in HBV-associated CAH. One possibility, not yet explored, is that the periportal cellular infiltrates contain cells known as activated lympho-

cyte killer (ALK) cells; under *in vitro* conditions, these originate either from NK cells or from T cells under the influence of lymphokines, including interleukin (IL)-2, but the role of such cells *in vivo* is unknown. It would be necessary for an initial immunological stimulus to generate release of IL-2 and attract and activate ALK cells.

4. Immunogenetics

These are strong genetic determinants of autoimmune CAH. The disease has long been known to occur most frequently among populations of northern European extraction. The female to male incidence ratio is ~8:1. Mackay and Morris (1972) described a high frequency (70%) of *HLA-B8* among patients with CAH, mostly of autoimmune type, and this was subsequently confirmed in several European and North American studies (Mackay, 1977, 1984). Selection of cases of CAH according to more precise diagnostic markers for the autoimmune type of disease appeared to result in an even higher association of *HLA-B8* (85%) with a relative risk of 15 (Freudenberg *et al.*, 1977). Subsequently there was established an association with autoimmune CAH of *HLA-Dw3* (Opelz *et al.*, 1977) and -DR3; moreover it was found among cases of autoimmune CAH that *HLA-B8* and *HLA-DR3* were in high linkage disequilibrium due to the presence of genes existing as a haplotype on the one chromosome, as judged by family studies (Mackay and Tait, 1980). It is curious that *HLA-B8, -DR3* is strongly associated with various other immunopathological diseases, but there is little overlap between these and autoimmune CAH. There are two explanations for the *HLA-B8, -DR3* association with immunopathological disease. One is that a gene product (Ia molecule) expressed on the surface of antigen-presenting cells becomes readily associated with the disease-provoking antigen (or autoantigen) so as to stimulate a response by T cells. The other explanation is that the *HLA-B8, -DR3* specificities are phenotypic markers for weak immunoregulatory activity (Eddleston and Williams, 1978; Ambinder *et al.*, 1982), and that other genetic or environmental factors determine the organ or tissue susceptibility among the *HLA-B8-* or *-DR3*-associated immune-mediated diseases. In addition there may be modifying effects of other *HLA* alleles such as *HLA-B12*, the presence of which may be prejudicial to survival (Eddleston and Williams, 1978). Studies relevant to associations between *HLA* and Gm and immunopathological diseases including CAH have been reviewed (Mackay, 1984; Whittingham *et al.*, 1984a); some arguments on the genetic associations are listed in Table I.

In addition to *HLA-B8, -DR3*, there is another weak although interest-

TABLE I

Summary of Arguments on *HLA-B8, -DR3, -Dw3* Haplotype[a]

Observation	Inference
A number of different immunopathological disorders are associated with *HLA-B8, -DR3,* and *-Dw3*	One (or more) immunoregulatory alleles are in linkage disequilibrium with *HLA-B8, -DR3, -Dw3*
Celiac disease and other disorders tend to "breed true" in families and segregate with *HLA-B8, -DR3, -Dw3*	At least one modifying allele may be linked to *-B8, -DR3, -Dw3* haplotype
Diabetes mellitus of juvenile onset (type 1) is most concordant in HLA-identical siblings with *HLA-DR3, -DR4*	At least one modifying allele can be linked to *-DR4* haplotype
Autoimmune chronic active hepatitis shows	
HLA-B8, -DR3 association	See above
Effects related to B12.....	Suggests modifying genes on *HLA-B12* haplotypes
Effects related to Gm	Suggests modifying genes on Gm haplotypes
Little familial aggregation	Suggests that there are several modifying loci

[a] From Mackay *et al.* (1980).

ing genetic association with autoimmune CAH, the Gm allotypic marker G1m (1,2), with a relative risk of ~2 (Whittingham *et al.*, 1981a); however the presence of both *HLA-B8* and G1m (1,2) in the one patient appears to confer a 40-fold added risk for disease, suggesting that the association of *HLA-B8* and G1m (1,2) is interactive (synergistic) rather than additive (Whittingham *et al.*, 1981a). The explanation may be that genes for *HLA* and Gm are each associated with genes that specify antigen recognition by immunocytes: *HLA* is associated with the product of *Ir* genes, or according to Klein (1982), in fact may itself be this product, and genes for Gm allotypes are in linkage with *V* genes (Whittingham *et al.*, 1984a).

Given this immunogenetic predisposition to autoimmune CAH, it is curious that there is a paucity of published reports on coexistences within families of autoimmune CAH, or CAH with other *HLA-B8*-associated diseases; the few examples include pedigrees in which family members had either CAH or various other nonhepatic immune-mediated diseases (Joske and Laurence, 1970; Whittingham *et al.*, 1970). On the other hand, a high frequency of ANA reactions among relatives of cases of lupoid

hepatitis (MacLachlan *et al.*, 1965) and a high frequency of hyperglobulinemia among relatives in a single pedigree (Cavell and Leonhardt, 1965) and among relatives of patients with cirrhosis possibly consequential to chronic hepatitis (Elling *et al.*, 1966), were reported. Finally, of relevance are the results of two family studies based in England (Galbraith *et al.*, 1974) and in Finland (Salaspuro *et al.*, 1976) in which the occurrence of disease and a range of autoantibodies were examined in first-degree relatives of cases of CAH or PBC: While in both studies there was an increased frequency of certain autoantibody specificities in the relatives, there was not a close agreement in the two studies on the specificities involved (Mackay, 1984). The most that can be said is that all of the above-cited data point to an inheritance of a rather weakly expressed defect in self-tolerance, perhaps based on a genetic defect in suppressor-cell function as has been reported for cases of SLE (Miller and Schwartz, 1979).

5. Laboratory Diagnosis

The issues to be considered under this heading are first, the diagnosis of the disease CAH, and second, the differentiation of autoimmune CAH from other types, of relevance to the delivery of appropriate therapy (see below).

The diagnosis of CAH depends predominantly on experienced interpretation of biochemical tests of liver function, including raised levels of transaminase enzymes, a decreased level (in later stages) of serum albumin, and a raised level of serum globulin, exclusively of IgG class. Skilled histological examination of a liver biopsy specimen is essential. The diagnosis is based on the characteristic pattern of periportal necrosis and accumulation of lymphoid cells, including many pyroninophilic cells in portal and periportal areas, and almost invariably the histopathological pattern is that of chronic active rather than chronic persistent hepatitis, although the latter may be seen after successful treatment with prednisolone. Clear specification of the autoimmune versus other subtypes is usually, but not invariably, possible by use of serological tests. The most useful markers are hypergammaglobulinemia at levels of at least 20 g/liter, and autoantibodies.

It must be appreciated that serum autoantibodies are present as continuous variables that can be measured over a quantitative range and with (at present) a regrettable lack of standardization of procedures for their detection and titration. Moreover, disease-relevant autoantibodies (and this applies particularly to liver disease) tested for by immunofluorescence

may be quite heterogeneous in regard to antigenic specificities. The diagnostically relevant autoantibodies for autoimmune CAH are ANA and ASMA, but there are variant types of CAH marked by positive reactions for antimitochondrial antibody and antimicrosomal antibody of the type reactive with liver and kidney microsomes and known as LKM. Results should be calibrated against accepted reference standards: Standards for ANA exist, and ASMA has been "submitted pending WHO review" as a reference preparation. It is essential that titration is performed, although there is no consensus on diagnostically positive results: Cut-off titers cited for positivity range from 1:20 to 1:80. A useful although not widely used ANA reaction is that with specificity for the granulocyte nucleus; granulocyte-specific ANA to high titer is with few exceptions confined to autoimmune CAH and rheumatoid arthritis.

Assays for liver-specific autoantibodies remain cumbersome and have not yet obtained any degree of diagnostic precision or usefulness. Antibody to LSP does not specify autoimmune types of CAH, and antibody to LMAg has diagnostic specificity for autoimmune CAH, but measurement is not routinely applicable; a simplified dot-blot assay for anti-LMAg could prove useful (Mackay *et al.*, 1984). There are few diseases in which diagnostic HLA typing is cost effective, but some weight can be attached to the detection of *HLA-B8* (and *-DR3*) in CAH.

6. Relationship of Autoimmune Hepatitis to Other Types of CAH

In this chapter, consideration has been limited to that type of CAH with autoimmune features. Of numerically greater significance in global terms is the type associated with infection with HBV in the liver. The HBV-associated type of CAH (CAH-B) is similar to autoimmune CAH in many respects, clinically and histologically, but differs according to (a) the absence of serological markers of autoimmunity, (b) the genetic background, since *HLA-B8* is not increased in CAH-B, (c) geography and ethnic structure of populations affected, and (d) the low responsiveness to treatment with corticosteroid drugs. The degree to which the pathogenesis differs awaits better understanding of determinants of immune damage in the liver. It is assumed that in CAH-B there is expression in the liver cell membrane of antigens specified by the genome of the hepatitis B virus, HBsAg or HBcAg, but which predominates and which is the more provocative of immunological attack is at present uncertain.

There exist other types of CAH, defined histologically, but the only one that bears any clinical resemblance to autoimmune CAH is that induced

by certain drugs. In all other types the serological and genetic features of autoimmune CAH are lacking, suggesting that the histological characteristic of periportal piecemeal necrosis does not have any etiological specificity.

F. Treatment and Outcome

Historically, a clear indication that treatment with corticosteroid drugs was effective in autoimmune hepatitis was given in 1959 by the immediate effect on raised levels of transaminase enzymes, together with other evidence of hepatic functional improvement including reduction in serum bilirubin (O'Brien *et al.*, 1958). Although most writers in the early 1960s considered that prednisolone ameliorated the clinical expression of the disease, the experience was based on limited case numbers and uncontrolled observations or relatively short-term follow-up. A widely held view was that corticosteroid drugs, despite conferring short-term benefit, might not influence the outcome of the disease, and long-term randomized treatment trials were required. In view of a conviction that corticosteroids were of benefit at least in the early stages of the disease, we considered it unjustifiable to initiate an untreated control series to assess the long-term effect (Mackay and Wood, 1961, 1962); however, in a long-term treatment study (Mackay, 1968) in which the pretreatment state of the patient was the control index, it was shown that, after institution of long-term prednisolone or of prednisolone in lower dosage with azathioprine, there was freedom from relapses, transaminase levels remained at or near normal levels, and all indices of liver function synchronously improved within 3 months and remained significantly improved compared with pretreatment values. In this study azathioprine alone also appeared to exert control over relapses.

However, randomized paired treatment versus nontreatment trials were still required to assess survival, and three were completed around 1970 (Copenhagen Study group, 1969; Cook *et al.*, 1971; Soloway *et al.*, 1972): All testified to a greatly improved survival of prednisolone-treated cases. The trial of 49 patients by Cook *et al.* (1971) was convincing in that deaths among the 27 nontreated patients during the period of observation numbered 15 (56%) and three of the remaining 12 had relapses, in marked contrast to outcome in the 22 treated cases. Although prednisolone became a well-accepted therapy for CAH after 1970, problems still remain. These include (a) the identification of prednisolone-responsive cases, (b) withdrawal strategies, and (c) the use of azathioprine as adjunctive treatment. Azathioprine in combination with prednisolone has the advantage of allowing lower doses of prednisolone in addition to having a possible

independently suppressive activity (Mackay, 1968; Schalm, 1982); the long-term oncogenic risk of azathioprine is yet to be fully evaluated.

In regard to selection of cases for treatment, only those patients meeting diagnostic criteria for autoimmune CAH, including hypergammaglobulinemia and serological markers, will give an unequivocal response; indeed, such patients are steroid dependent, since, within 2–8 weeks after prednisolone has been reduced or withdrawn, a rise in serum transaminase levels usually occurs, followed by clinical relapse. Indicators of a weak or marginal response to prednisolone include positively for serological markers of HBV infection or advanced end-stage cirrhosis; for further comment on selection for treatment of CAH, the reviews of Czaja (1984) and Schalm (1982) may be consulted. In regard to treatment withdrawal, a suggested policy is to maintain treatment for an initial period of 3 years, at levels that maintain normality of transaminase enzymes; then, if a trial withdrawal of treatment is unsuccessful, treatment should be maintained for a further 3-year period. Only some 10% of patients will fail to enter remission after 3 years of treatment, but some 60–90% relapse after treatment withdrawal (Hegarty *et al.*, 1983; Czaja, 1984). Survival after 10 years for adequately treated patients is about 70% (Steven *et al.*, 1979; Czaja, 1984), as indicated in Fig. 7. Of note in terms of survival,

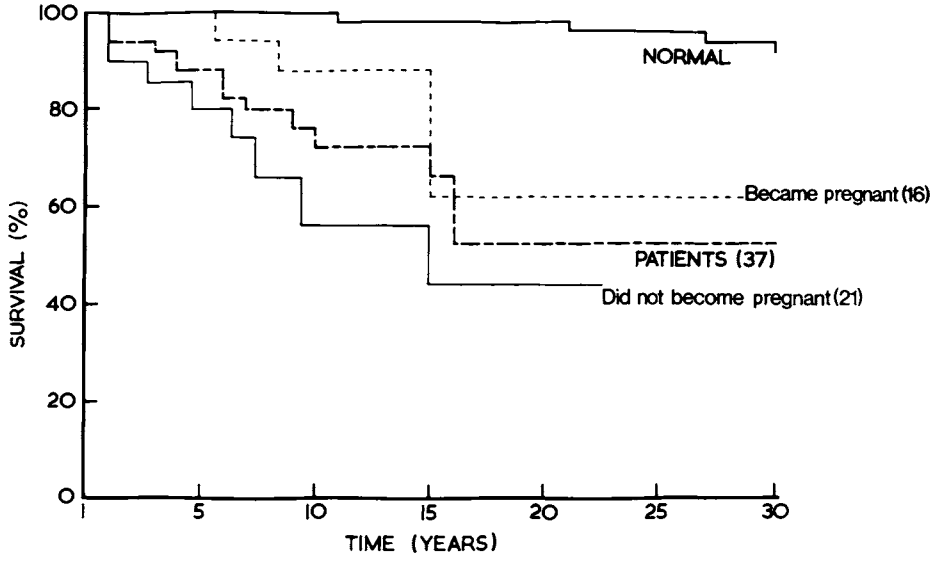


FIG. 7. Survival experience up to 30 years of 37 patients with autoimmune CAH compared with that of the age- and sex-matched normal population: the capacity to become pregnant in women was associated with better survival. (From Steven *et al.*, 1979.)

development of cirrhosis seems to be of lesser importance than persisting activity of disease (Davis *et al.*, 1984).

G. Concluding Remarks: Future Prospects

Autoimmune CAH has been under immunological scrutiny for some 30–35 years. However, despite the marked serological abnormalities and the striking histological evidence of immunopathology, explanations for the mode of origin and apparent self-perpetuation of this disease are still lacking. One reason is the lateness of the clinical expression of the disease, as evident from the name adopted, *chronic* active hepatitis. There are few clues to the initiating process of events on day 1 in the history of the disease, nor can it be stated whether the onset of autoimmune CAH is better attributable to a sequel of liver cell damage or a fault arising *de novo* in immune recognition or regulation. The autoantibodies characteristic of autoimmune CAH, including antibody to LSP and LMAg, can be detected in cases of chronic liver damage from various causes and so may be a consequence of this; yet, on the other hand, autoimmune CAH has not been reported to occur after liver damage due to acute virus hepatitis A.

Once established, autoimmune CAH is remarkably tenacious, with predisposition to recurrence even after prolonged drug-induced suppression. Evidence for disordered immune regulation is provided by functional indices showing impaired T-cell-suppressor activity and numerical indices using counts of T-cell subsets in blood, although there are discrepancies between these. Also, despite a number of studies in CAH showing a preponderance of T lymphocytes adjacent to damaged liver cells, the mediators of hepatocyte destruction are unknown: the lack of a good animal model is a decided handicap. Cytolsis may be effected by T8+ cytotoxic lymphocytes, by injurious lymphokines released by activated T4+ lymphocytes or by the newly recognized class of activated lymphocyte killer cells: Whichever cell may operate, its activity could be potentiated by induction of Ia antigen expression by lymphokines including interferon α; application of more comprehensive panels of monoclonal antibodies to tissue sections may prove informative here.

A plea must be made for rigorous specification in research studies of subtypes of CAH, since there are now sufficient histological, serological, and genetic markers to do this. Thus, pathogenetic mechanisms applicable to autoimmune CAH may or may not be applicable to HBV-associated or other types of disease. Moreover identification of the type of CAH is of major importance in selecting cases for long-term treatment, since only those with autoimmune CAH appear to derive benefit. There remains in

most case studies a substantial group of cases designated as cryptogenic CAH, of which some may be related to antecedent infection with HBV or NANB viruses and others to an antecedent autoimmune process; there will be other causes also. The development of more sensitive serological markers, or the applications of molecular biology technics to identify host-integrated HBV, should be of assistance here.

H. Summary

Autoimmune CAH became recognized as a disease around 1950, was associated with stigmata of autoimmunity in 1960, and was distinguished from other types of CAH in 1970. It is identifiable by clinical and histological features, disease-specific serological markers, the HLA-B8, -DR3 phenotype, and responsiveness to corticosteroid drugs. There is no close animal model. The serological markers used for diagnosis include autoantibody to nuclei and smooth muscle, but the actual specificity for the latter reactivity is the cytoskeletal filament, actin; autoantibodies to microsomes and mitochondria are present in variant forms of the disease. Autoantibodies to liver-specific protein and liver membrane antigen remain undefined: There are multiple reactants in hepatocytes for each of these specificities, and their pathogenetic significance is uncertain. There are various immunoregulatory abnormalities in CAH, but their role is still unclear. The periportal cellular infiltrate in the liver has been defined by monoclonal antibodies as predominantly T cellular, but cell counts do not strongly implicate either the helper or the cytotoxic subset of T lymphocytes. Treatment with prednisolone and azathioprine is highly effective in CAH; initially, treatment should be given for 3 years and may be needed indefinitely; survival after 10 years for adequately treated patients is ~70%.

II. PRIMARY BILIARY CIRRHOSIS

A. Introduction: Historical Background

The relationship between biliary obstruction and hepatic damage has been recognized since the mid-eighteenth century; it was Hanot, a "sad, but considerable figure" according to Mann (1974), who in 1875 defined our present concept of PBC, then described as hypertrophic biliary cirrhosis. Thereafter, a variety of terms were applied to this disease, but it was the superb clinical study of Ahrens and colleagues (1951) that led to the clear differentiation of liver damage due to intrahepatic and extrahe-

patic obstruction of bile ducts and to the definitive naming of the disease as primary biliary cirrhosis (PBC), but with no cause specified. In the early descriptions, much emphasis was placed on the lipid disturbances resulting from prolonged biliary obstruction, which is expressed clinically as cutaneous xanthomata (hence "xanthomatous" biliary cirrhosis), but this characteristic feature has become diluted out due to the recognition, by biochemical and serological tests, of milder expressions of the disease.

The first hint of immunological disorder in PBC came with the recognition that the serum of patients with the disease gave a high-titer reaction in the AICF test, using cytoplasmic homogenate as antigen (Mackay, 1958). The introduction of immunofluorescence allowed mitochondria to be identified as the source of antigen (Walker *et al.*, 1965), and it was soon recognized that antimitochondrial reactivity was highly specific for the disease. Subsequently, evidence accumulated for a multiplicity of immunological disorders in PBC (see Section II,E), and for the likelihood that the disease had a multisystem expression but with the intrahepatic biliary system as the main target. Yet the origin of the disease still remains an enigma, as does the serological reactivity for mitochondria in all tissues of the body yet with the singling out of the biliary system for immunological attack.

B. General Description, Classification, and Animal Models

Primary biliary cirrhosis can be defined as a chronic progressive liver disease due to inflammatory damage to intrahepatic bile ducts, with interference with bile secretion, that results in periportal hepatitis, fibrosis, and eventual cirrhosis. The underlying histopathological process is described as a "chronic nonsuppurative destructive cholangitis": either this term or "primary autoimmune cholangitis" would be more appropriate, since "cirrhosis" is a relatively late feature of the disease, but "primary biliary cirrhosis" is too entrenched a term to be supplanted. It is stated that PBC can be recognized in all population groups, and that no racial predilection exists. In a study from northeast England (Hamlyn *et al.*, 1983), the diagnosis rate was 1.0 per 100,000 (1.8 per 100,000 for females aged 15 or older), and the estimated prevalences for rural and industrial urban areas were 3.7 and 14.4 per 100,000, respectively, but this difference could be due to more asymptomatic cases being recognized in urban areas. In recent years there has been recognition that PBC may have a very wide disease expression in terms of severity and tempo, and distinct clinical stages are described (see Section II,D), but there appear to be no grounds on which to subclassify the disease, at least in relation to immu-

nological aspects. There is, however, awareness that mild asymptomatic disease is more frequent than hitherto believed, and that such disease is associated with a survival little different from that of the normal population (James *et al.*, 1981; Roll *et al.*, 1983).

The constellation of events necessary to set PBC on course, while still a mystery, appears to contain at least three components: genetic predisposition, an environmental trigger, and an immunoregulatory disturbance (James, 1983): Evidence relating to these components is presented in following sections. No disease that stimulates PBC occurs naturally, and an experimental model of PBC has not been induced by immunizing procedures.

C. Clinical Presentations

1. Features of PBC

Primary biliary cirrhosis has a marked predilection for women, the sex ratio being ~9 to 1, and the age of onset is usually 35–75 years, with a mean ~50 years. Age-specific onset rates were shown to rise linearly between ages 35 and 65 (Hamlyn *et al.*, 1983). The clinical features are almost entirely attributable to biliary obstruction; the earliest symptom is pruritus, which is followed by mild but progressive and unremitting jaundice. Hyperlipidemia is a classic feature and is expressed by lipid deposits (xanthelasmata) in eyelids, palmar creases, and other cutaneous sites. Clinical features in the later stages are those consequent on intestinal malabsorption and the development of cirrhosis with portal hypertension and hepatocellular failure. Clinicians and pathologists have found it useful, for prognosis and assessment of treatment, to assign four stages (see Section II,D, Histopathology). The biochemical features include the characteristically highly elevated levels of serum alkaline phosphatase, usually >500 IU/liter, hyperbilirubinemia, and moderate increases in hepatic transaminase enzymes, in addition to the immunological features (see Section II,E).

2. Associated Diseases

It is stated that 84% of cases with PBC have one associated autoimmune disease, and that 40% have two or more (Culp *et al.*, 1982). One of the most frequently cited diseases associated with PBC is Sjögren's syndrome (Alarcón-Segovia *et al.*, 1973; Golding *et al.*, 1973; Culp *et al.*, 1982) of secondary type, since the anti-La (SS-B) antibody characteristic

of primary Sjögren's syndrome is lacking (Bernstein *et al.*, 1984). Other associations include arthritis (Crowe *et al.*, 1980b), autoimmune hypothyroidism (Crowe *et al.*, 1980a; Elta *et al.*, 1983), scleroderma (Reynolds *et al.*, 1971), polymyositis (Benoist *et al.*, 1977), or SLE (Hall *et al.*, 1984); multisystem pathology may be expressed by lymphocytic accumulations in extrahepatic sites, kidney, and stomach (Mackay, 1960). A proposed similarity between the lesions of graft-versus-host (GVH) disease and those of PBC (Epstein *et al.*, 1980) led to the idea that PBC may be determined by a reaction against MHC antigens, but this idea has not gained popularity.

Of particular interest is the mixed form or overlap disease with coexisting features of PBC and CAH. This is seen in middle-aged to elderly women, and may represent some 10% of all cases diagnosed either as PBC or CAH. This syndrome is associated with multisystem features, particularly those associated with rheumatoid arthritis, Sjögren's syndrome, or scleroderma. Biochemically there are increased levels of both alkaline phosphatase and transaminase enzymes, and histologically there is bile duct injury and granulomata as in PBC and periportal necrosis as in CAH. Serologically, the overlap is sustained by the presence of antimitochondrial antibodies (see Section II,E,1,b) and the antibodies characteristic of CAH. However, there is no increase in HLA-B8, -DR3, as seen in autoimmune CAH.

D. Histopathology

The essential lesion of PBC involves the interlobular and septal bile ducts, which are ~100 μm in diameter, and four stages are recognized (Rubin *et al.*, 1965). In early lesions, in stage I, the ductular cells are damaged, the ducts are disorganized, and there are periductular accumulations of mononuclear cells, often accompanied by typical granulomas with multinucleate giant cells. With progression, stage II, biliary ductules are destroyed and their number becomes reduced, and, presumably as a regenerative activity, there is proliferation of small aberrant ductules, the portal tracts are expanded with mononuclear cells, and fibrosis is evident (Fig. 8); there are secondary changes affecting hepatic parenchymal cells, particularly areas of necrosis—feathery degeneration and bile infarcts—in regions of cholestasis where bile ducts are undergoing destruction. In stage III, there are invasive fibrosis of the hepatic lobule from portal tracts, pronounced disappearance of bile ductules, and accumulation of copper in periportal hepatocytes. The terminal stage IV is marked by the above features together with a macronodular cirrhosis.

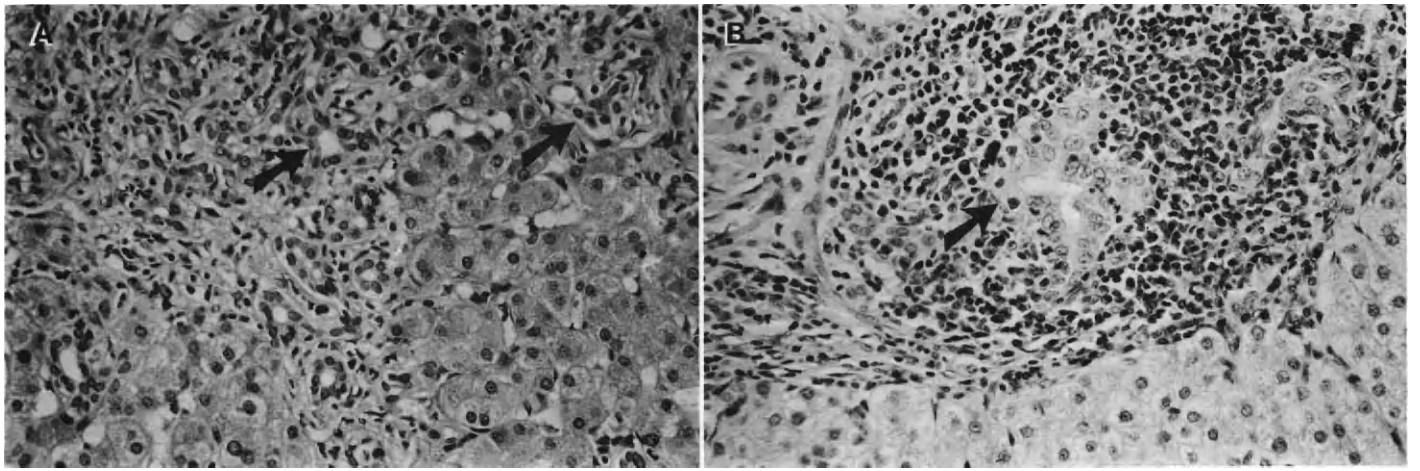


FIG. 8. The histological lesions of primary biliary cirrhosis showing (B) early stage 1 lesion with a periductular lymphoid infiltrate and biliary ductular damage (arrows), and (A) stage 3 lesion with obliterated bile ducts and duct cell hyperplasia (arrow) and portal fibrosis. (The assistance of Dr. John Dowling with the photomicrographs is gratefully acknowledged.)

E. Immunology

1. Immunological Features

a. Hypergammaglobulinemia. There is a moderately elevated level of gamma globulin, 15–30 g/liter with IgG and IgM represented (Feizi, 1968). There are unusual aspects to serum IgM in PBC, since IgM monomers may be present (Fakunle *et al.*, 1979), and the IgM has increased preciptability and capacity for fixation of complement without demonstrable antigen-binding activity (Lindgren *et al.*, 1981). Some of the antimitochondrial antibody is represented by the IgM isotype.

b. Antimitochondrial Antibody. This characteristic reactivity is against mitochondria from all mammalian cells, and also against mitochondria from lower organisms. In the standard immunofluorescence test system using unfixed cryostat sections in a composite tissue block (human kidney, thyroid, and stomach), the mitochondrial reactivity appears as a coarse granular cytoplasmic fluorescence; the HEp2 cell line is also a suitable substrate (Fig. 9). Screening tests should be performed at 1:10 dilutions of serum because some undiluted sera may give nonspecific fluorescence, and strongly positive sera may show a prozone effect (Berg and Baum, 1980). The frequency of AMA reactions in PBC is at least 90% (Walker *et al.*, 1965; Klatskin and Kantor, 1972; Christensen *et al.*, 1980), and AMA reactivity occurs in few other diseases, so that sensitivity and specificity for the diagnosis of PBC is extremely high. Antimitochondrial antibody reactivities may have several different specificities (Berg and Baum, 1980; Berg *et al.*, 1982), and each is associated with a particular clinical condition (Table II). These specificities are known as M1 to M6, but only the M2 and M4 antibodies are associated with liver disease: M2 which reacts with a trypsin-sensitive antigen is the true AMA of PBC, and M4 which reacts with a trypsin-insensitive antigen characterizes the CAH-PBC overlap syndrome.

Polyacrylamide gel electrophoresis of human and other mitochondrial preparations and immunoblotting onto nitrocellulose, followed by probing with PBC sera, were used to identify reactive mitochondrial antigens by Frazer and colleagues (1985). All AMA-positive sera reacted with two trypsin-sensitive human mitochondrial polypeptides, with apparent MWs of 70 and 45 kD. There were counterparts of the 70 kD antigen in mitochondrial preparations from rat and mouse with MWs of 65 to 70 kD, and in yeast and *E. coli* with MW of ~55 kD. There were counterparts of the 45-kD antigen in all mitochondrial preparations tested except for *E. coli*. The reactive polypeptide autoantigens identified by immunoblotting did not correspond, by MW at least, with other proposed mitochondrial anti-

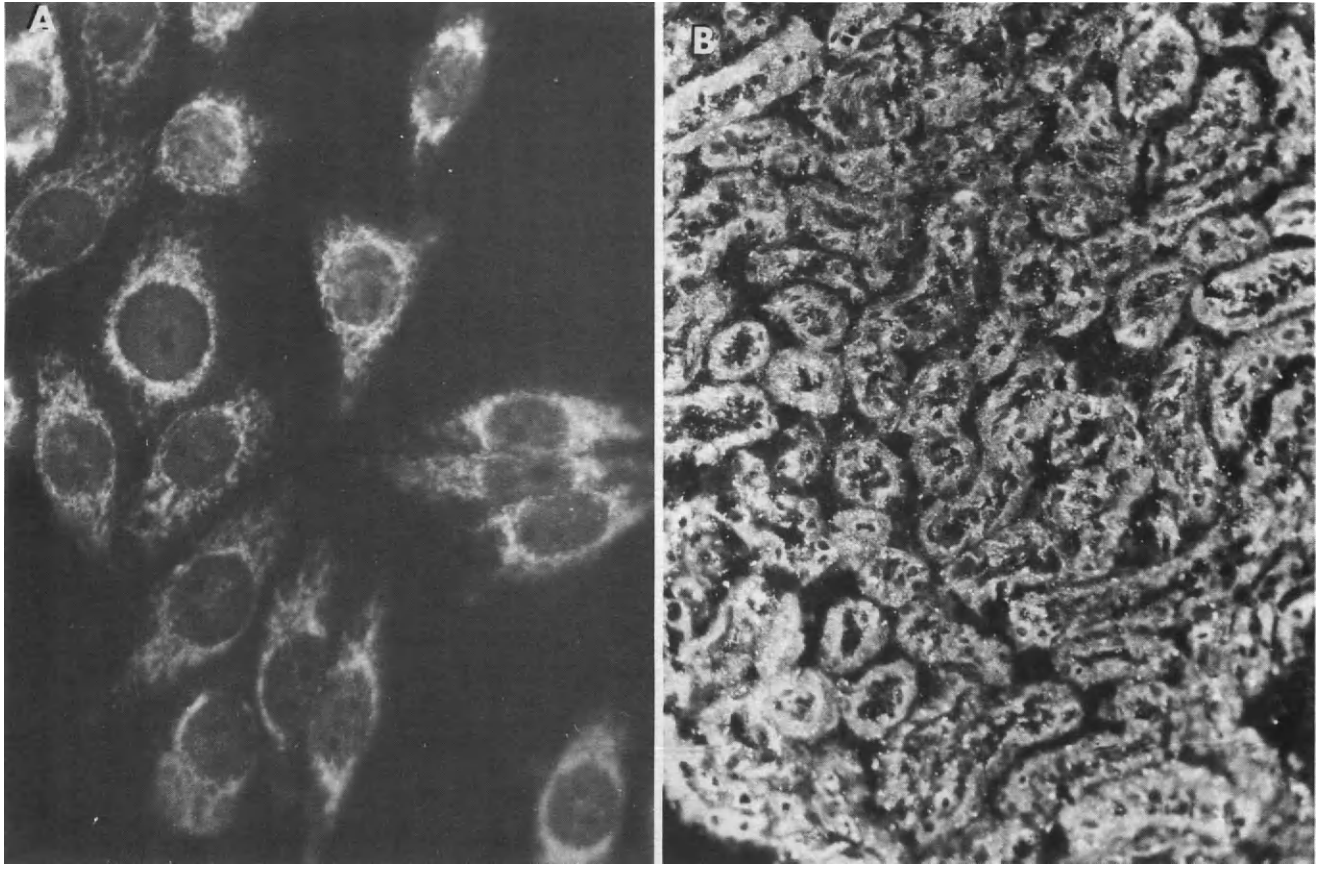


FIG. 9. Mitochondrial antibody reactivity of sera by immunofluorescence is seen as coarse granular cytoplasmic staining using (B) frozen section of kidney, and (A) cells of the HEp2 cell line.

TABLE II
The Antimitochondrial Antibodies

M1	Anticardiolipin, predominantly on inner mitochondrial membrane: "Wasserman-type" antibody
M2	Antibody to antigen on inner membranes of mitochondria, trypsin-sensitive: specifies PBC
M3	Antibody to antigen on outer mitochondrial membranes of mitochondria, trypsin-insensitive; provoked by a multiple-drug preparation "Venocuran" that contains phenopyrazone; associated with "pseudolupus" syndrome
M4	Antibody to antigen on outer membranes of mitochondria; trypsin-insensitive: detected (with M2) in the CAH-PBC "mixed form" overlap disease
M5	Antibody to antigen on mitochondrial membranes, reactive predominantly with proximal renal tubules: detected in "collagenlike" disorders including SLE
M6	Antibody to mitochondria associated with iproniazid-induced hepatitis

gens, that is, the subunit of ATPase of the adenine nucleotide translocator. The identification of the reactive mitochondrial polypeptide autoantigens identified by immunoblotting should provide much needed insight into the pathogenesis of PBC.

c. Other Antibodies. There is some disagreement as to whether PBC is marked by a frequency beyond normal of autoantibodies other than AMA (e.g., ANA and ASMA) and autoantibodies of the thyrogastric group. Those reporting an increase include Doniach *et al.* (1966) and Christensen *et al.* (1980); those finding no increase or only low-titer reactivity include Kaplowitz *et al.* (1973), Berg and Baum (1980), and my laboratory (unpublished). According to Berg and Baum (1980), the true M2-associated PBC is not associated with other autoantibodies, whereas in the mixed form, M4-associated, ASMA (antiactin) and ANA are demonstrable. Use of the HEp2 nuclear substrate, which is very suitable for speckled patterns of staining, has disclosed a diverse range of nonhomogeneous nuclear patterns, including anticentromere in cases of PBC with scleroderma (Bernstein *et al.*, 1984).

Anti-LSP. Antibodies reacting with human LSP by radioimmunoassay were detected in 19 (51%) of 37 patients with PBC, but only in the advanced stages, III and IV, in which their incidence was high, and there was close correlation between the degree of reactivity and the extent of piecemeal necrosis (Tsantoulas *et al.*, 1980). This reactivity is presumably secondary to hepatocellular damage, although a contribution to pathogenesis is possible.

d. Bile Duct-Specific Antigens. Earlier indications were that there might be a serum antibody reacting specifically with bile ducts (Paronetto *et al.*, 1967) or bile canaliculi (MacSween *et al.*, 1973), and a leucocyte

migration-inhibition assay was used to show cell-mediated reactivity by peripheral blood lymphocytes with a biliary protein in cases of PBC and also sclerosing cholangitis (McFarlane *et al.*, 1979). Although specific reactivity with a biliary antigen would satisfactorily explain the pathogenesis of PBC, it is not yet established that this is in fact demonstrable.

e. Complement Abnormalities. Levels of total serum complement are normal in PBC, but there is a marked alteration in complement turnover, as judged by the increase in plasma of complement activation products (C3b) and increased Clq and C3 catabolism (Potter *et al.*, 1976; Wands *et al.*, 1978). The explanation for these changes may be the fixation of complement products on immune complexes (see Section II,E,1,f). It is stated that there may be activation of the classical pathway (Thomas *et al.*, 1977) or the alternate pathway (Wands *et al.*, 1978). A detailed description of complement metabolism in PBC is given by Thomas and Epstein (1980) and James (1983).

f. Immune Complexes. The recognition of complement activation in PBC led to the notion that immune complexes may be implicated in the disease process (Thomas *et al.*, 1977). The presence of immune complexes in serum in large amounts was shown by various procedures [e.g., Clq binding, by Thomas *et al.* (1978), or Raji cell binding, by Wands *et al.* (1978)]. These complexes were of large size, contained substantial amounts of IgM, appeared to activate C3 *in vitro*, and also were highly susceptible to cryoprecipitation. There are reports that the immune complexes in PBC contain antigens derived from bile or mitochondria (cited by James, 1983), but these have not been confirmed. Thomas and Epstein (1980) suggested that the complexes were the result of associations of globulin–antiglobulin, C3b–anti-C3b, or idiotype–antiidiotype. It is curious that there is usually no systemic expression of immune complex disease in PBC, although Clq binding was correlated with complaints of arthritis (Crowe *et al.*, 1980b), and that there is no evidence for the hepatic lesions being due to formation or deposition of immune complexes in the liver (Krogsgaard *et al.*, 1981). Finally, there is a report in which five assays for immune complexes were used; few patients with PBC were positive in all five (Goldberg *et al.*, 1982).

g. Depressed Cellular Immunity. There is a general depression of T-lymphocyte function in PBC, illustrated by cutaneous anergy and failure to develop delayed-type hypersensitivity after immunization (Fox *et al.*, 1973); other abnormalities including inhibitory serum factors were reviewed by Thomas and Epstein (1980). Impaired CMI is further exemplified by a deficiency of peripheral blood T cells to proliferate in response to autologous B cells in the autologous mixed-lymphocyte reaction (AMLR)

(James *et al.*, 1977); the responsive subpopulation of T cells for the AMLR are precursors of suppressor T lymphocytes.

2. Immunological Derangements

Several studies published on numbers of lymphocyte subsets in PBC (James, 1983) suggested a decrease in the helper (T4) subpopulation. More recently, Miller *et al.* (1984) considered whether changes present were a cause or effect by studying subsets at the four stages of PBC, using normal controls and patients with other types of advanced liver disease; patients with early (stage I) PBC had significantly increased counts of suppressor (Leu 2a+) cells, but normal counts for pan-T lymphocytes and helper (Leu 3a+) cells, whereas in advanced disease counts for all subsets were lower than normal and at levels comparable with those in other types of cirrhosis. On the other hand, functional assays have yielded solid evidence for decreased function of suppressor T cells, perhaps at variance with increase in the putative suppressor subset (above). One of the assays used (James *et al.*, 1980; James, 1983) depended on the ability of T lymphocytes from patients with PBC to help or suppress pokeweed-induced immunoglobulin synthesis, with attention to ratios of T and B cells in the system; normally, with increasing T/B ratios >1.0, suppressor effects predominate, whereas in PBC this effect was not evident, pointing to relative hypofunction of the suppressor subset. It is uncertain whether this would be due to an excess of helper influences, decreased suppressor cell numbers (e.g., in advanced disease and perhaps related to deficiency in the AMLR), or to other types of failure in induction of suppressor cells.

The other major derangement is the diversion of the immune response to self antigens of the mitochondrial membrane. It has been speculated that the origin of cell mitochondria was from bacterial cells, and it has been noted that the PBC-specific antigen is present in all mitochondria so far tested, including mitochondria of microorganisms (Berg and Baum, 1980; Thomas and Epstein, 1980; Frazer *et al.*, 1985). Accordingly, PBC might be explained by the occurrence, in subjects with a congenital or acquired defect of immune regulation, of autoimmune induction by cross-reactivity with mitochondria-like antigens or microorganisms, or conceivably, food antigens.

3. Mechanisms of Damage

There is virtually no information on mechanisms of damage to biliary ductules in PBC. Candidate mechanisms would include intrahepatic immune complexes [although this has been negated (see Section II,E,1,f)

and cytotoxic effects of T cells among the portal ilfiltrates. Antibody-dependent cellular cytotoxicity (ADCC) of peripheral blood cells was found not to differ from normal (Vierling *et al.*, 1977). However, a discriminatory assay would require specific antibodies and target cells of biliary origin, and such would be difficult to establish; moreover, since there is little evidence for a bile duct-specific antibody in PBC, there would be no specific direction of ADCC mechanisms toward biliary epithelium.

4. Immunogenetics

Female sex is one unequivocal component to genetic predisposition to PBC, since the sex ratio is ~9:1 in all case series reported. While hormonal influences are called on to explain female predisposition in many of the autoimmune diseases that occur in early life, such would not appear to pertain in PBC, which usually presents in females after the menopause and virtually never in young females. Although published case series usually do not contain examples of intrafamilial PBC, and neither do population surveys (Triger, 1980), an impressive number of familial cases have been reported involving siblings (eight examples) and parent–child (three examples), and other familial clusterings have been cited (Mackay, 1984). Cales *et al.* (1983) reported a family study based on two brothers with PBC and a sister with granulomatous hepatitis; of the nine subjects examined in two generations of this family, immune abnormalities were observed in six, but *HLA* haplotype-associated transmission of PBC was not observed.

In addition to coincident cases of PBC in families, there is evidence in relatives of a high frequency of AMA (Feizi *et al.*, 1972) or other immune-mediated diseases including thyroiditis, rheumatoid arthritis, Sjögren's disease, and others (Galbraith *et al.*, 1974). These findings could be due to various genetic influences, including (a) predisposition to a specific immunological abnormality that is basic to PBC, (b) predisposition to development of chronic liver damage with cirrhosis, and (c) a general weakness of immunological regulation. The latter possibility is raised by two studies that showed that relatives of cases of PBC had a higher frequency than did controls of certain autoantibodies (cited by Mackay, 1984). Further, a study of suppressor-cell function based on recruitment of suppressor cells by Con-A showed significant impairment of IgG suppression in 13 of 16 patients with PBC, and in 6 of 23 healthy female relatives, but not in patients with other types of cirrhosis or in healthy controls (Miller *et al.*, 1983).

If there were immunogenetic abnormalities critical to the development

of PBC, a strong association with HLA types might be expected, but reports are variable and in general negative (Mackay, 1984). Thus studies on *HLA-A* and *-B* locus frequencies in PBC have not shown differences from controls. There is one report of an increase in *HLA-DR3*, an abstract claiming an increase in *-DR4* [50% versus 21% in controls, relative risk (RR) 3.7], and a report on Japanese with PBC showing a higher frequency of *HLA-DR2* than in controls (68 versus 30%, RR 5.0), but the number of cases in each of these reports was low.

5. Laboratory Diagnosis

The diagnosis of PBC is strongly supported by demonstration of the M2 type of AMA, and this, together with other procedures (liver biopsy, cholangiography), now obviates the need for confirmatory laparotomy. Since the sensitivity and specificity of the AMA reaction is almost but not entirely absolute, the problems of negative reactions in confirmed cases and positive reactions in the absence of PBC require consideration. Given that the usually cited figure for positivity of AMA in PBC is 90%, can the apparent 10% of reported false-negative AMA reactions be accounted for? Some may have been due to an erroneous clinical diagnosis and others to laboratory error (e.g., prozone effect if a serum titration had not been done); it would be of interest to review a series of AMA-negative cases of PBC for ascertainable clinical differences from seropositve cases.

The converse situation, AMA positivity in the absence of PBC, is receiving much attention as clinical immunology laboratories now routinely screen large numbers of sera for multiple autoimmune reactions. For example, Triger *et al.* (1982) detected AMA in 69 (1.47%) of 4200 sera referred for serological testing; 9 had unequivocal PBC, 6 had CAH (or possibly the overlap syndrome), 10 had abnormal liver tests not suggestive of PBC, and 44 had no evidence of liver disease. The conclusion was that, in the absence of clinical liver disease, the AMA test lacked specificity for the diagnosis of PBC. There are several problems to consider here in relationship to positive AMA tests and the diagnosis of PBC. (i) The use of rat tissue can give a misleading positive test due to the presence of heteroantibody, and hence a positive test if rat tissue is used should be confirmed using a human tissue substrate. (ii) Since only the M2 and M4 types of AMA are specific for liver disease, AMA reactions due to M1, M3, M5, and M6 type antibodies must be excluded. (iii) The M4 type of AMA has a specificity for a separate type of cholestatic liver disease, described above. (iv) Given the current recognition of the frequency of mild or early types of PBC (see Section II,B), it may be impossible to

exclude a clinically featureless example of PBC without a liver biopsy, yet this would seem unjustified in an otherwise healthy person.

To conclude, attention is directed to the study of Berg *et al.* (1982), which can be taken as the current benchmark. Serum from 91 (97%) of 94 patients with PBC reacted with M2 AMA, and when submembranous particles from mitochondria were added to the test system, reactivity reached 100%. Conversely, 417 patients with other hepatic and nonhepatic disorders were tested and only 4 had M2 AMA, the diseases being SLE (1), systemic sclerosis (2), and Sjögren's syndrome (1); of these, there was histological evidence of PBC in 2.

The nature of the relationship between PBC and the mixed CAH–PBC overlap syndrome remains uncertain, but could be clarified as antimitochondrial reactivites become better defined. (Frazer *et al.*, 1985).

F. Treatment and Outcome

Therapy in PBC has three aims, the first two being irrelevant to immunological aspects: (a) to prevent or treat complications of cholestasis, (b) to alleviate manifestations of hepatic decompensation, and (c) to arrest the pathological process in the liver, aiming either at the primary immunopathological process or at secondary effects such as copper overload in the liver.

The viewpoint taken here is that no treatment yet is of sufficiently established effectiveness to justify probable adverse effects when administered on a long-term basis. The background to this opinion is the summarized review by Jones (1983). The case against prednisolone is the mediocre response and the likely exacerbation of bone disease. Azathioprine has been evaluated against placebo in two randomized trials, and, although clear effects on various indices were not discernible, survival data slightly favored the azathioprine-treated groups (Heathcote *et al.* 1976; Christensen *et al.*, 1984). Data for combined prednisolone and azathioprine from the setting of a controlled trial are not available. Other regimens reviewed by Jones (1983) included cyclophosphamide, chlorambucil, and cyclosporin, but data favoring long-term benefit are not to hand.

There have been extensive studies on D-penicillamine, a copper-chelating agent, because of the recognition of extensive deposits of copper in hepatocytes. This is presumably secondary to biliary obstruction, but the excess copper could itself potentiate the primary injury. Penicillamine, in addition to copper chelation, might favorably affect various immunological abnormalities associated with PBC (Jones, 1983). In the event, the

results of trials have proved equivocal and side effects are too frequent for this therapy to be recommended.

Opinion on the outcome of PBC has undergone considerable revision with the recognition that mild asymptomatic cases constitute a substantial proportion of the spectrum of PBC (James *et al.*, 1981; Roll *et al.*, 1983). Although an average survival time can be put at 10 years, the range will be wide due to great differences among patients in the inherent tempo of the disease. Whether immunological factors can be related to such differences in tempo has not yet been ascertained.

G. Concluding Remarks: Future Prospects

As with many autoimmune diseases, PBC gives much promise of providing important leads to the basis of autoimmunity, yet displays many anomalous and some possibly irrelevant features. The identification of known mitochondrial enzymes as targets of the characteristic autoantibody specificity, and the knowledge that structurally similar enzymes are constituents of almost all mitochondrial systems, seem attractive hints that the disease may be initiated by an immunological exposure to microorganisms (cross-reactivity) under conditions in which tolerance breakage can occur; this is evidently more likely to occur in females, and, given the age incidence of the disease, hormonal factors may be less important than in SLE, for example. The paradox of generalized antimitochondrial activity with biliary epithelium as a predominant target is tantalizing.

The idea has been proposed, based on similarities between PBC and GVH disease, than an immune response to MHC specificities could be implicated in PBC, but direct evidence for this is lacking. The peculiar complement abnormalities might contribute to part of the pathology, but do not seem to be part of the primary pathology. The range of diseases coexisting with PBC, and the clinical settings of these associations, are worthy of much further consideration. The genetic predisposition to PBC is provocative, but determinants of this need further exploration from the standpoint of family studies and predictive markers. Finally, regimens of treatment that may arrest the disease, at least in some of its phases, need investigation, keeping in mind the possibility (Berg and Baum, 1980) that PBC might consist of a group of etiologically different diseases, as pertains in CAH.

H. Summary

Primary biliary cirrhosis, like CAH, became recognized as a disease after 1950, and an autoimmune basis was established ~10 years later.

There are four stages identifiable in the 5- to 10-year evolution of PBC, from the initial nonsuppurative destructive cholangitis to the terminal cirrhosis; however, with more mild asymptomatic cases now being recognized, the survival time for PBC can be extended. The histological hallmark is an inflammatory obliterative destruction of biliary ductules in the liver with accompanying lymphoid aggregates and granulomata. There is no valid animal model. Serologically, there is a characteristic reaction with an antigen on the inner membrane of mitochondria and the target structure is likely to be a mitochondrial enzyme. The mitochondrial antigen with which PBC sera are specifically reactive is known as M2; diagnostically, the M2 type of antimitochondrial antibody has very high specificity and sensitivity for PBC. Other mitochondrial antigens, M1 and M3-6, exist; antibody to M4 specifies the mixed PBC–CAH overlap syndrome and antibodies to the remainder are irrelevant to liver disease. There are profound complement disturbances associated with the presence in serum of large immune complexes. Neither the nature of the postulated immunoregulatory defect in PBC nor the mechanism of damage to bile ducts is known. Although there is familial predisposition, no clear genetic markers including HLA have been identified. No therapy clearly ameliorates the course of the disease.

REFERENCES

A. Chronic Active Hepatitis

Alexander, G. J. M., and Williams, R. (1984). *Lab. Invest.* **50,** 247–249.

Alexander, G. J. M., Nouri, Aria, K. T., Eddleston, A. L. W. F., and Williams, R. (1983). *Lancet* **1,** 1291.

Allison, A. C., Denman, A. M., and Barnes, R. D. (1971). *Lancet* **2,** 135–140.

Ambinder, J. M., Chiorazzi, N., Gibofsky, A., Fotino, M., and Kunkel, H. (1982). *Clin. Immunol. Immunopath.* **23,** 269–274.

Andrews, P. M., and Coffey, A. K. (1983). *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **42,** 3046–3052.

Asherson, G. L. (1959). *Br. J. Exp. Pathol.* **40,** 209–215.

Bartholomaeus, W. N., Reed, W. D., Joske, R. A., and Shilkin, K. B. (1981). *Immunology* **43,** 219–226.

Becker, C. G., and Nachman, R. L. (1973). *Am. J. Pathol.* **71,** 1–22.

Bernuau, D., Rogier, E., and Feldmann, G. (1982). *Am. J. Pathol.* **109,** 310–320.

Bjarnason, I., Magnusson, B., and Bjornsson, S. (1982). *Acta Med. Scand.* **211,** 305–307.

Blumberg, B. S., Gerstley, B. J. S., Hungerford, D. A., London, W. T., and Sutnick, A. I. (1967). *Ann. Intern. Med.* **66,** 924–931.

Bottazzo, G. F., Florin-Christensen, A., Fairfax, A., Swana, G., Doniach, D., and Gröschel-Stewart, U. (1976). *J. Clin. Pathol.* **29,** 403–410.

Bretherton, L., Brown, C., Pederson, J. S., Toh, B. H., Clarke, F. M., Mackay, I. R., and Gust, I. D. (1983). *Clin. Exp. Immunol.* **51,** 611-616.

Burns, G. F., Librach, C. L., Frazer, I. H., Kronborg, I. J., and Mackay, I. R. (1982). *Clin. Immunol. Immunopathol.* **24,** 386–395.

Cavell, B., and Leonhardt, T. (1965). *Acta Med. Scand.* **177,** 751–759.

Cook, G. C., Mulligan, R., and Sherlock, S. (1971). *Q. J. Med.* **40,** 159–185.

Coovadia, H. M., Mackay, I. R., and d'Apice, A. J. F. (1981). *Clin. Immunol. Immunopathol.* **18,** 268–275.

Copenhagen Study Group for Liver Diseases (1969). *Lancet* **1,** 119–121.

Crapper, R. M., Bhathal, P. S., and Mackay, I. R. (1983). *Liver* **3,** 327–337.

Crapper, R. M., Bhathal, P. S., Mackay, I. R., and Frazer, I. H. (1984). *Digestion* (in press).

Cunningham, A., Mackay, I. R., Frazer, I. H., Brown, G., Pedersen, J. S., Toh, B. H., and Clarke, F. M. (1985). *Clin. Immunol. Immunopathol.* **34,** 158–164.

Czaja, A. J. (1984). *Semin. Liver Dis.* **4,** 1–12.

Davis, P., and Read, A. E. (1975). *Gut* **16,** 413–415.

Davis, G. L., Czaja, A. J., and Ludwig, J. (1984). *Gastroenterology* **87,** 1222–1227.

De Groote, J., Desmet, V. J., Gedigk, P., Korb, G., Popper, H., Poulsen, H., Scheuer, P. J., Schmid, M., Thaler, H., Uehlinger, E., and Wepler, W. (1968). *Lancet* **2,** 626–628.

Diederichsen, H. (1969). *Acta Med. Scand.* **186,** 299–302.

Eddleston, A. L. W. F., and Williams, R. (1978). *Br. Med. Bull.* **34,** 295–300.

Eliastam, M., and Holmes, A. W. (1971). *Am. J. Dig. Dis.* **16,** 1014–1018.

Elling, P., Ranlov, P., and Bildsoe, P. (1966). *Acta Med. Scand.* **179,** 527–533.

Fagraeus, A., Lidman, K., and Biberfeld, G. (1974). *Nature (London)* **252,** 246–247.

Finlayson, N. D. C., Krohn, K., Fauconnet, M. H., and Anderson, K. E. (1972). *Gastroenterology* **63,** 653–659.

Fravi, G., and Lindenmann, J. (1968). *Nature (London)* **218,** 141–143.

Frazer, I. H., and Mackay, I. R. (1984a). *J. Clin. Lab. Immunol.* **14,** 165–167.

Frazer, I. H., and Mackay, I. R. (1984b). *Clin. Exp. Immunol.* **57,** 429–437.

Frazer, I. H., Kronborg, I. J., and Mackay, I. R. *Clin. Exp. Immunol.* (1983). **54,** 213–218.

Frazer, I. H., Mackay, I. R., Bell, J., and Becker, C. G. (1985). *Liver* (In press.)

Freudenberg, J., Baumann, H., Arnold, W., Berger, J., and Meyer zum Büschenfelde, K. H. (1977). *Digestion* **15,** 260–270.

Fujinami, R. S., Oldstone, M. B. A., Wroblewska, Z., Frankel, M. E., and Koprowski, H. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80,** 2346–2350.

Gabbiani, G., Ryan, G. B., Lamelin, J. P., Vassalli, P., Majno, G., Bouvier, C. A., Cruchaud, A., and Luscher, E. F. (1973). *Am. J. Pathol.* **72,** 473–488.

Galbraith, R. M., Smith, M., MacKenzie, R. M., Tee, D. E., Doniach, D., and Williams, R. (1974). *N. Engl. J. Med.* **290,** 63–69.

Hegarty, J. E., Aria, K. T. N., Portmann, B., Eddelston, A. L. W. F., and Williams, R. (1983). *Hepatology* **3,** 685–689.

Hodges, J. R., Millward-Sadler, G. H., Barbatis, C., and Wright, R. (1981). *N. Engl. J. Med.* **304,** 557–560.

Hodges, J. R., Millward-Sadler, G. H., and Wright, R. (1982). *Lancet* **1,** 550–552.

Hodgson, H. J. F., Wands, J. R., and Isselbacher, K. J. (1978). *Proc. Natl. Acad. Sci. U.S.A.* **75,** 1549–1553.

Holborow, E. J. (1972). *Br. Med. Bull.* **28,** 142–144.

Holdstock, G., Chastenay, B. F., and Krawitt, E. L. (1982). *Gastroenterology* **82,** 216–212.

Hopf, U., and Meyer zum Büschenfelde, K. H. (1974). *Br. J. Exp. Pathol.* **55,** 509–513.

Hopf, U., Meyer zum Büschenfelde, K. H., and Arnold, W. (1976). *N. Engl. J. Med.* **294,** 578–582.

International Group of Pathologists (1977). *Lancet* **2,** 914–919.
Jain, S., Markham, R., Thomas, H., and Sherlock, S. (1976). *Clin. Exp. Immunol.* **26,** 35–41.
Jensen, D. M., McFarlane, I. G., Portmann, B. S., Path, M. R., Eddleston, A. L. W. F., and Williams, R. (1978). *N. Engl. J. Med.* **299,** 1–7.
Johnson, G. D., Holborow, E. J., and Glynn, L. E. (1965). *Lancet* **2,** 878–879.
Joske, R. A., and King, W. E. (1955). *Lancet* **2,** 477–480.
Joske, R. A., and Laurence, B. H. (1970). *Gastroenterology* **59,** 546–552.
Kakumu, S., Arakawa, Y., Goji, H., Kashio, T., and Yata, K. (1979). *Gastroenterology* **76,** 665–672.
Kenneally, D., Mackay, I. R., and Toh, B. H. (1984). *J. Clin. Lab. Immunol.* **14,** 13–16.
Klein, J. (1982). *Adv. Cancer Res.* **37,** 234–317.
Koffler, D., Carr, R. J., Agnello, V., Feizi, T., and Kunkel, H. G. (1969). *Science* **166,** 1648–1649.
Kurki, P., and Virtanen, I. (1984). *J. Immunol. Methods* **67,** 209–223.
Kurki, P., Miettnen, A., Linder, E., Pikkarainen, P., Vuoristo, M., and Salaspuro, M. P. (1980). *Gut* **21,** 878–884.
Lam, K. C., Lai, C. L., Wu, P. C., and Todd, D. (1980). *J. Chronic Dis.* **33,** 375–381.
Leevy, C. M., Popper, H., and Sherlock, S. (1976). *Fogarty Int. Cent. Proc. No. 22, DHEW Publ.* No. (NIH) 76–725. Lidman, K., Biberfeld, G., Fagraeus, A., Norberg, R., Torstensson, R., Utter, G., Carlsson, L., Luca, J., and Lindberg, U. (1976). *Clin. Exp. Immunol.* **24,** 266–272.
McFarlane, I. G., Wojcika, B. M., Zueker, G. M., Eddleston, A. L. W. F., and Williams, R. (1977). *Clin. Exp. Immunol.* **27,** 381–390.
Mackay, I. R. (1968). *Q. J. Med.* **37,** 379–392.
Mackay, I. R. (1972). *Ann. Intern. Med.* **77,** 649–651.
Mackay, I. R. (1975). *In* "Frontiers of Gastrointestinal Research (L. van der Reis, ed.), Vol. 1, pp. 142–187. Karger, Basel.
Mackay, I. R. (1977). *In* "HLA and Disease" (J. Dausset and A. Svejgaard, eds.), pp. 186–195. Munksgaard, Copenhagen.
Mackay, I. R. (1978). *In* "Immunological Diseases" (M. Samter, D. W. Talmage, B. Rose, K. F. Austin, and J. H. Vaughan, eds.), 3rd ed., pp. 1454–1477. Little, Brown, Boston, Massachusetts.
Mackay, I. R. (1981). *In* "Systemic Effects of HBsAg Immune Complexes" (E. Bartolo, L. Chiandussi, and S. Sherlock, eds.), pp. 300–308. Piccin Medical Books, Padova.
Mackay, I. R. (1982). *In* "Diseases of the Liver" (L. Shiff and E. Shiff, eds.), 5th ed., Chapter 20. Lippincott, Philadelphia, Pennsylvania.
Mackay, I. R. (1983). *Hepatology* **3,** 724–728.
Mackay, I. R. (1984). *Semin. Liver Dis.* **4,** 13–25.
Mackay, I. R., and Frazer, I. H. (1984). *In* Advances in Hepatitis Research (F. V. Chisari, ed.), pp. 179–189. Masson, New York.
Mackay, I. R., and Gajdusek, D. C. (1958). *Arch. Intern. Med.* **101,** 30–46.
Mackay, I. R., and Larkin, L. (1958). *Aust. Ann. Med.* **7,** 251–258.
Mackay, I. R., and Morris, P. J. (1972). *Lancet* **2,** 793–795.
Mackay, I. R., and Wood, I. J. (1961). Autoimmunity in liver disease. *In* "Progress in Liver Disease" (H. Popper and F. Schaffner, eds.). Grune & Stratton, New York.
Mackay, I. R., and Wood, I. T. (1962). *Q. J. Med.* **31,** 485–507.
Mackay, I. R., and Tait, B. D. (1980). *Gastroenterology* **79,** 95–98.
Mackay, I. R., Taft, L. I., and Cowling, D. C. (1956). *Lancet* **2,** 1323–1326.
Mackay, I. R., Whittingham, S., Mathews, S. D., and Tait, B. D. (1980). *Springer Semin. Immunopathol.* **3,** 285–296.

Mackay, I. R., Frazer, I. H., and Jordan, T. W. (1984). "Viral Hepatitis and Liver Disease" (G. N. Vyas, J. L. Dienstag, and J. H. Hoofnagle, eds.), pp. 181–190. Grune & Stratton, Orlando, Florida.
Mackay, I. R., Frazer, I. H., Toh, B. H., Pederson, J. S., and Alter, H. J. (1985). *Clin. Exp. Immunol.* (In press.)
MacLachlan, M. J., Rodnan, G. P., Cooper, W. M., and Fennell, R. H. (1965). *Ann. Intern. Med.* **62,** 425–462.
Meyer zum Büschenfelde, K. H., and Hopf, U. (1974). *Br. J. Exp. Pathol.* **55,** 498–503.
Meyer zum Büschenfelde, K. H., and Hütteroth, T. H. (1979). *In* "Immune Reactions in Liver Disease" (A. L. W. F. Eddleston, J. C. P. Weber, and R. Williams eds.), pp. 12–20. Pitman, London.
Meyer zum Büschenfelde, K. H., and Manns, M. (1984). *Semin. Liver Dis.* **4,** 26–35.
Meyer zum Büschenfelde, K. H., and Miescher, P. A. (1972). *Clin. Exp. Immunol.* **10,** 89–102.
Meyer zum Büschenfelde, K. H., and Schrank, C. (1966). *Klin. Wochenschr.* **44,** 654–656.
Meyer zum Büschenfelde, K. H., Knolle, J., and Berger, J. (1974). *Klin. Wochenschr.* **52,** 246–248.
Meyer zum Büschenfelde, K. H., Manns, M., Hütteroth, T. H., Hopf, U., and Arnold, W. (1979). *Clin. Exp. Immunol.* **37,** 205–212.
Miller, K. B., and Schwartz, R. S. (1979). *N. Engl. J. Med.* **301,** 803–809.
Mondelli, M., Naumov, N., and Eddleston, A. L. W. F. (1984). *In* "Advances in Hepatitis Research" (F. V. Chisari, ed.), pp. 144–151. Masson, New York.
Nouri-Aria, K. T., Hegarty, J. E., Alexander, G. J. M., and Williams, R. (1982). *N. Engl. J. Med.* **307,** 1301–1304.
O'Brien, E. N., Goble, A. J., and Mackay, I. R. (1958). *Lancet* **1,** 1245–1249.
Odièvre, M., Maggiore, G., Homberg, J. C., Saadoun, F., Couroucè, A. M., Yvart, J., Hadchouel, M., and Alagille, D. (1983). *Hepatology* **3,** 407–409.
Opelz, G., Vogten, A. J. M., Summerskill, W. H. J., Schalm, S. W., and Terasaki, P. I. (1977). *Tissue Antigens* **9,** 36–40.
Page, A. R., and Good, R. A. (1960). *J. Dis. Child.* **99,** 288–314.
Paronetto, F., Rubin, E., and Popper, H. (1962). *Lab. Invest.* **11,** 150–158.
Pedersen, J. S., Toh, B. H., Mackay, I. R., Tait, B. D., Gust, I. D., Kastelan, A., and Hadzic, N. (1982). *Clin. Exp. Immunol.* **48,** 527–532.
Reynolds, T. B., Peters, R. L., and Yamada, S. (1971). *N. Engl. J. Med.* **280,** 813–820.
Riisom, K., and Diederichsen, H. (1983). *Gastroenterology* **85,** 1271–1276.
Rizzetto, M., Swana, G., and Doniach, D. (1973). *Clin. Exp. Immunol.* **15,** 331–344.
Runyon, B. A., La Brecque, D. R., and Anuras, S. (1980). *Am. J. Med.* **69,** 188–194.
Salaspuro, M. P., Laitinen, O. I., Lehtola, J., Makkonen, H., Rasanen, J. A., and Supponen, P. (1976). *Scand. J. Gastroenterol.* **11,** 313–320.
Schalm, S. W. (1982). *Liver* **2,** 69–76.
Smalley, M. J., Mackay, I. R., and Whittingham, S. (1968). *Aust. Ann. Med.* **17,** 28–32.
Smeenk, R., van der Lelij, G., and Swaak, T. (1982). *Arthritis Rheum.* **25,** 631–638.
Soloway, R. D., Summerskill, W. H. J., Baggenstoss, A. H., Geall, M. G., Gitnick, G. I., Elveback, L. R., and Schoenfield, L. J. (1972). *Gastroenterology* **63,** 820–833.
Steven, M. M., Buckley, J. D., and Mackay, I. R. (1979). *Q. J. Med.* **48,** 519–531.
Toh, B. H. (1979). *Clin. Exp. Immunol.* **38,** 621–628.
Weeds, A. (1982). *Nature (London)* **296,** 811–816.
Weigle, W. O. (1977). *In* "Autoimmunity, Genetic, Immunologic, Virologic and Clinical Aspects" (N. Talal, ed.), pp. 141–170. Academic Press, New York.

Whittingham, S., Mackay, I. R., and Irwin, J. (1966). *Lancet* **1,** 1333–1336.
Whittingham, S., Mackay, I. R., and Kiss, Z. S. (1970). *Gut* **11,** 811–816.
Whittingham, S., Mathews, J. D., Schanfield, M. S., Tait, B. D., and Mackay, I. R. (1981a). *Clin. Exp. Immunol* **43,** 80–86.
Whittingham, S., Morstyn, G., Wilson, J. W., and Vadas, M. (1981b). *Blood* **58,** 768–771.
Whittingham, S., Mackay, I. R., and Mathews, J. D. (1984a). *Clin. Immunol. Allergy* **4,** 626–640.
Whittingham, S., Propert, D. N., and Mackay, I. R. (1984b). *Immunogenetics* **19,** 295–299.

B. Primary Biliary Cirrhosis

Ahrens, E. H., Rayne, M. A., Kunkel, H. G., Eisenmenger, W. J., and Blondheim, S. H. (1950). *Medicine (Baltimore)* **29,** 299–364.
Alarcón-Segovia, D., Diaz-Jouanen, E., and Fishbein, E. (1973). *Ann. Intern. Med.* **79,** 31–36.
Benoist, M., Henin, D., Kahn, M. F., and Benhamou, J.-P. (1977). *Nouv. Presse Med.* **6,** 2427–2429.
Berg, P. A., and Baum, H. (1980). *Springer Sem. Immunopathol.* **3,** 355–373.
Berg, P. A., Doniach, D., and Roitt, I. M. (1967). *J. Exp. Med.* **126,** 277–293.
Berg, P. A., Klein, R., Lindenborn-Fotinos, J., and Klöppel, W. (1982). *Lancet* **2,** 1423–1425.
Bernstein, R. M., Neuberger, C. C., Bunn, M. E., Callender, G. R., Hughes, G. R. V., and Williams, R. (1984). *Clin. Exp. Immunol.* **55,** 553–560.
Cales, P., Calot, M., Voigt, J. J., Oksman, F., Cassignuel, J., Vinel, J. P., and Pascal, J. P. (1983). *Gastroenterol. Clin. Biol.* **10,** 777–784.
Christensen, E., Crowe, J., Doniach, D., Popper, H., Ranek, L., Rodis, J., Tygstrup, N., and Williams, R. (1980). *Gastroenterology* **78,** 236–246.
Christensen, E., Neuberger, J., Crowe, J., Popper, H., Portmann, B., Doniach, D., Ranek, L., Tygstrup, N., and Williams, R. (1984). *Liver* **4,** 81.
Crowe, J. P., Christensen, E., Butler, J., Wheeler, P., Doniach, D., Keenan, J., and Williams, R. (1980a). *Gastroenterology* **78,** 1437–1441.
Crowe, J. P., Molloy, M. G., Wells, I., Thompson, B. R., Holborow, E. J., Hamilton, E., and Williams, R. (1980b). *Gut* **21,** 418–422.
Culp, K. S., Fleming, C. R., Duffy, J., Baldus, W. P., and Dickson, E. R. (1982). *Mayo Clinic Proc.* **57,** 365–370.
Doniach, D., Roitt, I. M., Walker, J. G., and Sherlock, S. (1966). *Clin. Exp. Immunol.* **1,** 237–262.
Elta, G. H., Scpersky, R. A., Goldberg, M. J., Connors, C. M., Miller, K. B., and Kaplan, M. M. (1983). *Dig. Dis. Sci.* **28,** 971–975.
Epstein, O., Thomas, H. C., and Sherlock, S. (1980). *Lancet* **1,** 1166–1168.
Fakunle, Y. M., Aranguibel, F., DeVilliers, D., Thomas, H. C., and Sherlock, S. (1979). *Clin. Exp. Immunol.* **38,** 204–210.
Feizi, T. (1968). *Gut* **9,** 193–198.
Feizi, T., Naccarato, R., Sherlock, S., and Doniach, D. (1972). *Clin. Exp. Immunol.* **10,** 609–622.
Fox, R. A., Dudley, F. J., and Sherlock, S. (1973). *Clin. Exp. Immunol.* **14,** 473–480.
Frazer, I. H., Mackay, I. R., Jordan, T. W., Whittingham, S., and Marzuki, S. (1985). *J. Immunol.* (in press).

Galbraith, R. M., Smith, M., MacKenzie, R. M., Tee, D. E., Doniach, D., and Williams, R. (1974). *N. Engl. J. Med.* **290,** 63–69.
Goldberg, M. J., Kaplan, M. M., and Mitamura, T. (1982). *Gastroenterology* **83,** 677–683.
Golding, P. L., Smith, M., and Williams, R. (1973). *Am. J. Med.* **55,** 772–782.
Hall, S., Axelsen, P. H., Larson, D. E., and Bunch, T. W. (1984). *Ann. Intern. Med.* **100,** 388–389.
Hamlyn, A. N., Macklon, A. F., and James, O. (1983). *Gut* **24,** 940–945.
Heathcote, J., Ross, A., and Sherlock, S. (1976). *Gastroenterology* **70,** 656–660.
James, O., Macklon, A. F., and Watson, A. J. (1981). *Lancet* **1,** 1278–1281.
James, S. P. (1983). *Ann. Intern. Med.* **99,** 500–512.
James, S. P., Elson, C. O., Waggoner, J. G., Jones, E. A., and Strober, W. (1977). *J. Clin. Invest* **66,** 1305–1310.
James, S. P., Elson, C. O., Jones, E. A., and Strober, W. (1980). *Gastroenterology* **79,** 242–244.
Jones, E. A. (1983). *Ann. Intern. Med.* **99,** 500–512.
Kaplowitz, N., Finlayson, N. D. C., Walmsley, P., Thurnherr, N., and Javitt, N. B. (1973). *Am. J. Med.* **54,** 725–730.
Klatskin, G., and Kantor, F. S. (1972). *Ann. Intern. Med.* **77,** 533–541.
Krogsgaard, K., Tage-Jensen, U., Wantzin, P., Aldershvile, J., and Hardt, F. (1981). *J. Clin. Pathol.* **34,** 1076–1079.
Lindenborn-Fotinos, J., Sayers, T. J., and Berg, P. A. (1982). *Clin. Exp. Immunol.* **50,** 267–274.
Lindgren, S., Eriksson, S., Lofberg, H., and McKay, J. (1981). *Acta Med. Scand.* **210,** 317–320.
McFarlane, I. G., Wojcika, B. M., Tsantoulas, D. C., Portmann, B. C., Eddleston, A. L. W. F., and Williams, R. (1979). *Gastroenterology* **76,** 1332–1340.
Mackay, I. R. (1958). *N. Engl. J. Med.* **258,** 185–188.
Mackay, I. R. (1960). *Lancet* **2,** 521–523.
Mackay, I. R. (1984). *Semin. Liver Dis.* **4,** 13–25.
MacSween, R. N. M., Armstrong, E. M., Gray, K. G., and Mason, M. (1973). *Lancet* **1,** 1419.
Mann, W. N. (1974). *Guy's Hosp. Rep.* **123,** 197–219.
Miller, K. B., Sepersky, R. A., Brown, K. M., Goldberg, M. J., and Kaplan, M. M. (1983). *Am. J. Med.* **75,** 75–80.
Miller, K. B., Elta, G. H., Rudders, R. A., and Kaplan, M. M. (1984). *Ann. Intern. Med.* **100,** 385–387.
Paronetto, F., Schaffner, F., and Popper, H. (1967). *J. Lab. Clin. Med.* **69,** 979–989.
Potter, B. J., Elias, E., and Jones, E. A. (1976). *J. Lab. Clin. Med.* **88,** 427–439.
Reynolds, T. B., Denison, E. K., Frankl, H. D., Liebermann, F. L., and Peters, R. L. (1971). *Am. J Med.* **50,** 302–312.
Roll, J., Boyer, J. L., Barry, D., and Klatskin, G. (1983). *N. Engl. J. Med.* **308,** 1–7.
Rubin, E., Schaffner, F., and Popper, H. (1965). *Am. J. Pathol.* **46,** 387–407.
Schultheiss, H.-P., Berg, P., and Klingenberg, M. (1983). *Clin. Exp. Immunol.* **54,** 648–654.
Thomas, H. C., and Epstein, O. (1980). *Springer Semin. Immunopathol.* **3,** 375–384.
Thomas, H. C., Potter, B. J., and Sherlock, S. (1977). *Lancet* **2,** 1261–1263.
Thomas, H. C., De Villiers, D., Potter, B. J., Hodgson, H., Jain, S., Jewell, D. P., and Sherlock, S. (1978). *Clin. Exp. Immunol.* **31,** 150–157.
Triger, D. R. (1980). *Br. Med. J.* **281,** 772–775.
Triger, D. R., Charlton, C. A. C., and Ward, A. M. (1982). *Gut* **23,** 814–818.

Tsantoulas, D., Perperas, A., Portmann, B., Eddleston, A. L. W. F., and Williams, R. (1980). *Gut* **21,** 557–560.

Vierling, J. M., Nelson, D. L., Strober, W., Bundy, B. M., and Jones, E. A. (1977). *J. Clin. Invest.* **60,** 1116–1128.

Walker, J. G., Doniach, D., Roitt, I. M., and Sherlock, S. (1965). *Lancet* **1,** 827–831.

Wands, J. R., Dienstag, J. L., Bhan, A. K., Feller, E. R., and Isselbacher, K. J. (1978). *N. Engl. J. Med.* **298,** 233–237.

CHAPTER 13

Autoimmune Kidney Diseases

BERNICE NOBLE
Department of Microbiology and Department of Pathology
School of Medicine
State University of New York at Buffalo
Buffalo, New York

JAN R. BRENTJENS
Department of Pathology, Department of Microbiology, and Department of Medicine
School of Medicine
State University of New York at Buffalo
Buffalo, New York

GIUSEPPE A. ANDRES
Department of Microbiology, Department of Pathology, and Department of Medicine
School of Medicine
State University of New York at Buffalo
Buffalo, New York

I. INTRODUCTION

Techniques of immunocytochemistry, immunofluorescence, and electron microscopy have been invaluable for the analysis and evaluation of the contribution of immunological processes to the pathogenesis of kidney disease. The application of those techniques to the study of human

biopsy specimens has made it possible to trace the course of immunopathogenetic events in many human nephropathies of autoimmune origin. In addition, animal models of immunologically mediated kidney disease have been essential for the development of current concepts of kidney immunopathology. In a number of instances, the recognition of immunological processes in human autoimmune kidney diseases has been preceded by observations made in the laboratory with appropriate animal models. Therefore, this chapter will review salient features of the most important examples of autoimmune kidney disease in laboratory animals as well as in man. For reasons of convenience and tradition, glomerulonephritides and tubulointerstitial nephritides will be described separately, although that division is often somewhat artificial.

It will become apparent that there is extensive and convincing documentation implicating mechanisms of humoral immunity in the pathogenesis of autoimmune kidney diseases. In contrast, relatively little evidence has been provided to show that the effectors of cellular immunity contribute significantly to kidney damage either in animals or in man. While it seems that antibody-mediated injury is the more important mode of immunological attack on the kidney, it is also possible that our disproportionate appreciation of the role of humoral factors in kidney pathology may reflect biases imposed by the available analytical methods. In any event, the relative importance of mechanisms of delayed hypersensitivity in kidney disease remains to be assessed.

II. ANTIBODIES AND AUTOIMMUNE KIDNEY DISEASES

There are two mechanisms by which antibodies may accumulate in tissues. An antibody may have specificity for antigens located and fixed within a particular tissue. In that case, the antibody will be retained at the site of the antigen as the result of a unique reaction of immunological recognition. Alternatively, antibodies may be directed against soluble antigens that are present in body fluids. The immunologically specific reaction will occur in a fluid compartment of the body; the products of the reaction may deposit or lodge in body tissues. In the latter case, the site of accumulation of an antibody molecule is unrelated to its immunological specificity.

For many years it was believed that these two mechanisms could be unequivocally distinguished by direct immunofluorescence tests that reveal the pattern of distribution of immune reactants within a tissue. Binding of antibodies to structural antigens, especially basement membranes,

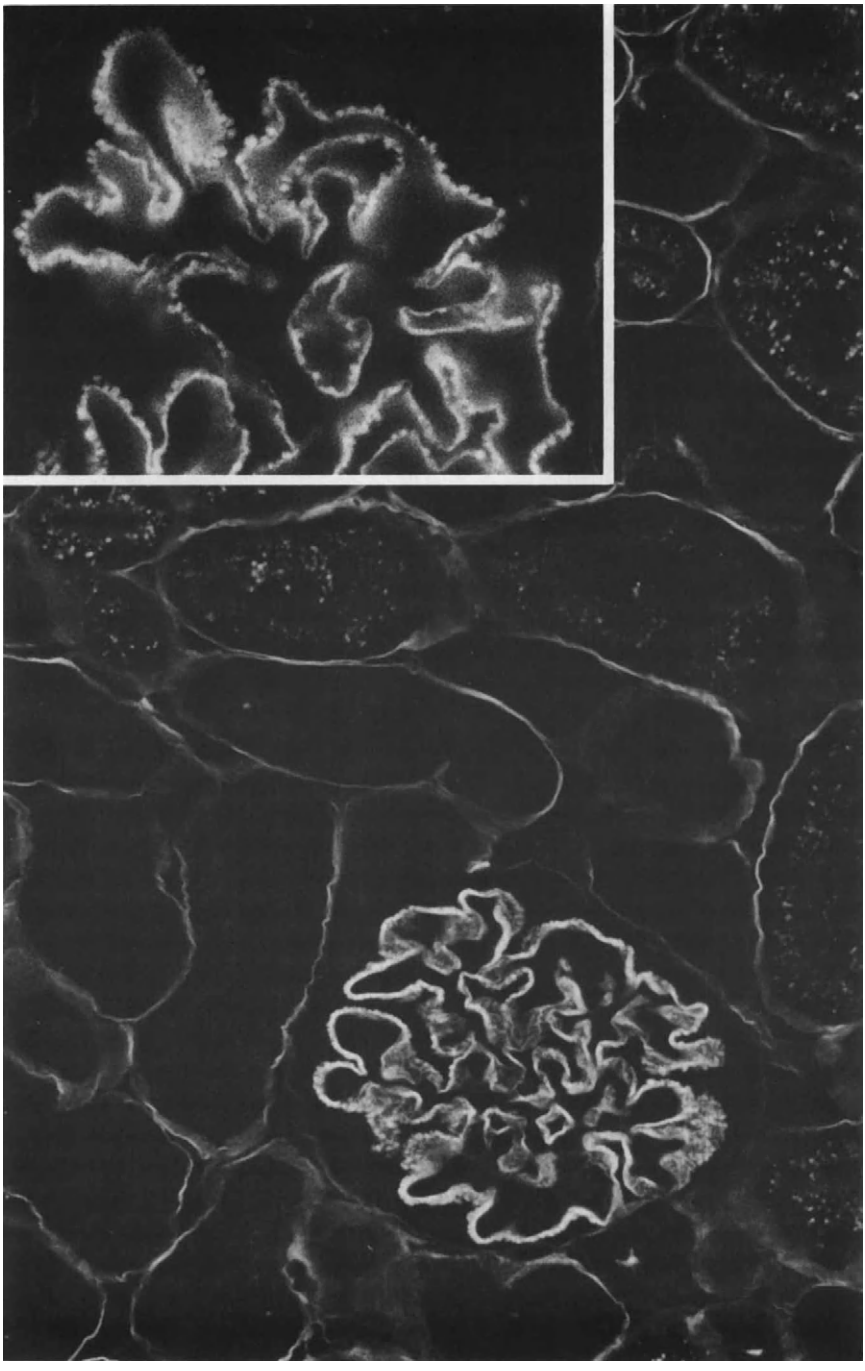
was recognized by a continuous, uniform, finely linear staining pattern that contrasted with the discrete granular distribution of antibodies attributed to the deposition of immune complexes from the circulation. In the animal model, Heymann nephritis, it has recently been shown that reaction of antibodies with structural antigens of the glomerular capillary wall may also produce an irregular, uneven distribution (Van Damme *et al.*, 1978; Couser *et al.*, 1978; Makker and Moorthy, 1981; Neale and Wilson, 1982). Thus, the reaction *in situ* of an antibody with a structural antigen could lead to an immunofluorescence pattern that is discretely granular, resembling the pattern observed with deposited circulating immune complexes. Although it has not yet been demonstrated that a similar event occurs in human kidney disease, it is clear that identification of mechanisms of antibody-mediated injury can no longer depend only on the observable pattern of distribution of immunoglobulin molecules.

A. Diseases Mediated by Antibodies Reacting with Structural Antigens of the Kidney

1. Glomerulonephritides

a. Animal Models

Antiglomerular Basement Membrane (GBM) Nephritis. The classic example of autoimmune disease produced by antibodies to structural antigens of the kidney is anti-GBM glomerulonephritis. Autologous antibodies to antigens of the GBM were elicited first in sheep by immunization with heterologous or homologous GBM administered in Freund's adjuvant (Steblay, 1962). Antibodies formed against the foreign GBM cross-react with autologous antigens to produce nephritis. The course of autoimmune nephritis in sheep is rapidly progressive, leading to renal failure and death. Clinical symptoms of the disease appear 1–3 months after the first immunization, although immune reactants may actually be present in the kidney for several weeks before the manifestation of clinical signs of disease. Direct immunofluorescence tests on kidney tissue from affected animals show immunoglobulin and complement to be distributed in a linear pattern along the GBM. Antibodies in the eluates of diseased kidneys also bind *in vitro* to normal GBM in a linear pattern (Rudofsky and Steblay, 1968). Severe proliferative lesions are observed in glomeruli of animals with anti-GBM nephritis; epithelial crescents occur with high frequency and thus are the histological hallmark of the disease. Neutrophils, macrophages, and fibrin deposits may also be found in glomeruli.



The importance of circulating anti-GBM antibodies to the pathogenesis of this form of autoimmune nephritis has been rigorously demonstrated by passive transfer experiments (Steblay, 1964; Lerner and Dixon, 1966). Normal recipients of serum containing anti-GBM antibodies develop a glomerulonephritis that is indistinguishable from the disease of the actively immunized serum donor.

Sheep autoantibodies that bind to basement membranes of glomeruli will also react *in vitro* with basement membranes in the lung (Rudofsky and Steblay, 1968). Furthermore, immunization of sheep with preparations of alveolar basement membranes produces anti-GBM nephritis (Steblay and Rudofsky, 1968). As anti-basement-membrane antibodies in kidney eluates can be removed by absorption with basement membrane antigens that have been purified from either lung or kidney, it appears that the autoantibodies recognize determinants common to lung and kidney. However, despite the presence of circulating antibodies that react *in vitro* with alveolar basement membranes, sheep immunized with basement membrane antigens do not exhibit signs of lung disease. Other factors in addition to the autoantibodies are probably essential for the clinical expression of autoimmune lung disease. Evidence to support that idea has been provided by passive transfer experiments in which anti-basement-membrane pneumonitis was detected only in recipients exposed earlier to sublethal doses of pure oxygen (Jennings *et al.*, 1981). In that case, an increased permeability of the alveolar capillary, mediated by oxygen, permitted accessibility of antibodies to the alveolar basement membrane.

Immunization of Swiss–Webster mice with human GBM leads to a glomerulonephritis that is characterized, in an early stage, by linear deposits along the GBM and tubular basement membrane (TBM) of antibodies of all IgG subclasses (Bolton *et al.*, 1978). With progression of the disease, the distribution of many IgG subclasses assumes a granular pattern, although the pattern of binding of IgG_1 remains linear (Fig. 1). Eluates of the kidneys of mice with this form of nephritis contain IgG of all subclasses. However, antibodies staining normal GBM *in vitro* are of the IgG_1 class alone. Examination of diseased kidneys by electron micros-

FIG. 1. Direct immunofluorescence test on a frozen section of kidney taken from a Swiss–Webster mouse 5 months after immunization with whole human glomeruli. The section has been stained for mouse IgG. Deposits of IgG in a finely granular-to-ribbonlike pattern are seen along the glomerular capillary wall and in a linear pattern along the basement membrane of tubules. Cells of some proximal tubules contain droplets of IgG (250×). Inset: At a higher magnification, IgG can be seen to be present in the GBM and in subepithelial protrusions emanating from the GBM (800×).

copy reveals a unique lesion of the GBM that is characterized by an expansion of the lamina rara externa in a "beaded" pattern.

Guinea pigs may also develop autoimmune anti-GBM nephritis after immunization with human GBM (Couser *et al.*, 1973). Immunopathology and pathophysiology of glomeruli occur by a mechanism that is independent of complement activation. The role of complement has not been evaluated in other models of autoimmune anti-GBM nephritis. Factors influencing the accumulation and/or proliferation of cells in glomeruli also have not been analyzed in anti-GBM nephritis.

Mercuric Choloride Nephropathy. Regular frequent injections of small doses of $HgCl_2$ induce a biphasic autoimmune kidney disease in rabbits and in Brown-Norway rats (Roman-Franco *et al.*, 1976; Sapin *et al.*, 1977). In the early weeks after the start of injections of $HgCl_2$, deposition of IgG in a linear pattern can be seen along basement membranes in the kidney, muscles, and vessels. Later in the course of the disease, the distribution of immunoglobulins in the kidney and in other organs assumes a discrete granular pattern (Roman-Franco *et al.*, 1978; Druet *et al.*, 1978). Anti-GBM antibodies may be recovered from the kidneys of animals with $HgCl_2$ nephropathy by acid elution. These observations are consistent with the hypothesis that, in the first phase of the disease, $HgCl_2$ induces the formation of autoantibodies to basement membranes. The second phase of disease, characterized by a granular pattern of fluorescence, may be mediated by immune complexes, although the relevant antigen(s) has not been identified. It has been suggested that basement membrane antigens, altered during the early phase, are involved in the formation of the granular immune deposits that are characteristic of the second phase.

It is not known whether both phases of $HgCl_2$ nephropathy are actually separate diseases or different expressions of the same underlying process. Susceptibility to the first phase of $HgCl_2$ nephropathy (anti-GBM antibody production) is a dominant trait linked to the major histocompatibility complex. A possible genetic basis of susceptibility to the second phase appears to be more complex (Sapin *et al.*, 1982).

Heymann Nephritis. It was first reported by Heymann and co-workers (1959) that nephrotic syndrome could be produced in rats by immunization with an extract of rat kidney. A crude extract of rat kidney tubules, designated Fx1A, is adequate to produce Heymann nephritis in rats of susceptible inbred strains and has been used in most studies (Edgington *et al.*, 1968). The nephritogenic antigen in the Fx1A extract has been identified as a plasma membrane glycoprotein that is present in large amounts

on the brush border of proximal tubules (Miettinen *et al.*, 1980; Kerjaschki and Farquhar, 1982). The antigen can also be found in the glomerular capillary wall, although the demonstration of its presence in glomerular sites is much more difficult (Van Damme *et al.*, 1978; Couser *et al.*, 1978; Makker and Moorthy, 1981; Neale and Wilson, 1982; Kerjaschki and Farquhar, 1983).

Nephrotic syndrome develops ~2 months after the first immunization with Fx1A. Examination of kidney tissue from rats with Heymann nephritis reveals the classic picture of membranous nephropathy. By immunofluorescence microscopy, discrete granular deposits of immunoglobulin and complement are observable; corresponding dense deposits, detected with the electron microscope, are located at subepithelial sites in glomeruli (Grupe and Kaplan, 1969; Schneeberger and Grupe, 1976). Loss of epithelial cell foot processes and thickening of the GBM are the most prominent alterations of the ultrastructure of glomeruli in Heymann nephritis. Sera from rats with Heymann nephritis stain the brush border of proximal tubules in indirect immunofluorescence tests; antibodies with a similar specificity can be eluted from the glomeruli of affected animals (Grupe and Kaplan, 1969). The granular deposits in glomeruli appear to be composed of the brush border antigen(s), antibodies to the brush border antigen(s), and complement (Edgington *et al.*, 1967; Glassock *et al.*, 1968).

As it was long considered an axiom of immunopathology that a granular distribution of immune reactants could result only from the deposition, in tissues, of immune complexes formed in circulation, the earliest hypotheses to explain the immunopathogenesis of Heymann nephritis invoked the formation of brush border–anti-brush-border complexes outside the kidney. It was proposed that the accumulation in glomeruli of those complexes, containing a kidney-specific antigen, was fortuitous and occurred as a consequence of the filtration function of the glomerular capillary wall. The weight of recent evidence suggests that a different mechanism may account for the accumulation of immune deposits in glomeruli in Heymann nephritis. It has been shown that perfusion of isolated rat kidneys with heterologous or autologous anti-Fx1A antibodies produces a fine granular pattern of distribution of immunoglobulin along the glomerular capillary wall, with formation of small electron-dense deposits at subepithelial sites (Van Damme *et al.*, 1978; Couser *et al.*, 1978; Makker and Moorthy, 1981; Neale and Wilson, 1982). These observations have led to the formulation of a proposal that immune deposits in glomeruli may form by a reaction *in situ* of circulating autoantibodies with a structural glomerular antigen.

Monoclonal antibodies raised against a highly purified preparation of

the nephritogenic antigen of Heymann nephritis have been employed to determine the precise location of that antigen within the glomerular capillary wall (Kerjaschki and Farquhar, 1983). The antigen is associated with epithelial cells; it can be demonstrated in the endoplasmic reticulum, Golgi apparatus, and multivesicular bodies, and at the cell surface. The accumulation of large immune deposits at subepithelial sites could be explained by a mechanism in which complexes formed at the plasma membrane are subsequently "shed" between the GBM and the base of visceral epithelial cells, where they become trapped at the site of highest fluid flow, namely in the filtration slits.

Susceptibility of different rat strains to Heymann nephritis has been shown to be influenced by many genes, of which at least one is linked to the major histocompatibility locus (Stenglein *et al.*, 1978). Other genetic and environmental factors probably have an additional influence on the expression of nephritis. The involvement of complement in antibody-mediated injury has traditionally been thought to be associated with the accumulation of inflammatory cells within the lesion. In Heymann nephritis, the development of proteinuria depends on the deposition of complement in the capillary wall, even though inflammatory cells are entirely absent from glomeruli (Edgington *et al.*, 1968; Salant *et al.*, 1980; Noble *et al.*, 1983). Rats that fail to develop proteinuria after immunization with Fx1A may have deposits of immunoglobulin in glomeruli that are indistinguishable in distribution and amount from those found in rats with proteinura. It is the presence of complement in the immune deposits that differentiates rats with kidney disease from those with normal protein excretion.

Despite the persistence of proteinuria, active immunological injury does not continue throughout the life of a rat with Heymann nephritis (Noble *et al.*, 1982b). Injury occurs only during a rather limited period in the course of disease when antibodies to brush border are present in circulation. However, restimulation of the antibody response to brush border antigens in rats with Heymann nephritis is accompanied by a significant exacerbation of functional and structural lesions in the kidney. After reimmunization, new immune deposits are formed in glomeruli. Those deposits appear only at the subepithelial aspect of the very thickened GBM, a site of accumulation that favors the idea of their formation by a reaction occurring *in situ*.

Rats that fail to develop Heymann nephritis after primary immunization with Fx1A have an abnormal susceptibility to develop clinically significant autoimmune kidney disease on reexposure to the nephritogenic antigen later in life (Noble *et al.*, 1983). In those animals, an autoantibody response inadequate to produce disease leads to subclinical immuno-

pathology and specific immunological memory, which can predispose the animal to an unusually rapid expression of autoimmune kidney disease.

b. Human Diseases

Anti-GBM Nephritis. Less than 2% of patients suffering from glomerulonephritis have a disease that results from autoantibodies with specificity for basement membrane constituents (Lerner *et al.*, 1967). The majority of those patients suffer from Goodpasture's disease, in which glomerular nephritis and hemorrhagic pneumonitis occur together. In those patients, linear deposits of antibody are found along both granular and alveolar basement membranes (Wilson and Dixon, 1973). In the remaining cases, pathology seems to be limited to the kidney. Anti-GBM nephritis may have a sudden onset and is usually characterized by rapid deterioration of kidney function, although milder forms of the disease can also occur. The lesion in glomeruli is one of severe proliferation; extensive crescent formation is common. In affected lungs, intraalveolar hemorrhages, infiltration of the pulmonary interstitium with leukocytes, and accumulation of iron-containing macrophages in alveoli may be observed.

In this disease rather rigorous proof has been provided of the pathogenicity of autoantibodies to GBM (Wilson, 1981). First, it has been shown that patients produce antibodies that stain basement membranes *in vitro*. Those autoantibodies are usually not specific for basement membranes of the kidney, but often cross-react *in vitro* with basement membranes of many organs. In some patients the autoantibodies may not be detectable in circulation during an acute phase of illness, but they may be recovered from serum after bilateral nephrectomy. Second, immunoglobulins eluted from diseased kidneys, which also stain GBM *in vitro*, will elicit a comparable glomerulonephritis when passively transferred to monkeys. Third, glomerulonephritis may recur in a kidney graft if transplantation surgery is performed when autoantibodies are present in the circulation of the recipient.

It seems likely that the same pathogenetic mechanism accounts for both lung and kidney damage in Goodpasture's syndrome (Wilson, 1981). However, lung disease may be absent in many patients with anti-GBM nephritis, despite the presence of circulating antibodies that stain alveolar basement membranes *in vitro*. Furthermore, there is little correlation of the severity of lung histopathology with the serum titer of anti-basement membrane autoantibodies. These discrepancies have led to the hypothesis that factors besides autoantibodies alone are required to initiate damage in the lung. These factors, which supposedly increase the permeability of the alveolar capillary wall, include viral or bacterial infections, fluid overload, uremia, and toxic fumes.

Anti-GBM antibodies appear to react predominantly with the subendothelial aspect of the GBM (Fish *et al.*, 1979). The GBM is a mixture of collagenous and noncollagenous proteins; nephritogenicity is associated with noncollagenous components, but multiple determinants may be involved.

Membranous Nephropathy. Antigens of the brush border of proximal tubules similar to those responsible for Heymann nephritis have been identified in glomerular immune deposits in a few patients with membranous nephropathy (Naruse *et al.*, 1973; Douglas *et al.*, 1981). However, most other attempts to identify that particular antigen–antibody system in the pathogenesis of human membranous nephropathy have met with failure. It seems likely that brush border antigens are only rarely the cause of human glomerulonephritis.

Because glomerular damage in Heymann nephritis, the animal model most often compared to human idiopathic membranous glomerulonephritis, has been shown to result from autoantibodies directed against a structural component of the glomerular capillary wall, it must now be considered that many cases of idiopathic membranous nephropathy in humans could have a similar immunopathogenesis (Fig. 2).

2. Tubulointerstitial Nephritides

a. Animal Models

Anti-TBM Nephritis. Active immunization of laboratory animals with antigens of the TBM produces a severe and often fatal tubulointerstitial nephritis (Steblay and Rudofsky, 1971). This form of experimental nephritis was first described in guinea pigs, but it has also been produced in mice, rats, and goats. The guinea pig model has been studied in great detail in a number of laboratories and is one of the best analyzed of the experimental autoimmune kidney diseases.

Several weeks after immunization of guinea pigs with TBM antigens and adjuvant, clinical manifestations of the disease, such as glucosuria and uremia, are detectable. By the indirect immunofluorescence technique, the sera of immunized animals are seen to contain autoantibodies against TBM. The autoantibodies are able to react *in vivo* to form linear deposits of IgG along the basement membrane of proximal tubules. Deposition of the antibodies along TBM precedes the development of histological lesions (Fig. 3).

In an early stage of anti-TBM nephritis, the kidney cortex is diffusely infiltrated with mononuclear cells, principally monocytes and macrophages (Andres *et al.*, 1979). In addition, small numbers of epithelioid cells appear to be engaged in active pinocytosis and phagocytosis. In the

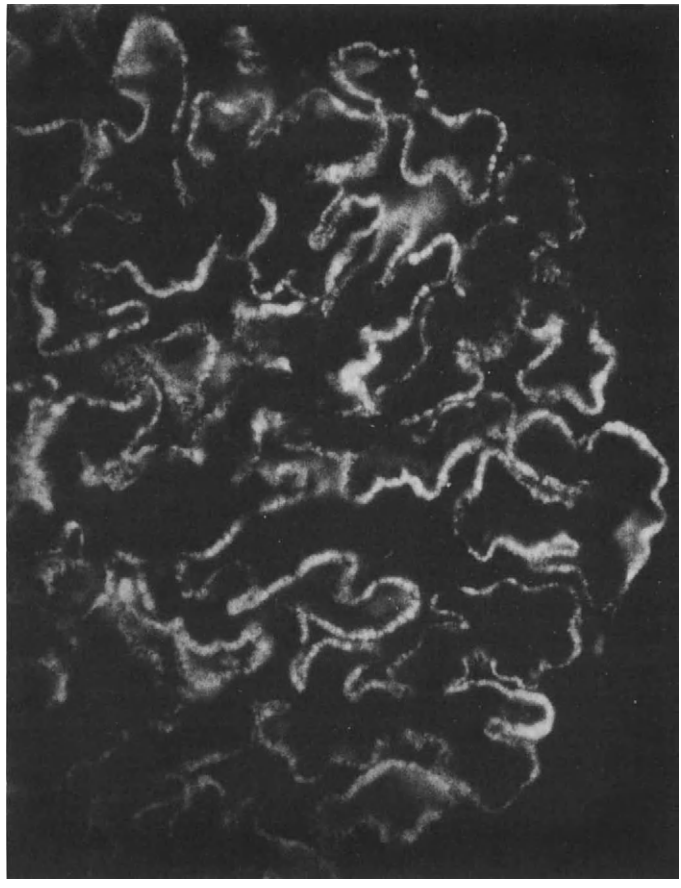


FIG. 2. Direct immunofluorescence test showing granular deposits of IgG along the glomerular capillary wall in a glomerulus from a patient with idiopathic membranous nephropathy. Mesangial areas are free of deposits (600×).

severe and final stage of the disease, which develops 1 to 2 months after the first immunization, the TBM becomes thin and fragmented. The TBM may be missing from the base of many tubules. Multinucleated giant cells, formed by fusion of adjacent epithelioid cells, are distributed throughout the cortex and are often in direct contact with the TBM. Near the point of contact the cytoplasm of the giant cell is homogenous, containing bundles of microfilaments but not cytoplasmic organelles. Direct contact of giant cells with the TBM, rather than secretory activity, is probably a prerequisite for TBM destruction. It is interesting to note that the peritubular giant cells that appear to be important in the destruction of the TBM in guinea

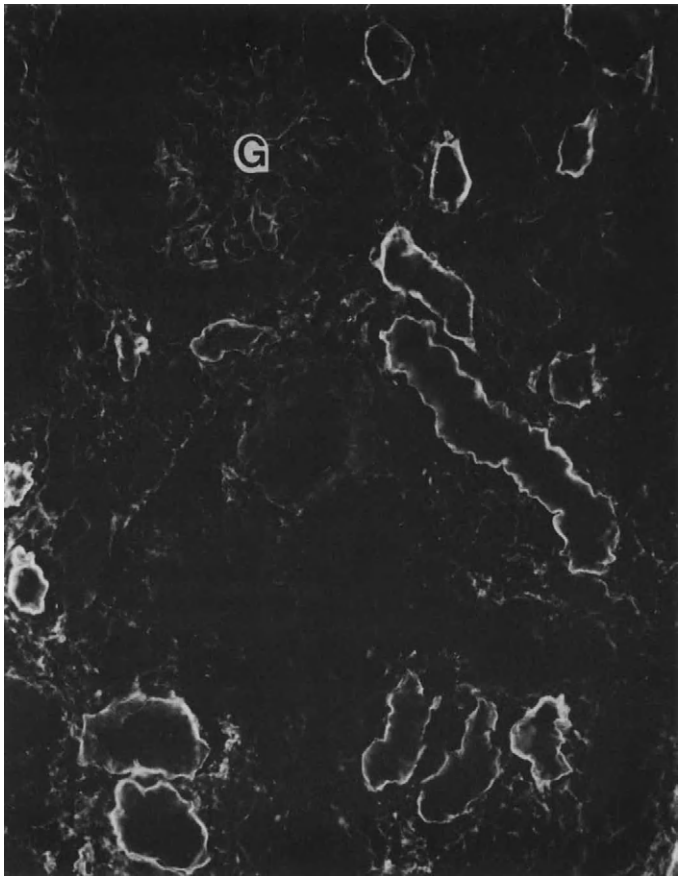


FIG. 3. Direct immunofluorescence test showing the distribution of IgG in the kidney of a guinea pig with severe anti-TBM nephritis. Although some tubules exhibit heavy linear deposits of IgG along the basement membrane, in many others those deposits have disappeared because of dissolution and destruction of the basement membrane. A minimal accumulation of IgG in an interrupted, linear pattern is visible along the GBM in a glomerulus (G) that is contained within an enlarged Bowman's space (250×).

pigs are not a conspicuous feature of the histopathology of the interstitial inflammation that develops in rats immunized with TBM antigens (Sugisaki *et al.*, 1973; Lehman *et al.*, 1974b). Species differences in the histopathology of interstitial lesions associated with TBM deposits of immunoglobulin may reflect differences in underlying pathogenetic mechanisms. On the other hand, it is also possible that, although having

similar mechanisms, patterns of mediator production and inflammatory cell accumulation may vary from species to species.

Conclusive evidence of the important pathogenetic role of the anti-TBM autoantibodies in anti-TBM nephritis has been provided by the successful passive transfer of disease to normal animals with sera or kidney eluates from animals actively immunized with TBM antigens (Steblay and Rudofsky, 1973). However, although anti-TBM antibodies are essential for the development of tubulointerstitial nephritis, antibodies alone are not sufficient effectors of the disease. Activation of complement by either the classical or alternative pathway is required to achieve full expression of the most severe form of nephritis (Rudofsky *et al.*, 1974, 1975).

Despite strong evidence that autoantibodies play a central role in the pathogenesis of anti-TBM nephritis, it cannot be denied that the histological appearance of the interstitial infiltration is consistent with a cell-mediated reaction. For that reason there has been considerable effort expended to assess the contribution of specific cellular effector mechanisms to the pathogenesis of the disease. It has been shown that T cells from guinea pigs with anti-TBM nephritis respond *in vitro* to the crude rabbit TBM preparation with which they have been immunized (Neilson and Phillips, 1979). Furthermore, radiosensitive, bone marrow-derived cells must be present in the recipient for successful transfer of anti-TBM nephritis with serum (Rudofsky and Pollara, 1975). However, treatment of donors with niridazole, a compound that suppresses delayed hypersensitivity, does not prevent the transfer of disease with serum (Rudofsky and Pollara, 1977).

To explain these somewhat confusing and contradictory observations, it has been proposed that an antibody-dependent cell-mediated reaction could be the important mode of tissue damage (Andres *et al.*, 1979). The requirement for radiosensitive, bone marrow-derived cells as well as specific antibody would be explained by such a mechanism. In addition, analogy has been noted between histological features of the interstitial lesions in anti-TBM nephritis, including the apparent direct lytic action of giant cells on the TBM, and the reactions of antibody-dependent lymphocytotoxicity studied *in vitro*.

Other aspects of humoral immunity are involved in the expression and regulation of anti-TBM responses. The phenomenon of "autoimmune amplification" that was first described in autoimmune hemolytic anemia may be relevant in the progressive course of anti-TBM nephritis. Active immunization of guinea pigs with rabbit TBM elicits production of anti-TBM antibodies of both IgG_1 and IgG_2 isotypes. The separate transfer of either isotype induces nephritis, but also stimulates the synthesis by the recipi-

ent of anti-TBM autoantibodies of both immunoglobulin subclasses (Hall *et al.*, 1977). The mechanism by which transferred antibody might stimulate the synthesis of autoantibody of the same reactivity is not understood. One possibility is that the injected antibodies cause sufficient modification and/or release of TBM antigens to initiate an autoimmune response by the recipient.

It has been reported that significant inhibition of tubulointerstitial nephritis in the guinea pig can be achieved by the intraperitoneal administration of small amounts of an antiidiotypic antiserum at the time of active immunization with TBM antigens. The antiidiotypic antibodies obtained from rabbits are directed against guinea pig autoantibodies to TBM (Brown *et al.*, 1979). Although the mechanism of suppression of nephritis in these circumstances has not been explained, comparable effects in other systems have been ascribed to the clonal deletion of B cells or the stimulation of T suppressor cells.

In several studies, the nature of the nephritogenic TBM antigen(s) has been analyzed. Because TBM shares antigens with the GBM, immunization of certain strains of mice or rats with heterologous GBM antigens results in both anti-GBM and anti-TBM nephritis (Robertson *et al.*, 1977; Bolton *et al.*, 1978). Guinea pigs immunized with a bovine TBM preparation form antibodies against GBM as well as TBM. The larger portion of antibodies eluted from diseased kidneys of these animals appears to be directed against noncollagenous antigens of the TBM; a smaller fraction reacts with collagenous antigens of the GBM and TBM (Lehman *et al.*, 1974a). There is some indication that the nephritogenic antigens are noncollagenous proteins, which explains why the disease is accompanied only by minimal glomerular abnormalities. Despite the great similarity between TBM and GBM, there are antigenic as well as chemical differences (Wakashin *et al.*, 1981). By a combination of immunochemical and physicochemical procedures, a soluble nephritogenic TBM antigen with a molecular weight of 30,000 has been isolated from human kidney. This antigen appears to be unique for the TBM. Animals immunized with the antigen form antibodies against TBM but not GBM, and they develop only lesions of the tubules and interstitium.

In mice and in guinea pigs the susceptibility to develop anti-TBM nephritis appears to be linked, as a dominant trait, to the major histocompatibility complex. Responder and nonresponder strains of mice have similar antibody responses to TBM antigens (Rudofsky *et al.*, 1980). Therefore, factors influencing the expression of disease have not yet been identified in the mouse model of anti-TBM nephritis. In rats, the antibody response to TBM antigens is strain dependent. In addition, some rat

strains lack the nephritogenic antigen and are resistant to the disease (Sugisaki *et al.*, 1973; Lehman *et al.*, 1974b; Hart and Fabre, 1980).

Heymann Nephritis. Heymann nephritis is produced by immunization with antigens of the proximal tubule brush border (see Section II,A,1,a). As a consequence of the increased permeability that results from the membranous lesion in glomeruli, anti-brush border antibodies gain access to the proximal tubule lumen. Shortly after the onset of proteinuria in an early phase of Heymann nephritis, deposits of immunoglobulin and complement may be found along the luminal membrane of proximal tubules as well as in glomeruli. Examination by light and electron microscopy of kidney tissue, fixed by perfusion *in situ* to ensure optimal preservation of tubule architecture, reveals that antibody deposition results in an extensive pathology of proximal tubules (Mendrick *et al.*, 1980). Aspects of tubule injury seen in the early proteinuric stage of Heymann nephritis include extensive loss of microvilli, flattening of epithelial cells, increased mitotic activity of the tubule epithelium, accumulation of cells within tubule lumina, and reduction in the number of pinocytotic vesicles. Focal sites of mononuclear cell infiltration of interstitium may also be recognized.

In rats with active Heymann nephritis, the proximal tubule epithelium is exposed chronically to specific antibody. With chronic exposure, luminal deposits of immunoglobulin in proximal tubules become focal, weak, or absent, attesting to the loss of antigen from the plasma membrane. Indirect immunofluorescence tests with anti-brush border antiserum confirm that the antigen has disappeared. Faint, finely granular deposits detectable at the basal portion of many proximal cells in an early, preproteinuric stage of Heymann nephritis become prominent heavy granular accumulations of immunoglobulin. By electron microscopy those deposits can be seen to lie between the basement membrane and the basal membrane of the cells. Their presence is associated with wrinkling and thickening of the TBM, flattening of the epithelium, and loss of basal infoldings. The subepithelial location of those immune deposits and their strict limitation to proximal tubules suggest strongly that a specific immune reaction accounts for their formation (Mendrick *et al.*, 1980). If the nephritogenic antigen were present in small amounts on the basolateral membrane of proximal tubule cells, subepithelial deposits along the TBM might form by a mechanism similar to that proposed to account for the subepithelial accumulation of deposits along the GBM in Heymann nephritis. The antigen–antibody complexes could accumulate after formation on the plasma membrane by a process analogous to the "capping"

and "shedding" of immunoglobulin–antiimmunoglobulin complexes from suspensions of B cells *in vitro*.

In the natural course of Heymann nephritis, anti-brush border antibodies eventually fall to undetectable or insignificant concentrations in serum. Despite persistent proteinuria, in the final, chronic stage of the disease, proximal tubules are no longer exposed to antibody. Although immunoglobulin deposits are present in the kidney they are limited to glomeruli; proximal tubules are completely free of those deposits. As a result, a partial recovery of the normal architecture of the proximal tubule epithelium occurs. The cell height becomes normal; brush border is present, if sparse and short, on many epithelial cells. The active inflammatory infiltration of the interstitium is replaced by fibrosis. A new TBM forms below the basal membrane of the epithelial cells, to which characteristic infoldings are restored. An accumulation of electron-dense debris remains between the new TBM and the original, thickened, and damaged TBM. Thus, in a late stage of Heymann nephritis, there may be no evidence in the form of immune deposits at the brush border or the basement membrane to suggest the pathogenesis of tubulointerstitial damage (Mendrick *et al.*, 1980; Noble *et al.*, 1982a).

In this model of antibody-mediated injury to the kidney, as in many others, direct evidence of the pathogenicity of autoantibodies has been obtained from passive transfer experiments (Noble *et al.*, 1981a). Passive transfer of rat anti-brush border antiserum from actively immunized animals to recipients with increased glomerular permeability produces deposition of large amounts of rat immunoglobulin along the luminal border of proximal tubules. The pathology produced in previously normal proximal tubules by transferred antibodies is similar to that seen in active Heymann nephritis. Prominent features of the damage include loss of microvilli and proliferation of epithelial cells. From passive transfer experiments it has become clear that the majority of cells seen within the tubule lumen in Heymann nephritis are not blood borne. These cells are derived instead from the tubule epithelium as a consequence of the dramatic proliferation that follows antibody deposition on the brush border.

It should be stressed that, although Heymann nephritis has long been considered a valuable model of idiopathic membranous nephropathy in man, the close and inevitable association in Heymann nephritis of tubulointerstitial damage with membranous lesions of the glomerular capillary wall is quite different from the human disease.

Interstitial Nephritis in Rabbits. Rabbits immunized with nonglomerular components of rabbit kidney develop an interstitial nephritis that is characterized by focal lymphocytic infiltration, tubule degeneration,

and extensive fibrosis of the interstitium (Klassen *et al.*, 1971, 1977). The immunization elicits production of autoantibodies that stain the cytoplasm of normal proximal tubule cells in indirect immunofluorescence tests. In the kidneys of affected rabbits, immune deposits are found in a granular pattern along the TBM of proximal tubules. Passive transfer of serum from rabbits with interstitial nephritis to normal rabbits results in focal deposition of immunoglobulin along TBM in a pattern identical to that seen in the donors. Eluates of the kidneys of immunized rabbits stain antigens located primarily in proximal tubules of normal kidneys. In addition, the eluates contain antibodies that stain antigens present in the granular TBM deposits. Observations with this animal model led to the original formulation of the hypothesis that granular immune deposits in tissues may form by the local reaction of circulating antibodies with antigen(s) present on or "leaking" from cells of a specific tissue.

Tubulointerstitial Nephritis Elicited by Immunization with Tamm-Horsfall Protein. Tamm-Horsfall protein is an antigen present in the membrane of cells of the thick ascending limb of Henle's loop. An infiltration of leukocytes around the thick ascending limb can be found in rats that have been immunized with Tamm-Horsfall protein (Hoyer, 1980). Granular immune deposits containing immunoglobulin, complement, and Tamm-Horsfall protein are also located at the base of the epithelium of the thick ascending limb. Those deposits are believed to arise by the reaction *in situ* of circulating antibodies with molecules whose distribution is restricted to Henle's loop.

b. Human diseases

Anti-TBM Nephritis. Antibodies to TBM are a rare cause of human disease (Andres and McCluskey, 1975). The observation of antibodies to TBM as an isolated phenomenon comparable to the animal models is very unusual. Antibodies to TBM are detected most frequently in patients with concomitant anti-GBM glomerulonephritis and occur in 50 to 70% of these patients (Lehman *et al.*, 1975). The reactivity of circulating or eluted anti-TBM antibodies in anti-GBM/anti-TBM disease is not always restricted to the basement membrane of proximal tubules. In some cases the antibodies only react with the basement membrane of more distal segments of the nephron (Noble *et al.*, 1982a). This may explain why extensive binding *in vivo* of anti-TBM antibodies is observed in some patients, whereas in others the binding is confined to just a few tubules.

In an early stage infiltration of the kidney interstitium with polymorphonuclear leukocytes may be seen. As the disease progresses, the tubules become heavily damaged. Large numbers of lymphocytes, monocytes,

macrophages, and plasma cells accumulate in the interstitium. Inflammatory cells penetrate between epithelial cells, and focal or diffuse destruction of TBM is frequently observed. Proliferation of tubule cells is also seen. The tubulointerstitial lesions in kidneys of patients with anti-GBM/anti-TBM nephritis are more frequent and severe than those in kidneys of patients with isolated anti-GBM disease or other forms of crescentic nephritis. It is likely, therefore, that those lesions are mediated, at least partially, by anti-TBM antibodies.

In a few patients, mostly children, anti-TBM antibodies, interstitial nephritis, and severe tubular dysfunction have been observed to accompany immune complex glomerulonephritis. In one instance, formation of anti-TBM antibodies followed acute poststreptococcal glomerulonephritis (Morel-Maroger *et al.*, 1974). The development of interstitial lesions in that patient was correlated with the appearance of anti-TBM antibodies in circulation and with their deposition along TBM. In other reports, nephrotic syndrome associated with granular deposits along GBM and TBM was accompanied by anti-TBM antibodies in the kidney and/or circulation, as well as with Fanconi's syndrome (Tung and Black, 1975; Levy *et al.*, 1978). It has been proposed that immune complex-mediated injury to the tubules may stimulate an autoimmune anti-TBM response. One patient with linear and granular immune deposits along TBM developed pulmonary symptoms; antibodies reacting with alveolar basement membrane were found in circulation. This finding suggests that anti-TBM antibodies cross-reacting with alveolar basement membrane may occasionally cause pulmonary lesions in man. Anti-TBM antibodies may occur in association with systemic immune complex disease. They have been noted in mice with a lupuslike syndrome, as well as in a patient with systemic lupus erythematosus (Makker, 1980; Rudofsky, 1980).

Anti-TBM antibodies have been found in a few cases of drug-related interstitial nephritis (Baldwin *et al.*, 1968; Border *et al.*, 1974). It has been proposed that drugs, when bound to the TBM as a hapten–protein conjugate, stimulate the production of antibodies that subsequently deposit on the TBM. However, the pathogenetic role of these antibodies is doubtful. Titers of anti-TBM antibodies in the serum may not correlate with disease activity. Circulating antibodies may fail to bind *in vivo* to the TBM (Ooi *et al.*, 1978; Grussendorf *et al.*, 1981). In the majority of patients with tubulointerstitial nephritis caused by drug hypersensitivity, anti-TBM antibodies are not detectable (Ooi *et al.*, 1975; Galpen *et al.*, 1978; Linton *et al.*, 1980). In those cases the pathogenesis of the nephritis is is still entirely a matter of speculation.

Antibodies to Brush Border Antigens. Granular deposits of immunoglobulin along the basement membrane of proximal tubules have

been described in one patient with a membranous glomerulopathy that appeared to be analogous to Heymann nephritis (Douglas *et al.*, 1981). It is possible, therefore, that the tubule deposits seen in the biopsy from those patients were also comparable to those found in Heymann nephritis.

The transient nature of the tubule deposits of immunoglobulin and complement seen in several animal models may explain the failure to identify immune mechanisms in human tubule pathology with great frequency. If clues to immunopathogenetic events in the tubules and interstitium in the form of deposits detectable by immunofluorescence are present for only short periods in the natural course of a disease, they could be overlooked or entirely absent in tissue samples obtained from patients with advanced kidney disease. It is hoped that careful study of the available animal models will provide other clues of histopathology, function, or serology to enhance our ability to recognize immunologically mediated lesions of kidney tubules and interstitium in man.

B. Diseases Mediated by the Deposition of Immune Complexes

1. Glomerulonephritides

a. Animal Models

Experimental Serum Sickness. Serum sickness, resulting from an immune response against foreign serum proteins, is the classic example of tissue injury induced by immune complexes. Although the antigens used to elicit experimental serum sickness in laboratory animals are heterologous proteins, our present appreciation of the contribution of circulating immune complexes to the pathogenesis of certain autoimmune kidney diseases, particularly the nephritis of systemic lupus erythematosus (SLE), depends on insights gained from the study of serum sickness.

Chronic immunization with bovine serum albumin can produce kidney disease in rabbits, rats, and mice (Albini *et al.*, 1979; Arisz *et al.*, 1979; Steward, 1979; Noble *et al.*, 1980; Wilson and Dixon, 1981). The clinical severity of the disease may be influenced by the amount of antibody to albumin that is produced and by the quality of that antibody. The onset of proteinuria is associated with a decline in the serum titer of anti-albumin antibody. In rabbits with chronic serum sickness, a wide range of abnormalities may be found in the kidney; the severity of disease varies from animal to animal. There may be, in order of severity, isolated proliferation of mesangial cells, membranous nephropathy, membranoproliferative changes, or fulminating necrotizing glomerulonephritis. By immuno-

fluorescence microscopy, coarse granular deposits containing bovine serum albumin, IgG, and complement can be found in glomeruli. In rabbits showing histological changes limited to the mesangium, immune complex deposits are found only in mesangial areas of the glomerulus.

In cases of more severe nephritis, immune complex deposits are located predominantly in the glomerular capillary wall, where they may be present at subendothelial, subepithelial, and transmembranous sites. Chronic serum sickness may be produced in rats by means of an immunization protocol that causes all the animals to develop severe, proliferative glomerulonephritis (Arisz *et al.*, 1979; Noble *et al.*, 1981b). In that model, three stages in the development of kidney disease can be identified by criteria of function that correspond to distinct categories of kidney immunopathology (Van Liew *et al.*, 1983). In mild chronic serum sickness, kidney function is normal, and immunopathology is limited to the mesangium. As moderate disease develops, proliferative changes are correlated with abnormal protein excretion. However, it is only in severe serum sickness, when diffuse, necrotizing proliferative glomerulonephritis is present, that significant compromise of whole kidney function occurs. The transitions from mild to moderate and moderate to severe nephritis are not gradual but occur as rather abrupt events.

Rabbits immunized with heterologous or altered homologous thyroglobulin may develop autoimmune thyroiditis. In some cases immune complexes containing thyroglobulin and antibodies to thyroglobulin may deposit in the glomeruli and can lead to mild inflammation with proteinuria (Weigle and Nakamura, 1969). Immune complex deposits have also been found in glomeruli of vasectomized rabbits that had antibodies to sperm together with immune complex-mediated orchitis. Immunoglobulins eluted from kidneys of those rabbits stained the acrosomes of normal sperm, providing strong evidence that the glomerular lesion had a pathogenesis similar to that of serum sickness (Bigazzi *et al.*, 1976).

Spontaneous Immune Complex Glomerulonephritis. Immune complex glomerulonephritis is a major feature of a genetic autoimmune disease that develops spontaneously in several inbred mouse strains (Andrews *et al.*, 1978; Theofilopoulos, 1980). The strains include NZB, NZB × NZW (NZBxW)F_1 hybrids, MRL/1, and BxSB, of which NZB and NZBxW animals have been studied most extensively. Because many aspects of the disease are very similar to human SLE, this lupuslike syndrome of mice is considered to be a valuable animal model with which to analyze the pathogenesis of SLE. Significant and progressive proteinuria develops when the mice are several months old, and affected individuals have a significantly shortened life span. In an early stage of the disease, deposits of immunoglobulin and complement in the glomerular mesan-

gium are accompanied by focal increases in mesangial cellularity. Later the glomerulonephritis becomes more severe, with diffuse cell proliferation and fibrinoid necrosis. In an advanced stage of nephritis, immune deposits are observed in the glomerular capillary wall; corresponding electron-dense deposits may be found at subendothelial as well as subepithelial sites. Serological features of the disease that are common to all strains include elevated concentrations of immunoglobulin, circulating immune complexes, antibodies to nuclear antigens, DNA, and murine retrovirus antigen gp70 (Lambert and Dixon, 1968; Seegal *et al.*, 1969; Andrews *et al.*, 1978). The immune deposits in glomeruli contain DNA and retrovirus antigens as well as the corresponding antibodies.

An immunopathogenetic mechanism similar to that of serum sickness is considered to produce the lupuslike glomerulonephritis of these animals. Among the postulated etiological factors of the underlying autoimmune disease are retroviruses, thymic atrophy or failure, antithymocyte antibodies, hyperreactivity of B cells, deficiency of suppressor T cells, and abnormalities of function of the mononuclear phagocyte system. Because such a large number of immunological and serological abnormalities has been detected, it remains difficult to distinguish significant etiopathogenetic factors from epiphenomena. A genetically determined hyperreactivity of B cells is common to all strains of mice that develop the disease and may be a decisive determinant.

Major differences among the strains in the expression of lupuslike autoimmune disease are the amounts and specificities of autoantibodies, the age of onset and rapidity of progress, and the influence of sex on the course of disease (Andrews *et al.*, 1978). In NZBxW mice, sex hormones may modulate the expression of glomerulonephrits as well as other aspects of autoimmunity, with androgens exerting a protective influence (Roubinian *et al.*, 1977). Administration of prostaglandin E_1 inhibits the development of immune complex-mediated glomerulonephritis and increases the life span of NZBxW and MRL/1 mice (Kelley *et al.*, 1981). Fat content of the diet also affects the natural course of the disease. Animals receiving a high-fat diet develop more severe nephritis and exhibit increased mortality when compared to mice on low-fat diets (Levy *et al.*, 1982).

b. Human Diseases

Systemic Lupus Erythematosus. Immune complex formation is thought to be the main pathogenetic mechanism producing the glomerulonephritis that is an important manifestation of SLE. Clinical signs of kidney involvement develop in ~40% of patients within the first year of the disease and in a much higher percentage during the course of the

illness. Immune deposits accumulate in the kidneys of practically all patients with SLE, including those showing no clinical signs of kidney impairment. Four categories of lupus nephritis are recognized: minimal or mesangial, mild or focal proliferative, severe or diffuse, and membranous lupus nephritis.

In minimal or mesangial lupus nephritis, deposits of Ig and C are detectable in mesangial regions. The histology of glomeruli may be normal, or mild hypercellularity of the mesangium may be noted. In mild or focal proliferative lupus nephritis, a segmental increase in cellularity occurs in some glomeruli; others may be normal. The cell increase is a consequence of the proliferation of glomerular cells as well as the influx of blood-borne inflammatory cells. Immune deposits in mild lupus nephritis are not confined to the mesangium but may also be found in a granular pattern along the peripheral glomerular capillary wall.

The lesions observed in severe or diffuse proliferative lupus nephritis are basically the same as those in the focal proliferative form, but are more severe and widespread. Nearly all glomeruli show segmental or global hypercellularity, resulting from proliferation of endothelial and mesangial cells and infiltration of polymorphonuclear leukocytes and mononuclear cells. The glomeruli are frequently enlarged and have a lobulated aspect. In addition, necrosis, karyorrhexis, intracapillary hyaline thrombi, and irregular thickening and splitting of the GBM are found. Characteristic are "wire loops," representing glomerular capillary walls that are thickened as a result of accumulation of hyaline material at the endothelial side of the GBM. Hematoxylin bodies, when present, are pathognomonic for lupus nephritis. In kidney sections stained with hematoxylin and eosin, these bodies appear as ill-defined structures of variable size, usually associated with areas of necrosis. The hematoxylin bodies originate from nuclei of polymorphonuclear leukocytes, which have been phagocytized by mesangial cells, probably after reacting with antinuclear antibodies. In the most severe cases of proliferative lupus nephritis, crescent formation and arteritis are seen. The most extensive accumulation of immune deposits is found in this form of lupus nephritis. In glomeruli the deposits are located in the mesangium, and in a granular-to-ribbonlike pattern at both sides of the GBM and in the GBM. Large subendothelial immune deposits give rise to the aforementioned "wire loops" (Fig. 4).

The membranous form of lupus nephritis is characterized by a diffuse, uniform thickening of the glomerular capillary wall that is a consequence of the presence of granular immune deposits located only at the epithelial side of the GBM. In some patients, the immunopathological appearance may be indistinguishable from that of idiopathic membranous glomerulopathy. However, the following findings in the kidney biopsy should raise

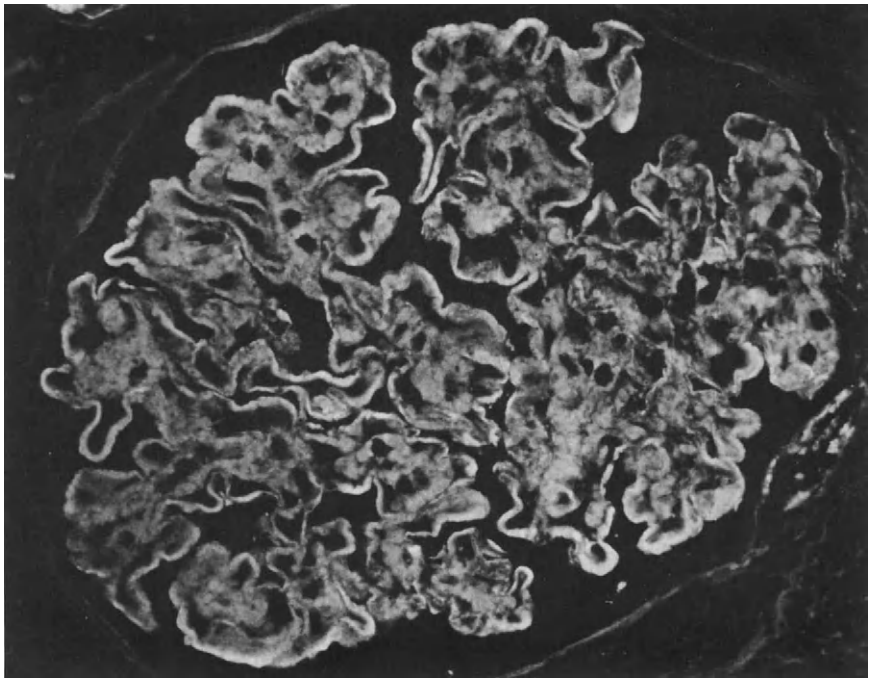


FIG. 4. Direct immunofluorescence test for IgG on a glomerulus from a patient with diffuse proliferative lupus nephritis is seen. The lobulated glomerulus contains ribbonlike deposits along the glomerular capillary wall. In addition, there is an accumulation of IgG in mesangial areas (300×).

the strong suspicion that membranous nephropathy is a manifestation of SLE: mesangial hypercellularity, immune deposits in mesangial areas or at extraglomerular sites in the kidney, and/or IgA in the glomerular immune deposits.

Proliferative and membranous lupus nephritis may eventually progress to end-stage kidney disease. During this evolution, sclerosing kidney lesions become increasingly prominent. At the same time, the immune deposits become less numerous and the intensity of staining for Ig and C fades.

The most important immune complex systems in lupus nephritis appear to involve DNA antigens. DNA antigens and/or corresponding antibodies have been demonstrated in kidney immune deposits in lupus nephritis (Andres *et al.*, 1970; Agnello *et al.*, 1973). It is noteworthy that the presence of DNA–anti-DNA immune complexes in the circulation of patients with SLE has not yet been unequivocally demonstrated (Izui *et al.*, 1977).

Other immune complex systems are suspected to have pathogenetic significance in SLE. The relevant antigens include nuclear ribonucleoprotein, a cytoplasmic glycoprotein called Ro, and antiglobulins (for review, see Wilson and Dixon, 1981; Glassock and Cohen, 1981). There is a positive correlation between disease activity and the concentration of both circulating immune complexes and anti-DNA antibodies. In active disease, serum complement activity may decrease (Swaak *et al.*, 1979). The significance of those phenomena for the pathogenesis of nephritis in SLE is still far from clear.

Immune Complex Glomerulonephritides in which Nuclear Antigens Are Not Involved. Several endogenous antigens in addition to DNA have been reported to give rise to immune complex glomerulonephritis in humans; these include erythrocyte antigens, thyroglobulin, and immunoglobulins. Glomerulonephritis in association with autoimmune thyroid disease may be caused by the deposition of thyroglobulin–antithyroglobulin immune complexes (O'Regan *et al.*, 1976; Ploth *et al.*, 1978; Jordan *et al.*, 1978). Both antigen and antibody have been detected in the kidneys of some patients; thyroglobulin was identified as a component of immune complexes isolated from serum of one patient (Jordan *et al.*, 1981).

Although deposits of immunoglobulins and complement can be demonstrated readily in many human kidney diseases, the identification of specific antigens has been possible only in a few instances. Failure to identify specific exogenous antigens with ease has led to speculation that formation of nephritogenic immune complexes in glomeruli may result from specific interactions between immunoglobulin molecules, with little involvement of exogneous antigens except in the initial phase of disease. According to that idea, small antigen–antibody immune complex deposits in glomeruli could stimulate production of large amounts of autoantibodies directed against the antibody present in the deposits. In that event the original exogenous antigen could be "buried" within a large deposit consisting mostly of immunoglobulins and antiglobulins, including rheumatoid factors.

Conditions in which large amounts of antibodies and/or immune complexes are present either in circulation or in tissue deposits are often associated with production of rheumatoid factors. Polyclonal B-cell activation or B-cell hyperreactivity associated with autoimmunity or resulting from persistent infection could also lead to synthesis of antiglobulins.

Several lines of evidence indicate that antibodies to immunoglobulins may be incorporated among the immunoglobulins deposited in the kidneys of some patients with cryoglobulinemia, SLE, idiopathic membra-

nous glomerulonephritis, and poststreptococcal glomerulonephritis (Koffler and Cochrane, 1973; McIntosh *et al.*, 1975; Rossen *et al.*, 1977). However, an active role of antiglobulins in the pathogenesis of tissue damage has not yet been determined. Although antiglobulins and rheumatoid factors might be present only as an epiphenomenon, it is tempting to speculate that they could play an active role in the immunopathogenesis of some glomerulonephritides. This hypothesis awaits confirmation.

Antiglobulins and/or rheumatoid factors in kidney immune deposits may interfere with direct immunofluorescence testing for specific antigens (Maggiore *et al.*, 1981). False-positive tests for specific antigens resulting from such interference could lead to faulty interpretation of the pathogenesis of glomerular lesions. It has been proposed recently that some cases of glomerulonephritis attributed originally to immune complexes containing hepatitis B antigens were actually diseases in which antiglobulins were involved.

2. Tubulointerstitial Nephritides

a. Animal Models

Chronic Serum Sickness. As we have mentioned earlier (Section II,B,1), experimental chronic serum sickness, induced by daily injection of heterologous serum proteins, is the animal model from which most has been learned about the potential of circulating immune complexes to damage the kidney tubules and interstitium as well as glomeruli. In animals with chronic serum sickness, immune complex deposits are found disseminated throughout the body (Brentjens *et al.*, 1975; Arisz *et al.*, 1979; Noble *et al.*, 1981b). Extraglomerular kidney deposits are distributed in the walls of peritubular capillaries, in the interstitium, along Bowman's capsule, and along the TBM. The TBM deposits are not restricted to any particular tubule segment, and may be found in the medulla as well as the cortex. Deposition of immune complexes in the interstitium could be mediated and/or facilitated by the receptors for the Fc fragment of IgG that are normally present. In severe chronic serum sickness, cells of the tubule epithelium are frequently damaged and may become atrophic; the TBM is often thickened and split. The accumulation of immune deposits in the interstitium is often accompanied by mononuclear cell infiltration and interstitial fibrosis. A similar distribution of extraglomerular immune deposits has also been detected in mice with spontaneous autoimmune lupuslike syndrome. Interstitial and tubular deposits in those mice increase in frequency with age (Andrews *et al.*, 1978). Mononuclear cell infiltration of the kidney interstitium of NZBxW mice reaches a maximum

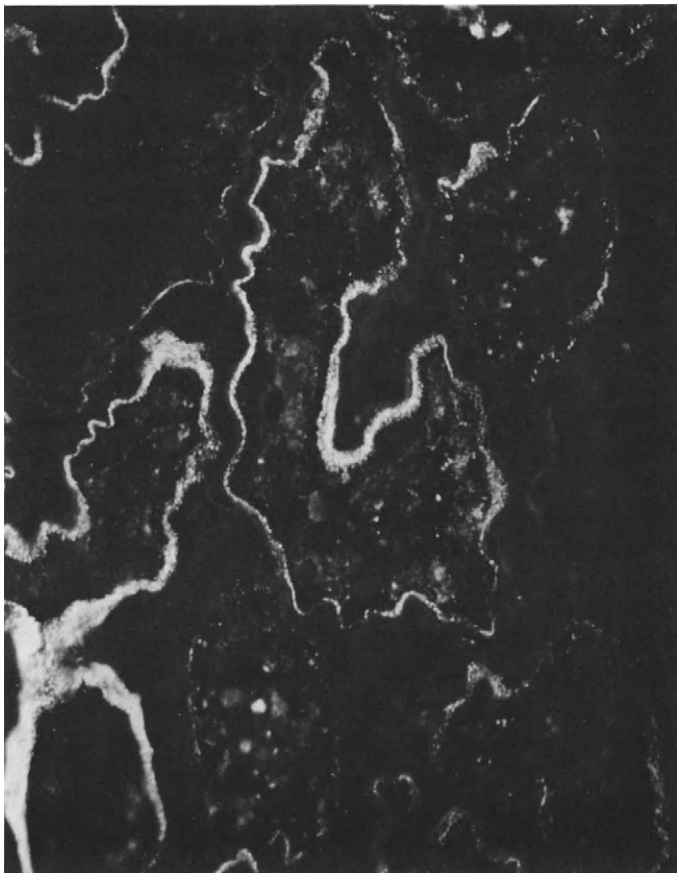


FIG. 5. Direct immunofluorescence test for IgG on a section from the kidney of a patient with tubulointerstitial lupus nephritis. A heavy accumulation of IgG in a granular pattern is seen along the basement membrane of tubules. The interstitium is free of those deposits. Some tubule cells contain droplets of IgG (300×).

at 9 months of age and is associated with an increased rate of tubule cell proliferation (Hurd and Ziff, 1978). However, no correlation of histopathology with the immunological findings has been made in that model.

b. Human Diseases

Systemic Lupus Erythematosus. Tubulointersitial nephritis associated with extraglomerular immune complexes is found in a high percentage of patients with SLE. Tubulointerstitial involvement occurs most frequently in patients with severe or diffuse proliferative lupus glomerulonephritis. However, a few patients have been reported in which tubu-

lointerstitial nephritis was the predominant lesion, associated with only minimal glomerular pathology (Cunningham *et al.*, 1978; Torn *et al.*, 1979). Granular-to-ribbonlike immune deposits may be present at many sites including the TBM of all tubule segments, the walls of peritubular capillaries and larger vessels, and the interstitium. In a few cases these deposits have been shown to contain DNA–anti-DNA complexes (Brentjens and Andres, 1982). As in chronic serum sickness, the deposits in lupus nephritis are not confined to one particular tubule segment, but occur at the base of tubules in the cortex as well as medulla. There is a fair correlation between the severity of tubulointerstitial histopathology in SLE and the presence of immune deposits in the interstitium. This observation supports the contention that immune complexes play a pathogenetic role.

An even better relationship has been demonstrated between the presence of the membrane attack complex (composed of C5b, C6, C7, C8, and C9) along the TBM and tubulointerstitial lesions (Biesecker *et al.*, 1981). It has been proposed that the membrane attack complex, generated by immune complexes and possibly also by nonimmune mechanisms, may function as a direct mediator of tissue injury in lupus nephritis. From a practical point of view it is important to note that, because they are extremely rare in other nephropathies, the detection of extraglomerular immune deposits in a kidney biopsy are so characteristic of SLE that their demonstration is of diagnostic significance (Brentjens and Andres, 1982) (Fig. 5).

III. CELL-MEDIATED AUTOIMMUNE INJURY OF THE KIDNEY

A. Glomerular Injury

Lymphocyte proliferation, as stimulated by kidney antigens, has been measured *in vitro* using peripheral blood lymphocytes from some patients with glomerulonephritis (Bendix, 1968). Specific production of migration inhibition factor has also been demonstrated in patients with anti-GBM glomerulonephritis (Rocklin *et al.*, 1970). Demonstration *in vitro* of specific T-cell reactivity to kidney antigens depends on use of the appropriate antigen(s) (Fillit *et al.*, 1978). Lymphocytes of some patients were found to proliferate when exposed to glycosidase-treated GBM, but not to native GBM. The importance of mechanisms of delayed hypersensitivity for the cell proliferation seen in many autoimmune diseases of glomeruli is not known.

B. Tubulointerstitial Injury

The prototype of tubulointerstitial injury mediated predominantly by cellular immunity is first-set kidney allograft rejection. Aside from allograft rejection, we have very little information about the contributions of T-cell effector mechanisms to lesions in the kidney tubules and interstitium. It has been demonstrated that under appropriate (if highly artificial) conditions inflammatory responses analogous to delayed hypersensitivity skin reactions can be elicited in the kidney (Van Zwieten *et al.*, 1977). Furthermore, in that experimental model, lymph node cells and not serum transfer specific sensitivity to appropriate donors.

Specific T-cell responses to kidney antigens, including blast transformation and macrophage migration inhibition, have been measured *in vitro* with lymphocytes taken from rats with autoimmune kidney disease produced by immunization with kidney extracts (Grupe, 1968; Litwin *et al.*, 1971). In other experiments with the same model, transfer of lymphocytes from immunized donors to normal recipients produced focal mononuclear cell infiltration of the kidney cortex of the recipient (Sugisaki *et al.*, 1980).

The mononuclear cell composition of the interstitial infiltration found in many examples of human and experimental interstitial nephritis is consistent with the possibility that specific cell-mediated responses contribute significantly to the pathogenesis of the lesions. However, most of the available data reviewed here support the view that humoral factors are of primary pathogenetic importance in the well-studied human and animal diseases. A good laboratory model of cell-mediated kidney damage is needed for further clarification of this issue.

IV. SUMMARY

Despite intensive study of autoimmune kidney diseases, important questions remain unanswered. Mechanisms of humoral immunity involving autoantibodies have been implicated in the immunopathogenesis of many human glomerulonephritides. However, in other cases of which idiopathic membranous nephropathy is the most notable example, a great deal of careful research has not yet led to an identification of the specificity or specificities of antibodies present in glomeruli. Although animal models suggest that the humoral mechanisms that cause autoimmune glomerular disease appear to be able to produce injury in kidney tubules and interstitium, these mechansism have been implicated in only a small fraction of human tubulointerstitial nephritides. Antibody-mediated tubulointerstitial nephritis in humans may be relativity rare, but it seems likely

that an immune pathogenesis has not yet been identified in some instances in which it is important.

Some aspects of the immunopathogenesis of diseases involving anti-basement-membrane antibodies have been clarified, but the etiology of those autoimmune nephropathies, and of nearly all others, remains unknown. Useful animal models of kidney injury mediated by immediate or delayed hypersensitivity do not exist. That is one reason why the possible contribution of those mechanisms to autoimmune kidney diseases remains an elusive problem.

Another important feature of the pathogenesis of autoimmune kidney disease that remains to be clarified is the relationship of immunologically mediated damage of kidney structures to the compromise of function of glomeruli and tubules. Finally, study of the production of soluble mediators by various immune mechanisms and the evaluation of the role these mediators play in kidney inflammation has only just begun.

REFERENCES

Agnello, V., Koffler, D., and Kunkel, H. G. (1973). *Kidney Int.* **3,** 90–99.

Albini, B., Brentjens, J. R., and Andres, G. A. (1979). *In* "The Immunopathology of the Kidney" (J. Turk, ed.), pp. 62–75. Arnold, London.

Andres, G. A., and McCluskey, R. T. (1975). *Kidney Int.* **7,** 271–289.

Andres, G., Accinni, L., Beiser, S., Christian, C., Cinotti, G., Erlanger, B., Hsu, K., and Seegal, B. (1970). *J. Clin. Invest.* **49,** 2106–2118.

Andres, G. A., Szymanski, C., Albini, B., Brentjens, J. R., Milgrom, M., Noble, B., Ossi, E., and Steblay, R. (1979). *Am. J. Pathol.* **96,** 21–29.

Andrews, B. S., Eisenberg, R. A., Theofilopoulos, A. N., Izui, S., Wilson, C. B., McConahey, P. J., Murphy, E. D., Roths, J. B., and Dixon, F. J. (1978). *J. Exp. Med.* **148,** 1198–1215.

Arisz, L., Noble, B., Milgrom, M., Brentjens, J. R., and Andres, G. A. (1979). *Int. Arch. Allergy Appl. Immunol.* **60,** 80–88.

Baldwin, D. S., Levine, B. B., and McCluskey, R. T. (1968). *N. Engl. J. Med.* **279,** 1245–1252.

Bendix, G. (1968). *Acta Med. Scand.* **184,** 99–103.

Biesecker, G., Katz, S., and Koffler, D. (1981). *J. Exp. Med.* **154,** 1779–1794.

Bigazzi, P. L., Kosuda, L., Hsu, K. C., and Andres, G. A. (1976). *J. Exp. Med.* **143,** 382–404.

Bolton, W. K., Benton, F. R., and Sturgill, B. C. (1978). *Clin. Exp. Immunol.* **33,** 463–473.

Border, W., Lehman, D., Egan, J., Sass, H., Glode, J., and Wilson, C. (1974). *N. Engl. J. Med.* **291,** 381–384.

Brentjens, J. R., and Andres, G. A. (1982). *Arthritis Rheum.* **25,** 880–886.

Brentjens, J. R., O'Connell, D., Albini, B., and Andres, G. A. (1975). *Ann. N. Y. Acad. Sci.* **254,** 603–615.

Brown, C. A., Carey, K., and Colvin, R. B. (1979). *J. Immunol.* **123,** 2102–2107.

Couser, W. G., Stilmant, M., and Lewis, E. J. (1973). *Lab. Invest.* **29,** 236–243.

Couser, W. G., Steinmuller, D. R., Stilmant, M. M., Salant, D. J., and Lowenstein, L. M. (1978). *J. Clin. Invest.* **62,** 1275–1287.
Cunningham, E., Provost, T., Brentjens, J., Reichlin, M., and Venuto, R. C. (1978). *Arch. Int. Med.* **138,** 1560–1561.
Douglas, M. F. S., Rabideau, D., Schwartz, M. M., and Lewis, E. J. (1981). *N. Engl. J. Med.* **305,** 1326–1329.
Druet, P., Druet, E., Potdevin, F., and Sapin, C. (1978). *Ann. Immunol.* **129C,** 777.
Edgington, T. S., Glassock, R. J., and Dixon, F. J. (1967). *Science* **152,** 1432–1434.
Edgington, T. S., Glassock, R. J., and Dixon, F. J. (1968). *J. Exp. Med.* **127,** 555–572.
Fillit, H., Read, S. E., Sherman, R. L., Zabriskie, J. B., and van den Rÿn, I. (1978). *N. Engl. J. Med.* **298,** 861–868.
Fish, A. J., Carmody, K. M., and Michael, A. F. (1979). *J. Lab. Clin. Med.* **94,** 447–457.
Galpin, J. E., Shinaberger, J. H., Stanley, T. M., Blumenkrantz, M. J., Bayer, A. S., Friedman, G. S., Montgomerie, J. Z., Guze, L. B., Coburn, J. W., and Glassock, R. J. (1978). *Am. J. Med.* **65,** 756–765.
Glassock, R. J., and Cohen, A. H. (1981). *In* "The Kidney" (B. M. Brenner and F. C. Rector, Jr., eds.), pp. 1493–1570. Saunders, Philadelphia, Pennsylvania.
Glassock, R. J., Edgington, T. S., Watson, J., and Dixon, F. J. (1968). *J. Exp. Med.* **127,** 573–587.
Grupe, W. E. (1968). *Proc. Soc. Exp. Biol. Med.* **127,** 1217–1222.
Grupe, W. E., and Kaplan, M. (1969). *J. Lab. Clin. Med.* **74,** 400–409.
Grussendorf, M., Andrassy, K., Waldherr, R., and Ritz, E. (1981). *Am. J. Nephrol.* **1,** 105–111.
Hall, C. H., Colvin, R. B., Carey, K., and McCluskey, R. (1977). *J. Exp. Med.* **146,** 1246–1260.
Hart, D. N. J., and Fabre, J. W. (1980). *Clin. Exp. Immunol.* **40,** 111–119.
Heymann, W., Hackel, D. B., Harwood, S., Wilson, S. G., and Hunter, J. L. (1959). *Proc. Soc. Exp. Biol. Med.* **100,** 660–664.
Hoyer, J. R. (1980). *Kidney Int.* **17,** 284–292.
Hurd, E. R., and Ziff, M. (1978). *Clin. Exp. Immunol.* **32,** 1–11.
Izui, S., Lambert, P.-H., and Miescher, P. A. (1977). *Clin. Exp. Immunol.* **30,** 384–392.
Jennings, L., Roholt, O. A., Pressman, D., Blau, M., Andres, G. A., and Brentjens, J. R. (1981). *J. Immunol.* **127,** 129–134.
Jordan, S. C., Johnston, W. H., and Bergstein, J. M. (1978). *Arch. Pathol. Lab. Med.* **102,** 530–533.
Jordan, S., Buckingham, B., Sakai, R., and Olson, D. (1981). *N. Engl. J. Med.* **304,** 1212–1215.
Kelley, V. E., Winkelstein, A., Izui, S., and Dixon, F. J. (1981). Clin. *Immunol. Immunopath.* **21,** 190–203.
Kerjaschki, D., and Farquhar, M. G. (1982). *Proc. Natl. Acad. Sci. U.S.A.* **79,** 5557–5561.
Kerjaschki, D., and Farquhar, M. G. (1983). *J. Exp. Med.* **157,** 667–686.
Klassen, J., McCluskey, R. T., and Milgrom, F. (1971). *Am. J. Pathol.* **63,** 333–350.
Klassen, J., Milgrom, F., and McCluskey, R. T. (1977). *Am. J. Pathol.* **88,** 135–141.
Koffler, D., and Cochrane, C. G. (1973). *Adv. Immunol.* **16,** 185–264.
Lambert, P.-H., and Dixon, F. J. (1968). *J. Exp. Med.* **127,** 507–522.
Lehman, D. H., Marquardt, H., Wilson, C. B., and Dixon, F. J. (1974a). *J. Immunol.* **112,** 241–248.
Lehman, D. H., Wilson, C. B., and Dixon, F. J. (1974b). *Kidney Int.* **5,** 187–195.
Lehman, D. H., Wilson, C. B., and Dixon, F. J. (1975). *Am. J. Med.* **58,** 765–786.
Lerner, R. A., and Dixon, F. J. (1966). *J. Exp. Med.* **124,** 431–442.

Lerner, R., Glassock, R. J., and Dixon, F. J. (1967). *J. Exp. Med.* **126,** 989–1004.
Levy, J. A., Ibrahim, A. B., Shirai, T., Ohta, K., Nagasawa, R., Yoshida, H., Estes, J., and Gardner, M. (1982). *Proc. Natl. Acad. Sci. U.S.A.* **79,** 1974–1978.
Levy, M., Gagnadoux, M. F., Beziau, A., and Habib, R. (1978). *Clin. Nephrol.* **10,** 158–165.
Linton, A. L., Clark, W. F., Driedger, A. A., Turnbull, I., and Lindsay, R. M. (1980). *Ann. Int. Med.* **93,** 735–741.
Litwin, A., Adams, L. E., Levy, R., Cline, S., and Hess, E. B. (1971). *Immunology* **20,** 755–766.
McIntosh, R. M., Griswold, W. R., Chernack, W. B., Williams, G., Strauss, J., Kaufman, D. B., Koss, M. N., McIntosh, J. R., Cohen, R., and Veill, R., III (1975). *Q. J. Med.* **174,** 285–307.
Maggiore, Q., Bartolomeo, F., L'Abbate, A., and Misefari, V. (1981). *Kidney Int.* **19,** 579–586.
Makker, S. (1980). *Am. J. Med.* **69,** 949–952.
Makker, S. P., and Moorthy, D. (1981). *Lab. Invest.* **44,** 1–5.
Mendrick, D., Noble, B., Brentjens, J., and Andres, G. A. (1980). *Kidney Int.* **18,** 328–343.
Miettinen, A., Tornroth, T., Tikkanen, I., Virtanen, I., and Linder, E. (1980). *Lab. Invest.* **43,** 547–555.
Morel-Maroger, L., Kourilsky, O., Mignon, F., and Richet, G. (1974). *Clin. Immunol. Immunopath.* **2,** 185–194.
Naruse, I., Kitamura, K., Miyakawa, Y., and Shibata, S. (1973). *J. Immunol.* **110,** 1163–1166.
Neale, T. J., and Wilson, C. B. (1982). *J. Immunol.* **128,** 323–330.
Neilson, E. G., and Phillips, S. M. (1979). *J. Immunol.* **123,** 2373–2378.
Noble, B., Olson, K., Milgrom, M., and Albini, B. (1980). *Clin. Exp. Immunol.* **42,** 255–262.
Noble, B., Mendrick, D. L., Brentjens, J. R., and Andres, G. A. (1981a). *Clin. Immunol. Immunopath.* **19,** 289–301.
Noble, B., Milgrom, M., Van Liew, J. B., and Brentjens, J. (1981b). *Clin. Exp. Immunol.* **46,** 499–507.
Noble, B., Brentjens, J. R., and Andres, G. A. (1982a). *In* "Immunopathology: VIIIth International Symposium" (F. J. Dixon and P. A. Miescher, eds.), pp. 305–330. Academic Press, New York.
Noble, B., Van Liew, J. B., Brentjens, J. R., and Andres, G. A. (1982b). *Lab. Invest.* **47,** 427–436.
Noble, B., Van Liew, J., Brentjens, J., and Andres, G. A. (1984). *Clin. Immunol. Immunopath.* **30,** 241–254.
Ooi, B. S., Jao, W., First, M. R., Mancilla, R., and Pollak, V. E. (1975). *Am. J. Med.* **59,** 614–629.
Ooi, B. S., Ooi, Y. M., Mohini, R., and Pollak, V. E. (1978). *Clin. Immunol. Immunopath.* **10,** 330–334.
O'Regan, S., Fong, J. S. C., Kaplan, B. S., de Chadarevian, J. P., Lapointe, N., and Drummond, K. N. (1976). *Clin. Immunol. Immunopath.* **6,** 341–346.
Ploth, D. W., Fitz, A., Snetzler, D., Seidenfeld, J., and Wilson, C. B. (1978). *Clin. Immunol. Immunopath.* **9,** 327–334.
Robertson, J. L., Hill, G. S., and Rowlands, D. T., Jr. (1977). *Am. J. Pathol.* **88,** 53–68.
Rocklin, R. E., Lewis, E. J., and David, J. R. (1970). *N. Engl. J. Med.* **283,** 497–501.
Roman-Franco, A. A., Turiello, M., Albini, B., Ossi, E., and Andres, G. (1976). *Kidney Int.* **10,** 549.
Roman-Franco, A. A., Turiello, M., Albini, B., Ossi, E., Milgrom, F., and Andres, G. A. (1978). *Clin. Immunol. Immunopath.* **9,** 464–481.

Rossen, R. D., Rickaway, R. H., Reisberg, M. A., Singer, D. B., Schloeder, F. X., Suki, W. N., Hill, L. L., and Eknoyan, G. (1977). *Arthritis Rheum.* **20,** 947–961.
Roubinian, J. R., Papoian, R., and Talal, N. (1977). *J. Clin. Invest.* **59,** 1066–1070.
Rudofsky, U. H. (1980). *Clin. Immunol. Immunopath.* **15,** 200–212.
Rudofsky, U., and Pollara, B. (1975). *Clin. Immunol. Immunopath.* **4,** 425–439.
Rudofsky, U. H., and Pollara, B. (1977). *Clin. Exp. Immunol.* **27,** 522–525.
Rudofsky, U., and Steblay, R. W. (1968). *Nature (London)* **218,** 1269–1270.
Rudofsky, U. H., McMaster, P. R., Ma, W.-S., Steblay, R. W., and Pollara, B. (1974). *J. Immunol.* **112,** 1387–1393.
Rudofsky, U., Steblay, R. W., and Pollara, B. (1975). *Clin. Immunol. Immunopath.* **3,** 396–407.
Rudofsky, U. H., Dilwith, R. L., and Tung, K. S. (1980). *Lab. Invest.* **43,** 463–470.
Salant, D., Belok, S., Medaio, M., and Couser, W. G. (1980). *J. Clin. Invest.* **66,** 1339–1350.
Sapin, C., Druet, E., and Druet, P. (1977). *Clin. Exp. Immunol.* **28,** 173–179.
Sapin, C., Mandet, C., Druet, E., Gunster, E., and Druet, P. (1982). *Clin. Exp. Immunol.* **48,** 700–704.
Schneeberger, E. E., and Grupe, W. E. (1976). *Lab. Invest.* **34,** 298–305.
Seegal, B. C., Accinni, L., Andres, G. A., Beiser, S. M., Christian, C. L., Erlanger, B. F., and Hsu, K. C. (1969). *J. Exp. Med.* **130,** 203–216.
Steblay, R. W. (1962). *J. Exp. Med.* **116,** 253–272.
Steblay, R. W. (1964). *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **23,** 449. (Abstr.)
Steblay, R. W., and Rudofsky, U. (1968). *Science* **160,** 204–206.
Steblay, R., and Rudofsky, U. H. (1971). *J. Immunol.* **107,** 589–594.
Steblay, R., and Rudofsky, U. H. (1973). *Science* **130,** 966–968.
Stenglein, B., Thoenes, G., and Gunther, E. (1978). *Clin. Exp. Immunol.* **33,** 88–94.
Steward, M. (1979). *Clin. Exp. Immunol.* **38,** 414–423.
Sugisaki, T., Klassen, J., Milgrom, F., Andres, G. A., and McCluskey, R. T. (1973). *Lab. Invest.* **28,** 658–671.
Sugisaki, T., Yoshida, T., McCluskey, R. T., Andres, G. A., and Klassen, J. (1980). *Clin. Immunol. Immunopath.* **15,** 33–40.
Swaak, A. J. G., Aarden, L. A., Statius van Eps, T., and Feltkamp, T. E. W. (1979). *Arthritis Rheum.* **22,** 226–235.
Theofilopoulos, A. N., McConahey, P. J., Izui, S., Eisenberg, R. A., Pereira, A. B., and Creighton, W. D. (1980). *Clin. Immunol. Immunopath.* **15,** 258–278.
Torn, F., Canerval, D., and Droz, D. (1979). *Am. J. Med.* **67,** 529–532.
Tung, K., and Black, W. (1975). *Lab. Invest.* **32,** 696–700.
Van Damme, B. J. C., Fleuren, G. J., Bakker, W. W., Vernier, R. L., and Hoedemaeker, Ph. J. (1978). *Lab. Invest.* **38,** 502–509.
Van Liew, J. B., Brentjens, J. R., and Noble, B. (1983). *Kidney Int.* **24,** 160–169.
Van Zwieten, M., Leber, P. D., Bhan, A. K., and McCluskey, R. T. (1977). *J. Immunol.* **118,** 589–593.
Wakashin, Y., Takei, I., Ueda, S., Mori, Y., Iesato, K., Wakashin, M., and Okuda, K. (1981). *Clin. Immunol. Immunopath.* **19,** 360–371.
Weigle, W. O., and Nakamura, R. M. (1969). *Clin. Exp. Immunol.* **4,** 645–657.
Wilson, C. B. (1981). *Immunol. Rev.* **55,** 257–297.
Wilson, C. B., and Dixon, F. J. (1973). *Kidney Int.* **3,** 74–89.
Wilson, C. B., and Dixon, F. J. (1981). *In* "The Kidney" (B. M. Brenner and F. C. Rector, Jr., eds.), pp. 1237–1350. Saunders, Philadelphia, Pennsylvania.

CHAPTER 14

Ocular Diseases of Presumed Autoimmune Origin

ROBERT B. NUSSENBLATT
Clinical Ophthalmic Immunology Section
National Eye Institute
National Institutes of Health
Bethesda, Maryland

ARTHUR M. SILVERSTEIN
The Wilmer Ophthalmological Institute
The Johns Hopkins University School of Medicine
Baltimore, Maryland

I. AUTOIMMUNE DISEASES OF THE EYE: INTRODUCTION

The growing awareness of the contributions of immunopathological processes to human disease has been paralleled by the finding that the same processes may contribute to blinding eye disease. The ophthalmologist and experimental ocular pathologist have paid increasingly greater attention in recent years to, among the immunological mechanisms in-

volved, autoimmune reactions, which now appear to contribute significantly to a growing number of ophthalmic problems. This is particularly noteworthy, since historically one of the first demonstrations of an autoimmune response was made by Uhlenhuth (1903), using lens antigen. Thus, by virtue of antigens restricted to this organ, the eye makes several unique additions, such as sympathetic ophthalmia and phacoantigenic uveitis, to the list of autoimmune diseases. But the eye also shares a variety of antigens with the rest of the body, so that it may be involved to a greater or lesser degree in a number of systemic autoimmune processes. Because of this significant overlap of interest, it is important that both the internist and ophthalmologist maintain contact with the conceptual and practical developments in the other's field.

II. GENERAL IMMUNOPATHOLOGY OF THE EYE

Before considering the specific features of ocular autoimmune diseases, the general reader ought to be aware of some special characteristics of the eye that may influence the course of immunological reactions. These include such features as the anatomical isolation of certain ocular antigens, the lack of lymphatic drainage of the tissues inside the globe, and the avascularity of several structures within the optical pathway. Each of these features can exert a profound effect on the normal workings of the afferent and efferent arcs of the immune response.

The eye is, in many respects, a microcosm in which all types of immunological interaction and all forms of immunopathological disease may be found, in either clinical or experimental situations. Thus, the conjunctiva is a favorite site for IgE-mediated allergic reactions; allograft rejection may be seen in corneal transplants; delayed hypersensitivity-type inflammation is a frequent accompaniment of intraocular infections; some of the earliest studies of immune complex disease were performed in the rabbit cornea; and the uveal tract is able to support exuberant antibody formation. The eye has proved to be especially useful in both clinical and experimental studies of these reactions because the well-defined ocular layers and relatively small number of cellular constituents, coupled with the normal transparency of the conjunctival epithelium, the cornea, and the lens, render these processes readily visible with the ophthalmoscope.

The normal course of immunological interactions may be profoundly modified by a number of factors peculiar to the eye. Among these, the most immediately obvious is the degree of immunological privilege enjoyed by several ocular tissues. Thus, tissue allografts, which would induce sensitization and invite rejection anywhere else in the body, may be

practiced in the cornea with relative freedom from the rejection process (Khodadoust and Silverstein, 1972; Porter and Knight, 1973). The normally avascular cornea apparently does not participate in the usual lymphocyte surveillance traffic of the body, so that corneal allografts are more or less spared due to impediments to host sensitization as well as to inhibition of effector mechanisms.

The anterior chamber of the eye also enjoys a degree of immunological privilege, depending in part on the existence of a blood-aqueous barrier analogous to the blood–brain barrier. In addition, there is no lymphatic drainage of the uveal tract and retina, so that immunogens in these locations can only stimulate the host systemically via the blood. These factors contribute to the lengthened survival of allogeneic transplants of tumors and certain endocrine tissues that would quickly be rejected at other sites. In addition, these anatomical factors appear to contribute to an immune deviation of antigens introduced into the anterior chamber, so that the formation of suppressor T cells is favored over the more normal forms of active immune responses (Kaplan and Streilein, 1978).

Finally, the tissues of the eye and especially the uveal tract can support local antibody formation to such an extent that in certain ocular infections the titer of circulating antibody found in the aqueous may exceed that in the blood (Witmer, 1964). Unlike other sites of ectopic antibody formation, the uveal tract has the almost unique ability to retain for long periods the immunological memory of an earlier experience with antigen (Segal *et al.*, 1933), so that rechallenge with the same antigen causes recurrence of the earlier uveitis. This booster antibody response is similar to that seen in regional lymph nodes and has been implicated in the pathogenesis of recurrent anterior uveitis (Silverstein, 1974).

III. AUTOIMMUNE DISEASES, PROVED OR PRESUMED

A. Phacoantigenic Uveitis

1. Clinical Manifestations

Intraocular inflammatory disease induced by lens protein has been recognized since the early part of this century by Straub (reported by van der Hoeve, 1920). This condition occurs only after disruption of the lens capsule, whether by surgery, by trauma, or spontaneously. The disease can take the form of a massive ocular inflammatory response that is devastating to the eye. Some of the features of this disease that can

compromise vision are a clouding of the ocular media, glaucoma, and adhesions (synechiae) to the lens of the iris and inflammatory debris (Fig. 1).

Irvine and Irvine (1952) made the distinction between phacoantigenic uveitis (earlier called endophthalmitis phacoanaphylactica) and phacogenic glaucoma (earlier called phacotoxic uveitis). In the former, the histological response is predominantly that of neutrophils invading the lens substance; the invasion is frequently surrounded by a zone of giant and epithelioid cells and then by lymphocytes and plasma cells. The phacogenic form of the disease may occur in eyes with a mature cataract that is presumably leaking protein. The inflammation may be moderate, and histologically a nongranulomatous picture is evident, with occasional plasma cells and large eosinophilic macrophages. Rupture of the lens in most patients does not usually yield such an exuberant response (Yanoff and Scheie, 1968), suggesting that other factors are involved.

Phacoantigenic uveitis can be a bilateral condition, especially in cases where the second eye has advanced cataractous changes. It is often associated with sympathetic ophthalmia (Blodi, 1959; Easom and Zimmerman, 1964), which is also a bilateral condition. An important differential diagnostic point to help distinguish sympathetic ophthalmia from pha-

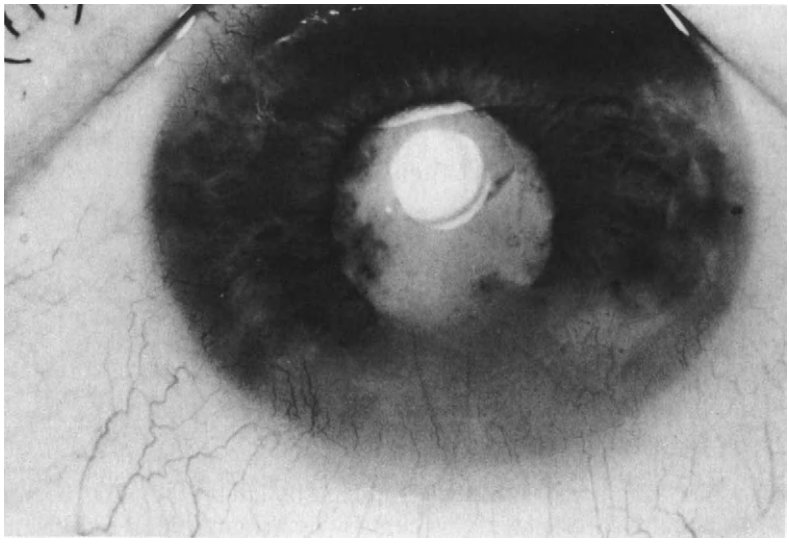


FIG. 1. Intraocular inflammation with corneal disease induced by phacoanaphylaxis. (Courtesy of Dr. D. Cogan.)

coantigenic uveitis was proposed by de Veer (1953). In sympathetic ophthalmia, both eyes are intensely inflamed simultaneously, while in the lens-induced condition, the initial responding eye is usually quiet.

2. Immunological Manifestations

Lens antigens are thought to be isolated from the systemic circulation early in embryonic development, prior to maturation of the immune system. This sequestration is thought to explain the vigorous immune response by the body when the lens capsule is disrupted. The lens has been shown to have antigens that are very strongly organ specific, but it has few species-specific antigens (Halbert *et al.*, 1961); this has been noted using both *in vitro* antibody (Halbert and Manski, 1963) and cell-mediated (Nussenblatt *et al.*, 1979) responses. Although skin reactivity to lens antigens was shown in patients with phacoantigenic uveitis (Verhoeff and Lemoine, 1922), more recent data suggested that a delayed hypersensitivity mechanism is not occurring. Lens antibodies have been reported in 50% of normal individuals (Hackett and Thompson, 1964), and Sandburg and Closs (1979a) found similar responses, but only to α crystallins. Sandburg and Closs (1979b) has also shown the presence of lens crystallin in the aqueous of patients with cataracts. The level of circulating antibody has been noted to increase after extracapsular lens extraction (Wirostko and Spalter, 1967).

In a series of articles, Marak and co-workers induced a disease in rats similar to the disease seen in humans (Marak *et al.*, 1974,1976a,b,1977). The intraocular inflammation is induced by hyperimmunization and subsequent disruption of the lens capsule. The histological picture differs somewhat depending on whether homologous or heterologous lens antigen is employed. The lesion can be passively transferred using immune serum, and the inflammatory response can be markedly attenuated by treating animals with cobra venom. Furthermore, the injured lenses will show the deposition of immunoglobulin and C3. Gery and colleagues (1981) showed that homologous, but not heterologous, lens immunization induced high titers of specific antibody but essentially no cellular immune response. Goldschmidt *et al.* (1982) recently found no correlation between disease activity and the level of circulating antibody to α lens crystallin in their animal model. The accrued evidence would strongly suggest a noncellular mechanism for this disease.

Marak and colleagues (1979) speculated that T-cell tolerance is normally maintained by low amounts of circulating lens antigen. This normal state of tolerance might be altered by trauma, with resultant disease. It is

also possible that lens injury may create an adjuvant effect, thus circumventing T-cell control and resulting in the activation of the B-cell arm of the immune system.

3. Therapeutic Modalities

Severe lens-related inflammatory reactions may be initially treated with both systemic and topical corticosteroids. However, the treatment of choice for prophylaxis or for an ongoing inflammation is removal of the inciting agent, whether that be lens fragments or a mature, leaking, cataractous lens.

B. Sympathetic Ophthalmia

1. Clinical Manifestations

Sympathetic ophthalmia is, fortunately, a rare disease, but it may progress almost inexorably to bilateral blindness when it does occur. It typically follows a perforating injury of the globe in which the uveal tract is involved but can occur after ocular surgery, such as retinal detachment repair and vitrectomy. After a lag of some weeks, the injured eye may develop what appears to be a typical granulomatous endophthalmitis. Mutton-fat keratic precipitates form on the posterior corneal surface, and the lesion is one of diffuse and often massive lymphocytic infiltration of uveal (predominantly choroidal) tissues, with areas of epithelioid and giant cell infiltration (Fig. 2). At any time from about the third week to several years following the trauma or surgery, a spontaneous lesion may appear in the "sympathizing" eye that is similar in all respects to that of the first eye. While sympathetic ophthalmia is primarily an inflammatory disease of the uveal tract, intraretinal inflammation may occur in up to 42% of the cases, and plasma cells are also present (Lubin *et al.*, 1980).

2. Immunological Manifestations

An allergic theory of the pathogenesis of sympathetic ophthalmia was advanced as early as 1910 by Elschnig, who demonstrated that uveal emulsions were autoantigenic and suggested that uveal pigment was the offending antigen. The postulate gained important support from the investigation of Woods (1921, 1956), who was able to identify antiuveal antibodies in patients with perforating injuries of the globe. Woods also showed that extracts of uveal tissue would elicit a positive intradermal

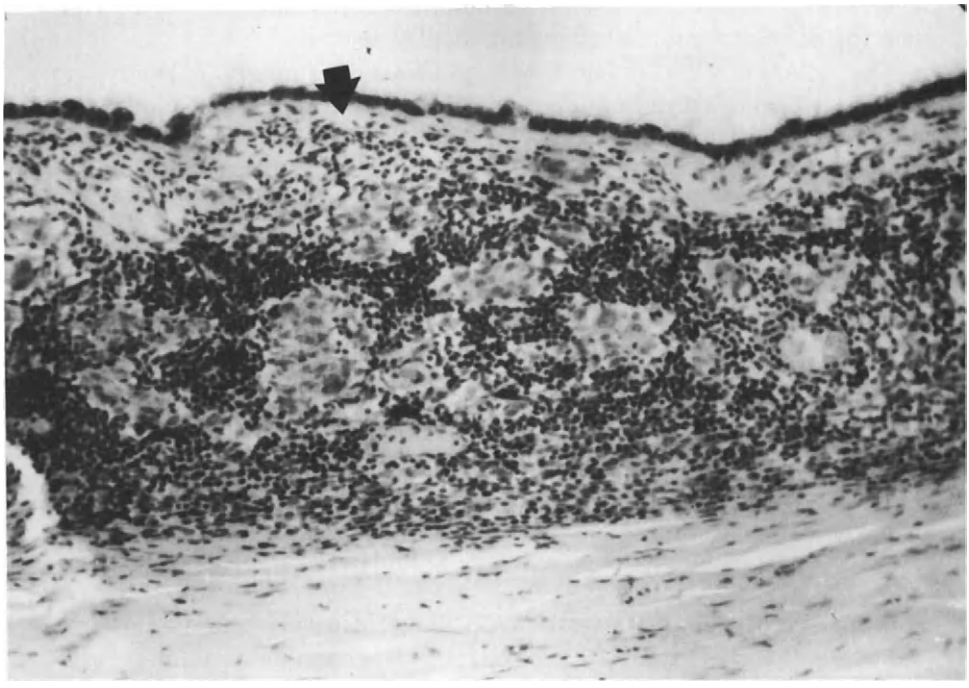


FIG. 2. Photomicrograph showing the typical choroidal infiltration in sympathetic ophthalmia. The choroid is filled with large numbers of lymphocytes interspersed with epithelioid cells. The choriocapillaris (arrow) is characteristically spared. (Courtesy of Dr. D. Cogan.)

hypersensitivity test in these patients, and Friedenwald (1934) showed that the history of the skin test response was strikingly similar to that of the ocular lesion.

Recently, Jakobiec and colleagues (1983) reported that the choroidal infiltrate was composed mainly of OKT8+ T lymphocytes, and that the histocytes and epithelioid cells were Ia+ and OKM1+. Further, the Dalen–Fuchs nodules were composed of OKM1+, Ia+ histiocytes and Ia−, OKMI− retinal pigmented epithelial cells.

The earliest attempts to develop an experimental model for sympathetic ophthalmia were those of Collins (1949, 1953). Collins injected emulsions of crude uveal tissue in Freund's adjuvant mixture into guinea pigs and observed the development of diffuse, predominantly nongranulomatous inflammatory reactions in the choroid of many of these animals. Further clinical and laboratory studies by Wacker and Lipton (1965) and Henley

et al. (1972) made it increasingly certain that retinal antigens might stimulate the development of autoimmune ocular disease.

The isolation of the retinal S antigen (Wacker *et al.*, 1977; Dorey and Faure, 1977) has permitted the establishment of a reproducible model of autoimmune inflammation of the eye. This organ-specific glycoprotein has a molecular weight of ~55,000 (Wacker *et al.*, 1977). It routinely causes bilateral uveitis in lower mammals. The disease induced is immune mediated (Nussenblatt *et al.*, 1980a), and varying doses of the antigen injected at a site far from the eye will induce a spectrum of disease ranging from mild nongranulomatous inflammation to severe granulomatous disease similar to that seen in the human (Wacker *et al.*, 1977; Rao *et al.*, 1979).

Autoimmune uveitis has also been induced in nonhuman primates with S-antigen immunization, accompanied by clinical, fluorescein angiographic, and histological alterations very suggestive of the human condition (Nussenblatt *et al.*, 1981a,b; Faure and de Kozak, 1981) (Fig. 3). Further, suggestive evidence for a role of the S antigen in human disease was the demonstration of *in vitro* cell-mediated responses to this antigen by leukocytes from uveitis patients (Nussenblatt *et al.*, 1980b,1982b; Tanoe *et al.*, 1983).

Using cyclosporin A, the disease could be totally prevented or abrogated when therapy is started 1 week after S-antigen immunization, despite the maintenance of high anti-S-antigen titers in the protected animals (Nussenblatt *et al.*, 1981d,1982a). Because of the known T-cell specificity of cyclosporin A (White, 1982), the results suggested the mandatory T-cell participation in this disease model, and by extension, in some cases of human uveitis.

3. Therapeutic Modalities

Enucleation of the injured or exciting eye before the development of inflammatory disease begins in the sympathizing eye has been the traditional therapy for this disease. Furthermore, Lubin *et al.*, (1980) reported in their review that enucleation even early after the onset of disease in the sympathizing eye may be beneficial. However, newer surgical techniques permit restoration of many injured globes to a reasonable function and appearance, making enucleation a less commonly performed procedure. Systemic corticosteroids have been utilized with reasonable success in this condition, while severe steroid-resistant cases have been successfully treated with cytotoxic agents (Wong, 1969; Dinning and Perkins, 1975.)

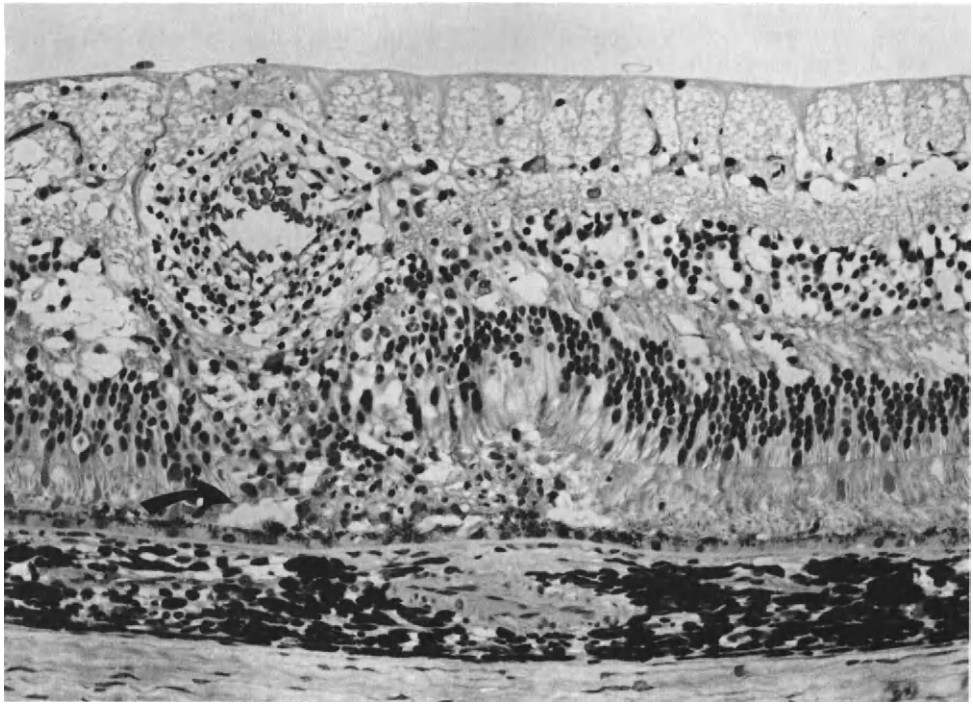


FIG. 3. Photomicrograph of the retina and choroid of a monkey with retinal S-antigen-induced uveitis. In addition to the retinal periphlebitis, a focal dropout of the photoreceptor region is seen (arrow), the anatomical site of the S-antigen. (Reproduced, with permission; *Archives of Ophthalmology* **99**, p. 1090–1092, 1981.)

C. Ocular Cicatricial Pemphigoid

1. Clinical Manifestations

Cicatricial pemphigoid, or benign mucosal pemphigoid, is a chronic nonrelenting disease of the mucosal surfaces and skin. The oral mucosa and conjunctiva are the most common sites of involvement, the former affected in 90% and the latter in 66% of patients with the disease (Lever, 1979). The disease is rare with an incidence of perhaps 1 in 20,000 ophthalmic cases (Duke-Elder, 1965). Although it has no known racial predilection, it is essentially a disease of older age groups, with the average age of onset being 58 years (Hardy *et al.*, 1971). Although more women are affected, this may rather reflect their greater life span as compared to

men. The ocular components of this progressive disease include conjunctival shrinkage and symblepheron formation, leading to trichiasis, xerosis, and blindness due to corneal opacification (Fig. 4). Blindness has been reported in 16% of patients with conjunctival disease (Lever, 1979). This disease should not be confused with bullous pemphigoid, which has similar histological and immunopathological features but rarely if ever affects the conjunctiva (Friedlaender, 1979). Histologically, subepithelial bullae can be seen, but the formation of submucosal fibrosis and its subsequent contracture is probably of prime importance to the ocular character of this disease. Bullae are almost never seen clinically.

This disease may initially affect only one eye and manifest itself in the second after as long as 2 years (Mondino, 1981). Patients will often complain of ocular burning or irritation and be diagnosed as having a chronic conjunctivitis. This condition can often have secondary bacterial infections superimposed on the basic process. Cicatricial pemphigoid is frequently not recognized in its early stage, but becomes evident once symblepheron and inferior fornicial shortening occur (Fig. 4). Ultimately, the fournices may be totally obliterated. The development of entropion and a decrease and destabilization of the tear film are characteristically seen with ongoing submucosal fibrosis. Although the course of the disease generally is slow and chronic, an acute form of the disease has been

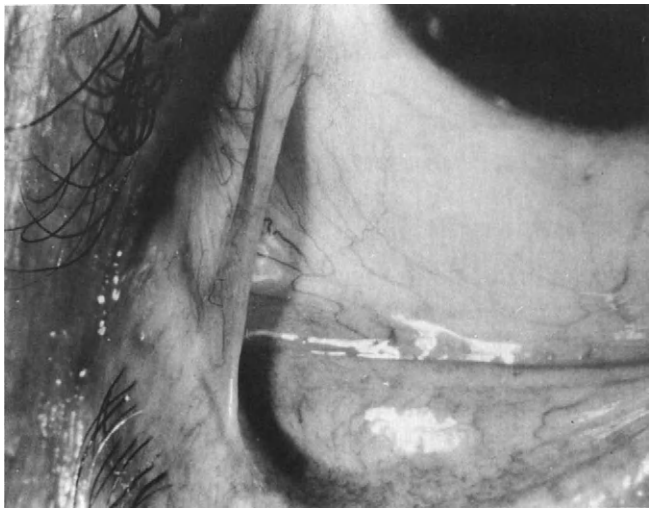


FIG. 4. Ocular changes of cicatricial pemphigoid symblepharon formation and shortening of the fornix are evident.

described by Mondino and colleagues (1979a). In this form, the activity can manifest itself as a localized or diffuse intensely hyperemic area associated with chemosis. This may be in response to local trauma, such as a conjunctival biopsy. This acute stage will lead to a rapid shrinkage and scarring of the conjunctiva. Corneal opacification and impaired vision will ensue because of the abnormal blink reflex, entropion, decreased tear production, exposure, and secondary ocular infections that occur in these patients.

Several reports have now appeared that locally administered medications may induce a clinical entity similar to that of cicatricial pemphigoid. The best documented has been the effect of echothiophate iodide, an antiglaucomatous medication (Patten *et al.*, 1976).

2. Immunological Manifestations

The disease process of cicatricial pemphigoid has been thought to be mediated by a type II hypersensitivity reaction and has been associated with HLA-B12 (Mondino *et al.*, 1979b). The deposition of immunoglobulin in a linear fashion along the basement membrane zone (BMZ) in both skin and mucosa is characteristic of both bullous and cicatricial pemphigoid (Griffith *et al.*, 1974; Lever, 1979). Using direct immunofluorescence techniques, a deposition of immunoglobulin has been demonstrated in the conjunctiva of scarring pemphigoid patients. Furey and colleagues (1975) reported finding IgA or IgM deposited on the BMZ of the conjunctiva of 3 of 10 patients, while Mondino and associates (1977) reported 8 of 12 patients with either IgA, IgG, or IgM deposition. Four of these tissue specimens demonstrated the presence of C3 as well. This was in contrast to the staining pattern seen in adjacent conjunctival tissue not involved in the disease process, where immunoglobulin but not complement could be found. Rogers *et al.* (1977) reported the presence of components of both the classic and alternate complement pathways in oral specimens of pemphigoid they have examined. Circulating antibodies to the BMZ have been reported in cicatricial pemphigoid, but at low titers (Dabelsteen *et al.*, 1974; Furey *et al.*, 1975).

In vitro evidence would suggest that these autoantibodies can mediate pathological processes. Gammon and colleagues (1980) demonstrated that circulating autoantibodies from bullous (but not cicatricial) pemphigoid patients will fix to normal skin, and with the addition of complement will cause peripheral white blood cells to attach along the BMZ. However, the passive transfer of sera from bullous patients does not produce disease, although binding to the BMZ by the antibody could be demonstrated (Sams and Gleich, 1971). However, injections of IgG fractions from bul-

lous pemphigoid patients into the corneal stroma of rabbits will induce histological alterations compatible with the disease (Anhalt *et al.*, 1981). The identification of the antigen provoking the immune response in cicatricial pemphigoid still remains to be determined. In bullous pemphigoid, a component of the BMZ produced by human and mouse epidermal cells having a molecular weight of 220 kdaltons has been identified (Stanley *et al.*, 1981), and this antigen is not laminin (Stanley *et al.*, 1982).

The role of T cells in the ocular component of cicatricial pemphigoid remains to be fully explored. Bhan and colleagues (1982) have shown that the majority of T cells present in the conjunctiva during the active phase of the disease are OKT4+; this monoclonal antibody preparation defines the helper/inducer T-cell subpopulation. Characterization of the T-cell subsets after therapy showed that their distribution had returned to a proportion more like that of normal conjunctiva.

3. Therapeutic Modalities

Although the disease is called "benign" mucous pemphigoid, the ocular complications are extremely serious and potentially sight threatening. Persistent and meticulous lid hygiene is mandatory in caring for patients with this disease. The administration of local antibiotics and wetting agents is also important as is the epilation of ingrowing cilia. However, systemic therapy is necessary to arrest the course of the disease. Systemic corticosteroids have been used with initial success, but long-term results are less impressive. Nonsteroidal immunosuppressive and antiinflammatory agents, particularly cyclophosphamide, are the treatments of choice in this disease. Foster (1980) was able to totally arrest the progression of disease in two pemphigoid patients with cyclophosphamide. Rogers *et al.* (1982) treated 24 cicatricial pemphigoid patients with dapsone, and 20 had a partial or complete control of their inflammatory activity. Of these, 17 had ocular involvement, and 5 of these had complete control of their disease; another 7 showed partial control, with minimal inflammation.

D. Sjögren's Syndrome

1. Clinical Manifestations

The triad described by the ophthalmologist Henrik Sjögren in 1933 is that of keratoconjunctivitis sicca, xerostomia, and rheumatoid arthritis or any other connective tissue disorder. The diagnosis is usually made if any two of the three features of the disease are present (Cummings *et al.*,

1971). The oral and ocular disease may precede the arthritic component of the syndrome in ~10% of cases (Talal, 1968), and ~30% of the 3–4 million patients with rheumatoid arthritis will have this syndrome, so that it is not in any way rare (Cummings *et al.*, 1971).

Patients with this disease are overwhelmingly female and are usually in their sixth decade of life. They will often complain of photophobia, burning, irritation, itching, redness, and a foreign body sensation in their eyes. The symptomatology is due to the lymphocytic infiltration of the lacrimal gland and a disruption in the tear composition. The tear film is made up of three layers. The outermost or oily phospholipid layer is produced by the meibomian glands of the lid. The middle or aqueous layer provides a smooth optical surface and is secreted in the main by the lacrimal gland. The layer closest to the cornea is the mucous layer, produced by goblet cells in the conjunctiva and to some degree by the lacrimal glands. It is these last two layers that are particularly affected by Sjögren's syndrome (Medical Staff Conference, University of California, 1975).

Examination of the cornea will often reveal a filamentary keratitis. Additionally, the tear break-up time will be markedly reduced in these patients. Three tests have been utilized to support the clinical diagnosis. The Shirmer tear strip can be placed to measure the aqueous tear flow. Tear lysozymes are an important component of the aqueous layer and are decreased in Sjögren's syndrome. This can be measured with the use of *Micrococcus* sp., an organism sensitive to tear lysozyme. Finally, the internist as well as the ophthalmologist can utilize a rose bengal dye to detect abnormal or dead epithelium. This test is frequently positive in this disease, and it can be performed without the aid of a slit lamp. It should be emphasized that systemic medication can mimic the dry-eye syndrome. Systemic antihistamines, atropine, and L-dopa have been implicated in such cases (Medical Staff Conference, University of California, 1975; Lemp, 1982).

Patients with dry eyes need to be followed carefully for complications of their disease. These include corneal ulceration, vascularization, and even perforation. This condition also makes patients more susceptible to secondary bacterial infections.

2. Immunological Manifestations

Little is understood about the underlying cause of the initial lymphocytic infiltrate into the lacrimal and salivary glands that leads to destruction and scarring. A genetically distinct subgroup of patients with Sjögren's syndrome can be identified. In one series (Chused *et al.*, 1977), 84% of patients with this syndrome without rheumatoid arthritis bore

HLA-Dw3 as compared to 24% of controls. An increased incidence of this antigen was not found in rheumatoid arthritis patients with the disorder. What this may imply about basic differences in mechanisms remains unknown.

Antibody to salivary gland duct cells are commonly detected in these patients, with one series reporting 53% (Feltkamp *et al.*, 1968). A polyclonal hypergammaglobulinemia is frequently seen. A monoclonal disorder, invariably Waldenstrom's macroglobulinemia, and lymphomas and other neoplastic disorders have been noted as well (Friedlaender, 1979).

In the peripheral blood, an increase in B cells and a modest decrease in T cells have been reported (Talal *et al.*, 1974). Furthermore, the production of IgM, IgG, and rheumatoid factor by lymphocytes infiltrating the salivary glands will occur (Friedlaender, 1979). Alterations in the cellular component of the immune system in patients with Sjögren's syndrome have also been found. Abnormal *in vitro* responses to mitogen stimulation have been seen, but peripheral blood lymphocytes from these patients will have positive *in vitro* cellular responses to salivary gland extracts. It has been suggested that the release of lymphokines from these activated cells may induce the pathological changes that are manifested (Friedlaender, 1979).

3. Therapeutic Modalities

Treatment for this condition is usually topical and almost always involves the application of wetting agents. In severe cases, surgically stenosing the puncta and the wearing of goggles to create a moist chamber have been tried. Tabbara (1979) suggested that alternate-day steroid therapy might help reduce symptoms and increase tear production.

E. Mooren's Ulcer

1. Clinical Manifestations

Mooren's ulcer is a chronic peripheral corneal disease. The condition is painful and is bilateral in at least 25% of cases. The hallmark of the ailment is an ulceration that moves toward the center of the cornea and has a steep, undermined border (Fig. 5). The ulcer can destroy all but the deepest portions of the cornea, sparing only Descemet's membrane and some overlying corneal stroma. Healing occurs from the periphery, so that the extremely thinned cornea will be highly vascularized and covered by ingrowing conjunctiva. This leaves the cornea opaque and the patient with severe visual loss.

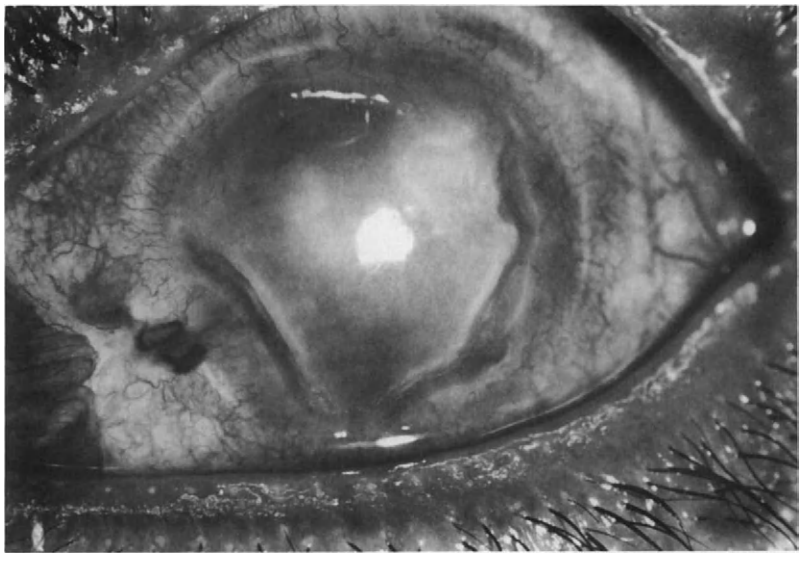


FIG. 5. Typical appearance of Mooren's ulcer. Lesion begins peripherally and moves toward the center of the cornea. New vessel growth and corneal opacification are evident.

Wood and Kaufman (1971) suggested that the disease may present in two forms. The first is seen in older individuals in which the course of the disease is relatively benign and usually unilateral, and in whom ocular surgery can be reasonably well tolerated. The second is a bilateral form seen more frequently in younger individuals. This disease is relentless and responds poorly to surgical or standard medical management. Keitzman (1968) reported a high incidence of the latter type among Nigerians, with involvement of the surrounding sclera that often resulted in perforation of the involved corneal sites. Previous trauma to ocular tissue has been noted in several patients with Mooren's ulcer, and a similar entity has been reported to occur after standard intracapsular cataract extraction (Arentsen *et al.*, 1976).

The ulcerated corneal lesion has leucocytes in its leading edge, while the most posterior lamellae of the corneal stroma and the corneal endothelium are not involved in the disease process (Mondino, 1981). In the adjacent conjunctival tissue, both lymphocytes and plasma cells are found (Brown, 1975a,b).

The diagnosis of Mooren's ulcer should be made after a careful evaluation to rule out the presence of an associated systemic disease with an ocular manifestation. Perilimbal complications can be seen in several systemic collagen vascular disorders of presumed autoimmune origin. We-

gener's granulomatosis, with ocular involvement occurring in some 47% of patients (Haynes *et al.*, 1977), and periarteritis nodosa are potentially lethal diseases that have a particular predilection for peripheral corneal complications. In addition, patients with herpes zoster involving the eye have been reported to present with peripheral corneal ulcers similar to Mooren's; one bilateral case led to destruction of both corneas (Mondino *et al.*, 1978a).

2. Immunological Manifestations

Evidence of immune system activation has been demonstrated in patients with Mooren's ulcer. Brown (1975a,b) showed that plasma cells in the conjunctiva adjacent to disease activity produced collagenolytic and proteoglyconolytic enzymes. Although no HLA association could be identified in a series of 16 patients, 63% had serum IgA levels that were 2 SD above normal (Mondino, 1981).

Circulating autoantibodies to corneal epithelium have been detected in this disease (Schaap *et al.*, 1969). Using direct immunofluorescence methods, Brown and colleagues (1976) found immunoglobulin localized to the conjunctival epithelium in three Mooren's ulcer patients and the third component of complement deposited in the intercellular spaces and in the cell cytoplasm of surface epithelium of biopsy material in two patients with active disease. IgM was found deposited along the conjunctival basement membrane in one specimen. Autoantibodies to both corneal and conjunctival eipthelium were also detected.

Lymphocytes from Mooren's ulcer patients have also been noted to demonstrate positive *in vitro* cellular immune responses to ocular antigens. Mondino and colleagues (1978b) found that lymphocytes of six of seven Mooren's patients produced macrophage inhibition factor *in vitro* in the presence of crude corneal antigen. Foster and associates (1979) reported that the lymphocytes of a 42-year-old Mooren's ulcer patient had a positive blastogenic response to corneal stroma; we have seen such a response to corneal epithelium in a West African patient with bilateral progressive disease (unpublished result). A T-cell subset imbalence has been described in the conjunctiva of one patient with this condition (Bhan *et al.*, 1982). The interpretation of the evidence suggesting an autoimmune mechanism to this disease remains speculative. Brown *et al.* (1976) and Mondino and associates (1978a) suggested that Mooren's ulcer may be a secondary phenomenon that follows corneal injury in a patient prone to autoimmune disease, rather than a primary autoimmune event.

Recently, Langerhans cells have been postulated to play an important role in corneal and epithelial immune responses (Friedlaender, 1981). These HLA-D–bearing cells were found to migrate into the guinea pig

cornea after chemical injury (Hendkind, 1965). Although probably not present in the central cornea (Streilein, *et al.*, 1980), they can be present in the periphery. Langerhans cells increase dramatically in the cornea under various stimuli and decrease in number after topical corticosteroid administration (Gillette *et al.*, 1982). The fact that larger molecules in the complement cascade are restricted to the peripheral cornea may explain why most corneal diseases of presumed autoimmune origin occur in this location (Mondino, 1983). Furthermore, the cornea, particularly the epithelium, may have unique immunological characteristics. The corneal epithelium is both a strong immunogen and antigen (Nussenblatt *et al.*, 1979), and a corneal epithelial cell-derived thymocyte-activating factor (CETAF) has recently been described (Grabner *et al.*, 1982). The addition of CETAF significantly enhanced *in vitro* thymocyte proliferation. One might speculate that this phenomenon *in vivo* may help to augment or perpetuate a corneal immune response.

3. Therapeutic Modalities

The treatment of this disorder with local antiinflammatory therapy has not been rewarding, nor has the use of anticollagenolytic agents combined with systemic corticosteroids. Brown (1975b) reported a beneficial result when the adjacent conjunctiva was excised, presumably due to the removal of immunoreactive cells in the tissue thereby permitting healing of the tissue. A similar effect was seen when the conjunctiva was frozen (Aviel, 1972). Foster (1980) reported treating two patients who had the progressive form of the disease with cytotoxic agents, methotrexate and cyclophosphamide, resulting in a slowing or halting of the progression of ulceration.

F. Scleritis

1. Clinical Manifestations

This inflammatory condition is centered in the sclera, the outermost layer of the globe. The disease has been classically referred to as an extraarticular manifestation of rheumatoid arthritis. In reviewing 4210 patients with rheumatoid arthritis seen at a rheumatic disease center, McGavin and colleagues (1976) found that 28 patients (0.067%) had scleritis. However, 9 of 27 (33%) of patients with scleritis seen at an eye clinic had rheumatoid arthritis.

Watson and Hayreh (1976) classified this disease into anterior and posterior scleritis. In the anterior group, one is able to distinguish a diffuse, nodular, and necrotizing pattern to the scleritis. The necrotizing form of

the disease can occur with inflammation or without; this latter condition is termed scleromalacia perforans. The disease is generally seen in patients who are in the 40- to 70-year-old age group, and females more commonly have the disorder by a 8:5 ratio (Watson, 1982). Watson and Hayreh (1976) noted in their survey that only women presented with scleromalacia perforans.

Scleritis is a painful condition, with patients frequently describing the discomfort as boring in character (Watson and Hayreh, 1976). The exception to this is scleromalacia perforans. This entity is accompanied with little inflammation and frequently little pain. Although most commonly noted in rheumatoid arthritis patients, it has been reported in conjunction with systemic lupus erythematosus, Wegener's granulomatosus, polyarteritis nodosa, herpes zoster, Crohn's disease, and porphyria (Yanoff and Fine, 1975). Scleral inflammatory disease is unlike that of episcleritis, where severe ocular pain is not a feature, although the sclera may be tender to the touch. The differentiation between episcleritis, an inflammation of the superficial episcleral plexus, and scleritis, where the deep episcleral plexus is involved, is important, since the former is usually without sequelae.

Clinically, the hallmark of scleritis is scleral edema associated with the deep episcleral vessel inflammation, and this is best seen with a slit lamp. Histologically, the lesions are similar to that of a rheumatoid nodule. A granulomatous inflammatory response can be seen surrounding necrotic sclera (Yanoff and Fine, 1975). However, the ocular installation of 10% phenylephrine or 1:1000 epinephrine will blanch the more superficial vessels of the episcleritis, and the deeper enflamed vascular plexus of the scleritis will be more readily visualized.

Sight-threatening complications are of real concern in patients with scleritis, particularly with the necrotizing form of the disease; some 40% of patients suffer a decrease in visual acuity. An avascular zone of episclera or sclera is a particularly ominous finding. An accompanying sclerosing keratitis or other corneal changes can be noted in up to 29% of patients (Friedlaender, 1979), with even keratolysis, or a melting of the corneal stroma, having been noted. The brawny diffuse type of scleritis has been reported to have a positive radioactive phosphorus uptake and ultrasound indistinguishable from that classically described for ocular melanoma (Feldon *et al.*, 1978).

2. Immunological Manifestations

Little is known about the immune mechanisms involved in inflammatory scleral disease. Because of the fibrinoid necrosis noted in some le-

sions, an immune complex-mediated disease process can be hypothesized. Rao and associates (1983) described an animal model for scleritis in guinea pigs that appears histologically similar to the human condition. They also found circulating immune complexes containing IgM and positive *in vitro* blastogenic responses to scleral antigens in patients with idiopathic necrotizing scleritis. The association of scleritis with systemic collagen vascular diseases, particularly rheumatoid arthritis, would lead one to consider that shared antigens may be present in various sites of inflammation. Poole and colleagues (1982) reported that when using monospecific antibodies to bovine nasal cartilage proteoglycan and link protein, positive immunofluorescent reactions could be observed in the sclera and other ocular tissues; this suggests that they contain molecules immunologically related to those seen in cartilage.

3. Therapeutic Modalities

Although oxyphenbutazone was shown to be effective in the treatment of episcleritis (Watson *et al.*, 1966), the sight-threatening lesions due to scleritis are more problematic. Systemic corticosteroids, gold, and indomethacin frequently will not control the ongoing destructive process. Grafts of donor sclera to areas of the globe threatened by perforation run the risk of becoming engulfed in the destructive process. Foster (1980) reported the arrest of ongoing scleral destruction in five rheumatoid arthritis patients (four with scleromalacia perforans) by use of cyclophosphamide.

IV. OCULAR INVOLVEMENT IN SYSTEMIC DISEASES

A. Anterior Uveitis

1. Clinical Manifestations

Inflammations in the anterior portion of the eye can be readily observed. Although termed uveitis, the inflammatory response need not be located in the uvea. The major features of an anterior uveitis are inflammatory cells and an increase in the protein content (flare) evident in the anterior chamber. If an acute process, the eye will have ciliary injection, and the patient will complain of tearing, photophobia, pain, and decreased vision. Other features of the disease include iris nodules (as seen in sarcoidosis) and keratitic precipitates (KPs), collections of inflammatory

cells on the corneal endothelium. Dr. Alan Woods (1961) divided the anterior inflammatory response into granulomatous and nongranulomatous reactions, based on the type of KPs present. The KPs in granulomatous disease are the large "mutton-fat" type.

Systemic disease may present with an anterior uveitis. The type and binocular or uniocularity of the process is helpful at times in distinguishing which systemic ailment may be involved. A bilateral granulomatous uveitis is characteristically noted in sarcoidosis, tuberculosis, syphilis, and brucellosis, the last being quite rare in the United States. Lesions primarily involving the posterior segment of the eye, such as ocular toxoplasmosis and sympathetic ophthalmia, will also manifest a granulomatous response. Behcet's disease, if involving the anterior segment of the eye, will characteristically present with an hypopyon. Ankylosing spondylitis is classically a nongranulomatous anterior segment inflammation which, although being seen in both eyes, will be uniocular in its acute presentation. Series have demonstrated that from 23 to 35% of patients with acute anterior uveitis may have ankylosing spondylitis (Friedlaender, 1979). Childhood uveitis occurs in ~5 to 10% cases seen in an eye clinic (Perkins, 1966). It is a significant complication of pauciarticular arthritis in children, with a variable prognosis (Key and Kimura, 1975).

The sequelae of anterior segment inflammatory disease are potentially serious, and combined with the posterior uveitis group are the cause of ~10% of the visually handicapped in the United States (National Institutes of Health, 1976). These complications include band keratopathy, glaucoma, and cataracts.

2. Immunological Manifestations

Intraocular inflammatory disease have been shown to have HLA associations (Rahi, 1979; Nussenblatt, 1980). Anterior uveitis in Caucasians has been associated with HLA-B27 (Brewerton *et al.*, 1973; Ehlers *et al.*, 1974; Mapstone and Woodrow, 1975; Zervas *et al.*, 1977). However, HLA-B8 and -D5 have been found to be present in Black American patients with anterior uveitis (Nussenblatt and Mittal, 1981; Berg and Kaplan, 1982). In addition, Behcet's disease, with a severe anterior ocular inflammation, is associated with HLA-B5 both in Japanese (Ohno *et al.*, 1973, 1979) and in Europeans (Bloch-Michel *et al.*, 1979). α_1 Antitrypsin has been reported by Brewerton and co-workers (1978) to be strongly associated with acute anterior uveitis, but this could not be corroborated by Saari and colleagues (1981).

To date, definitive evidence of an autoimmune mechanism to anterior uveitis is lacking. Both T and B cells have been found in the aqueous of

uveitis patients, but their specificity was not determined (Belfort *et al.*, 1982). It has been suggested that molecular mimicry may be mechanistically important in human disease. Inflammatory bowel disease and uveitis have been noted to be associated (Korelitz and Coles, 1967). Ebringer and colleagues (1979) noted a relationship between acute anterior uveitis in ankylosing spondylitis and *Klebsiella pneumoniae* in the stool, although others have not made this correlation (Warren and Brewerton, 1980). Avakian and colleagues (1981) demonstrated that *Klebsiella* may carry antigenically similar determinants to those found in the vitreous, and Welsh and co-workers (1981) further suggested that the uveitis seen in ankylosing spondylitis may be attributed to anti-gram-negative antibodies binding to cross-reacting ocular antigens. The possible role of gram-negative components in the induction or perpetuation of this disease was supported by Rosenbaum and colleagues (1980) in their demonstration that endotoxin or lipopolysaccharide given into the footpad or intraperitoneally will produce an inflammatory response in eyes of Lewis rats, with no histological alterations found in any of the other major organs.

3. Therapeutic Modalities

The treatment of anterior uveitis can be frustrating because of the tendency to recurrence. The usual initial approach to therapy is the use of mydriatics and topical corticosteroids (Nussenblatt, 1983). Severe cases might require periocular or systemic corticosteroid administration. Prostaglandin inhibitors or other nonsteroidal antiinflammatory agents have not been found to be particularly effective in treating this condition (Coles, 1967).

B. Posterior Uveitis

1. Clinical Manifestations

Posterior uveitis denotes primarily the anatomical area of greatest inflammatory activity in the eye, and multiple clinical syndromes have been noted (Maumenee, 1970; Ryan and Maumenee, 1980). Causes of these entities presenting with a granulomatous uveitis were felt to be identifiable in all cases in a study performed in 1941, while in 1969 at least one-fourth of the cases had undetermined presumptive diagnoses (Schlaegel, 1969). These diseases are generally painless at the onset, and the initial symptom may be a haze or debris floating before the patient's line of vision. This can then be followed by a marked decrease in visual acuity.

The cause of the visual loss may be due to several factors, including vitreal haze, destruction of retinal tissue, and chronic cystoid edema of the macula region of the retina.

Posterior uveitis can also be a manifestation of systemic disease. Sarcoidosis can involve all the layers of the posterior segment of the eye. One of the most potentially devastating ocular inflammations is that of Behcet's disease, in which uveitis is one of the major criteria in the diagnosis (Japanese Research Committee of Behcet's Disease, 1974).

2. Immunological Manifestations

A genetic propensity has been defined in some posterior uveitic entities of presumed autoimmune origin. Behcet's disease has already been mentioned as being associated with HLA-B5, while Vogt–Kayanagi–Harada's syndrome associates with HLA-Bw22 (Tagawa *et al.*, 1976) and the supertypic HLA-DR antigen MT3, an *HLA-D*-linked B-cell alloantigen in the Japanese (Ohno, 1981). HLA-A29 was found in 85% of Caucasians with birdshot retinochoroidopathy tested, giving it one of the highest relative risks (50) thus reported (Nussenblatt *et al.*, 1982b).

Kaplan and associates (1984) looked at cell-surface markers from vitrectomy samples obtained from uveitis patients, with the proportion of T and B cells present varying considerably depending upon the underlying presumptive diagnosis. *In vitro* functional abnormalities of suppressor cells have been noted in patients with posterior but not anterior uveitis, in which two populations of suppressor cells were defined (Nussenblatt *et al.*, 1980c). Ohno (1981) and coworkers (1982a,b) studied abnormalities of interferon production in Behcet's disease. Serum levels of α interferon were elevated when compared to controls. In further studies, the interferon levels were noted to decrease in the exacerbation stage, and the T-cell subset bearing the Fc receptor for IgG was found to be spontaneously producing α interferon *in vitro*. Although immune complexes have been suggested as a mechanism for the ocular destruction noted in uveitis (Dernouchamps *et al.*, 1977; Char *et al.*, 1979), Dumonde and associates (1982) recently questioned whether this is true in all cases. Their evidence would suggest that in some cases of retinal vasculitis, immune complexes are idiotypic–antiidiotypic and protective in nature.

3. Therapeutic Modalities.

Although transfer factor and cryoablation have been suggested, the use of systemic corticosteroids has been generally the therapy of choice for posterior uveitic conditions with a presumed autoimmune origin. Periocular corticosteroids can be used for uniocular conditions or if there is

concern about the systemic effects of high-dose corticosteroids (Nussenblatt, 1983). The International Uveitis Study Group (1980, unpublished) agreed that the treatment of choice for the ocular manifestations of Behcet's disease should be cytotoxic agents; those usually used have been cyclophosphamide and chlorambucil. Cytotoxic agents have been used in patients with disease refractory to corticosteroid therapy or in those patients with disease refractory to corticosteroid therapy or in those patients with complications of steroids not permitting continuance of this therapeutic approach (Wong, 1969; Wong *et al.*, 1971; Godfrey *et al.*, 1974). Recently, a pilot project using cyclosporin A in the treatment of uveitis patients with positive *in vitro* responses to the S antigen has begun, with very encouraging initial findings (R. B. Nussenblatt *et al.*, 1983).

REFERENCES

Anhalt, G. J., Bahn, C. F., Labib, R. S., Voorhees, J. J., Sugar, A., and Diaz, L. A. (1981). *J. Clin. Invest.* **68,** 1097–1101.

Arentsen, J. J., Christiansen, J. M., and Maumenee, A. E. (1976). *Am. J. Ophthalmol.* **81,** 194–197.

Avakian, H., Abuknesha, R., Welsh, J., and Ebringer, A. (1981). *Br. J. Ophthalmol.* **65,** 315–322.

Aviel, E. (1972). *Br. J. Ophthalmol.* **56,** 48–51.

Belfort, R., Jr., Moura, N. C., and Mendes, N. F. (1982). *Arch. Opthalmol.* **100,** 465–467.

Berg, J., and Kaplan, H. J. (1982). *Ophthalmology (Rochester, Minn.), Suppl.* **89,** 104.

Bhan, A. K., Fujikawa, L. S., and Foster, C. S. (1982). *Am. J. Ophthalmol.* **94,** 205–212.

Bloch-Michel, E., Campinchi, R., Muller, J. Y., Binaghi, M., and Sales, J. (1979). *In* "Proceedings of the Second International Symposium on Immunology and Immunopathology of the Eye" (A. M. Silverstein and G. R. O'Connor, eds.), pp. 10–14. Masson, New York.

Blodi, F. C. (1959). *Trans. Am. Acad. Ophthalmol. Otolaryngol.* **63,** 642–649.

Brewerton, D. A., Caffrey, M., Nicholls, A., Walters, D., and James, D. C. O., (1973). *Lancet* **2,** 994–996.

Brewerton, D. A., Webley, M., Murphy, A. H., and Milford-Ward, A. (1978). *Lancet* **1,** 1103.

Brown, S. I. (1975a). *Br. J. Ophthalmol.* **59,** 670–674.

Brown, S. J. (1975b). *Br. J. Ophthalmol.* **59,** 675–682.

Brown, S. I., Mondino, B. J., and Rabin, B. S. (1976). *Am. J. Ophthalmol.* **82,** 835–840.

Char, D. H., Stein, P., Masi, R., and Christensen, M. (1979). *Am. J. Ophthalmol.* **87,** 678–681.

Chused, T. M., Kassan, S. S., Opelz, G., Moutsopoulos, H. M., and Terasaki, P. I. (1977). *N. Engl. J. Med.* **296,** 895–897.

Coles, R. S. (1967). *In* "Ocular Therapy: Complications and Management" (I. H. Leopold, ed.), Vol. 2, pp. 115–121. Mosby, St. Louis, Missouri.

Collins, R. C. (1949). *Am. J. Ophthalmol.* **32,** 1687–1699.

Collins, R. C. (1953). *Am. J. Ophthalmol.* **36,** (Part II), 150–162.

Cummings, N. A., Schall, G. L., Asofsky, R., Anderson, L. G., and Talal, N. (1971). *Ann. Int. Med.* **75,** 937–950.

Dabelsteen, E., Ullman, S., Thomsen, K., and Rygaard, J. (1974). *Acta Derm.-Venereol.* **54,** 189–192.

Dernouchamps, J. P., Vaerman, J. P., Michiels, J., and Masson, P. L. (1977). *Am. J. Ophthalmol.* **84,** 24–31.

de Veer, J. A. (1953). *AMA Arch. Ophthalmol.* **49,** 607–632.

Dinning, W. J., and Perkins, E. S. (1975). *Br. J. Ophthalmol.* **59,** 397–403.

Dorey, C., and Faure, J.-P. (1977). *Ann. Immunol. (Inst. Pasteur)* **128C,** 229–232.

Duke-Elder, S. (1965). *In* "Systems of Ophthalmology" Vol. 8, pp. 502–510. Henry Kimpton, London.

DuMonde, D. C. Kasp-Grochowska, E., Graham, E., Sanders, M. D., Faure, J.-P, de Kozak, Y., and van Tuyen, V. (1982). *Lancet* **2,** 787–792.

Easom, H. A., and Zimmerman, L. E. (1964). *Arch. Ophthalmol.* **72,** 9–15.

Ebringer, R., Cawdell, D., and Ebringer, A. (1979). *Br. Med. J.* **i,** 383.

Ehlers, N., Kissmeyer-Nielsen, F., Kjerbye, K. E., and Lamm, L. U. (1974). *Lancet* **1,** 99.

Faure, J.-P., and de Kozak, Y. (1981). *In* "Immunology of the Eye" (R. J. Helmsen, A. A. Suran, I. Gery, and R. B. Nussenblatt, eds.), Workshop II, pp. 33–48. Information Retrieval, Inc., Washington, D. C.

Feldon, S. E., Sigelman, J., Albert, D. M., and Smith, T. R. (1978). *Am. J. Ophthalmol.* **85,** 781–787.

Feltkamp, T. E. W., and Van Rossun, A. L. (1968). *Clin. Exp. Immunol.* **3,** 1–16.

Foster, C. S. (1980). *Ophthalmology (Rochester, Minn.)* **87,** 140–149.

Foster, C. S., Kenyon, K. R., Greiner, J., Greineder, D. K., Friedland, B., and Allansmith, M. R. (1979). *Am. J. Ophthalmol.* **88,** 149–159.

Friedenwald, J. S. (1934). *JAMA, J. Am. Med. Assoc.* **17,** 108–1018.

Friedlaender, M. H. (1979). "Allergy and Immunology of the Eye." Harper and Row, Hagerstown, Maryland.

Friedlaender, M. H. (1981). *Immunol. Allergy Prac. November/December,* pp. 193–197.

Furey, N., West, C., Andrews, T., Paul, P. D., and Bean, S. F. (1975). *Am. J. Ophthalmol.* **80,** 825–831.

Gammon, W. R., Lewis, D. M., Carlo, J. R., Sams, W. M., and Wheeler, C. E. (1980). *J. Invest. Dermatol.* **75,** 334–339.

Gery, I., Nussenblatt, R., and BenEzra, D. (1981). *Invest. Ophthalmol. Visual Sci.* **20,** 32–39.

Gillette, T. E., Chandler, J. W., and Greiner, J. V. (1982). *Opthalmology (Rochester, Minn.)* **89,** 700–710.

Godfrey, W. A., Epstein, M. V., O'Connor, G. R., Kimura, S. J., Hogan, M. J., and Nozik, R. A. (1974). *Am. J. Ophthalmol.* **78,** 415–428.

Goldschmidt, L., Goldbaum, M., Walker, S. M., and Weigle, W. O. (1982). *J. Immunol.* **129,** 1658–1662.

Grabner, G., Luger, T. A., Smolin, G., and Oppenheim, J. J. (1982). *Invest. Ophthalmol. Visual Sci.* **23,** 757–763.

Griffith, M. R., Fukuyama, K., Tuffanelli, D., and Silverman, S. (1974). *Arch. Dermatol.* **109,** 195–199.

Hackett, E., and Thompson, A. (1964). *Lancet* **2,** 663–666.

Halbert, S. P., and Manski, W. (1963). *In* "Progress in Allergy" (P. Kallos and B. Waksman, eds.), Vol. 7, pp. 107–186. Karger, Basel.

Halbert, S. P., Manski, W., and Auerbach, T. (1961). *In* "The Structure of the Eye" (G. K. Smelser ed.), p. 249. Academic Press, New York.

Hardy, K. M., Perry, H. O., Pingree, G. C., and Kirby, T. J., Jr. (1971). *Arch. Dermatol.* **104,** 467–475.

Haynes, B. F., Fishman, M. L., Fauci, A. S., and Wolff, S. M. (1977). *Am. J. Med.* **63,** 131–141.

Henkind, P. (1965). *Exp. Eye Res.* **4,** 42–47.

Henley, W. L., Okas, S., and Leopold, I. H. (1973). *Invest. Ophthalmol. Visual Sci.* **12,** 520–524.

Irvine, S. R., and Irvine, A. R. Jr. (1952). *Am. J. Ophthalmol.* **35,** 489–499.

Jakobiec, F. A., Marboe, C. C., Knowles, D. M., II, Iwamoto, T., Harrison, W., Chang, S., and Coleman, D. J. (1983). *Ophthalmology (Rochester, Minn.)* **90,** 76–95.

Japanese Research Committee of Behcet's Disease (1974). *Jpn. J. Ophthalmol.* **18,** 291.

Kaplan, H. J., and Streilein, J. W. (1978). *J. Immunol.* **120,** 689–693.

Kaplan, H. J., Waldrep, J. C., Nicholson, J. K. A., and Gordon, D. (1984). *Arch. Ophthalmol.* **102,** 572–575.

Keitzman, B. (1968). *Am. J. Ophthalmol.* **65,** 679–685.

Key, S., and Kimura, S. J. (1975). *Am. J. Ophthalmol.* **80,** 425–429.

Khodadoust, A. A., and Silverstein, A. M. (1972). *Invest. Ophthalmol. Visual Sci.* **11,** 137–148.

Korelitz, B. I., and Coles, R. S. (1967). *Gastroenterology* **52,** 78–82.

Lemp, M. L. (1982). *In* "Clinical Ophthalmology" (T. D. Duane and E. A. Jaeger, eds.), Vol. 4, p. 3. Harper and Row, Philadelphia, Pennsylvania.

Lever, W. F. (1979). *J. Am. Acad. Dermatol.* **1,** 2–31.

Lubin, J. R., Albert, D. M., and Weinstein, M. (1980). *Ophthalmology (Rochester, Minn.)* **87,** 109–121.

McGavin, D. D. M., Williamson, J., Forrester, J. V., Faulds, W. S., Buchanan, W. W., Dick, W. C., Lee, P., MacSwein, R. N. M., and Whaley, K. (1976). *Br. J. Ophthalmol.* **60,** 192–226.

Mapstone, R., and Woodrow, J. C. (1975). *Br. J. Ophthalmol.* **59,** 270–275.

Marak, G. E., Jr., Font, R. L., Czawlytko, L. N., and Alepa, F. P. (1974). *Exp. Eye Res.* **19,** 311–316.

Marak, G. E., Jr., Font, R. L., and Alepa, F. P. (1976a). *Mod. Probl. Ophthalmol.* **16,** 75–79.

Marak, G. E., Jr., Font, R. L., and Alepa, F. P. (1976b). *Ophthalmic Res.* **8,** 117–120.

Marak, G. E., Jr., Font, R. L., and Alepa, F. P. (1977). *Ophthalmic Res.* **9,** 162–170.

Marak, G. E., Jr., Font, R. L., and Weigle, W. O. (1979). *In* "Immunology and Immunopathology of the Eye" (A. M. Silverstein and C. R. O'Connor, eds.), pp. 135–137. Masson, New York.

Maumenee, A. E. (1970). *Trans. Am. Acad. Ophthalmol. Otolaryngol.* **74,** 473–504.

Medical Staff Conference, Univ. of Calif. (1975). *West J. Med.* **122,** 50–58.

Mondino, B. (1981). *In* "Immunology of the Eye" (R. J. Helmsen, A. A. Suran, I. Gery, and R. B. Nussenblatt, eds.), Workshop II, pp. 77–90. Information Retrieval, Inc., Washington, D. C.

Mondino, B. J. (1985). *In* "Immunology and Immunopathology of the Eye" (G. R. O'Connor and J. W. Chandler, eds.), pp. 194–198. Masson, New York.

Mondino, B. J., Ross, A. N., Rabin, B. S., and Brown, S. I., (1977). *Am. J. Ophthalmol.* **83,** 443–450.

Mondino, B. J., Brown, S. I., and Mondzelewski, J. P. (1978a). *Am. J. Ophthalmol.* **86,** 611–614.

Mondino, B. J., Brown, S. I., and Rabin, B. S. (1978b). *Am. J. Ophthalmol.* **85,** 788–791.

Mondino, B. J., Brown, S. I., Lempert, S., and Jenkins, M. S. (1979a). *Ophthalmology (Rochester, Minn.)* **86,** 543–552.

Mondino, B. J., Brown, S. I., and Rabin, B. S. (1979b). *Arch. Ophthalmol.* **97,** 479.

National Institutes of Health (1976). Interim Report of the National Advisory Eye Council, Dept. of Health, Education, and Welfare, pp. 20–22.

Nussenblatt, R. B., (1980). *In* "Immunology of the Eye" (G. M. Steinberg, I. Gery, and R. B. Nussenblatt, eds.), Workshop 1, pp. 25–42. Information Retrieval Inc., Washington, D.C.

Nussenblatt, R. B. (1983). *In* "Current Therapy in Allergy and Immunology" (A. S. Fauci and L. M. Lichtenstein, eds.), pp. 224–227. Dekker, New York.

Nussenblatt, R. B., and Mittall, K. K. (1981). *Br. J. Ophthalmol.* **65,** 329–332.

Nussenblatt, R. B., Gery, I., and BenEzra, D. (1979). *In* "Immunology and Immunopathology of the Eye" (A. M. Silverstein and G. M. O'Connor, eds.), pp. 145–150. Masson, New York.

Nussenblatt, R. B., Gery, I., and Wacker, W. B., (1980a). *Invest. Ophthalmol. Visual Sci.* **19,** 686–690.

Nussenblatt, R. B., Gery, I., Ballintine, E. J., and Wacker, W. B. (1980b). *Am. J. Ophthalmol.* **89,** 173–179.

Nussenblatt, R. B., Cevario, S. J., and Gery, I. (1980c). *Lancet* **2,** 722–724.

Nussenblatt, R. B., Kuwabara, T., de Monasterio, F. M., and Wacker, W. B. (1981a). *Arch. Ophthalmol.* **99,** 1090–1092.

Nussenblatt, R. B., Gery, I., Kuwabara, T., de Monasterio, F. M., and Wacker, W. B. (1981b). *In* "Immunology of the Eye" (R. J. Helmsen, A. A. Suran, I. Gery, and R. B. Nussenblatt, eds.), Workshop II, pp. 49–66. Information Retrieval, Inc., Washington, D. C.

Nussenblatt, R. B., Rodrigues, M. M., Wacker, W. B., Cevario, S. J., Salinas-Carmona, M. C., and Gery, I. (1981c). *J. Clin. Invest.* **67,** 1228–1231.

Nussenblatt, R. B., Rodrigues, M. M., Salinas-Carmona, M. C., Gery, I., Cevario, S. J., and Wacker, W. B. (1982a). *Arch. Ophthalmol.* **100,** 1146–1149.

Nussenblatt, R. B., Mittal, K. K., Ryan, S., Green, W. R., and Maumenee, A. E. (1982b). *Am. J. Ophthalmol.* **94,** 147–158.

Nussenblatt, R. B., Palestine, A. G., Rook, A. H., Scher, I., Wacker, W. B., and Gery, I. (1983). *Lancet* **ii,** 235–238.

Ohno, S. (1981). *Trans. Ophthalmol. Soc. U. K.* **101,** 335–341.

Ohno, S., Aoki, K., Sugiura, S., Nakayama, E., Itakura, K., and Aizawa, M. (1973). *Lancet* **2,** 1383–1384.

Ohno, S., Sugiura, S., Ohguchi, M., and Aoki, K. (1979). *In* "Proceedings of the Second International Symposium on Immunology and Immunopathology of the Eye" (A. M. Silverstein and G. R. O'Connor, eds.), pp. 15–17. Masson, New York.

Ohno, S., Kato, F., Matsuda, H., Fujii, N., and Minagawa, T. (1982a). *Ophthalmologica* **185,** 187–192.

Ohno, S., Kato, F., Matsuda, H., Fujii, N., and Minagawa, T. (1982b). *Infect. Immun.* **36,** 202–208.

Patten, J. T., Cavanagh, H. D., and Allansmith, M. R. (1976). *Am. J. Ophthalmol.* **82,** 272–276.

Perkins, E. S. (1966). *Br. J. Ophthalmol.* **50,** 169–185.

Poole, A. R., Pidoux, I., Reiner, A., Coster, L., and Hassell, J. R. (1982). *J. Cell Biol.* **93,** 910–920.

Porter, R., and Knight, J. (1973). "Corneal Graft Failures." Elsevier Biomedical, New York.

Rahi, A. H. (1979). *Br. J. Ophthalmol.* **63,** 283–292.
Rao, N. A., Wacker, W. B., and Marak, G. E., Jr. (1979). *Arch. Ophthalmol.* **97,** 1954–1958.
Rao, N. A., Phillips, T. M., Wong, V. G., Sliwinski, A. J., and Marak, G. E., Jr. (1985). *In* "Immunology and Immunopathology of the Eye" (A. M. Silverstein and G. R. O'Connor, eds.), pp. 54–57. Masson, New York.
Rogers, R. S., Perry, H. O., Bean, S. F., and Jordan, R. E. (1977). *J. Invest. Dermatol.* **68,** 39–43.
Rogers, R. S., Seehafer, J. R., and Perry, H. O. (1982). *J. Am. Acad. Dermatol.* **6,** 215–223.
Rosenbaum, J. T., McDevitt, H. O., Guss, R. B., and Egbert, P. R. (1980). *Nature (London)* **286,** 611–613.
Ryan, S. J., and Maumenee, A. E. (1980). *Am. J. Ophthalmol.* **89,** 31–45.
Saari, K. M., Solja, J., Sirpela, M., Frants, R. R., and Eriksson, A. W. (1981). *Albrecht von Graefes Arch. Klin. Ophthalmol.* **216,** 205–207.
Sams, W. H., and Gleich, G. J. (1971). *Proc. Soc. Exp. Biol. Med.* **136,** 1027–1031.
Sandberg, H. O., and Closs, O. (1979a). *In* "Immunology and Immunopathology of the Eye" (A. M. Silverstein and G. R. O'Connor, eds.), pp. 325–330. Masson, New York.
Sandberg, H. O., and Closs, O. (1979b). *Exp. Eye Res.* **28,** 601–604.
Schaap, O. L., Feltkamp, T. E. W., and Brecbaart, A. C. (1969). *Clin. Exp. Immunol.* **5,** 365–370.
Schlaegel, T. F., Jr. (1969). "Essentials of Uveitis." Little, Brown, Boston, Massachusetts.
Segal, B. C., Seegal, D., and Kharazo, D. (1933). *J. Immunol.* **25,** 207–220.
Silverstein, A. M. (1974). *Trans. Ophthalmol. Soc. U. K.* **94,** 496–517.
Sjögren, H. (1933). *Acta Ophthal. (Suppl. 2)* **11,** 1–151.
Stanley, J. R., Hawley-Nelson, P., Yuspa, S. H., Shevach, E. M., and Katz, S. I. (1981). *Cell* **24,** 897–903.
Stanley, J. R., Hawley-Nelson, P., Yaar, M., Martin, G. R., and Katz, S. I. (1982). *J. Invest. Dermatol.* **78,** 456–459.
Streilein, J. W., Toews, G. B., and Bergstresser, P. R. (1980). *Nature (London)* **282,** 326–327.
Tabbara, K. F. (1979). *Invest. Ophthalmol. Visual Sci. (Suppl.) April,* p. 70.
Tagawa, Y., Sugiura, S., Yakura, H., Wakishaka, A., and Aizawa, M. (1976). *N. Engl. J. Med.* **295,** 173.
Talal, N. (1968). *In* "Current Diagnosis—2" (H. F. Conn, R. B. Conn, eds.), p. 241. Saunders, Philadelphia.
Talal, N., Sylvester, R. A., Daniels, T. E., Greenspan, J. S., and Williams, R. C., Jr. (1974). *J. Clin. Invest.* **53,** 180–189.
Tanoe, A., Faure, J.-P., Le Hoang, P., and Bloch-Michel, E. (1985). *In* "Proceedings of the Third Symposium on Immunology and Immunopathology of the Eye" (G. R. O'Connor and J. W. Chandler, eds.), pp. 250–254. Masson, New York.
Uhlenhuth, L. (1903). *In* "Festschrift zum Sechzigsten Geburtstage von Robert Koch" pp. 49–74.
Van der Haeve, J. (1920). *Br. J. Ophthalmol.* **4,** 195–198.
Verhoeff, F., and Lemoine, A. N. (1922). *Am. J. Ophthalmol.* **5,** 737–745.
Wacker, W. B., and Lipton, M. (1965). *Nature (London)* **206,** 253–254.
Wacker, W. B., Donoso, L. A., Kalsow, C. M., Yankeelov, J. A., Jr., and Organisciak, D. T. (1977). *J. Immunol.* **119,** 1949–1958.
Warren, R. E., and Brewerton, D. A. (1980). *Ann. Rheum. Dis.* **39,** 37–44.
Watson, P. G. (1982). *In* "Clinical Ophthalmology" (T. D. Duane and E. A. Jaeger, eds.), Vol. 4, pp. 1–39. Harper and Row, Hagerstown, Maryland.
Watson, P. G., and Hayreh, S. S. (1976). *Br. J. Ophthalmol.* **60,** 163–191.

Watson, P. G., Lobascher, D., Sabiston, D., Lewis-Faning, E., Fowler, P. D., and Jones, B. R. (1966). *Br. J. Ophthalmol.* **50,** 463–481.

Welsh, J., Avakian, H., and Ebringer, A. (1981). *Br. J. Ophthalmol.* **65,** 323–328.

White, D. J. G., ed. (1982). "Cyclosporin A." Elsevier Biomedical, Amsterdam.

Wirostko, E., and Spalter, H. F. (1967). *Arch. Ophthalmol.* **78,** 1–7.

Witmer, R. (1964). *In* "Immunopathology of Uveitis" (A. E. Maumenee and A. M. Silverstein, eds.), pp. 111–128. Williams & Wilkins, Baltimore, Maryland.

Wong, V. G. (1969). *Arch. Ophthalmol.* **81,** 628–637.

Wong, V. G., Anderson, R., and O'Brien, P. J. (1971). *Am. J. Ophthalmol.* **72,** 960–966.

Wood, T. O., and Kaufman, H. E. (1971). *Am. J. Ophthalmol.* **71,** 417–422.

Woods, A. C. (1921). *JAMA, J. Am. Med. Assoc.* **77,** 1217–1322.

Woods, A. C. (1956). "Endogenous Uveitis." Williams & Wilkins, Baltimore, Maryland.

Woods, A. C. (1961). "Endogenous Inflammations of the Uveal Tract." Williams & Wilkins, Baltimore, Maryland.

Yanoff, M., and Fine, B. S. (1975). *In* "Ocular Pathology" pp. 316–317. Harper and Row, Hagerstown, Maryland.

Yanoff, M., and Scheie, H. G. (1968). *Arch. Ophthalmol.* **80,** 166–170.

Zervas, J., Tsokos, G., Papadakis, G., Kabouklis, E., and Papadopoulos, D. (1977). *Br. J. Ophthalmol.* **61,** 699–701.

CHAPTER 15

Multiple Sclerosis, Allied Central Nervous System Diseases, and Immune-Mediated Neuropathies

BARRY G. W. ARNASON
Department of Neurology
The University of Chicago
Chicago, Illinois

I. MULTIPLE SCLEROSIS

A. Introduction

Multiple sclerosis (MS) is an inflammatory disease of the white matter of the central nervous system (CNS) that results in progressive neurological deficits most often manifest as impaired vision, spasticity, ataxia, and sensory abnormalities. Multiple sclerosis has been recognized for >100 years, but its cause remains unknown; current opinion holds that it will probably prove to be an autoimmune process, an infection, or some combination of the two. Certain facts concerning MS will have to be incorporated into any coherent theory of its pathogenesis. These include (1) a curious geographical distribution; (2) an inflammatory response in the CNS with *in situ* synthesis of IgG; (3) a selective destruction of myelin and of oligodendrocytes, the myelin-synthesizing cells; (4) a genetically determined predisposition to develop the disease; and (5) abnormalities in circulating immunocytes.

B. Clinical Presentations and Epidemiology

In the northern United States, in Canada, and in northern Europe, 1 adult in 1000 has MS. Onset peaks in young adult life; onset in childhood is rare. Six of 10 cases occur in females. Two of 3 cases are characterized by flare-ups that last for days to several weeks. These are followed by periods of arrest of progression or of improvement known as remissions. No trigger for attacks is known; their onset is capricious and unpredictable. Frequency of attacks is greatest during the first few years of disease. One of 3 cases progresses steadily from the onset. Beyond age 45–50, even those cases in which the earlier course had been marked by flare-ups and remissions are likely to slide into a progressive course.

The clinical features of MS are protean. The picture is determined by the sites of foci of demyelination, and these are highly variable. Common features include impaired vision because of involvement of optic nerves; decreased perception of vibration and position, which bespeaks lesions in the posterior columns of the spinal cord; ataxia and intention tremor,

evidencing damage to spinocerebellar pathways or to cerebellar outflow tracts; and weakness or paralysis of one or more limbs, which together with spasticity and bladder problems point to involvement of pyramidal pathways. Diagnosis requires proof of more than one episode of deficit involving white matter tracts and of lesions at more than one site in the CNS. Examination of the spinal fluid usually reveals a modest increase in lymphocyte number and an increase in IgG content, with oligoclonal IgG bands detected when the fluid is examined electrophoretically. Evoked potential testing of visual, auditory, and somatosensory pathways may reveal delayed conduction. The laboratory abnormalities just mentioned, while supportive of a diagnosis of MS, are not specific for this disease and should be viewed as aids in diagnosis rather than as diagnostic tests.

Multiple sclerosis is primarily a disease of north Europeans and those whose forebears were north Europeans. The disease is rare in East Asia, and has not been recorded among blacks in Africa. The prevalence in American blacks is one-half that encountered in the American Caucasian population. High-risk groups carry risk for MS when they migrate to areas where the risk for indigenous populations is low, although risk may be somewhat attenuated. Conversely, MS remains rare among East Asians who have migrated to Europe or to America. The disease is also excessively rare in American Indians, if it occurs at all.

Within the United States, a gradient in prevalence of MS from north (higher) to south (lower) has been documented within a seemingly homogeneous population and medical system. This gradient exists for both whites and blacks. A similar gradient has been reported for Australia, MS prevalence being greatest in the south. To some, the data suggest that an environmental hazard, tied to climate and possibly encountered early in life, triggers MS. This explanation has deficiencies. The climates of much of Japan, Korea, China, and Siberia are as rigorous as those of northern Europe or North America, yet MS is rare among Japanese, Koreans, Chinese, and the native populations of Siberia, although not among Caucasian Russians living in Siberia. I conclude that an environmental factor, while relevant to MS, does not suffice to cause it. This environmental factor is probably present both in higher and in lower risk climates. The environmental (? infectious) thesis finds additional support from twin studies. Identical twins show a 25% concordance for MS; this finding excludes the possibility that a single gene (or for that matter, multiple genes) determines irrevocably who will develop MS. On the other hand, there is firm evidence (reviewed subsequently) for a genetic component in susceptibility to MS.

Surveys seeking to pinpoint a single environmental factor that MS patients share have failed to turn up any common thread. It is also of interest

that no increased incidence of any other disease has been documented in the MS population. This situation is to be contrasted with that pertaining in several autoimmune diseases, for example, myasthenia gravis and thyroiditis, both of which occur in the same individual more often than chance would predict.

C. Histopathology

The cardinal feature is the presence of multiple scattered discrete areas of demyelination. These vary in diameter from a millimeter or so to several centimeters and are called plaques. Large plaques form by coalescence of smaller ones and by expansion of their margins. Sites of predilection for plaques exist; these include the paraventricular regions of the cerebrum and subpially, the optic nerves and chiasm, and the posterior and lateral columns of the spinal cord. Plaque topography corresponds to that of the venous drainage of the brain and spinal cord. Peripheral nerves are spared.

Lymphocyte and monocyte–macrophage aggregates surround venules within plaques and are prominent at plaque margins where myelin destruction is occurring (Prineas and Wright, 1978). The inflammatory cells are derived from the blood. T cells (chiefly helper/lymphokine-producing cells) are seen throughout MS plaques, at the margins of active plaques where myelin destruction is ongoing, and in otherwise normal-appearing white matter just outside plaques (Traugott *et al.*, 1983). T suppressor/cytotoxic cells are also found but are outnumbered by the helper cells. Macrophages are prominent in MS plaques and are sometimes the predominant cell type. Macrophages are thought to be the final vectors of myelin destruction. How they are "turned on" is not known; antibodies secreted by plasma cells, which are abundant in established MS lesions, lymphokines secreted by T cells, or both, may be the cause.

The inflammatory response seen in the active MS lesion is almost surely directed against an antigen or antigens relevant to the pathogenesis of the disease. This antigen has defied detection. Some years ago it was thought likely that the MS antigen would be identical with that against which the immune response is directed in the animal model disease experimental allergic encephalitis (EAE). Subsequent experience has not borne out this hope. The antigen responsible for EAE has been shown to be myelin basic protein (MBP). The consensus at present is that sensitivity to MBP is unusual in MS.

Why is the MS antigen, be it a virus or a tissue component, so elusive? No answer can be provided, but possibilities include: (1) there is no single antigen, and (2) the antigen is a minor tissue component that has escaped

detection. This latter notion has some attraction, since in other autoimmune processes the antigen has proven to be a cell-surface receptor with only a few thousand copies present on any individual cell. An example is provided by myasthenia gravis, an autoimmune disease in which the immune response is directed against the acetylcholine receptor; this receptor might not have been purified to this day save for the findings that certain snake venoms bind it and that the electric organs of electric fishes are enriched for it.

In inactive plaques, a mat of fibrillary gliosis throughout the area of myelin loss is usual and oligodendrocytes are depleted (Lumsden, 1970). Gliosis gives a toughened appearance to old MS lesions. It was this feature that led to the use of the term plaque to describe them. There is evidence (Fontana *et al.*, 1980) that the morphology of astrocytic glia is altered by exposure to lymphokines, and it is conceivable that the prominence of gliosis in MS plaques reflects this fact.

Axons tend to be spared, although this is far from absolute. At autopsy, plaques are invariably found that were silent during life. From this it is evident that axons denuded of myelin still can conduct nerve impulses. They do so inefficiently, however, and tend to fail on sustained effort. This forms the basis for the inordinate fatigue so frequently complained of by victims of MS.

D. Immunology

1. Immunological Features

a. Spinal Fluid. The spinal fluid in MS contains an abnormally high content of IgG. A similar increase in IgG is seen in chronic infections of the nervous system. Thus, the finding is not unique for MS. When spinal fluid from MS patients is resolved electrophoretically or by isoelectric focusing, distinct IgG bands are observed in 90% of established cases. These bands bespeak expansion of selected clones of plasma cells in MS brain. The antigen or antigens against which these clones are directed is unknown, and the basis for IgG banding in MS remains unresolved. There are two hypotheses. (1) The bands constitute monoclones of antibody directed against an unknown MS antigen. The argument rests on a presumed analogy between MS and chronic CNS infections. In subacute sclerosing panencephalitis, a chronic measles virus infection of the brain, oligoclonal bands of IgG are detected in spinal fluid and can be shown to be antimeasles antibody. Cross-idiotypy can be shown in a substantial proportion of subacute sclerosing panencephalitis specimens as expected when response to a common antigen is occurring (Ebers *et al.*, 1979).

There is no cross-idiotypy between IgG bands from different MS patients (Ebers *et al.*, 1979; Gerhard *et al.*, 1981), however, and despite vigorous effort, all attempts to find an antigen against which MS bands react have failed. (2) Some and perhaps all of the IgG produced in MS brain is "nonsense antibody" in terms of disease pathogenesis, that is, a nonspecific activator of B cells is present in MS brain. Antibodies to several common viruses such as measles, mumps, and herpes are synthesized in MS brain. While any one of these viruses could be the elusive MS antigen, it is improbable that all can be, and many patients make antibody to several viruses. The antiviral antibodies are not found in the major monoclonal peaks, rather, they are minor monoclones only resolvable from the background of polyclonal IgG (also synthesized in brain in MS) by special techniques (Vartdal *et al.*, 1980). Aliquots of IgG eluted from different plaques from the same autopsy specimen contain different monoclones (Mattson *et al.*, 1980). These findings are consistent with the nonsense antibody hypothesis. The hypothesis begs the issue of what initially draws B cells into MS lesions.

Lymphoid cells are modestly increased in the spinal fluid in MS. Helper T cells predominate particularly when disease is ongoing, and many of them are "activated" not only when disease is progressing but also during periods of seeming quiescence (Noronha *et al.*, 1980).

b. Blood. Nonspecific T-suppressor-cell function as measured *in vitro* is subnormal during MS attacks, recovers as attacks end, and is normal when disease is in remission (Arnason and Antel, 1978; Antel *et al.*, 1979). T-suppressor-cell number (counted as cells binding the monoclonal antibodies OKT5 or OKT8) is low in progressive MS, with selective loss of cells binding large amounts of monoclonal antibody (Reinherz *et al.*, 1980; Bach *et al.*, 1980; Reder *et al.*, 1983). The correlation between T-suppressor-cell counts and T-suppressor-cell function measured *in vitro* is poor, suggesting that quantitative variations in T-suppressor-cell number fail to account for the deficiencies in function.

The findings just outlined could reflect some epiphenomenon; for example, a noxious substance released from damaged brain could compromise T-suppressor function. Alternatively, a loss of T-suppressor activity occurring for whatever reason could, by perturbing immune homeostasis, permit an autoimmune process to break free from the checks that otherwise hold it in abeyance. A third possibility is that suppressor cells could be lost from the blood because they have entered the brain. Suppressor-cell percentages fall in the spinal fluid at the onset of MS attacks, a finding difficult to reconcile with this proposal (Oger *et al.*, 1982; Cashman *et al.*, 1982). Lastly, both suppressor cells and myelin (and/or oligodendrocytes)

could be targets in MS, that is, the disease while appearing to be exclusively one of the CNS could in fact be a disease of both nervous and lymphoid systems. This formulation presupposes an antigen shared by brain and T-suppressor cells. There is already evidence that natural killer cells and myelin share an antigenic determinant (Helfand *et al.*, 1983) and that Ia and monocyte antigens are found on glia (Hauser *et al.*, 1983). An answer to the questions just posed would advance our understanding of the MS process.

Total T-cell number in the circulation (measured as cells binding the OKT3 monoclonal antibody) is reduced when MS is active, but T-helper-cell number (OKT4+ cells) is normal as is helper-cell function measured *in vitro*.

The number of monocytes in the circulation is increased in MS, and monocyte function is abnormal during acute flares of disease (Dore-Duffy and Zurier, 1981). B-cell capacity to synthesize IgG *in vitro* in response to stimulation with pokeweed mitogen is increased in MS (Goust *et al.*, 1982; Oger *et al.*, 1980). This response is influenced by T-helper and T-suppressor-cell function and may reflect no more than the T-suppressor-cell abnormalities already discussed. Consistent with this view is the observation that when B cells from MS patients are driven with salmonella antigen (which does not require T-cell help), the response is if anything subnormal (Wrabetz *et al.*, 1982). There is some evidence that natural killer-cell function may be subnormal when MS is active (Benczur *et al.*, 1980) and that K-cell function is increased in MS attacks (reviewed in Oger *et al.*, 1983).

2. Immunogenetics

Among northern Caucasians with MS, the *A3*, *B7*, *DR2*, and *Dw2* histocompatibility (*HLA*) alleles are overrepresented. Relative risks for MS are 1.5–2.0 for *A3*, 2.5–3 for *B7*, and 4–5 for *DR2,Dw2* (Jersild *et al.*, 1975; Batchelor *et al.*, 1978). Linkage disequilibria exist between A3, B7, and DR2;Dw2. The data have usually been interpreted as follows: An MS susceptibility immune-response allele is linked to *HLA* and maps closer to the *D* locus than to the *A* or *B* locus. The HLA association is far from absolute. This suggests that there may be more than one MS susceptibility allele or that in some cases the MS susceptibility allele is on haplotypes lacking *B7*, *DR2*, or *Dw2*. In addition, it is possible that in some MS victims a strong environmental stimulus causes a breakthrough to disease in the absence of a disease susceptibility allele. *B8* and *DR3* are also overrepresented in MS populations; it has been proposed that an allele sited between *B8* and *DR3* (*B8* and *DR3* are in linkage disequilibrium)

augments susceptibility to MS; that is, that gene complementation may be in play. Representation of *A2, B12,* and *DR7,Dw7* (these alleles are in linkage disequilibrium) is decreased in MS. It has been suggested that a protective allele on the *A2–B12–DR7,Dw7* haplotype may exist (Madigand *et al.,* 1982).

Among American blacks, MS is one-half as common as among American Caucasians. It is unknown among black Africans in whom *B7* and *DR2* are not uncommon alleles; *Dw2* in contrast is rare (i.e., in this group *Dw2* is not in linkage disequilibrium with *B7* or *DR2*). Among American blacks with MS, a marked increase in *Dw2* has been noted (i.e., a Caucasian gene is found in this black population) (Dupont *et al.,* 1977). Caucasian genes are more frequent in northern than in southern American blacks; this could bear on the greater frequency of MS among them.

In Japan, MS is rare, as is *B7* (1% of the population), and in the Japanese *DR2* is not linked to *Dw2*. Among Arabs with MS, a strong association with BT102 (a marker for *DR4,DR7*) has been reported with no association with *DR2* (Kurdi *et al.,* 1977). Data at hand fail to indicate whether a single MS susceptibility allele is linked to different markers in Oriental and Occidental populations or whether there may be more than one.

Among first-degree relatives of MS patients, risk for MS is increased some 15- to 20-fold over that of the population at large. Ten percent of MS patients have an affected family member. Analysis of linkage between disease and *HLA* haplotypes of sibling pairs, both with MS, has revealed a link between MS and *HLA* endowment (Stewart *et al.,* 1981). At the same time MS does occur in siblings lacking any shared haplotype, suggesting that a susceptibility allele may not always be essential for development of MS.

The Gm 1,17:21 phenotype of IgG is found more frequently in MS victims than in controls (Pandey *et al.,* 1981). It will be of interest to determine whether *HLA* type and *gm* allotype analysis can be combined to provide a better predictor of risk for MS than does either marker alone.

3. Laboratory Diagnosis

a. Spinal Fluid. The protein content of the spinal fluid is low in health. Normal values for total protein range from 15 to 45 mg/100 ml for fluid obtained from the lumbar space. Most spinal fluid protein is a transudate from serum; the low levels of protein ordinarily found in CSF attest to the efficacy of the blood–brain barrier (BBB) in minimizing entry of proteins

into the brain extracellular space and into the spinal fluid with which brain extracellular space communicates. Entry of serum proteins into the CSF is a function of their hydrodynamic radii, which correlates roughly with molecular weight (MW). For albumin (MW 67,000), the ratio between spinal fluid and serum is normally ~1 : 200; for IgG (MW 154,000) the ratio is ~1 : 400. Given these ratios, it should be evident that spinal fluid protein content is directly dependent on serum protein levels. If serum albumin is low, spinal fluid albumin will also be low. If serum IgG is elevated, the spinal fluid-IgG content may rise to what would ordinarily be construed as a pathological level, even while the 1 : 400 spinal fluid/serum ratio for IgG remains preserved. For these reasons, measurement of serum protein levels should be performed in conjunction with spinal fluid analysis.

Spinal fluid protein begins to rise slightly at about age 40. This continues with advancing age, even though serum protein levels change minimally with age. For this reason, a mildly abnormal result in terms of spinal fluid-IgG in an older person should be interpreted cautiously.

When spinal fluid-IgG is elevated beyond certain normative values, pathology can be assumed provided serum IgG is normal. Spinal fluid-IgG can be elevated under three circumstances: (1) when the BBB is disrupted; (2) when local synthesis of IgG within the CNS is occurring; and (3) when both barrier disruption and local IgG synthesis occur simultaneously. In MS, the chief value of spinal fluid-IgG analysis lies in establishing that IgG is being synthesized within the CNS, that mechanism (2) or (3) is operating rather than mechanism (1).

Numerous formulas have been derived to distinguish among the three mechanisms that can account for an increased IgG in CSF. The simplest formula is the spinal fluid-IgG/spinal fluid-total protein ratio. In normals, this ratio does not exceed 0.08. When the total protein level in spinal fluid does not exceed 45 mg/100 ml, the BBB can be presumed to be intact, and an elevated CSF-IgG/CSF-total protein ratio points to *in situ* IgG production, provided the serum IgG level is normal. The ratio is abnormal in 70% of MS patients. Problems in interpretation may arise when the BBB is disrupted and spinal fluid total protein is elevated; this occurs in 30% of MS patients. Of MS patients, 5% have a total protein >90 mg/100 ml, 0.5%, a total protein >110 mg/100 ml. In the face of a BBB disruption and the raised spinal fluid total protein which this entails, the preferential exclusion of IgG over lower MW proteins from the CSF may no longer hold. Under this circumstance, the spinal fluid-IgG/spinal fluid-total protein ratio may be skewed, and for this reason, subject to misinterpretation. The problem is not trivial; erroneous diagnoses have been made because this consideration was ignored.

A second commonly employed formula is that of spinal fluid-IgG/spinal fluid-albumin. This ratio, which should not exceed 0.16, is abnormal in 80% of MS patients. The formula does not correct for increased serum IgG (occasionally present in MS) or for lowered serum albumin (also occasionally present in MS), both of which can skew this index.

A third formula is the spinal fluid-IgG index. This is calculated as

$$\frac{\text{Spinal fluid-IgG (mg/100 ml)/spinal fluid-albumin (mg/100 ml)}}{\text{Serum IgG (mg/100 ml)/serum albumin (mg/100 ml)}}$$

Spinal fluid-IgG index values greater than 0.58 point to IgG synthesis within the CNS. An elevated spinal fluid-IgG index is found in 85% of MS patients. The index corrects for increased serum IgG, for lowered serum albumin and also to some extent for BBB damage.

Spinal fluid-IgG tends to be higher in patients with severe and extensive disease, although exceptions occur. During a first attack of MS, the IgG content may not be elevated. There is no clear correlation between IgG level and disease activity. In a suspected case of MS, a diagnosis of MS is probable if any of the indexes given above are grossly abnormal.

Electrophoretic analysis has revealed that the spinal fluid-IgG in MS has unusual characteristics helpful in diagnosis. Normally, CSF-IgG migrates electrophoretically as a broad smear, but in 85–90% of MS patients, discrete oligoclonal bands superimposed on a background of homogeneous IgG are detected. The finding, while not unique for MS, is suggestive of that diagnosis if the clinical presentation fits. Similar oligoclonal banding patterns are noted in chronic infections of the CNS. Also, 5% of spinal fluid samples obtained from normal individuals show oligoclonal bands. Whether or not this reflects past infection of the CNS, subclinical or forgotten, is unknown. During the first year or two of MS, band number may be lower than later in the illness, but even early most patients show bands. After the first 2 to 3 years of illness, total band number does not change appreciably, but band intensity may fluctuate over time, and some bands may vanish to be replaced by new ones. There is no relation between disease activity and band number or the appearance of new bands. Each patient has a unique band pattern.

Spinal fluid can also be analyzed by isoelectric focusing. Samples of CSF are subjected to an electric field in a gel that spans a pH gradient of 3 to 10. Proteins in this field focus at their isoelectric point. IgG moves cathodally to lie between pH 7.0 and 9.5, and in MS spinal fluid it focuses into a dozen or more discrete bands. The test is positive in 90+% of MS patients, and when combined with the extremely sensitive immunoperoxidase method that identifies IgG exclusively, can be performed on as little

as 50 μl of unconcentrated spinal fluid. Each patient has a unique focusing pattern.

In MS, the cell count in spinal fluid is modestly elevated; 5–10 cells/mm^3 are commonly found. Higher cell counts should alert the clinician to the possibility of some other diagnosis. The cells are chiefly T lymphocytes. Subset analysis has not proven useful in terms of clinical diagnosis.

b. Blood. As discussed, T-cell-subset enumeration reveals abnormalities in MS populations, as does *in vitro* study of T-suppressor-cell function. These tests have not proven of value in the analysis of individual cases.

c. Scanning. A small proportion (20–30%) of MS patients show demyelinating lesions in brain on a CAT scan if a double dose of dye is infused and delayed examination is performed. Preliminary experience with nuclear magnetic resonance scanning has given a high yield of visible lesions within brain in MS (Buonanno *et al.*, 1983). Unfortunately, the test is not available generally.

d. Evoked Potentials. These are abnormal in 80% of MS patients. A detailed discussion of this and other electrophysiological techniques relevant to MS lies beyond the scope of this chapter (for discussion, see Eisen, 1983).

E. Treatment

There is no established treatment for MS. Administration of ACTH has been shown to ameliorate the symptoms of acute flares of disease, but this exerts no long-term beneficial effect (Rose *et al.*, 1970). Prednisone, beginning with 10–15 mgm every 6 h and tapering over 2–3 weeks, will also lessen symptoms during an exacerbation. Long-term glucocorticoid treatment is seldom if ever warranted. Several immunosuppressive regimens have been advocated in MS. Included are azathioprine (2–3 mg/kg/day) administered over the long term and cyclophosphamide (150 mg every 6 h) for 10 days to 2 weeks. These treatments should be viewed as experimental at the present time.

In addition to therapy aimed at the presumed underlying immune disorder, there are several drugs that may ameliorate specific symptoms. Baclofen may lessen spasticity; clonazepam may lessen tremor; and tricyclic antidepressants are useful in managing depression. For a more detailed discussion of the treatment of MS, the interested reader may consult Oger and Arnason (1980).

II. PERIVENOUS ENCEPHALOMYELITIS (PVE)

A. Introduction

This disease, also known as acute disseminated encephalomyelitis, postinfectious encephalomyelitis, and postvaccinial encephalomyelitis, is a multifocal process of abrupt onset characterized by symptoms and signs pointing to damage chiefly to white matter of the brain and spinal cord; it is nearly always monophasic. The process may be severe, even fatal, or mild and evanescent. Formerly PVE was encountered most commonly following vaccinations against rabies or smallpox and as a complication of the exanthems, especially measles. Presently, PVE most commonly follows upper respiratory infection; it also occurs without any evident antecedent. The cause is believed to be a hypersensitivity to MBP, and the disease is held to be the human counterpart of the acute demyelinating disease induced in animals by immunization with MBP and known as experimental allergic encephalomyelitis (EAE).

B. Clinical Description

Perivenous encephalomyelitis begins abruptly. Headache and delirium give way to lethargy and coma. Seizures may occur at onset, and stiffness of the neck may be present, as may fever. Focal signs may be engrafted onto this encephalitic picture; spinal cord involvement with flaccid paralysis of the legs or of all four limbs with loss of sphincter control are not infrequent. The spinal fluid in most cases shows an increase in protein (50–100 mg/100 ml), and lymphocytes are present, ranging from a few to several hundred cells. Less commonly, the spinal fluid is normal.

Diagnosis is straightforward when there is a history of rabies, of smallpox vaccination, or of measles. In the vast majority of cases without such a history, distinction of PVE from viral encephalitis, from Reye's syndrome, or from acute MS may not be possible. Perivenous encephalomyelitis is less common now than formerly. Replacement of nervous tissue containing rabies vaccines with duck embryo- or tissue culture-based preparations has drastically reduced this complication of rabies prophylaxis. Eradication of smallpox has ended the need to vaccinate against it, and measles vaccination has markedly reduced what used to be the largest group of PVE cases.

C. Histopathology

The picture is one of innumerable minute foci of perivenular mononuclear cell infiltration, all of like age, and with demyelination, the topogra-

phy of which corresponds to that of the inflammatory infiltrates (DeVries, 1960). Foci of demyelination resemble a string of beads along the course of a venule. The cellular infiltrate consists of lymphocytes and macrophages. Axons are relatively spared in regions of demyelination; ultimately, small glial scars form.

D. Immunology

The lesions of PVE duplicate those of acute EAE (see Section V). In EAE, sensitivity of T lymphocytes to MBP can be shown. Lymphocytes from EAE animals proliferate on exposure *in vitro* to MBP. Proliferation of peripheral blood lymphocytes on exposure to human MBP has been reported during acute PVE; this finding argues for the similarity, if not identity, of the two processes (Lisak *et al.*, 1974).

Formerly, rabies vaccines contained brain tissue plus rabies virus, but PVE has followed inoculation of noninfected brain material. This observation indicates that brain antigen rather than virus was responsible for PVE. Brisk inflammatory reactions of cell-mediated immunity type at sites where vaccine had been injected have occurred coincident with the onset of PVE, indicating a response to an antigen present both in the vaccine and in the CNS. The marked reduction in neuroparalytic accidents that followed introduction of duck embryo-based rabies vaccines in the 1960s reinforces this conclusion. Neuroparalytic accidents following rabies vaccination were most frequent in young adults, peak age of occurrence duplicating that of the onset of MS. This finding indicates that both host and vaccine were important in the genesis of neuroparalytic accidents.

In former times, smallpox vaccination was complicated by PVE in perhaps 1 in 5000 vaccine recipients, with marked differences between vaccine lots and vaccination programs. The complication almost always occurred in conjunction with a primary take rather than with a booster response. It was never seen in vaccinees aged <2 years (DeVries, 1960). Most often, PVE followed the peak of the vaccination response by a few days to a week or more but occasionally preceded it.

One case of measles in 1000 is complicated by neurological signs and symptoms. The latent period between the rash and PVE varies from -2 to $+13$ days; 90% of cases develop between $+2$ and $+7$ days. Mortality averages 20%; half the survivors are left with significant residual damage. The complication is more frequent in older subjects of the exanthem, that is, those over 10 years of age. The severity of PVE does not relate to the severity of the measles itself. Sensitivity to MBP has been shown in measles-associated PVE (R. Johnson, personal communication).

E. Treatment

Administration of high doses of glucocorticoids every 4 to 6 h is the treatment of choice, although controlled trials of this treatment have not been carried out.

III. ACUTE NECROTIZING HEMORRHAGIC LEUKOENCEPHALITIS (ANHL)

Very rarely a catastrophic tissue-destructive CNS syndrome with unique features occurs within a few days of an upper respiratory tract infection, usually "influenzalike." The illness resembles PVE save for the intensity and speed of progression of the process, which can lead to death within 48 h. Fever and neutrophilic leucocytosis are the rule, and the spinal fluid shows increased protein, up to 2000 polymorphonuclear cells and often red blood cells as well. Pathological examination reveals hemorrhagic areas, perivascular mononuclear and polymorphonuclear infiltrates, fibrin exudation, and frank necrosis of extensive areas of the white matter. The disease is thought to be a PVE variant with a Shwartzman-type reaction superimposed onto the basic perivenular demyelinating process. A model for it has been developed (see Section V,B,3).

IV. TRANSVERSE MYELITIS AND NECROTIZING MYELOPATHY

The term transverse myelitis is used operationally to describe a group of cases in which ascending and descending spinal tracts are interrupted relatively acutely by an intrinsic cord lesion of presumed demyelinating type. A transverse myelitis may occur in the course of MS, although at autopsy MS has seldom been found in cases presenting with an acute transverse myelitis as their initial complaint. Similarly, a transverse myelitis can be the major or only clinically apparent feature of PVE, but most cases of transverse myelitis and particularly those with extensive necrosis of the cord (which may involve gray matter as much as white and is characterized by macrophagic but not lymphocytic infiltrates) cannot be fit into the PVE category with assurance. Approximately 30–40% of cases will give a history of an antecedent infection, usually respiratory. Other associations have been noted. These include active pulmonary tuberculosis, heroin abuse, systemic lupus erythematosus, and occult malignancy,

particularly lymphoma and oat-cell carcinoma of the lung. All of the above, particularly the tie to SLE (the prototype autoimmune disease), and to cancer, several of the other remote effects of which have an immunological basis, suggest to me that an immunological process distinct from MS and from PVE is likely to be in play. Some progress has been made in developing a model for necrotizing myelopathy based on the EAE system.

V. EXPERIMENTAL ALLERGIC ENCEPHALOMYELITIS (EAE)

A. Introduction

Experimental allergic encephalomyelitis (EAE) was first described in 1933 (Rivers *et al.*, 1933) and during the ensuing 50 years several thousand studies of EAE have been published. Several different forms of EAE have been described. These mimic, to greater or lesser extent, the various human demyelinating diseases that have been discussed in the preceding sections. For clarity of exposition, the immunology of the various EAE models will be considered together in this section.

B. Immunology

1. Acute Monophasic EAE

Acute EAE is usually induced by immunization with MBP in Freund's complete adjuvant (a water in mineral oil emulsion containing killed tubercle bacilli). The immune response is T-cell mediated. T-cell clones that recognize epitopes on MBP are present in the lymphoid organs of all mammalian and at least some avian species. These clones proliferate upon exposure to MBP, and after expansion, MBP-activated T cells migrate, as blasts, through the circulation and across CNS venules into the nervous system parenchyma. In rats and mice the T effector cells responsible for EAE appear to belong to the helper/lymphokine-producing type (Holda and Swanborg, 1982; Pettinelli and McFarlin, 1981). The T cells in the brain attract monocytes that enter brain, transform there into macrophages, and are the final vectors of the destruction of myelin, which they strip from the axons, ingest, and digest. Some B cells enter the brain during acute EAE; the B cells remain confined to the perivascular spaces

and the meninges. They are the source of the immunoglobulin produced within the CNS in EAE. Oligoclonal IgG bands make their appearance in the spinal fluid as the acute process abates (Whitacre *et al.*, 1981). The bands are not antibody to MBP (Whitacre *et al.*, 1982). T cells predominate in the spinal fluid in acute EAE; B cells and macrophages are rare (Wilkerson *et al.*, 1978).

Acute EAE begins abruptly 8–20 days after immunization. Disease lasts for a few to several days, and unless death occurs in the acute phase, recovery follows. Recurrences are rare, and attempts to induce a second bout of disease by reimmunization fail.

The pathological feature is that of an acute multifocal perivenular inflammatory response, with small areas of demyelination at sites of inflammatory response. Sites of predilection vary between species. In rats, for example, a transverse myelitis-like picture is the norm. The pathology exactly duplicates that of PVE, as does the monophasic nature of the disease.

T cells from animals with acute EAE will adoptively transfer disease to virgin histocompatible recipients. In contrast, cells from recovered animals, when adoptively transferred, will (at least in the hands of some investigators) protect against active induction of disease in the recipients (Bernard, 1977). This last finding has been interpreted to mean that refractoriness to further attacks of EAE relates to development of a protective mechanism. The basis for protection is thought to depend on the expansion of a population of suppressor T cells, which hold in check the effector T cells responsible for tissue destruction in EAE (Welch *et al.*, 1980). There is some evidence that the suppressor cells recognize MBP (Hashim, 1981); others believe that they recognize the MBP receptors on T effector cells, that is, that antiidiotypy is in play (Ben-Nun *et al.*, 1981a,b). The two mechanisms are not mutually exclusive. Were a suppressor T cell-dependent mechanism comparable to that seen in EAE, to be operating in PVE, then the monophasic nature of PVE might find its logical explanation.

2. Hyperacute EAE

Hyperacute forms of EAE have been induced. Usually this requires administration of selected bacteria (e.g., pertussis) or of endotoxin in addition to MBP. When this is done, extensive fibrin deposition, frank vessel disruption, parenchymal hemorrhages, and areas of frank necrosis are engrafted onto the basic EAE lesion (Levine and Sowinski, 1977). The situation can be likened (in my opinion, although there is disagreement on

this point) to that seen in the Shwartzman reaction, and this provides an experimental model for ANHL in humans.

3. Myelomalacic EAE

Levine and Sowinski (1976) produced a model disease in rats that is reminiscent of necrotizing myelopathy. Lymph node cells from Lewis rats with EAE were transferred into virgin Lewis (genetically identical) recipients. Ordinarily, this maneuver leads to acute EAE in the recipients. The lesions in recipients consist of acute perivenular lymphoid infiltrates that appear, are present for a few days, and then regress. When recipient rats were treated with tilorone (an immunosuppressive drug) the day before transfer, acute passively transferred disease was no longer observed; rather, a progressive myelopathy was noted. Histologically, perivascular cuffs were few in number and scanty in content, but a massive influx of macrophages into the cord was noted, and cord softening in regions of macrophage infiltration was observed. The myelomalacia involved grey matter more than white; in some instances long lesions affecting the entire lower half of the cord, usually single, but at times multiple, were noted. It was shown, by killing recipients at various intervals after transfer, that monocyte infiltration preceded tissue necrosis.

Levine and Sowinski (1976) argued that perivascular lymphocytic cuffs, so characteristic of the usual forms of EAE, may be as protective as noxious. In adoptive transfer lesions of EAE, most perivascular cells are of host rather than of donor origin. Thus, they suggested that tilorone, by depleting host T cells, prevents their protective action and permits a necrotizing process to proceed. Whatever the explanation for their observation may ultimately prove to be, the data suggested that an immune basis for necrotizing myelopathy cannot be set aside simply because a lymphocyte response is lacking or minimal.

4. Chronic and Recrudescent EAE

Strain 13 guinea pigs develop acute monophasic EAE if immunized as adults with myelin or with MBP, but if immunized with myelin as juveniles they develop chronic and/or recrudescent EAE (Wisniewski and Keith, 1977; Raine, 1983). The lesions of the chronic disease resemble those of MS. Some animals develop an acute first attack in which cellular infiltration is prominent but demyelination is not. The picture here is that of acute EAE or of PVE. Remission follows but leads into recurrence, and with recurrence the morphological picture changes. Cellular infil-

trates are less intense than in the first attack, but demyelination becomes more extensive and gliosis increasingly evident (Lassmann and Wisniewski, 1978, 1979). Whether sensitivity to MBP is responsible for recurrent EAE in guinea pigs or whether some other antigen is in cause is unclear at present. In mice a chronic form of EAE in which sensitivity to MBP alone is said to be responsible has been reported (Lublin, 1983). Regardless, the model demonstrates that an autoimmune process can mimic MS.

5. Immunogenetics of EAE

Some strains of animals are susceptible to EAE, others resistant. Susceptibility links to the major histocompatibility complex (MHC) as a dominant trait, but MHC involvement does not totally determine susceptibility (Moore *et al.*, 1980; Bernard, 1976; Lando *et al.*, 1980). In the mouse, a second gene is implicated in susceptibility to EAE. This gene determines response to injected histamine, mice with the appropriate MHC and histamine sensitivity alleles being maximally susceptible to EAE (Linthicum and Frelinger, 1982). Other genes may also influence susceptibility to EAE. Female animals, for example, develop more severe EAE than males, and there is evidence that naturally occurring suppressor cells may, in some mouse strains, override the influence of a *MHC* gene that would otherwise permit susceptibility.

C. Treatment

Most work has been done with acute EAE. Administration of MBP in incomplete adjuvant will protect against subsequent attempts to actively induce disease. The treatment leads to the generation of T suppressor cells. T-cell clones reactive with MBP, if irradiated and given in Freund's complete adjuvant, protect against subsequent attempts to induce EAE in the cell recipients (Ben-Nun *et al.*, 1981a,b). The effect has been held to depend on the generation of an antiidiotypic response.

A wide range of immunosuppressive regimens has been shown to prevent or attenuate acute EAE. Included here are cyclophosphamide, azathioprine, cyclosporin A, and interferon. All four of these agents are currently under test as therapies for MS. Attempts to induce antigen-specific suppressor cells and antiidiotypic protective mechanisms in MS will have to await the discovery of an MS-relevant antigen.

For more comprehensive discussions of EAE than can be attempted here, the reader might profitably consult the reviews of Paterson (1976), Wisniewski *et al.* 1982, and Raine (1983).

VI. DEMYELINATING NEUROPATHIES: ACUTE INFLAMMATORY DEMYELINATING POLYRADICULONEUROPATHY (AIDP)

A. Introduction

This entity is most commonly known as the Guillain–Barré syndrome. It occurs sporadically throughout the world and at all ages. The clinical picture is of a subacute, primarily motor neuropathy with raised spinal fluid protein and few cells. Nerve condition velocities are slowed in most cases. Two-thirds of cases come on within days to weeks of an acute infection, usually a banal upper respiratory infection. Some patients mention an acute dysenteric illness. In most instances the offending agent is not known. Of known antecedents, cytomegalovirus, Epstein–Barr virus, and *Mycoplasma pneumoniae* infections are the most frequently encountered (Dowling and Cook, 1981). Other viral infections that may be associated with AIDP include smallpox–vaccinia, varicella–zoster, hepatitis, and mumps. Five percent of cases follow surgery (Arnason and Asbury, 1968), and a flurry of cases followed A New Jersey swine influenza vaccination in the United States in 1976 (Schonberger *et al.*, 1979). Several cases occurring on a background of lymphoma, particularly Hodgkin's disease, have been put on record (Lisak *et al.*, 1977). The disease is monophasic in the vast majority of cases.

B. Animal Models

Several models for AIDP exist. The best studied is experimental allergic neuritis (EAN), a disease induced by immunization with peripheral nerve myelin or with the P2 protein of peripheral nerve myelin. Cell-mediated immunity to P2 protein is important for the genesis of lesions in this disease, and the clinical, pathological, and electrophysiological features of EAN faithfully correspond to those of AIDP (Waksman and Adams, 1955; Brostoff *et al.*, 1972). Experimental allergic neuritis can be adoptively transferred with lymphocytes but not with serum, and skin tests with P2 protein provoke a characteristic delayed hypersensitivity type response (Kadlubowski *et al.*, 1980). The data suggest that EAN is a T-cell-mediated disease and that, arguing by analogy, AIDP may be as well. Human cases of inflammatory neuritis that are clinically and pathologically indistinguishable from AIDP have followed administration of nerve-containing vaccines (Lopez-Adaros and Held, 1971). It seems probable that such instances, in fact, are EAN in humans. In AIDP itself,

however, CMI response to P2 protein is absent. Thus, the features of AIDP and EAN are similar save for the antigen involved.

A second model for AIDP is provided by Marek's disease of the chicken. This is a subacutely evolving demyelinating neuritis that occurs in birds infected with an oncogenic herpes virus. Marek's virus also causes lymphoid tumors, as does the Epstein-Barr virus in man. The virus matures to a cell-free infectious enveloped form in the epithelial cells of the feather follicle and is shed with the dander. Infection is thought to follow aspiration of infectious dander. In infected birds it is extremely difficult to demonstrate the virion in tissues other than the hair follicle epithelium. Virus, however, can ultimately be recovered from nerve explanted onto chicken kidney cell monolayers, that is, nerve is latently infected possibly by viral material carried into nerve by lymphocytes, themselves latently infected. If dorsal root ganglia from infected birds are maintained in tissue culture for several days, virions appear but they are confined to satellite cells, nonmyelinating Schwann cells, and lymphocytes. Virus is not seen in myelinated Schwann cells. Thus, although nerve is latently infected, the infection is in "the wrong cells." Since inflammatory infiltrates in nerve have been shown to precede the appearance of viral material, I conclude that something other than an antiviral defense response is in play in the genesis of the neuritis.

The pathological lesions of Marek's disease neuritis are indistinguishable from those of AIDP. The invading cells within nerve consist of small and medium-sized lymphocytes, blasts, monocytes, and macrophages. Myelin is stripped by the macrophages, and axons are largely spared. In Marek's disease, CMI and antibody responses to myelin can be demonstrated (reviewed in Stevens *et al.*, 1981). Marek's disease occurs in birds from which the bursa of Fabricius has been removed. Such birds cannot generate an antibody response but CMI responses are preserved. This finding establishes that a CMI response alone can suffice for an inflammatory demyelinating neuritis.

It is of interest that lymphocytes latently infected with the Marek's disease virus express a cell-surface antigen that is specific. Cell-mediated immune and antibody responses to this surface antigen are present in Marek's virus-infected birds. Whether the antinerve and antilymphocyte responses are directed against cross-reactive or distinct antigens is not known. Note that cytomegalo and Epstein–Barr viruses, both of which are tied to AIDP, are herpes viruses like Marek's virus, and that all three viruses cause latent infection of lymphocytes.

A third model demyelinating neuritis follows repeated immunization of rabbits with galactocerebroside (GC). Galactocerebroside is a component of both CNS and peripheral nerve myelin, but the disease is confined to

the peripheral nerves. GC-induced neuritis is antibody mediated. Macrophages are found in nerve but lymphocytes are absent, that is, the pathological features are *not* those of AIDP (Saida *et al.*, 1981). Antibodies to GC are not detected in AIDP.

C. Clinical Presentations

Acute inflammatory demyelinating polyradiculoneuropathy is an uncommon disease. Epidemiological data indicate that approximately one person per million population develops AIDP each month. Several hundred cases of AIDP occurred in the United States in 1976 as a complication of the national swine influenza vaccination campaign. The rate of AIDP following swine influenza vaccinations has not exceeded background levels in subsequent years and just why the vaccination program of 1976 had this complication in the United States, although not in other countries, remains a mystery.

The illness evolves over days to 2–3 weeks. Motor weakness predominates, and reflexes are characteristically lost. In 20% of cases, weakness becomes so profound that assisted respiration becomes necessary. Weakness tends to be symmetrical. No tie of susceptibility to HLA endowment has been found.

Ataxia may be prominent early in disease and sensory symptoms can rarely be the mode of presentation, but weakness invariably supervenes. Objective sensory loss is usually less extensive than subjective complaints would lead one to anticipate. Facial diplegia is present in half the cases. Autonomic function is often abnormal and may be expressed in multiple ways. For a more detailed review of the autonomic manifestations of AIDP, the reader may consult a recent review (Arnason, 1983). For a set of diagnostic criteria, Asbury *et al.*, (1978) should be consulted.

D. Histopathology

The pathological feature of AIDP is perivenular inflammatory infiltrates scattered randomly through the nerves with accompanying segmental demyelination (Asbury *et al.*, 1969; Prineas, 1981). Blood-derived lymphocytes and macrophages predominate among invading cells; macrophages appear to be the final vector of myelin destruction. The macrophages break through the basement membranes that surround Schwann cells, insert tongue-like processes between myelin lamellae, sequester myelin fragments, and ingest and digest them. The mechanism underlying myelinolysis by macrophages is uncertain; lymphokine-mediated cytotoxicity may be involved. Within 2–3 weeks in 50% of AIDP cases, myelin

destruction has ceased and repair begins. Repair is heralded by Schwann cell proliferation, which is followed by laying down of new myelin. The new myelin is thinner than that it replaces, and the internodal distances are shortened. In severe cases there may be substantial axonal disruption in addition to demyelination. The pathological features are identical with those of EAN.

E. Immunology

Immunology

a. Blood. In blood there is usually a moderate polymorphonuclear pleocytosis; T cells are reduced and B cells are increased, particularly early in the disease. One group (Goust *et al.*, 1978) has reported that T-suppressor-cell function is defective; T-suppressor-cell counts are normal (H. Wiener, personal communication). Activated lymphocytes are found in the blood; possibly they represent a response to infection. No genetically determined predisposition to develop disease has been detected in AIDP; histocompatibility profiles mirror those of the population at large.

Proliferative response of peripheral blood lymphocytes to crude nerve antigens has been demonstrated by several techniques. Two groups have reported CMI responses to the P2 protein of nerve (Sheremata *et al.*, 1975; Abramsky *et al.*, 1980), but three others have failed to find it (Iqbal *et al.*, 1981; R.A.C. Hughes, personal communication; R. Lisak, personal communication). The point is important since, as discussed, P2 induces experimental allergic neuritis (EAN), an animal disease that mimics AIDP in many ways. I conclude that P2 is not the AIDP neuritogen and that to the present the AIDP neuritogen remains unknown.

Serum immunoglobulins are increased in AIDP; IgG, IgA, IgM, and IgE all share in this response (Cook *et al.*, 1970; Huang, 1975). Antibodies that destroy myelin in nerve cultures have been detected by several groups. The antigen against which they are directed is not known. Antibody to P2 protein is not present (Iqbal *et al.*, 1981) nor is antibody to GC (Hughes, 1979).

Antibodies that lyse lymphocytes *in vitro* have been found in half the cases (Searles *et al.*, 1981). Similar lymphocytotoxic antibodies are found in other diseases, notably MS and systemic lupus erythematosus (SLE). In cerebral SLE, antibrain antibodies that cross-react with lymphocytes have been detected (Bluestein *et al.*, 1981). Whether the lymphocytotoxic antibodies present in AIDP cross-react with nerve is not known.

b. Spinal Fluid. Within a few days of onset the spinal fluid total protein is elevated. There is no disproportionate increase in IgG, and oligoclonal

bands of IgG are rare. The raised proteins bespeak damage to nerve roots with disruption of the blood–nerve barrier. Protein content rises over time, peaking at 3–4 weeks even as the patient is beginning to recover. In a few cases the spinal fluid protein remains normal. Cells in the spinal fluid are not ordinarily elevated, but up to 20 cells may be seen in some cases; T cells predominate.

c. Nerves. Immunoglobulin deposits in nerve and circulating antibodies that bind to myelin *in vitro* have been reported, using fluorescent methods, to be present in AIDP, but the specificity of the finding has been questioned (Hughes, 1979). Nyland and Aarli (1978) used an antiglobulin consumption test to demonstrate IgG binding (via the Fab portion of the molecule) to peripheral nerve in 15 of 30 AIDP cases. Controls were negative. Binding was detected early and was lost with recovery. Complement levels are not decreased in AIDP. Circulating immune complexes are found in some AIDP patients and have been invoked as responsible for the glomerulonephritis that may complicate the disease. Others have argued that renal complications in AIDP may be mediated by lymphokines (Froelich *et al.*, 1980).

As already discussed, GC-induced neuritis can be induced in rabbits by repeated immunization with GC. The disease is antibody mediated, and the morphological features of AIDP are not mimicked. The rat is refractory to actively induced GC neuritis, despite the fact that antibodies to GC are readily generated, presumably because the blood-nerve barrier is not readily breached in this species and the antibody is unable to reach nerve parenchyma. Lampert (1980) injected PPD intraneurally into tuberculin-sensitive rats. A CMI response in nerve ensued but demyelination was minimal. In tuberculin-sensitive rats immunized with GC, intense demyelination followed intraneural injection of tuberculin, that is, the CMI response to tuberculin and the antibody response to GC synergized. The possibility that two types of immune response, perhaps directed at two distinct neural antigens, may underly AIDP is only now beginning to be explored.

d. Laboratory Diagnosis. An elevation of total spinal fluid protein is usual but by no means specific for this disease. Nerve conduction velocity measurements show slowing in the vast majority of cases.

F. Treatment

There is no known specific treatment for AIDP. Formerly, glucocorticoids were advocated, but a controlled trial failed to reveal benefit

(Hughes *et al.*, 1978). The major concern is respiratory failure; treatment here is supportive, and assisted respiration may be required.

VII. DEMYELINATING NEUROPATHIES: RELAPSING AIDP

Approximately 3% of AIDP cases relapse months to years after recovery from the initial episode. Of relapsing cases, perhaps half have more than one relapse. Individual attacks do not differ from those of AIDP in their essentials, although the tempo of evolution may be somewhat slower, and with several attacks residual deficit may summate and the nerves become palpably enlarged. One attack may follow an infection, the next surgery, and vice versa, suggesting nonspecificity in the antecedents that provoke relapses. *HLA* typing has revealed an overrepresentation of the *A1*, *B8*, and *DRW3* alleles in relapsing AIDP and in CIDP (see below) (Stewart *et al.*, 1978; Adams *et al.*, 1979).

VIII. DEMYELINATING NEUROPATHIES: CHRONIC INFLAMMATORY DEMYELINATING POLYRADICULONEUROPATHY (CIDP)

This disease, like AIDP, is an inflammatory demyelinating polyradiculoneuropathy with raised spinal fluid protein and slowed nerve conduction. Cell-mediated immune response to nerve and antibody response to crude nerve have both been demonstrated in CIDP. In most instances CIDP is indolently progressive from the outset, but cases are on record in which AIDP or relapsing AIDP led into CIDP, and I view CIDP as an AIDP variant. The predominant pathological feature in regions of myelin destruction is one of lymphocytic and macrophagic infiltrates. Some cases respond to a greater or lesser extent to prolonged glucocorticoid treatment (Austin, 1958; Dyck *et al.*, 1982). There is a suggestion that antibody to nerve may have a major role in CIDP if reports of response of CIDP cases to plasmapheresis are taken at face value. Some cases do not respond to plasmapheresis.

IX. DEMYELINATING NEUROPATHY OF PLASMA CELL DYSCRASIA

Latov *et al.* (1980, 1981) drew attention to this entity. The picture is one of a mixed (motor and sensory) progressive neuropathy with slowed

nerve conduction. It has most often been associated with an IgM k gammopathy, and the IgM k monoclone binds to a 100,000-dalton myelin protein known as myelin-associated glycoprotein (MAG). This finding suggests that the monoclone is an autoantibody. There is evidence that patients may share idiotypy of their IgM clone (Dellagi *et al.*, 1979). IgM k deposits along myelin in nerve can be detected, and by electron microscopy a loosening of outer myelin lamellae can be shown (Propp *et al.*, 1975; Julien *et al.*, 1978; Khan *et al.*, 1980). There are no inflammatory cells. Some cases have responded to plasmapheresis or to treatment of their underlying gammopathy.

X. AXONAL AND NEURONAL NEUROPATHIES

A. Subacute Sensory Neuronopathy

This entity, which is characterized by progressive sensory loss and preserved strength, occurs most often on a background of cancer, usually oat-cell carcinoma of the lung (Horwich *et al.*, 1977). Spinal fluid protein is elevated in three-quarters of the patients, usually >100 mg/dl, and mild spinal fluid lymphocytosis is seen in a substantial proportion of cases. Pathological findings are distinctive with intense lymphocytic and macrophage infiltration of dorsal root ganglia and neuronal cell loss, often profound, in dorsal root ganglia. In burnt-out cases, neuronal depopulation and fibrosis are found in dorsal root ganglia (Horwich *et al.*, 1977). What is probably the same condition also occurs in the absence of tumor (Sherman *et al.*, 1980). In this regard, subacute sensory neuronopathy is not different from other paraneoplastic syndromes affecting the nervous system, each and all of which also occur in the absence of tumor. The oat-cell tumor is thought by some to be of neural crest origin. Certainly it can secrete neural products, and expresses surface antigens that cross-react with neural tissues (Bell and Seetharam, 1977). Possibly an immune response that controls the tumor (oat-cell cancers with paraneoplastic complications have a good prognosis insofar as the tumor is concerned) is misdirected against sensory neurons because of shared antigenic determinants. While most cases are subacute, indolently progressive forms are seen.

B. Lambert–Eaton Syndrome

In this rare condition, release of acetylcholine from the terminal axon is deficient. The clinical picture includes weakness and fatigability, particu-

larly of the muscles of the pelvic girdle and thighs, with associated dry mouth, limb paresthesiae, aching in the thighs, impotence, and weak or absent deep tendon reflexes. Electromyography reveals, with repetitive stimulation at rates >10 stimuli/sec, a diagnostic increase of muscle response. The condition responds to plasmapheresis, and injection of serum from afflicted patients into mice reproduces the electrophysiological features of the illness, suggesting that it may be antibody mediated (Newsom-Davis *et al.*, 1982). The antibody is thought to interfere with acetylcholine release by binding to the terminal axon. Seventy percent of cases have cancer, most commonly an oat-cell carcinoma. In cases with cancer, there is a 4:1 preponderance of males. In cases without tumor, no male preponderance is noted. Most patients respond to guanidine, a drug that favors acetylcholine release.

REFERENCES

Abramsky, O., Korn-Lubetsky, I., and Teitelbaum, D. (1980). *Ann. Neurol.* **8,** 117.

Adams, D., Festenstein, H., Gibson, J. P., *et al.* (1979). *J. Neurol. Neurosurg. Psychiat.* **42,** 184–186.

Antel, J. P., Arnason, B. G. W., and Medof, M. E. (1979). *Ann. Neurol.* **5,** 338–342.

Arnason, B. G. W. (1984). *In* "Peripheral Neuropathy" P. J. Dyck, P. K. Thomas, and E. Lambert, eds.), Saunders, Philadelphia, Pennsylvania. (In press.)

Arnason, B. G. W., and Antel, J. P. (1978). *Ann. Immunol. (Inst. Pasteur)* **129C,** 159–170.

Arnason, B. G. W., and Asbury, A. K. (1968). *Arch. Neurol.* **18,** 500.

Asbury, A. K., Arnason, B. G., and Adams, R. D. (1969). *Medicine (Baltimore)* **48,** 173–215.

Asbury, A. K., Arnason, B. G., Karp, H. R., and McFarlin, D. E. (1978). *Ann. Neurol.* **3,** 565–566.

Austin, J. H. (1958). *Brain* **81,** 157–192.

Bach, M. A., Phan-Din-Tuy, F., Tournier, E., Chatenoud, L., Bach, J. F., Martin, C., and Degos, J.-D. (1980). *Lancet* **2,** 1221–1222.

Batchelor, J. R., Compston, A., and McDonald, W. I. (1978). *Br. Med. Bull.* **34,** 279.

Bell, C. E., and Seetharam, S. (1977). *J. Immunol.* **118,** 826–831.

Benczur, M., Petranyi, G. G., Palffy, G., *et al.* (1980). *Clin. Exp. Immunol.* **39,** 657–662.

Ben-Nun, A., Wekerle, H., and Cohen, I. R. (1981a). *J. Immunol.* **128,** 1450–1457.

Ben-Nun, A., Wekerle, H., and Cohen, I. R. (1981b). *Nature (London)* **292,** 60–61.

Bernard, C. C. A. (1976). *J. Immunogenet.* **3,** 263–274.

Bernard, C. C. A. (1977). *Clin. Exp. Immunol.* **29,** 100–109.

Bluestein, H. G., Williams, G. W., and Steinberg, A. D. (1981). *Am. J. Med.* **70,** 240–246.

Brostoff, S., Burnett, P., Lampert, P., and Eylar, E. H. (1972). *Nature (London), New Biol.* **235,** 210–212.

Buonanno, F. S., Kistler, J. P., Lehrich, J. R., Noseworthy, J. H., New, P. F. J., and Brady, T. J. (1983). *In* "Neurologic Clinics" (J. Antel, ed.), Vol. 1, pp. 757–764. Saunders, Philadelphia, Pennsylvania.

Cashman, N., Martin, C., Eizenbaum, J. F., *et al.* (1982). *J. Clin. Invest.* **70,** 387–392.

Cook, S. D., Dowling, P. C., and Whitaker, J. N. (1970). *Neurology* **20,** 403.
Dellagi, K., Broset, J.-C., and Danon, F. (1979). *J. Clin. Invest* **64,** 1530–1534.
DeVries, E. (1960). *In* "Postvaccinial Perivenous Encephalitis" p. 181. Elsevier, Amsterdam.
Dore-Duffy, P., and Zurier, R. B. (1981). *Clin. Immunol. Immunopathol.* **19,** 303–313.
Dowling, P. C., and Cook, S. D. (1981). *Ann. Neurol.* **9** (Suppl.), 44–55.
Dupont, B., Lisak, R. P., Jersild, C., Hansen, C., Silberberg, D. A., Whitsett, D. H., Zweiman, B., and Ciongoli, A. K. (1977). *Transplant. Proc.* **9** (Suppl.), 181–185.
Dyck, P. J., O'Brien, P. C., Oviatt, K. F., Dinapoli, R. P., Daube, J. R., Bartleson, J. D., Mokri, B., Swift, T., Low, P. A., and Windebank, A. J. (1982). *Ann. Neurol.* **11,** 136–141.
Ebers, G. C., Zabriskie, J. B., and Kunkel, H. G. (1979). *Clin. Exp. Immunol.* **35,** 67.
Eisen, A. (1983). *In* "Neurologic Clinics" (J. Antel, ed.), Vol. 1, pp. 615–629. Saunders, Philadelphia, Pennsylvania.
Fontana, A., Grieder, A., Arrenbrecht, S., *et al.* (1980). *J. Neurol. Sci.* **46,** 55–62.
Froelich, C. J., Searles, R. P., Davis, L. E., and Goodwin, J. S. (1980). *Ann. Int. Med.* **93,** 563–565.
Gerhard, W., Taylor, A., Wroblewska, Z., *et al.* (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 3225–3229.
Goust, J. M., Chenais, F. C., Carnes, J. E., *et al.* (1978). *Neurology* **28,** 421–425.
Goust, J. M., Hogan, E. L., and Arnaud, P. (1982). *Neurology* **32,** 228–234.
Hashim, G. A. (1981). *J. Immunol.* **126,** 419–423.
Hauser, S. L., Bhan, A. K., Gilles, F. H., Hoban, C. J., Reinherz, E. L., Schlossman, S. F., and Weiner, H. L. (1983). *J. Neuroimmunol.* **5,** 197–205.
Helfand, S. L., McGarry, R., Eaton, L., *et al.* (1983). *Neurology* **33** (Suppl. 2), 107.
Holda, J. H., and Swanborg, R. H. (1982). *Eur. J. Immunol.* **12,** 453–455.
Horwich, J. S., Cho, L., Porro, R. S., and Posner, J. B. (1977). *Ann. Neurol.* **2,** 7–19.
Huang, J. T. (1975). *Ann. Allergy* **34,** 1–6.
Hughes, R. A. C. (1979). *In* "Clinical Neuroimmunology" (R. C. Rose, ed.), pp. 170–184. Blackwell, Oxford.
Hughes, R. A. C., Newsom-Davis, J. M., Perkins, G. D., *et al.* (1978). *Lancet* **1,** 750–753.
Iqbal, A., Oger, J.-F., and Arnason, B. G. W. (1981). *Ann. Neurol.* **9** (Suppl.), 65–69.
Jersild, C., Dupont, B., Fog, T., Platz, P., and Svejgaard, A. (1975). *Transplant. Rev.* **22,** 148.
Julien, J., Vital, C., Vallat, J.-M., *et al.* (1978). *Arch. Neurol.* **35,** 423–425.
Kadlubowski, M., Hughes, R. A. C., and Gregson, N. A. (1980). *Brain* **184,** 439–454.
Khan, S. N., Riches, P. G., and Kohn, J. (1980). *J. Clin. Pathol.* **33,** 617–621.
Kurdi, A., Ayesh, I., Abdallat, A., Maayala, N., McDonald, W. I., Compston, D. A. S., and Batchelor, J. R. (1977). *Lancet* **I,** 1123.
Lampert, P. W. (1980). *J. Neuropathol. Exp. Neurol.* **30,** 369.
Lando, Z., Teitelbaum, D., and Arnon, R. (1980). *Nature (London)* **287,** 551–552.
Lassmann, H., and Wisniewski, H. M. (1978). *Acta Neuropathol.* **43,** 35–42.
Lassmann, H., and Wisniewski, H. M. (1979). *Arch. Neurol.* **36,** 490–497.
Latov, N., Sherman, W. H., Nemni, R., *et al.* (1980). *N. Engl. J. Med.* **303,** 612–621.
Latov, N., Braun, P. E., Gross, R. B., *et al.* (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 7139–7142.
Levine, S., and Sowinski, R. (1976). *Am. J. Pathol.* **82,** 381–389.
Levine, S., and Sowinski, R. (1977). *Arch. Int. Pharmacodyn.* **230,** 309–318.
Linthicum, D. S., and Frelinger, J. A. (1982). *J. Exp. Med.* **156,** 31–40.

Lisak, R. P., Behan, O., Zweiman, B., *et al.* (1974). *Neurology* **24,** 560–566.
Lisak, R. P., Mitchell, M., Zweiman, B., Orrechio, E., and Asbury, A. K. (1977). *Ann. Neurol.* **1,** 72–78.
Lopez-Adaros, H., and Held, J. R. (1971). *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* **49,** 178–186.
Lublin, F. (1983). Oral presentation. American Academy of Neurology.
Lumsden, C. E. (1970). *In* "Handbook of Clinical Neurology" (P. J. Vinken and G. W. Bruyn, eds.), Vol. 9, pp. 217–309. Elsevier, Amsterdam.
Madigand, M. J., Oger, J. J., Fauchet, R., *et al.* (1982). *J. Neurol. Sci.* **53,** 519–529.
Mattson, D. H., Roos, R. P., and Arnason, B. G. W. (1980). *Nature (London)* **287,** 335–337.
Moore, M. J., Singer, D. E., and Williams, R. M. (1980). *J. Immunol.* **124,** 1815–1820.
Newsom-Davis, J., Lang, B., Wray, D., Murray, N., Vincent, A., and Gwilt, M. (1982). *Neurology* **32,** A221.
Noronha, A., Richman, D. P., and Arnason, B. G. W. (1980). *N. Engl. J. Med.* **303,** 713–717.
Nyland, H., and Aarli, J. A. (1978). *Acta Neurol. Scand.* **58,** 35–43.
Oger, J. J.-F., and Arnason, B. G. W. (1980). *In* "The Suppression of Experimental Allergic Encephalomyelitis and Multiple Sclerosis" pp. 275–285. (A. N. Davison and M. L. Cuzner, eds.), Academic Press, London.
Oger, J., Antel, J. P., Mariotti, S., and Arnason, B. G. W. (1980). *Neurology* **30,** 448.
Oger, J., Antel, J., and Arnason, B. G. W. (1982). *In* "Multiple Sclerosis, East and West" (Y. Kuroiwa and L. T. Kurland, eds.), pp. 372–381. Kyusha Univ. Press, Japan.
Oger, J., Roos, R., and Antel, J. P. (1983). *In* "Neurologic Clinics" (J. Antel, ed.), Vol. 1, pp. 655–679. Saunders, Philadelphia, Pennsylvania.
Pandey, J. P., Goust, J. M., Salier, J. P., *et al.* (1981). *J. Clin. Invest.* **67,** 1797–1800.
Paterson, P. Y. (1976). *In* "Textbook of Immunopathology" (P. A. Miescher and J. J. Mueller-Eberhard, eds.), 2nd ed., pp. 179–213. Grune & Stratton, New York.
Pettinelli, C. B., and McFarlin, D. E. (1981). *J. Immunol.* **127,** 1420–1423.
Prineas, J. W. (1981). *Ann. Neurol.* **9** (Suppl.), 6–19.
Prineas, J. W., and Wright, R. G. (1978). *Lab. Invest.* **38,** 409–421.
Propp, R. P., Means, E., Deibel, R., *et al.* (1975). *Neurology* **25,** 980–988.
Raine, C. S. (1983). *In* "Multiple Sclerosis—Pathology, Diagnosis and Management" (J. Hallpike, C. W. M. Adams, and W. W. Tourtellotte, eds.), pp. 413–460. Chapman & Hall, London.
Reder, A. T., Antel, J. P., Oger, J. J.-F., McFarland, T. A., Rosenkoetter, M., and Arnason, B. G. W. (1984). *Ann. Neurol.* **16,** 242–249.
Reinherz, E. L., Weiner, H. L., Hauser, S. L., Cohen, J. A., Distaso, J. A., and Schlossmann, S. G. (1980). *N. Engl. J. Med.* **303,** 125–129.
Rivers, T. M., Sprunt, D. H., and Berry, G. P. (1933). *J. Exp. Med.* **58,** 39–53.
Rose, A. S., Kuzma, J. W., Kurtzke, J. F., *et al.* (1970). *Neurology* **20,** 1–59.
Saida, T., Saida, K., Silberberg, D. H., and Brown, M. J. (1981). *Ann. Neurol.* **9** (Suppl.), 87–101.
Schonberger, L. B., Bregman, D. J., Sullivan-Bolyai, J. Z., *et al.* (1979). *Am. J. Epidemiol.* **110,** 105–123.
Searles, R. P., Davis, L. E., Hermanson, S., and Froelich, C. J. (1981). *Lancet* **1,** 273.
Sheremata, W., Colby, S., Karkhanis, Y., and Eylar, E. H. (1975). *Can. J. Neurol. Sci.* **2,** 87–90.
Sherman, A. B., Schaumburg, H. H., and Asbury, A. K. (1980). *Ann. Neurol.* **9,** (Suppl.), 102–106.
Stevens, J. G., Pepose, J. S., and Cook, M. L. (1981). *Ann. Neurol.* **4,** 285–289.

Stewart, G. J., Pollard, J. D., McLeod, J. G., and Wolnizer, C. M. (1978). *Ann. Neurol.* **4,** 285–289.

Stewart, G. J., McLeod, J. G., Basten, A., and Bashir, H. V. (1981). *Hum. Immunol.* **3,** 13–29.

Traugott, U., Reinherz, E. L., and Raine, C. S. (1983). *Science* **219,** 308–310.

Vartdal, F., Vandvik, B., and Norrby, E. (1980). *Ann. Neurol.* **8,** 248–255.

Waksman, B. H., and Adams, R. D. (1955). *J. Exp. Med.* **102,** 213.

Welch, A. M., Holda, J. H., and Swanborg, R. H. (1980). *J. Immunol.* **125,** 186–189.

Whitacre, C. C., Mattson, D. H., Paterson, P. Y., *et al.* (1981). *Neurochem. Res.* **6,** 87–96.

Whitacre, C. C., Mattson, D. H., Day, E. D., *et al.* (1982). *Neurochem. Res.* **7,** 1209–1216.

Wilkerson, L. D., Lisak, R. P., and Zweiman, B. (1978). *Clin. Exp. Immunol.* **34,** 87–91.

Wisniewski, H. M., and Keith, A. B. (1977). *Ann. Neurol.* **1,** 144–148.

Wisniewski, H. M., Lassman, H., Brosnan, C. F., *et al.* (1982). *In* "Recent Advances in Clinical Neurology" (W. B. Mathews and G. H. Glaser, eds.), pp. 95–124. Churchill, Edinburgh.

Wrabetz, L. G., Antel, J. P., Oger, J., Arnason, B. G. W., Goust, J. M., and Hopper, J. E. (1982). *Cell Immunol.* **74,** 398–403.

CHAPTER **16**

Immunological Influences in Cardiovascular Disease

MATHEW A. VADAS

Clinical Research Unit of The Walter and Eliza Hall Institute of Medical Research and The Royal Melbourne Hospital
Melbourne, Victoria, Australia

I. INTRODUCTION

There are few areas of clinical medicine that have been so little influenced by immunological considerations as that of cardiovascular disease.

First, one should ask whether interest in this interface is justified? It has to be admitted that immunological mechanisms normally devoted to dealing with acute perturbations of the external milieu may not have, and in fact are designed not to have, effects on cardiovascular phenomena essential to homeostasis. Supporting this point of view is the scarcity of cardiovascular abnormalities described in animals immunodeficient as a result

of single gene mutations and the normality of the immunological system in cases of inborn errors of the cardivascular system (CVS). On the other hand, there are many difficulties in the design of experimental approaches to immunological disorders in cardiovascular disease. For example, there is a relative difficulty in obtaining cardiovascular tissue for immunological examination, and the large animals often used as models for cardiovascular research are not as yet well characterized and, in fact, are clumsy to study from the immunological point of view.

Despite these difficulties, there is already a modest body of evidence linking pathology of the CVS with immune perturbation, much of it arising by reason of the CVS being involved as a bystander in immunological and infective diseases: in particular, frank autoimmune phenomena are not prominent as causes of CVS disease. However, in the author's judgment, the real breakthroughs in the CVS–immunology interface are yet to come and many of those will be from spin-offs from the new "immunological methodology." For example, there are already elegant examples of the uses of antibodies of well-defined specificities in the diagnosis and treatment of myocardial disorders. In addition one can look forward to elucidating the effects of cell-derived factors, cytokines, on the conduction pathways in the heart, on endothelial as well as circulating cells, receptors for growth factors, and lipoproteins. Perhaps the most informative probes will be genetic ones that, by enabling analysis of restriction endonuclease patterns of DNA, will reveal genes linked to susceptibility to various CVS diseases. In the first instance these may be used to probe conditions with a dominant inheritance, such as idiopathic hypertrophic subaortic stenosis, in a manner similar to that which has been successful in Huntington's disease. However, there is no reason why diseases under the control of many genes should not be susceptible to this approach.

In this chapter, I shall confine discussion to a limited number of examples of cardiovascular diseases in which autoimmunity or immunological aberrations have been shown to be, or have a chance of becoming, important.

II. RHEUMATIC FEVER

Heart-reactive antibodies (HRA) have long been implicated in the pathogenesis of rheumatic fever (Van de Rinj *et al.*, 1977). Rheumatic fever and the subsequent rheumatic heart disease are initiated by an infection with a group A β-hemolytic *Streptococcus*. As a result of this infection, some individuals develop HRA, some develop disease, and some develop both. The central questions relate to the reasons for the suscepti-

bility to disease of certain individuals, for the development of HRA, and the possible relationships between these.

Doubtless a reason exists for the majority of patients with β-hemolytic streptococcal infections not getting rheumatic fever. Until now these were probed with relatively blunt tools, namely relating susceptibility to rheumatic fever to ABO blood groups or HLA antigens (Murray *et al.*, 1978), and no clear trends emerged. More recently, an HLA-linked B-cell alloantigen was related to susceptibility (Olsen, 1980), suggesting a genetic basis. This should lead to further analysis, but with the much finer tools of molecular biology, looking for the correlation between restriction fragment-length polymorphism and disease.

The development of HRA in rheumatic fever is a good example of "molecular mimicry." Prominent antigenic components of streptococci elicit antibodies in some individuals that cross-react with myocardial tissues. These antibodies tend to persist, and deposits of antibody and complement have been found on cardiac muscle at sarcolemmal and subsarcolemmal sites (Kaplan and Meyeserian, 1962) and in the pericardium (Persellin *et al.*, 1982; Russel and Persellin, 1982). The cross-reaction extends not only to human myocardium (Kaplan, 1979) and myocardial sarcolemmal antigens (Zabriskie and Freimer, 1966; Van de Rinj *et al.*, 1977), but also to heart valve glycoproteins (Goldstein *et al.*, 1968) and the myocardial conduction system (Kasp-Grochowska and Kingston, 1977). In addition, the serum of children with Sydenham's chorea react with cytoplasmic neuronal antigens in the subthalamic and caudate nuclei, and, since this reactivity can be absorbed by group A streptococcal membranes (Husby *et al.*, 1976), a strong implication for mimicry also applies here.

Thus, in the predisposed individual, streptococcal infection sets up a series of events culminating in disease. Besides the persistent elevation of antibodies to group A streptococcal antigens in the blood, there is also evidence for a cell-mediated immune attack, that is, the accumulation of OKT4-staining T lymphocytes (Raizada *et al.*, 1983) in valvular tissues.

Although the development of rheumatic fever is said to correlate positively with titers of HRA, there are examples of patients without rheumatic fever and with high levels of antibody, as well as the opposite association. Nevertheless, the body of evidence for the association is convincing (Williams, 1983), and investigations not only into the cause of development of these antibodies but also for the reasons of their persistence may be relevant. Besides the genetic reasons discussed above, another possible mechanism may derive from a coinfection with another agent. In this model a synergism between the streptococcus and another silent infectious agent is established, resulting in an adjuvant-like effect.

An example of this type of mechanism is the prolongation of immune responses to protein antigens by *Bordetella pertussis* toxin (Smith *et al.*, 1976).

III. MYOCARDITIS

A. Chagas' Disease

This is the commonest form of myocarditis in the world. Investigators have long been puzzled by the absence of *Trypanosoma cruzi* organisms from the hearts of individuals with active carditis. Recently, an explanation, again rooted in the mechanism of molecular mimicry, was provided (Cossio *et al.*, 1977; Teixera, 1974). An antibody was found in most patients with Chagas' heart disease that reacted both with *T. cruzi* (epimastigote stage) and with antigens present on heart endothelium, interstitium, and vessels. The addition of this cross-reactive antibody to living rat atrium resulted in (1) fixation of the antibody to heart muscle resulting in electron microscopic evidence of pathogenetic change, and (2) an increase in the atrial rate. The hypothesis that this antibody increased atrial rate by acting as an adrenergic agonist was supported by the observation that β-adrenergic blockers prevented the antibody-induced tachycardia, and that antibody prevented the expected effect of noradrenaline in the system.

Additional proof for the molecular mimicry in Chagas' disease comes from the description of a monoclonal antibody (MAb) raised against dorsal root ganglia that reacted with mammalian central and peripheral neuronal antigens and with *T. cruzi* (Wood *et al.*, 1982). All classes of neurons, including autonomic ganglia that undergo degeneration in Chagas' disease, reacted with the antibody, and the MAb with complement was also cytotoxic to cardiac cells in culture (Huber *et al.*, 1980).

These studies not only illustrate the possible pathogenesis of this disease involving the cardiac and nervous tissues but also suggest a mechanism for the protean clinical manifestations of other infectious agents.

B. Coxsackie Myocarditis

Cell-mediated immune reactions are implicated in the causation of Coxsackie B virus-induced myocarditis. Coxsackie B-3 and B-4 viruses have been shown to induce myocardial lesions in mice, and T lymphocytes from mice infected with Coxsackie virus are able to kill cardiac myofibers. In addition, subtypes of Coxsackie virus (for example, B-3o) that do

not produce myocarditis seem to differ from subtypes that do produce myocarditis (for example, B-3m) only by their inability to sensitize T lymphocytes (Cossio, 1977).

IV. CARDIOMYOPATHY

In general, immune processes have been implicated in this group of disorders by the discovery of deposition of antibody in myocardial tissues. It is not possible, however, to be certain that this deposition is the cause of injury rather than its result. Nevertheless there are two strands of evidence that suggest that immune mechanisms may be important in cardiomyopathy. First, there is a very high titer of anti-Coxsackie antibody in patients with recent-onset cardiomyopathy, especially in those whose illness started with fever (Cambridge *et al.*, 1979). In these individuals, an unusually prolonged immune reaction against the virus may be responsible for this chronic illness. Second, there exists an immune disturbance in congestive cardiomyopathy: The mononuclear suppressor-cell activity is lower than in control patients with cardiomyopathy due to ischemic heart disease. This deficiency may explain why the abnormally prolonged immune attack is sustained against Coxsackie virus in some patients. It may also explain why, after cardiac transplantation, lymphomas predominate among patients who had had congestive rather than ischemic cardiomyopathy: The former possibly had an impaired "immune surveillance" even before immunosuppressive treatment (Fowles *et al.*, 1979).

In hypertrophic cardiomyopathy, as exemplified by idiopathic hypertrophic subaortic stenosis, immune dysfunction has not been strongly implicated, in spite of the high incidence of antibody deposition in the myocardium (Das *et al.*, 1973). However, the autosomal dominant pattern of inheritance suggests a single aberrant gene function in this disease, although reports linking this gene to the *HLA* complex are unconvincing (Bloch *et al.*, 1980). Immunological mechanism that lead to the stimulation of myocardial growth by the induction of excess receptors on a subpopulation of myocardial cells, perhaps septal, may be considered.

V. HEART BLOCK

The immunological aspects of congenital heart block are truly remarkable. First, it is a recognized complication of maternal multisystem

autoimmune disease, notably systemic lupus erythematosus (Esscher and Scott, 1979; Maisch *et al.*, 1979a). Second, a strong association has been described between infantile heart block and the presence of anti-Ro (SS-A) autoantibodies in maternal serum (Alspaugh and Maddison, 1979; Scott *et al.*, 1983), even in the absence of overt disease. Although anti-Ro crosses the placenta (Vetter and Rashkind, 1983), it may not be the antibody that is directly involved in the etiology of the heart block. Nevertheless it does serve as a reliable marker for the identification of at-risk pregnancies.

The adult form of complete heart block has only tenuous immunological connections (Fairfax, 1976), but in Chagas' disease (see above) cross-reactive antibody may be involved. The possibility that other autoantibodies against conducting tissues change the threshold for conduction should be borne in mind in the search for the causes of some forms of this disease.

VI. POSTPERICARDIOTOMY SYNDROMES

Myocardial injury seems to lead to the development of HRA and to two clinical entities, the postpericardiotomy and the postmyocardial infarction syndromes.

In the postpericardiotomy syndrome, fever, leukocytosis and elevated ESR and percarditis develop in up to 20% of patients 1–6 weeks after myocardial surgery, and 84% of them (Maisch *et al.*, 1979b) develop antibodies not only against cardiac sarcolemmal antigens but also skeletal muscle, endothelial cell, and nuclear antigens. This is therefore a multisystem disease in which the heart plays a central clinical role. In spite of the antibodies not being heart specific [and unlike those in rheumatic fever not absorbed by streptococci (Zabriskie, 1977; Patarroyo *et al.*, 1979)], there is evidence for the severity of clinical disease being reflected in the magnitude of the immune response (Chang *et al.*, 1979). The other possibility of a cell-mediated immune effect being pathogenic has also been considered (Maisch *et al.*, 1983a), but no results bear directly on this question.

In the postmyocardial infarction syndrome, Dressler's syndrome, a similar clinical picture of pericarditis, fever, leukocytosis, and high ESR is seen 2–6 weeks after myocardial infarction in 1–4% of patients. Although the rise in titer of antiheart antibodies is associated in time with disease (Agrawal *et al.*, 1978), the association must for the moment be regarded as incidental.

VII. HEART TRANSPLANTATION

In spite of being a highly artificial system, heart transplantation provides a true model of immunological injury to heart tissue with evidence of conduction disturbance as well as muscle destruction. The observation that after cardiac transplantation lymphomas predominate among patients who had congestive rather than ischemic cardiomyopathy prompted the discovery of a defect of suppressor activity in the former group of patients, suggesting an underlying immune defect (Dudding and Ayoub, 1968).

VIII. SUBACUTE BACTERIAL ENDOCARDITIS (SBE)

Almost 100% of patients with SBE develop HRA (Maisch *et al.*, 1983a,b). Antiendocardial antibody is diagnostic of endocarditis, and anti-sarcolemmal and -myolemmal antibodies are probably related to the severity of the myocardial involvement in the endocarditis. For convenience, the chief diseases in which HRA occur are listed in Table I. Immune complexes are seen more frequently (85%) in patients with positive blood cultures (Maisch *et al.*, 1983a) and are associated with hypocomplementemia and hematuria, as well as a longer duration of illness and extravalvular manifestations (Mackay, 1979). This immunological phase of SBE is marked by an increase in plasma IgA and IgM and by the presence of IgM rheumatoid factor reacting with host IgG, itself directed

TABLE I
Heart-Reactive Antibodies

Disease	Probable mechanism	Site(s) of reaction
Acute rheumatic fever	Cross-reactive with Group A β-hemolytic *Streptococcus*	Myocardium, pericardium Cardiac conductive tissue CNS neuronal tissue
Chagas' disease	Cross-reactive with *Trypanosome cruzi*	Myocardium Cardiac conductive tissue Neuronal tissue
Postpericardiotomy and myocardial infarct	Immune reaction to damaged organ	Myocardium Skeletal muscle
Bacterial endocarditis	Immune reaction to damaged organ	Myocardium Endocardium

against bacterial antigens. Clearance of the infection is associated with the rapid disappearance of this rheumatoid factor.

It is interesting to note that some of the clinical features of SBE, including microscopic hematuria, arthritis, and splenomegaly may be mediated by circulating immune complexes, while Osler's nodes are believed to be due to septic emboli.

IX. IMMUNOREGULATORY LIPOPROTEINS

Since blood lipids are linked to cardiovascular disease, it is important to note that lipoproteins may be powerful regulators of the immune response. The effects of lipoproteins are as a rule inhibitory to the immune system. For example, the intravenous administration of high-density lipoproteins to mice results in significant depression of mitogenic responses, of primary immune responses to sheep red blood cells, and even of allograft immunity (Agrawal *et al.*, 1978). In addition, there is evidence that very low-density lipoprotein (VLDL) and a light-density fraction of low-density lipoprotein (LDL-In) in concentrations as low as 2.5–1.0 pmol/ml inhibit certain lymphocyte functions (Mathews, 1979). They act on lymphocytes in the early part of immune response, perhaps by preventing or delaying mitosis. Suppressor T cells are especially sensitive to the LDL-In immunoregulator. The concentrations of these substances at the site of antigen entry, where the initial immune responses take place, are not known, and therefore their physiological role is still in question. The relationships of levels of these immunoregulators with forms of hyperlipidemias are the subjects of ongoing research.

X. AUTOIMMUNE HYPERLIPIDEMIA

In this disease, lipoprotein clearance is inhibited by a circulating autoantibody to lipid-carrying protein. In addition to generalized atherosclerosis, xanthomas may form when circulating low-density lipoprotein complexed with an IgA autoantibody loses its ability to regulate cholesterol synthesis (Mathews, 1979).

XI. IMMUNOLOGICAL PROBES IN CARDIOVASCULAR DISEASE

Antibodies have great potential as specific probes. The injection of xenogeneic protein in general, and immunoglobulins (Ig) in particular,

however, elicits an immune response limiting their repeated injections. However, Fab fragments of Ig are less immunogeneic and are more rapidly excreted and thus are much more suitable for *in vivo* use. An even smaller antibody fragment, the F_v consisting of a single Ig domain (25,000 daltons), one that carries the full capacity to bind antigen, has an even greater clinical potential (Hochman *et al.*, 1976). A second problem in the use of antibodies is the presence of impurities. This can be overcome by the use of monoclonal antibodies. These have the added advantage of the capacity for mass production. They are still derived from xenogeneic species but this too may be capable of being altered by a genetic engineering approach of manufacturing hybrid antibodies composed mainly of human elements. Alternatively, given suitable human cell lines for hybridization, human autoantibodies of the right specificity may be derived and used therapeutically.

There are four areas in which MAb have already been shown to be useful:

1. Antirenin antibodies. Rabbit antirenin Fab has an immediate hypotensive affect in dogs rendered hypertensive as a result of renal artery constriction (Dzau *et al.*, 1980). Fab is much less immunogenic than whole IgG, and has the added advantage of equilibrating rapidly with extracellular fluid and having a shorter half-life. Monoclonal antibodies against renin have also been prepared, and their therapeutic role is under investigation.
2. Anti-β-Adrenergic receptor antibodies. This problem has been attacked by the raising of antibodies in rabbits against a β agonist and purifying by affinity chromatography. This antibody was then used for the raising of antiidiotypes that were found to bind membranes that bear the β receptor (Rockson *et al.*, 1980; Homey *et al.*, 1981; Schreiber *et al.*, 1980). This approach, and its monoclonal counterpart, will allow not only the isolation of the β receptor but also examination of its numbers and heterogeneity of various cell types.
3. Myosin-specific antibodies in myocardial infarct imaging. The unique covalent structure of cardiac myosin is normally protected from the extracellular fluid by the plasma membrane. With cell death, this barrier is broken and antibodies have access. This approach has been spectacularly successful in living dogs (Khaw *et al.*, 1980), resulting in high resolution. Its use in humans has obvious applications.
4. Antidigitalis antibodies. Digoxin-specific Fab was shown to be effective in treating digitalis toxicity in dogs (Curd *et al.*, 1971), and has been used in humans for this life-threatening intoxication. In each subject (ranging in age from infants to septuagenarians), dramatic reversal of

signs and symptoms of intoxication occurred (Smith *et al.* 1976; Aeberhard *et al.*, 1980; Hess *et al.*, 1979). Notable in these patients was the reduction in half-life of digitoxin, normally of the order of 3 days, to 12 h.

XII. ATHEROSCLEROSIS

Finally, it may be worth considering whether immunological mechanisms play a part in the common disorders of atherosclerosis and hypertension. There is already persuasive evidence that arterial injury can be caused by immunological mechanisms and that such injury leads to accelerated atherosclerosis. Indeed, autoimmune phenomena are found with an increased incidence in atherosclerotic patients, as judged by population surveys (Mathews *et al.*, 1974), and patients with autoimmune diseases such as systemic lupus erythematosus have an accelerated onset of atherosclerosis. In addition, cigarette smoke, implicated as a major risk factor in atherosclerosis, has also been found to elicit an immune reaction, and the majority of smokers demonstrate immediate skin sensitivity to component(s) of cigarette smoke.

Another way of implicating immunological mechanisms is through genetics. There are three chromosomal loci known to influence the immune responses. These are the *HLA* locus, encoded within the major histocompatibility complex, the *Gm* locus, which defines immunoglobulin allotypes, and a locus encoded within the *X* chromosome. Linkage has been demonstrated between atherosclerotic heart disease, as assessed clinically by coronary occlusion, and particular genotypes specified by these loci (Mathews *et al.*, 1980). Although all these findings make likely the involvement of immunological reactions in atherosclerosis, the question remains as to how important these reactions are in comparison with the other well-known risk factors. The answer may be forthcoming in the future by probing the genome with restriction endonucleases.

XIII. HYPERTENSION

There is evidence in hypertension for the involvement of specific immunological effects. These experiments in rats revealed the highly unexpected result that a single intravenous administration of spleen cells from hypertensive donors transferred hypertension into normotensive recipients; in addition, a delayed-type hypersensitivity response to homoge-

nized arterial wall was also transferred. Although startling, these experiments remain to be repeated by another group, and the mechanism of the effect understood. In humans, essential hypertension is associated with a higher frequency of autoantibodies and rheumatoid factor and higher levels of IgG and is linked with certain HLA specificities and *Gm* allotypes (Mathews *et al.*, 1980), suggesting the possibility of a "loose" link to immunological mechanisms.

XIV. SUMMARY

Although it is clear that the heart can be damaged by immunological mechanisms (an extreme example being seen in heart transplant) and that damage to the heart can result in an immune response against the myocardium (as seen after cardiac surgery), in most diseases the pathogenic role of the immune system has not been conclusively demonstrated. Nevertheless, the tools of immunology are likely to play an increasingly important part in the diagnosis and treatment of cardiovascular diseases, as exemplified by the treatment of digitalis toxicity by antidigitalis antibodies. Modern tools of immunology and molecular biology may shed more light on mechanisms of cardiovascular disease, especially those with genetically determined or infectious etiologies.

ACKNOWLEDGMENTS

I would like to thank Dr. Ian Mackay for helpful suggestions.

REFERENCES

Aeberhard, P., Butler, V. P., and Smith, T. W. (1980). *Arch. Mal. Coeur. Vaiss.* **73,** 1471–1478.

Agrawal, C. G., Gupta, S. P., Chaturvedi, U. C., Mitra, M., and Gupta, S. (1978). *Int. Arch. Allergy Appl. Immunol.* **57,** 246–252.

Alspaugh, M., and Maddison, P. (1979). *Arthritis Rheum.* **22,** 796–798.

Bloch, A., Crittin, J., Barras, C., and Jeannet, M. (1980). *N. Engl. J. Med.* **302,** 1033.

Cambridge, G. *et al.* (1979). *Br. Heart J.* **41,** 692–696.

Chang, K., Friedman, H., and Goldberg, H. (1979). *Adv. Exp. Med. Biol.* **121B,** 335–340.

Cossio, P. M. *et al.* (1977). *Am. J. Pathol.* **86,** 533–544.

Curd, J., Smith, T. W., Jaton, J.-C., and Haber, E. (1971). *Proc. Natl. Acad. Sci. U.S.A.* **68,** 2401–2406.

Das, S. K., Cassidy, J. T., Dodson, V. N., and Willis, P. W. (1973). *Br. Heart J.* **35,** 965–969.

Dudding, B. A., and Ayoub, E. M. (1968). *J. Exp. Med.* **128,** 1081–1088.
Dzau, V. J., Slater, E. E., and Haber, E. (1979). *Biochemistry* **18,** 5224–5228.
Dzau, V. J., Kopelman, R. I., Barger, A. C., and Haber, E. (1980). *Science* **207,** 1091–1093.
Esscher, E., and Scott, J. S. (1979). *Br. Med. J.* **1,** 1235.
Fairfax, A. J. (1976). *Clin. Exp. Immunol.* **23,** 1.
Fowles, R. E., Biebes, C. P., and Stinson, E. B. (1979). Circulation **59,** 483–491.
Goldstein, I., Rebeyrotta, P., Parlebas, J., and Halpern, B. (1968). *Nature (London)* **219,** 866–868.
Hess, T., Stucki, P., Barandun, S., Scholtysik, G., and Riesen, W. (1979). *Am. Heart J.* **98,** 767–771.
Hochman, J., Inbar, D., and Givol, D. (1976). *Biochemistry* **12,** 1130–1135.
Homey C., Rockson, S., and Haber, E. (1981). *Clin. Res.* **29,** 207A.
Hsu, K. H., Hiramoto, R. N., and Ghanta, V. K. (1982). *J. Immunol.* **128,** 2107–2110.
Huber, S. A., Job, K. P., and Woodruff, J. F. (1980). *Am. J. Pathol.* **98,** 681–694.
Husby, G., Van de Rijn, I., Zabriskie, J. B., Abdin, Z. H., and Williams, R. C., Jr. (1976). *J. Exp. Med.* **144,** 1094–1110.
Kaplan, M. H. (1979). *Rev. Infect. Dis.* **1,** 988–996.
Kaplan, M. H., and Meyeserian, M. (1962). *Lancet* **1,** 706–710.
Kaplan, M. H., and Svec, K. H. (1964). *J. Exp. Med.* **119,** 651–666.
Kasp-Grochowska, E., and Kingston, D. (1977). *Clin. Exp. Immunol.* **27,** 43–45.
Khaw, B. A., Fallon, J. T., Strauss, H. W., and Haber, E. (1980). *Science* **209,** 295–297.
Mackay, I. R. (1979). *In* "CRC Handb. Ser. Clin. Lab. Sci. Sect. F: Immunol. Vol. 1(2)" (A. Baumgarten and F. F. Richards, eds.), pp. 3–24. CRC, Boca Raton, Florida.
Maddison, P., Mogavero, H., Provost, T. T., and Reichlin, M. *J. Rheumatol.* **6,** 189–195.
Maisch, B., Berg, P. A., and Kochsiek, K. (1979a). *Clin. Exp. Immunol.* **38,** 189–197.
Maisch, B., Schuff-Werner, P., Berg, P. A., and Kochsiek, K. (1979b). *Clin. Exp. Immunol.* **38,** 198–203.
Maisch, B., Eichstädt, H., and Kochsiek, K. (1983a). *Am. Heart J.* **105,** 329–337.
Maisch, B., Mayer, E., Schubert, V., Berg, P., and Kochsiek, K. (1983b). *Am. Heart J.* **106,** 338–345.
Mathews, J. D. (1980). *In* "Atherosclerosis V. Proceedings of the 5th International Symposium," (A. M. Gotto, and L. C. Smith, eds.), pp. 324–329. Allen B. Springer, New York.
Mathews, J. D., Whittingham, S., and Mackay, I. R. (1974). *Lancet* **2,** 754–758.
Mathews, J. D., Mathieson, I. D., and Tait, B. D. (1980). Academic Press, London.
Murray, G. G., Monteil, M. M., and Persellin, R. H. (1978). *Arthritis Rheum.* **21,** 652–656.
Olsen, F. (1980). *Acta Pathol. Microbiol. Scand., Sect. C* **88,** 1–45.
Patarroyo, M. E., Winchester, R. J., Vejerano, A., Gibotsky, A., Chalem, F., Zabriskie, J. B., and Kunkel, H. G. (1979). *Nature (London)* **278,** 173–174.
Persellin, S. T., Ramirez, G., and Orme, E. C. (1982). *Arthritis Rheum.* **25,** 1054–1058.
Raizada, V., Williams, R. C., Jr., and Chopra, P. *et al.* (1983). *Am. J. Med.* **74,** 90–96.
Rockson, S., Homey, C., and Haber, E. (1980). *Clin. Res.* **46,** 808–813.
Russell, I., and Persellin, R. H. (1983). *In* "Immunology in Medicine" (E. J. Holborow and W. G. Reeves, eds.), pp. 273–284. Grune & Stratton, New York.
Schreiber, A. B., Couraud, P. O., Andre, C., Vray, B., and Strosberg, A. D. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 7385–7389.
Scott, J. S., Maddison, P. J., Taylor, P. V., Esscher, E., Scott, O., and Skinner, R. P. (1983). *N. Engl. J. Med.* **309,** 209–212.
Smith, T. W., Haber, E., Yeatman, L., and Butler, V. P., Jr. (1976). *N. Engl. J. Med.* **294,** 797–800.

Teixera, F. (1974). *J. Exp. Med.* **140,** 38.
Van de Rinj, I., Zabriskie, J. B., and McCarthy, M. (1977). *J. Exp. Med.* **146,** 579–599.
Vetter, V. L., and Rashkind, W. J. (1983). *N. Engl. J. Med.* **309,** 236–237.
Williams, R. C. (1983). *Am. J. Med.* **75,** 727–730.
Wood, J. M., Hudson, L., Jessell, T. M., and Yamamoto, M. (1982). *Nature (London)* **296,** 34–38.
Zabriskie, J. B. (1977). *Med. Times* **105:** (80) 11d-(79) 19d.
Zabriskie, J. B., and Freimer, E. H. (1966). *J. Exp. Med.* **124,** 661–678.

CHAPTER 17

Autoimmune Cutaneous Diseases*

LUIS A. DIAZ
GRANT J. ANHALT
HARISH P. PATEL
THOMAS T. PROVOST
Department of Dermatology
The Johns Hopkins University School of Medicine
Baltimore, Maryland

I. INTRODUCTION

Several dermatoses that are examples of organ-specific and organ-nonspecific autoimmune cutaneous diseases are listed in Table I. Organ-spe-

* This investigation was supported by the U.S. Public Health Service, grants (R01-AM-32081-01, 1-R23-AM-32079-01 and 2-R01-AM-25650) from the National Institutes of Health, and a gift from the Esteé Lauder Corporation.

TABLE I
Autoimmune Cutaneous Diseases

Organ-specific	Organ-nonspecific
Pemphigus vulgaris	Lupus erythematosus?
Bullous pemphigoid	Dermatomyositis?
Herpes gestationis	Scleroderma?
Epidermolysis bullosa acquisita?	Sjögren's syndrome?
Alopecia areata?	Dermatitis herpetiformis?
Vitiligo?	

cific models include pemphigus vulgaris and bullous pemphigoid; these diseases are characterized by a highly specific autoimmune reaction directed against the skin.

Pemphigus vulgaris is a potentially lethal bullous dermatosis characterized by a unique autoantibody that is specific for an epidermal cell-surface antigen normally present in differentiating keratinocytes of squamous epithelia ("pemphigus autoantibodies"). Normal individuals are tolerant to this epidermal antigen; loss of tolerance and autoantibody formation initiates pemphigus. The factors that trigger breakdown of immune tolerance to the antigen are unknown. Bullous pemphigoid and its variant herpes gestationis are also highly specific autoimmune dermatological diseases. These patients become sensitized to an antigen normally found within the cutaneous basement membrane zone (BMZ) and they produce anti-BMZ autoantibodies ("pemphigoid autoantibodies"). As in pemphigus vulgaris, these patients have lost their immune tolerance to this self-antigen.

The autoantibodies in pemphigus vulgaris and bullous pemphigoid, which can be detected both bound in and around lesions and circulating in the serum, have proven to be not only crucial for the accurate diagnosis of these conditions, but appear to be of primary importance in the pathogenesis of the diseases as well. There is not substantial evidence that pemphigus autoantibodies bind specifically to a cell-surface antigen on squamous epithelia and directly induce epithelial cell–cell detachment (acantholysis) of those cells through mechanisms that are currently being investigated. Current studies suggest that the pemphigoid autoantibodies bind with an antigen found in the lamina lucida of the BMZ of squamous epithelia and initiate a complex series of inflammatory events that eventuate in the formation of a subepidermal blister.

The organ-nonspecific autoimmune diseases are best represented by

systemic lupus erythematosus (SLE). Many antibody systems have been detected in the sera of these SLE patients, involving DNA, RNA, RNP, Sm, Ro, and La antibodies (Tan *et al.*, 1966; Provost, 1979b).

While there is a great deal of data to implicate autoimmune phenomena in the pathogenesis of SLE, a specific role for autoimmunity in the pathogenesis of the cutaneous lesions of these patients remains unproven. Recent exciting findings in a subpopulation of lupus patients who show anti-Ro antibodies in their sera may help to elucidate the autoimmune nature of the skin lesions in these patients (Provost *et al.*, 1977). Patients with this form of lupus manifest annular, infiltrated, erythematous lesions in sun-exposed skin (Provost *et al.*, 1977; Sontheimer *et al.*, 1979). Furthermore, similar lesions are seen in neonates born to mothers with Ro-positive serology (Franco *et al.*, 1981; Kephart *et al.*, 1981). In these infants, the skin lesions and the Ro antibodies disappear in the first 6 months of life. These data strongly suggest that the maternal anti-Ro antibodies are transplacentally transferred to the neonate, where they may cause the typical cutaneous lesions by an unknown pathogenetic mechanism(s).

This chapter reviews the evidence for the existence of autoimmunity in the pathogenesis of various bullous diseases and the cutaneous lesions of lupus erythematosus. Sections II and III discuss in detail the diseases pemphigus vulgaris and bullous pemphigoid. These are unique disorders in which there is convincing evidence that they are true autoimmune diseases. They are reviewed in depth, with special emphasis on immunopathological mechanisms. Sections IV and V discuss herpes gestationis and other bullous dermatoses in which there is highly suggestive, but not conclusive, evidence of their autoimmune nature. Similarly, Section VI discusses in detail the current concepts of the pathogenesis of cutaneous lupus erythematosus and recent evidence indicating that it too may be an autoimmune disorder.

II. PEMPHIGUS VULGARIS

A. Introduction

The term pemphigus refers to a group of chronic bullous diseases characterized by intraepidermal blister formation (Lever, 1965). These diseases include (a) pemphigus vulgaris; (b) pemphigus vegetans; (c) pemphigus foliaceus; and (d) a closely related endemic South American variant of pemphigus foliaceus, fogo selvagem. Each of these forms possesses dis-

tinctive histological and clinical features, but all are characterized by flaccid bullae arising on normal-appearing skin. On application of pressure, these bullae spread to involve large areas of skin (Nikolsky's sign) and rupture easily. In pemphigus vulgaris, there is blister formation above the epidermal basal cell layer. These blisters easily rupture and produce large denuded areas of skin that readily discharge serous or serosanguineous fluid. Pemphigus vegetans is a variant of pemphigus vulgaris characterized by the presence of vegetating lesions in flexural areas.

Pemphigus foliaceus is a benign superficial variant of pemphigus that is characterized by intraepidermal blister formation occurring at or near the epidermal granular cell layer. The Senear–Usher syndrome, or pemphigus erythematosus, represents a combination of the clinical and serological features of pemphigus foliaceus and systemic lupus erythematosus (Usher and Senear, 1976).

Although uncommon pemphigus vulgaris is found in all ethnic and racial groups, more commonly Jews. Pemphigus vulgaris in Jewish patients displays an increased prevalence of the HLA-DRw4 phenotype (Park *et al.*, 1979; Brautbar *et al.*, 1980). The precise etiology of all forms of pemphigus is unknown. Fogo selvagem (Brazilian pemphigus foliaceus) is thought to be produced by a virus, although there are no precise data. Recently, penicillamine has been implicated as a potential etiological factor in the induction of pemphigus in patients' receiving this drug (Tan and Powell, 1976).

All forms of pemphigus are characterized by epidermal cell–cell detachment (acantholysis) leading to intraepidemal vesiculation (Fig. 1, left) and by the existence of IgG antibodies directed against antigenic determinants present on the cell surfaces of differentiating keratinocytes (Fig. 1, left) (Beutner and Jordon, 1964). The underlying mechanisms involved in triggering this autoantibody formation are unknown.

B. Clinical Manifestations

Pemphigus vulgaris usually occurs in the fourth and fifth decades of life. It has, however, been described in the very old as well as in the very young. The disease may begin as isolated lesions on the scalp (frequently misdiagnosed as scalp pyoderma) or as oral erosions. These isolated lesions may persist for several months before generalized dissemination occurs (Fig. 1, left). All squamous mucous membranes (e.g., mouth, conjunctiva, nose, esophagus, and rectum) may be involved to varying degrees. The glabrous skin may show flaccid bullae that demonstrate poor healing. Prior to the introduction of antibiotics and steroids, ~50% of

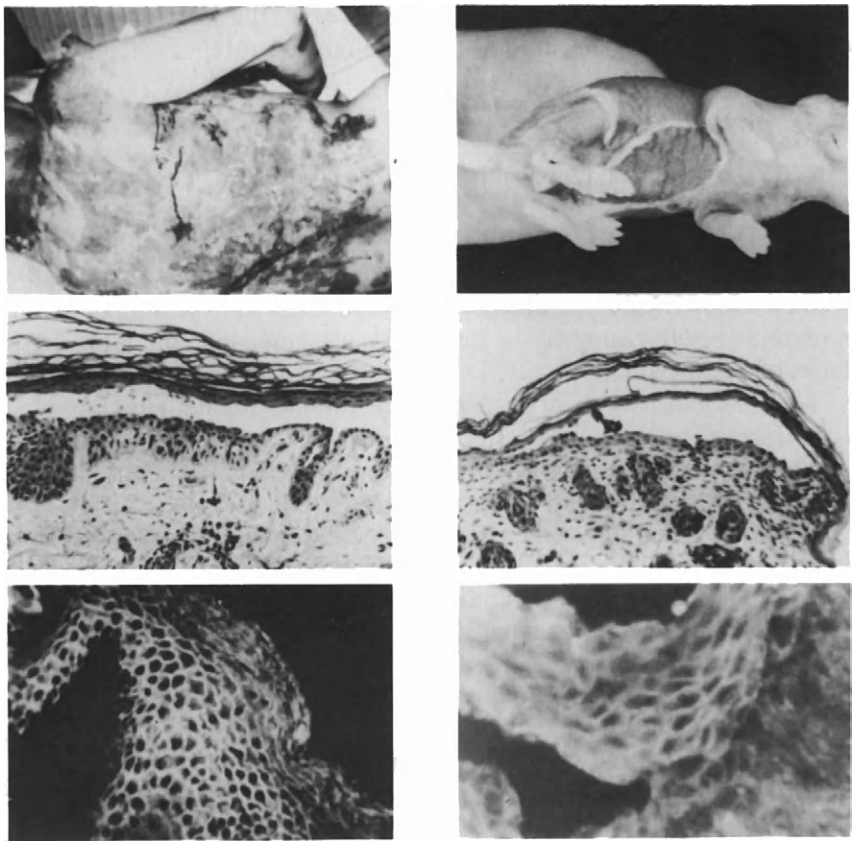


FIG. 1. Pemphigus vulgaris: Characteristic features of the disease. (Top to bottom) Extensive blisters and erosions, acantholytic intraepidermal blister formation, and deposition of IgG in the epidermal intercellular spaces. Left panels are human disease; right panels outline the disease reproduced in mice by passive transfer of human pemphigus immunoglobulins.

pemphigus vulgaris patients died during the first year of their disease (Lever, 1965). Death was usually secondarily caused by dehydration, electrolyte imbalance, malnutrition, and/or sepsis.

Pemphigus has been associated frequently with autoimmune disorders such as lupus erythematosus, rheumatoid arthritis, myasthenia gravis, pernicious anemia, Hashimoto's thyroiditis, and benign thymomas (Diaz *et al.*, 1980; Maize *et al.*, 1982). Usually the course of pemphigus is not affected by the presence of these disorders. Pemphigus has also been

associated with various malignancies; however, the incidence of malignancy is not high, and temporal relationships are rare.

C. Pathophysiology

The immunological abnormalities discovered two decades ago by Beutner and Jordon (1964) are consistently used as markers of the disease. These investigators showed by immunofluorescent (IF) techniques that the majority of patients with pemphigus have IgG-class autoantibodies directed against squamous epithelial cell surfaces. These autoantibodies are detected both bound to the diseased epithelium and circulating in the serum of these patients.

There is a large body of evidence that these epithelial autoantibodies are pathogenic in the disease. In the majority of patients with pemphigus, the titers of these antibodies as measured by indirect IF correlate with disease activity (Chorzelski *et al.*, 1979; Fitzpatrick and Newcomer, 1980), and plasmapheresis has been reported to induce short-term remissions in pemphigus patients (Rucco, 1977; Auerbach and Bystryn, 1979; Swanson and Dahl, 1981). Pemphigus has also affected neonates born to mothers with active pemphigus; the disease has resolved in these newborns after parturition (Moncada *et al.*, 1980; Storer *et al.*, 1982).

In vitro studies strongly support the pathogenic role of pemphigus antibodies in the induction of acantholysis. Bellone and Leone (1956), Schiltz and Michel (1976), and Morioka *et al.* (1981) demonstrated that pemphigus serum or IgG fractions from pemphigus serum induce acantholysis in human skin explants. Farb *et al.* (1978), Diaz and Marcelo (1978), and Woo *et al.* (1983) showed that pemphigus serum or IgG fractions from pemphigus serum, when added to murine primary epidermal cell cultures, will cause a highly specific epidermal cell detachment.

It has been shown (Farb *et al.*, 1978; Woo *et al.*, 1983) that the increased cell detachment observed in epidermal cultures treated with pemphigus IgG was abolished by certain protease inhibitors. This finding suggests that pemphigus antibodies, following their binding to epidermal cell-surface antigen(s), activate proteolytic enzymes; this was suggested by the studies of Farb *et al.* (1978) and Schlitz and Michel (1976). The released enzymes in turn may be responsible for the cell detachment induced by pemphigus IgG. These studies confirm early hypotheses postulated by Rothman (1957) and Lever (1965), who thought that injured epidermal cells in pemphigus may release pathogenic proteolytic enzymes. Hashimoto *et al.* (1982) identified a plasminogen activator (PA) in the supernatant of human cell cultures treated with pemphigus IgG. However, our laboratory recently demonstrated that in murine epidermal cell

cultures, pemphigus IgG promotes cell detachment independent of fibrinolytic (FA) or PA activities (Woo *et al.*, 1982). Removal of plasminogen from the growth media or treatment of the cultures with dexamethasone, which abolishes PA activity in the cultures, does not prevent the cell detachment induced by pemphigus antibodies. Furthermore, PA and FA activities in cultures treated with pemphigus IgG or normal human IgG were similar. These data obviously need further clarification.

Finally, there is also convincing *in vivo* evidence of the pathogenicity of pemphigus autoantibodies. Our laboratory demonstrated that IgG fractions purified from pemphigus serum, when given in sufficient doses, induce a disease in neonatal mice that reproduces the clinical, histological, and immunological features of the human disease (Fig. 1, right) (Anhalt *et al.*, 1982b). The extent of disease induced in these mice was dose dependent and correlated closely with the indirect IF titers for human pemphigus antibodies in the mouse serum. Despite these studies defining the pathogenicity of pemphigus antibodies, the mechanism by which these antibodies induce tissue injury is not clear.

There is some circumstantial evidence that activation of the complement system may be important in the development of tissue injury in pemphigus, but this is far from being clearly defined. Complement components Clq, C4, C3, and less frequently properdin and factor B, can be demonstrated in the intercellular space of early acantholytic lesions (Van Joost *et al.*, 1972; Jordon *et al.*, 1974; Jordon, 1976); pemphigus blister fluid shows a reduced total hemolytic complement activity, suggesting a local consumption of complement (Jordon and McDuffie, 1976). Despite the presence of complement components in lesional skin, it seems that most pemphigus antibodies do not fix complement directly (Jordan *et al.*, 1971). More significantly, those *in vitro* studies that demonstrated either acantholysis in skin explants (Schlitz and Michel, 1976) or cell detachment in primary epidermal cell culture (Diaz and Marcelo, 1978; Farb *et al.*, 1978; Woo *et al.*, 1984) that was induced by pemphigus antibodies were performed in complement-depleted media, implying that *in vivo* complement activation may be an epiphenomenon.

In lesions of pemphigus vulgaris, polymorphonuclear leucocytes infiltrate the acantholytic lesions to varying degrees. In some areas, acantholysis occurs with almost no inflammatory cells being present; in other areas, great numbers of neutrophils and eosinophils may be present. Knight *et al.* (1976) stated that in patients with pemphigus such inflammatory lesions occurred adjacent to noninflammatory lesions, even with the same biopsy specimens. Emmerson and Wilson-Jones (1968) applied the term "eosinophilic spongiosis" to such lesions and felt that they occurred most commonly early in the course of pemphigus. Our own studies

(Anhalt *et al.*, 1982b) in neonatal mice showed the same findings, that is, in some areas, intraepidermal vesicles were filled with polymorphonuclear cells and a few acantholytic epidermal cells, whereas the vesicles were devoid of neutrophils in other areas. The mechanism of neutrophil chemotaxis and the role that neutrophils subsequently play in the development of acantholysis have not been explored. It would seem highly probable that if the epidermal cells *in vivo* released extracellular proteases subsequent to pemphigus antibody binding [as they do *in vitro* (Schlitz and Michel, 1976; Morioka *et al.*, 1981)], these proteases would generate chemotactic fragments of C5 and C3 (Ward, 1967; Orr *et al.*, 1978). The generation of such complement fragments would recruit polymorphonuclear cells into the lesions. What subsequent role the neutrophil–epidermal cell interaction may play in the development or propagation of the lesions is unknown.

D. Diagnosis

In pemphigus, the characteristic histological lesion is an acantholytic intraepidermal blister (Fig. 1, left). Pemphigus autoantibodies almost always are detected by direct IF bound to the epidermal intercellular spaces of perilesional skin (Fig. 1, left), and only rarely will they fail to be shown in patients with pemphigus. Pemphigus antibodies are also detected in the circulation of ~80% of pemphigus patients. These serum antibodies appear to be highly specific for pemphigus, although pemphiguslike antibodies have been detected by indirect IF in certain individuals with high titers of isoagglutinins A and B, as well as in the sera of patients with burns (Cram *et al.*, 1974; Chorzelski *et al.*, 1979).

Clinically, the bullae of pemphigus must be distinguished from those of bullous pemphigoid and dermatitis herpetiformis. Usually there is no problem in making this differentiation; however, early lesions of pemphigus localized to the scalp or mouth may offer a challenge. In these circumstances, the immunological tests outlined above are extremely helpful.

E. Treatment

Therapy for pemphigus has changed radically during the past 30 years. It is now a controllable disease, whereas 30 years ago it was considered lethal (Lever, 1965; Lever and Schaumburg-Lever, 1977). The aim of therapy in pemphigus vulgaris is the complete elimination of pemphigus autoantibodies from the patient. The serum as well as the skin of these patients should become negative for pemphigus autoantibodies when test-

ing by indirect and direct IF techniques, respectively. It may take several weeks or months to induce serological remission in some patients.

The drugs that have had the greatest therapeutic success are the corticosteroids. Although they have produced a dramatic reduction in mortality, their chronic use has resulted in significant morbidity. The current mortality in pemphigus is 8–10% in 5 years, and is generally secondary to the complications of long-term steroid therapy (Rosenberg *et al.*, 1976). In an attempt to reduce steroid complications, investigators have successfully used immunosuppressive agents (e.g., azathioprine, methotrexate, and cyclophosphamide) and gold therapy in the treatment of pemphigus. This combined therapy is valuable in pemphigus and especially in those patients manifesting significant steroid complications. Preliminary studies suggest that plasmapheresis also may have a role in treatment of pemphigus (Rucco, 1977; Auerbach and Bystryn, 1979; Swanson and Dahl, 1981). It appears to be of greatest benefit as an adjunct to drug therapy in acute cases in which circulating autoantibody titers are highest.

III. BULLOUS PEMPHIGOID

A. Introduction

Bullous pemphigoid is a chronic, subepidermal, bullous dermatosis with a tendency to flexural distribution (Fig. 2, left). It occurs equally in males and females and has been described in all age groups, but it affects elderly persons far more commonly. There is no genetic predisposition or association with underlying malignant disease (Jordon *et al.*, 1967; Holubar *et al.*, 1975; Person *et al.*, 1976; Ahmed *et al.*, 1977a,b). The true incidence of bullous pemphigoid is unknown. About 5–10 new cases are seen yearly in a large hospital referral center.

The exact etiology has not been determined. Three striking findings that appear to be involved in the pathogenesis of the skin lesions are detected in these patients: (a) A blister cavity is formed as a result of detachment of epidermal basal cells from the basal lamina and is rapidly filled with fluid (Fig. 2, left). The cleavage of the dermal–epidermal junction (DEJ) or basement membrane zone (BMZ) occurs through the lamina lucida, as demonstrated by electron microscopy. (b) An acute inflammatory reaction is found in the papillary dermis adjacent to and beneath the blister cavity (Fig. 2, left); and (c) pemphigoid autoantibodies against the BMZ of squamous epidthelia (Jordon *et al.*, 1967) are found in lesional and perilesional skin (Fig. 2, left). Immunoelectron microscopic studies indicate

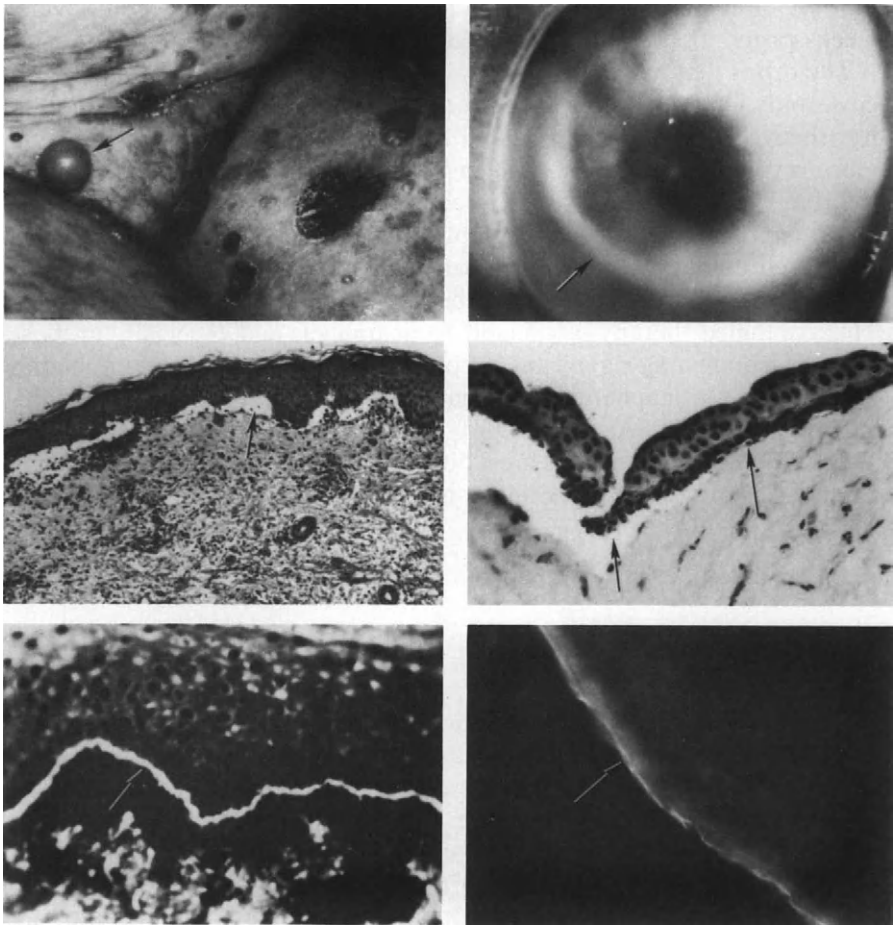


FIG. 2. Bullous pemphigoid: Characteristic features of the disease. (Top to bottom) Large, tense blisters and erosions in flexural areas, subepidermal blister formation, and deposition of IgG at the cutaneous basement membrane zone. Left panels, human disease; right panels show lesions produced in rabbit corneal epithelium by the local injection of human pemphigoid immunoglobulins. Corneal inflammation, subepithelial blister formation, and binding of human IgG occur at the epithelial–stromal junction.

that these antibodies are bound to antigen(s) localized in the lamina lucida of the BMZ (Holubar *et al.*, 1975).

B. Clinical Manifestations

Clinically, bullous pemphigoid represents a wide disease spectrum that ranges from a spontaneously remitting local bullous eruption to a generalized recurrent debilitating disease (Lever, 1965; Person *et al.*, 1976; Orr *et al.*, 1978). Itching is a variable symptom. Bizarre, gyrate urticarial plaques or erythema multiformelike lesions may preceed or follow the formation of bullae (Fig. 2, left). The Nikolsky sign is absent. Healing of the bullous lesions of localized pemphigoid usually results in scarring and secondary infection (Brunsting and Perry, 1957; Person *et al.*, 1976). The mucous membranes, including the esophagus, may be involved in 15–20% of patients. Although a history of true photosensitivity is absent, cutaneous lesions of bullous pemphigoid have been induced in these patients by ultraviolet irradiation (Cram and Fukuyama, 1972; Robinson *et al.*, 1978). The exact role of ultraviolet light in the genesis of lesions is unknown.

C. Pathophysiology

In 1967, Jordon *et al.* (1967) reported the presence of circulating anti-BMZ autoantibodies in the serum of patients with bullous pemphigoid. These antibodies, detected by direct IF techniques (Jordan *et al.*, 1967; Beutner *et al.*, 1970) were bound along the DEJ or BMZ of perilesional skin. Direct IF examination of perilesional skin in bullous pemphigoid patients also shows complement components, of both classical and alternative pathways, and fibrin, which is bound to the DEJ in a linear fashion (Provost and Tomasi, 1973, 1974; Jordon *et al.*, 1975b, 1978). Bullous pemphigoid autoantibodies from a great number of patients are known to fix complement *in vitro* (Jordon *et al.*, 1978). These findings suggest that *in situ* activation of complement may be an important step in the pathogenesis of the lesions of bullous pemphigoid. The acute inflammatory response seen in lesional skin may result from release of chemotactic factors for neutrophils and eosinophils. Complement activation and mast cell degranulation may be the origin of some of these chemotactic factors (Baba *et al.*, 1976; Diaz-Perez and Jordon, 1976; Wintroub *et al.*, 1976; Dvorak *et al.*, 1982).

There are *in vitro* studies that show that the circulating bullous pemphigoid autoantibodies may be pathogenic. Gammon *et al.* (1980) have shown that sections of normal human skin exposed to bullous pemphigoid

serum will fix bullous pemphigoid autoantibodies at the DEJ, and when treated with an exogenous source of complement and neutrophils, these sections will produce attachment of these cells along the DEJ. It seems that the attracted neutrophils can release proteases, which in turn may be responsible for the cleavage of the DEJ seen in these sections of skin and presumably in the lesions of patients with bullous pemphigoid. Recently, our laboratory has demonstrated that bullous pemphigoid IgG is pathogenic to rabbit corneal epithelium (Fig. 2, right). Bullous pemphigoid antigen is present at the epithelial–stromal (E–S) junction of both corneal epithelium and epidermis (Anhalt, 1982a). Briefly, these studies show that bullous pemphigoid antibodies injected into the corneal stroma will bind the bullous pemphigoid antigen at the E–S junction and trigger an intense inflammatory reaction in the rabbit cornea. This inflammatory reaction is associated with subepithelial vesiculation and migration of neutrophils toward the E–S junction. Interestingly, the pathogenic effects of several bullous pemphigoid IgG fractions tested in the rabbit cornea correlated with the dose and in the *in vitro* complement fixation properties of the injected bullous pemphigoid autoantibodies.

A recent study by Naito *et al.* (1982) showed that DEJ cleavage can be induced in human skin explants by incorporation of pemphigoid blister fluid in the growth media. The cleavage of the DEJ by pemphigoid blister fluid could be prevented by heating the fluid at 56°C for 30 min or by treatment of these samples with antisera against C1, C3, C4, and C5. Furthermore, α-2-macroglobulin was an effective inhibitor of this cleavage. These investigators were also able to induce subepidermal vesicles in the skin of guinea pigs injected with pemphigoid sera, which could be blocked by preincubation of the pemphigoid serum with antisera against human IgG. This investigation suggests that both serum and blister fluid of patients with bullous pemphigoid contain factors that are able to induce DEJ cleavage. The source and origin of these factors, other than antibodies and complement components, remain to be studied.

Early in the evolution of the bullous pemphigoid lesions, subepidermal microvacuoles coalesce and form large vacuoles and bullae (Lever and Schaumburg-Lever, 1975). The intact epidermis constitutes the roof of these lesions. A perivascular infiltrate and mild polymorphonuclear vasculitis are seen in lesional skin, and these cells infiltrate the blister cavity. Sequential histological studies on erythematous perilesional and lesional skin of bullous pemphigoid patients have shown histological changes compatible with an active mast cell degranulation process occurring in the upper dermis (Wintroub *et al.*, 1976; Dvorak *et al.*, 1982). Fibrin deposition is also detected throughout the papillary and reticular dermis of le-

sional skin (Dvorak *et al.*, 1982). The blister fluid from lesions of bullous pemphigoid has been shown to contain depressed values of total complement and individual components (C1–C7) (Jordon *et al.*, 1973), chemotactic factors for neutrophils (Diaz-Perez and Jordon, 1976) and eosinophils (Baba *et al.*, 1976; Wintroub *et al.*, 1976), complement fragments (Lohrisch *et al.*, 1980), prostaglandin E_2 (Standberg and Hagermark, 1977), and immune complexes (Jordon *et al.*, 1981). Recent studies (Mogavero *et al.*, 1982) showed that blister fluid from bullous pemphigoid patients contains elevated values of histamine and IgE as well as proteolytic enzymes such as prekallikrein activator, Hageman factor cleaver, and TAME-esterase-active substances. These various enzymes are thought to be of mast cell origin.

Although the immunological injury is thought to occur as a consequence of activation of the complement system following the binding of bullous pemphigoid autoantibodies to the BMZ, it is possible that bullous pemphigoid autoantibodies would cross-link basal cell-surface antigens and activate proteolytic enzymes, as demonstrated in other cell systems (Becker *et al.*, 1981) and especially in the pemphigus antibody–epidermal cell-surface reaction (Diaz and Marcelo, 1978; Farb *et al.*, 1978). The epidermal proteases may in turn activate the complement system with subsequent mast cell degranulation, migration of eosinophils and neutrophils, and release of other inflammatory mediators.

D. Diagnosis

Routine histological examination of a bulla reveals a subepidermal blister (Fig. 2, left). Detection of pemphigoid anti-BMZ antibodies is important in the evaluation of bullous pemphigoid patients and especially in evaluation of bullous pemphigoid patients with atypical lesions (Beutner *et al.*, 1970). Direct IF examination of perilesional skin usually will show linear deposition of IgG and C3 along the BMZ (Fig. 2, left). It is important to remember that perilesional skin biopsies may display only a linear C3 deposition along the BMZ in the apparent absence of IgG deposition (Ahmed *et al.*, 1977c). Examination of serum by indirect IF techniques demonstrates a circulating bullous pemphigoid antibodies in ~90% of patients. The individual titers, however, show no correlation with activity or extent of the disease.

The anti-BMZ antibodies found in bullous pemphigoid are highly specific. The presence of anti-BMZ antibodies in the absence of bullous pemphigoid is rare (Provost *et al.*, 1979). Bullous pemphigoid may be confused clinically and histologically with other vesiculobullous eruptions

such as pemphigus vulgaris, erythema multiforme, and dermatitis herpetiformis. The immunofluorescent studies outlined above, however, easily resolve these diagnostic problems.

E. Treatment

Therapy should be adjusted to the individual patient. Local eruptions may respond to use of wet dressings and topical steroids. A widespread bullous pemphigoid eruption is an indication for hospitalization and systemic steroid therapy. Dosages of steroid should be adapted to each patient. On disappearance of serum and *in vivo* bound BMZ antibodies, therapy has been successfully discontinued in a large proportion of bullous pemphigoid patients. The vast majority have remained in prolonged clinical and serological remission.

Bullous pemphigoid generally is a benign, self-limited disease. Before the use of steroids, a 30% mortality was reported (although this figure may be high). At present only a few patients die of bullous pemphigoid, and successful therapy induces prolonged clinical and serological remissions. Recurrences or relapses are infrequent and do not seem to correlate with the severity of the disease, as seen in the initial presentation, or its duration. Complications of steroid and immunosuppressive therapy, however, are always potential clinical problems. The morbidity and mortality in bullous pemphigoid are undoubtedly directly or indirectly related to the use of these therapeutic agents.

IV. HERPES GESTATIONIS

A probable variant of bullous pemphigoid is herpes gestationis. Herpes gestationis is a subepidermal blistering disease occurring during the second or third trimester of pregnancy. This disease process is characterized clinically by an intensely pruritic subepidermal blister occurring on an erythematous base. Unlike bullous pemphigoid, there seems to be a definite precipitation of the disease by pregnancy. There are recurrences with subsequent pregnancies or with the use of certain birth control pills (Lawley *et al.*, 1978).

These patients invariably demonstrate the deposition of C3 along the BMZ of perilesional skin (Katz *et al.*, 1976). The immunoreactants, as in bullous pemphigoid, are localized in the lamina lucida of the DEJ (Yaoita *et al.*, 1976). In addition, ~25% of these patients will show in their sera a circulating, avidly complement-fixing, anti-BMZ IgG.

The majority of patients with herpes gestationis fail to show the anti-BMZ IgG in their sera by routine indirect IF techniques. In 50% of these

patients, however, it is possible to demonstrate an IgG autoantibody (HG factor) directed against the DEJ by complement IF technique (Jordon *et al.*, 1976; Katz *et al.*, 1976). The HG factor shown by this technique is able to avidly fix C3 to the DEJ of normal skin sections.

Infants born to mothers with herpes gestationis may suffer a transient subepidermal blistering eruption (Lawley *et al.*, 1978). The skin of these neonates also demonstrates the deposition of C3 along the DEJ of perilesional skin. The cutaneous disease in these infants is thought to be triggered by the transplacental passage of the maternal anti-BMZ autoantibodies. Subsequent *in situ* activation of the complement sequence along the BMZ of the infant's skin may induce the inflammatory subepidermal vesicles. In these neonates, cutaneous lesions and immunoreactant deposits along the BMZ disappear within the first 6 months of life.

V. OTHER VESICULOBULLOUS DERMATOSES

It has recently been recognized that ~10–15% of patients diagnosed as having dermatitis herpetiformis (DH) have features that suggest they may have a different, possibly unrelated disorder. By IF techniques, they demonstrate a linear (not granular) deposit of IgA at the DEJ of perilesional skin, and some have circulating IgA anti-BMZ antibodies. Immunoelectron microscopy shows that these immunoglobulin deposits are within the lamina lucida (as in BP) in some patients and below the basal lamina in others (Lawley *et al.*, 1980). Unlike DH patients with granular IgA deposits in skin, those patients with linear IgA deposits do not have the gluten-sensitive enteropathy or the consistent association with the HLA-B8, -DW3 phenotype (Lawley *et al.*, 1980). Until the relationship of this subset of patients to DH is established more clearly, several investigators have proposed separate names for this disorder, such as linear IgA, DH, IgA bullous pemphigoid, or vesiculobullous dermatosis with linear IgA deposition (Chorzelski *et al.*, 1971; Jablonska *et al.*, 1976; Provost *et al.*, 1979).

Bullous pemphigoid and cicatricial pemphigoid sometimes have been difficult to differentiate from epidermolysis bullosa acquisita (EBA) because of the overlapping clinical and immunopathological findings of these conditions. In 1971, Roenigk *et al.* (1971) suggested the following four criteria for the diagnosis of EBA: (a) Clinical lesions consisting of trauma-induced bullae, occurring predominantly over the joints of the hands, feet, elbows, and knees, with atrophic scars, milia, and nail dystrophy; (b) onset of the disease after infancy; (c) no family history of hereditary epidermolysis bullosa; and (d) exclusion of other bullous dis-

eases. Two more diagnostic criteria have been added recently (Roenigk and Pearson, 1981): (e) The finding of linear deposition of IgG at the DEJ of perilesional skin by direct IF techniques, and (f) demonstration of a blister cavity occurring beneath the basal lamina. These studies have also shown, by immunoelectron microscopy, a zone of amorphous material just beneath the basal lamina that contains deposits of IgG (Yaoita *et al.*, 1981). Furthermore, several of these patients have demonstrated serum antibodies directed against an antigen(s) located in the dermal domain of the BMZ in the region of anchorage of the anchoring fibers (Yaoita *et al.*, 1981). These findings strongly support the hypothesis that EBA is a separate clinicopathological entity in which an associated autoimmune phenomenon may also be important in its pathogenesis.

VI. LUPUS ERYTHEMATOSUS

A. Introduction

There is now a great deal of data to document the existence of various autoimmune phenomena in systemic lupus erythematosus (SLE) patients. These patients make a variety of autoantibodies against various nuclear and cytoplasmic macromolecules, and they also frequently produce autoantibodies against various neutrophil, lymphocyte, and red blood cell-surface antigens. In addition, SLE patients commonly demonstrate serum immune complex formation. These immune complexes have been implicated in the vasculitis and glomerulonephritis commonly observed in SLE.

Lupus patients frequently develop cutaneous lesions. Approximately 20–25% of SLE patients will initially present with prominent cutaneous lesions. A number of the cutaneous lesions are so distinctive that they are viewed as being highly specific for SLE and are part of the American Rheumatism Association's diagnostic criteria for the disease (Tan *et al.*, 1982). These highly characteristic mucocutaneous lesions are malar dermatitis (butterfly), discoid lesions, photosensitivity, and oral ulcerations. Other cutaneous lesions, although not specific for lupus, occur frequently; these include cicatrical alopecia, urticaria-like lesions, panniculitis, vasculitic lesions, periunqual telangectasia, and Raynaud's phenomena.

A great deal of work has been performed investigating the pathogenesis of these cutaneous manifestations of lupus erythematosus. It seems highly probable that immune complex deposition plays a major, if not the dominant, role in the pathogenesis of some of these lesions (e.g., urtica-

ria-like lesions associated with vasculitis). These lesions are frequently detected in those SLE patients demonstrating serum immune complexes, anti-nDNA, Sm, nRNP, Ro(SSA), and/or La(SSB) antibodies, and frequently demonstrate immunoglobulin and complement deposits in the diseased blood vessels (O'Loughlin *et al.*, 1978; Provost *et al.*, 1980).

B. Pathophysiology of Cutaneous Lupus Lesions

The classic cutaneous lupus (discoid) lesion has also been extensively investigated, but as yet the pathogenesis of this distinctive lesion is unknown. This lesion is characterized as an annular, hyperkeratotic plaque, with telangectasia, follicular plugging, and atrophy. Hypo- and hyperpigmentation may be prominent (Fig. 3A).

This classic cutaneous lupus lesion can be the sole manifestation of lupus, in which case the condition is termed benign cutaneous lupus. The lupus lesion can also be seen in the presence of systemic disease (i.e., arthritis renal disease, etc.). Whether or not the cutaneous lesion is occurring in the setting of benign or systemic lupus cannot be determined with certainty by clinical or histological evaluation alone. A history and physical examination and serological examination of the patient are necessary to determine the cutaneous or systemic nature of the lupus process.

The histological features of these lesions are orthokeratosis, patchy epidermal atrophy, hydropic degeneration of the epidermal basal call layer, and a patchy mononuclear infiltrate hugging the DEJ and/or surrounding dermal appendages as hair follicles and sweat glands (Fig. 3B). The mononuclear infiltrate has been examined employing monoclonal antibodies and the peroxidase–antiperoxidase technique, and these studies have demonstrated prominent T cells (suppressor and helper) in this infiltrate. These T cells are Ia positive, suggesting they are present in an activated state. B cells and macrophages are present, but are less prominent (Synkowski and Provost, 1983).

Lupus lesions frequently demonstrate deposition of granular deposits of immunoglobulin and complement components along the cutaneous DEJ (Fig. 3C). The deposits are composed of IgG, IgM, IgA, properdin, properdin factor B, β-1-H globulin, and fibrin (Jordon *et al.*, 1975a; Carlo *et al.*, 1979). In addition, complement components Clq, C4, C3, and the membrane attack complex (C5–C9) are frequently detected (Biesecker *et al.*, 1982).

Despite the deposits of these various immunoreactants at the cutaneous DEJ of lupus lesions, the weight of the evidence suggests these deposits do not play a primary role in the pathogenesis of the lesions. For example, cutaneous lupus lesions can be induced in SLE patients by exposure to

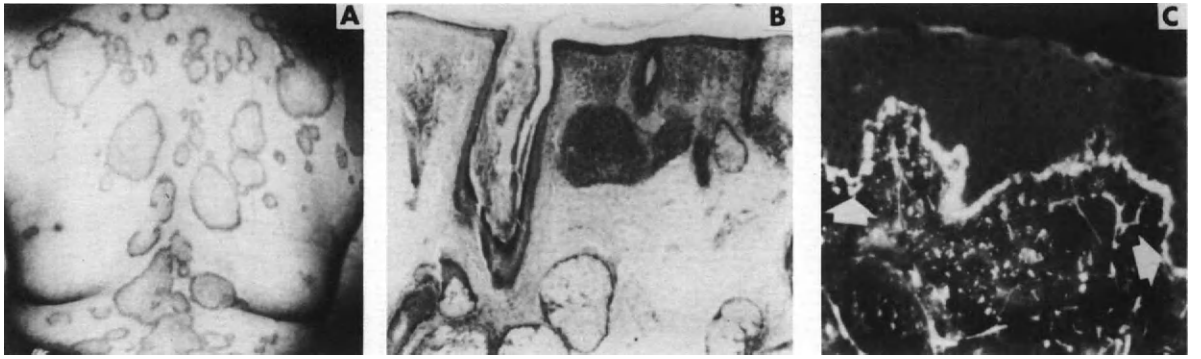


FIG. 3. Characteristic features of cutaneous lupus erythematosus. (A) Annular, erythematous lesions of trunk. (B) Lymphocytic dermal infiltration and vacuolar interface dermatitis. (C) Deposits of immunoreactants (IgG, IgM, C3) at the basement membrane zone of lesional skin.

ultraviolet light (UV) in the sunburn spectrum (Freeman *et al.*, 1969; Cripps and Rankin, 1973). These lupus lesions appear clinically and histologically similar to spontaneously occurring lupus lesions and develop in the absence of the immunoglobulin and complement deposits at the DEJ. In a few instances these deposits appeared, but only after the lupus lesion had been present for several weeks. In addition to these studies, similar immunoglobulin and complement components have been detected at the DEJ of normal-appearing skin of SLE patients (Provost and Tomasi, 1973; Carlo *et al.*, 1979). The demonstration of these deposits is termed the lupus band test. Despite a good deal of evidence demonstrating an association between renal disease (Gilliam *et al.*, 1974; Grossman *et al.*, 1974; Wertheimer and Barland, 1976; Provost *et al.*, 1980), hypocomplementemia, and anti-nDNA antibodies, these cutaneous deposits are not associated with the presence of skin disease. In fact, one study has carefully documented that these deposits can occur in the complete absence of any inflammatory infiltrate (Schrager and Rothfield, 1976).

Recent studies suggest that the membrane attack components of the complement sequence (C5–C9), which are present only in lesional and not in uninvolved skin of lupus patients, may be of pathological significance (Biesecker *et al.*, 1982). Much work must be done to examine this interesting observation. It is important to remember that prominent and classic cutaneous lupus lesions have been reported in various homozygous complement deficiency states (C2 and C4) in association with lupuslike syndromes (Agnello *et al.*, 1975; Tappeiner *et al.*, 1982).

The data cited above led Gilliam's group (1980) and ours (Provost, 1979a) to propose an alternative hypothesis to explain the pathogenesis of the lupus lesion. We believe that T cells may play the predominant role in the pathogenesis of the cutaneous lupus lesion. These T cells are envisioned to be specifically reactive against various DNA nucleotides.

As a corollary to this hypothesis, it is envisioned that UV light induces alterations in epidermal cell DNA. This UV-altered DNA, which contains UV-specific, single-strand, native DNA antigenic determinants, is extruded from the epidermal cells, presumably via the action of endonucleases. This denatured DNA then crosses the cutaneous BMZ, where it may bind to the collagen or gain entrance into the systemic circulation (Izui *et al.*, 1976). At the cutaneous DEJ, T cells are thought to bind the altered DNA, resulting in a cell-mediated immune response that produces the inflammatory cutaneous lesion we recognize as cutaneous lupus. We also hypothesize that the altered DNA absorbed into the circulation can result in the production of pathological immune complex formation.

Thus, we envision that the severe systemic and cutaneous flares of lupus associated with sunlight exposure (photosensitivity) result from the

generation of altered DNA by UV light exposure in a sensitive host. As support for such a concept, it is important to note that following sunlight exposure, free circulating DNA has been found in the serum of SLE patients (Tan *et al.*, 1966). Furthermore, it is known that peripheral blood cells of SLE are capable of responding with a proliferative response when cultured *in vitro* in the presence of DNA fractions (Patrucco *et al.*, 1967).

This hypothesis, however, has recently been challenged by studies of the rare neonatal lupus syndrome (Franco *et al.*, 1981; Kephart *et al.*, 1981). During the neonatal period these infants develop large annular lupus lesions (Fig. 4A,B). These lesions may be photo induced because of their relative predominance in light-exposed areas. These lupus lesions demonstrate classic histological features of lupus, generally show pronounced liquefaction degeneration of epidermal basal cells (Fig. 4C), and resemble the large annular polycyclic lesions of subacute cutaneous lupus erythematosus described by Sontheimer *et al.* (1979). The lupus lesions generally disappear without scar formation during the neonatal period, although hypopigmentation may transiently be present.

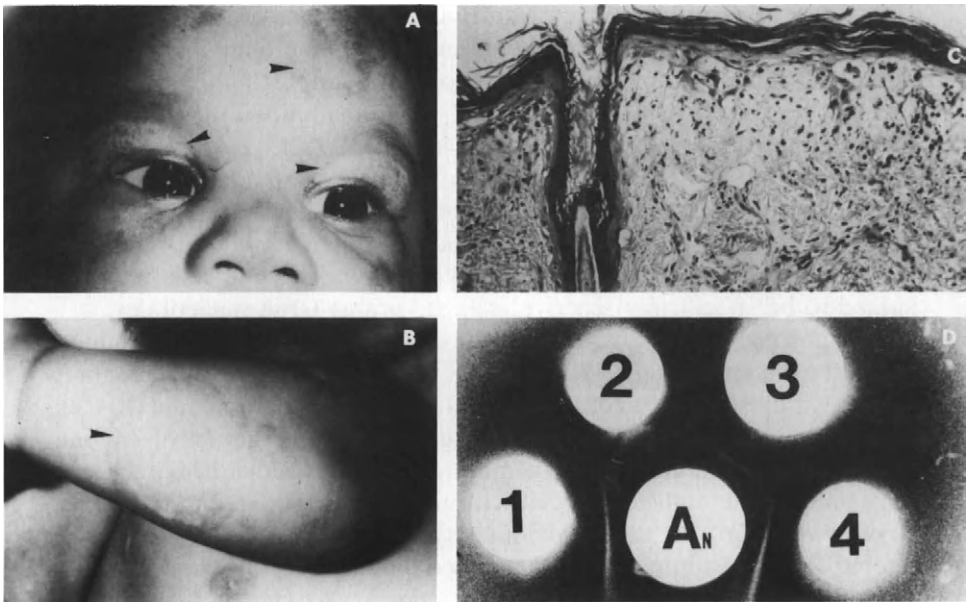


FIG. 4. Characteristic features of neonatal lupus. (A) and (B) Annular, infiltrated lesions of face and extremities (arrows). (C) Lymphocytic vacuolar interface dermatitis. (D) Demonstration of Ro-specific antibodies in serum of both mother (well 1) and child (well 2). Wells 3 and 4 are positive controls, and center well (An) contains Ro antigen.

In addition to these cutaneous lesions, these neonatal infants may demonstrate cardiac abnormalities, the most frequent of which are cardiac conduction defects (Chameides *et al.*, 1977). The severity of this abnormality ranges from first degree to complete heart block. Autopsies have demonstrated various cardiac abnormalities and generalized fibrosis of the cardiac conduction fibers. Infants with neonatal lupus also frequently develop hepatosplenomegaly, lymphadenopathy, and Coombs'-positive hemolytic anemia.

These infants and their mothers demonstrate antibodies reactive against La(SSB) and/or Ro(SSA) macromolecules (Fig. 4D) (Franco *et al.*, 1981; Kephart *et al.*, 1981). These antibodies are highly specific for SLE and Sjögren's syndrome, and they are rarely detected in symptomatic (nonconnective tissue disease) patients (Maddison *et al.*, 1979). These antibodies [especially the Ro(SSA) antibody] are seen in an SLE patient population that has an increased prevalence of photosensitivity, rheumatoid factor, and Sjögren's syndrome (Maddison *et al.*, 1979, 1981). In addition, these antibodies have been described with increased frequency in several distinctive forms of lupus, such as subacute cutaneous lupus, ANA-negative, and lupuslike syndromes in homozygous C2- and possibly C4-deficient patients (Sontheimer *et al.*, 1979; Maddison *et al.*, 1981; Provost *et al.*, 1982).

These antibodies in Sjögren's syndrome patients are associated with a markedly increased prevalence of such extraglandular manifestations as anemia, lymphopenia, neutropenia, lymphadenopathy, and vasculitis (Alexander *et al.*, 1982a). The vasculitis in Sjögren's patients possessing the La(SSB) and/or Ro(SSA) antibody systems can involve visceral organs, the skin, and the peripheral and central nervous system (Alexander *et al.*, 1982b; Alexander and Provost, 1983). Thus far, all 17 mothers and infants with neonatal lupus have demonstrated La(SSB) or Ro(SSA) antibodies (Franco *et al.*, 1981; Kephart *et al.*, 1981). These antibodies and the associated skin disease disappear during the neonatal period, but the cardiac conduction defects persist. The mothers of these infants are generally asymptomatic. Clinical features of SLE are demonstrated by 20% of mothers, and perhaps another 20% will subsequently develop clinical disease.

These data strongly suggest that the maternal antibodies La and/or Ro (or other antibodies) are passed across the placenta to the infant, where they produce the cardiac and cutaneous lesions. The exact pathophysiological mechanism(s) responsible is unknown. It is conceivable that antibody-dependent cellular cytotoxicity could play a pathological role. Studies by Norris *et al.* (1982) have shown that the Ro antibody is capable of participating in antibody-dependent cellular cytotoxicity. Furthermore, T

cells are very efficient effector cells in this Ro antibody-mediated cellular-dependent cytotoxicity.

Additional immunogenetic studies have demonstrated a statistically significant association of the Ro antibody with the HLA-Dr2 or HLA-DR3 phenotype (Bell and Maddison, 1980; Ahearn *et al.*, 1982). These studies, when applied to neonatal infants and their mothers, revealed some data potentially very important (Lee *et al.*, 1983). The mothers usually possess the HLA-DR3 phenotype, but the infants generally do not. This means that the Ro(SSA) antibody, undoubtedly of maternal origin, is associated with the DR3 phenotype, but that the clinical phenotype expression of the lupus lesion and the cardiac lesion are non-HLA linked.

These studies as outlined above suggest that several immune mechanisms may be responsible for the pathogenesis of the cutaneous lupus lesions. At present cell-mediated immunity and antibody-dependent cellular cytotoxicity are likely candidates.

REFERENCES

Agnello, V., Ruddy, S., Winchester, R. J., Christian, C. L., and Kunkel, H. G. (1975). *In* "Original Article Series" (D. Bergsma, R. A. Good, J. Finstad, and N. W. Paul, eds.), pp. 312–316. Sinauer, Sunderland, Massachusetts.

Ahearn, J. M., Provost, T. T., Dorsch, C. A., Stevens, M. B., Bias, W. B., and Arnett, F. C. (1982). *Arthritis Rheum.* **25,** 1031–1040.

Ahmed, A. R., Chu, T. M., and Provost, T. T. (1977a). *Arch. Dermatol.* **113,** 969.

Ahmed, A. R., Cohen, E., Blumenson, L., and Provost, T. T. (1977b). *Arch. Dermatol.* **113,** 1121.

Ahmed, A. R., Maize, J., and Provost, T. T. (1977c). *Arch. Dermatol.* **113,** 1043–1046.

Alexander, E., and Provost, T. T. (1983). *J. Invest. Dermatol.* **80,** 386–391.

Alexander, E., Arnett, F. C., Provost, T. T., and Stevens, M. B. (1982a). *J. Rheumatol.* **9,** 239–246.

Alexander, E., Provost, T. T., Stevens, M. B., and Alexander, G. E. (1982b). **61,** 247–257.

Anhalt, G. J., Bahn, C. F., Labib, R. S., Voorhees, J. J., and Diaz, L. A. (1982a). *J. Clin. Invest.* **68,** 1097–1101.

Anhalt, G. J., Labib, R. S., Voorhees, J. J., Beals, T. F., and Diaz, L. A. (1982b). *N. Engl. J. Med.* **306,** 1189–1196.

Auerbach, R., and Bystryn, J. (1979). *Arch. Dermatol.* **115,** 728–730.

Baba, T., Sonozaki, H., Seki, K., Uchiyama, M., Ikesawa, Y., and Torisu, M. (1976). *J. Immunol.* **116,** 112–116.

Becker, D., Ossowski, L., and Reich, E. (1981). *J. Exp. Med.* **154,** 385–396.

Bell, D. A., and Maddison, P. J. (1980). *Arthritis Rheum.* **23,** 1268–1273.

Bellone, A. G., and Leone, V. (1956). *G. Ital. Dermatol. Sifilol.* **97,** 97–109.

Beutner, E. H., and Jordon, R. E. (1964). *Proc. Soc. Exp. Biol. Med.* **117,** 505–510.

Beutner, E. R., Chorzelski, T. P., and Jordon, R. E. (1970). *In* "Autosensitization in Pemphigus and Bullous Pemphigoid" pp. 5–31. Thomas, Springfield, Illinois.

Biesecker, G., Lavin, L., Ziskind, M., and Koffler, D. (1982). *N. Eng. J. Med.* **306,** 264–270.

Brautbar, C., Moscovitz, M., Livshits, T., Haim, S., Hacham-Zadeh, S., Cohen, H. A., Sharon, R., Nelker, D., and Cohen, T. (1980). *Tissue Antigens* **16,** 238–243.

Brunsting, L. A., and Perry, H. O. (1957). *Arch. Dermatol.* **75,** 489–501.

Carlo, J. R., Rothfield, N. F., and Ruddy, S. (1979). *Arthritis Rheum.* **22,** 13–18.

Chameides, L., Truex, R. C., Vetter, V., Rashkind, W. J., and Galioto, F. M., Jr. (1977). *N. Engl. J. Med.* **297,** 1204–1207.

Chorzelski, T. P., Beutner, E. H., Jablonska, S., Blaszczyk, M., and Triftshauser, C. (1971). *J. Invest. Dermatol.* **56,** 373–380.

Chorzelski, T. P., Beutner, E. H., and Jablonska, S. (1979). *In* "Immunopathology of the Skin" (E. H. Beutner, T. P. Chorzelski, and S. F. Bean, eds.), 2nd ed., pp. 183–230. Wiley, New York.

Cram, D. L., and Fukuyama, K. (1972). *Arch. Dermatol.* **106,** 819–824.

Cram, D. L., Griffith, M. R., and Fukuyama, K. (1974). *Arch. Dermatol.* **109,** 235–238.

Cripps, D. J., and Rankin, J. (1973). *Arch. Dermatol.* **107,** 563–567.

Diaz, L. A., and Marcelo, C. L. (1978). *Br. J. Dermatol.* **98,** 631–637.

Diaz, L. A., Glamb, R. W., and Silva, J., Jr. (1980). *Arch. Dermatol.* **116,** 77–79.

Diaz-Perez, J. L., and Jordon, R. E. (1976). *Clin. Immunol. Immunopathol.* **5,** 360–370.

Dvorak, A. M., Mihm, M. C., Jr., Osage, J. E., Kawan, T. H., Austen, K. F., and Wintroub, B. U. (1982). *J. Invest. Dermatol.* **78,** 98–101.

Emmerson, R. W., and Wilson-Jones, E. (1968). *Arch. Dermatol.* **97,** 252–257.

Farb, R. M., Dykes, R., and Lazarus, G. S. (1978). *Proc. Natl. Acad. Sci. U.S.A.* **75,** 459–463.

Fitzpatrick, R. E., and Newcomer, V. D. (1980). *Arch. Dermatol.* **116,** 285–290.

Franco, H. L., Weston, W. L., Peebles, C., Forstot, S. L., and Phanuphak, P. (1981). *J. Am. Acad. Dermatol.* **4,** 67–72.

Freeman, R. G., Knox, J. M., and Ownes, D. W. (1969). *Arch. Dermatol.* **100,** 677–682.

Gammon, W. R., Lewis, D. M., Carlo, J. R., Sams, W. M., and Wheeler, C. E. (1980). *J. Invest. Dermatol.* **75,** 334–339.

Gilliam, J. N. (1980). *In* "Immunodermatology" (B. Safai, and R. A. Good, eds.), pp. 323–332. Plenum, New York.

Gilliam, J. N., Cheatum, D. E., Hurd, E. R., Stastny, P., and Ziff, M. (1974). *J. Clin. Invest.* **53,** 1434–1440.

Grossman, J., Callerame, M. L., and Condemi, J. J. (1974). *Ann. Intern. Med.* **80,** 496–500.

Hashimoto, K., Singer, K. H., Safran, K. M., and Lazarus, G. J. (1982). *Clin. Res.* **30,** 588A.

Holubar, K., Wolff, K., Donrad, K., and Beutner, E. H. (1975). *J. Invest. Dermatol.* **68,** 220–227.

Izui, S., Lambert, P., and Meischer, P. A. (1976). *J. Exp. Med.* **114,** 428–443.

Jablonska, S., Chorzelski, T. P., Beutner, E. H., Maciejowka, E., and Rzesa, G. (1976). *Arch. Dermatol.* **112,** 45–48.

Jordon, R. E. (1976). *J. Invest. Dermatol.* **67,** 366–371.

Jordon, R. E., and McDuffie, F. C. (1976). *Proc. Soc. Exp. Biol. Med.* **151,** 594–598.

Jordon, R. E., Beutner, E. H., Witebsky, E., Blumenthal, G., Hale, W. L. and Lever, W. F. (1967). *JAMA, J. Am. Med. Assoc.* **200,** 751–756.

Jordon, R. E., Sams, W. F., Jr., Diaz, G., and Beutner, E. H. (1971). *J. Invest. Dermatol.* **57,** 407–410.

Jordon, R. E., Day, N. K., Sams, W. M., Jr., and Good, R. A. (1973). *J. Clin. Invest.* **52,** 1207–1214.

Jordon, R. E., Schroeter, A. L., Rogers, R. S., III, and Perry, H. O. (1974). *J. Invest. Dermatol.* **63,** 256–259.

Jordon, R. E., Schroeter, A. L., and Winkelmann, R. K. (1975a). *Br. J. Dermatol.* **92,** 263–268.
Jordon, R. E., Schroeter, A. L., Good, R. A., and Day, N. K. (1975b). *Clin. Immunol. Immunopathol.* **3,** 307–314.
Jordon, R. E., Heine, K. G., Tappeiner, G., Buskell, L. L., and Provost, T. T. (1976). *J. Clin. Invest.* **57,** 1426–1433.
Jordon, R. E., Norby-McFarland, J. M., and Tappeiner, G. (1978). *J. Clin. Lab. Immunol.* **1,** 59–65.
Jordon, R. E., Struve, M. F., and Bushkell, L. L. (1981). *Clin. Exp. Immunol.* **45,** 29–36.
Katz, S. I., Hertz, K. C., and Yaoita, H. (1976). *J. Clin. Invest.* **57,** 1434–1441.
Kephart, D. C., Hood, A. F., and Provost, T. T. (1981). *J. Invest. Dermatol.* **77,** 331–333.
Knight, A. G., Black, M. M., and Delaney, T. J. (1976). *Clin. Exp. Dermatol.* **1,** 141–153.
Lawley, T. J., Stingl, G., and Katz, S. I. (1978). *Arch. Dermatol.* **114,** 552–555.
Lawley, T. J., Strober, W., Yaoita, H., and Katz, S. I. (1980). *J. Invest. Dermatol.* **74,** 9–12.
Lee, L. A., Weston, W. L., Norris, D. A., Huff, J. C., Bias, W. B., Arnett, F. C., and Provost, T. T. (1983). *Clin. Res.* **31,** 148A.
Lever, W. F. (1965). *In* "Pemphigus and Pemphigoid" (A. C. Curtis, ed.), pp. 15–72. Thomas, Springfield, Illinois.
Lever, W. F., and Schaumburg-Lever, G. (1975). *In* "Histopathology of the Skin" 5th ed., pp. 115–118. Lippincott, Philadelphia, Pennsylvania.
Lever, W. F., and Schaumburg-Lever, G. (1977). *Arch. Dermatol.* **113,** 1236–1241.
Lohrisch, I., Herkmann, K., and Haustein, U. F. (1980). *Br. J. Dermatol.* **102,** 135–143.
Maddison, P. J., Mogavero, H., Provost, T. T., and Reichlin, M. (1979). *J. Rheumatol.* **6,** 189–195.
Maddison, P. J., Provost, T. T., and Reichlin, M. (1981). *Medicine (Baltimore)* **60,** 87–93.
Maize, J. C., Green, D., and Provost, T. T. (1982). *J. Am. Acad. Dermatol.* **7,** 736–741.
Mogavero, H. S., Jr., Meier, H. L., Newball, H. H., and Provost, T. T. (1982). *Clin. Res.* **30,** 598A.
Moncada, B., Kettelsen, S., Hernandes-Moctezuma, J. L., and Ramirez, F. (1980). *Clin. Res.* **28,** 251A.
Morioka, S., Naito, K., and Ogawa, H. (1981). *J. Invest. Dermatol.* **76,** 337–341.
Naito, K., Morioka, S., and Ogawa, H. (1982). *J. Invest. Dermatol.* **79,** 303–306.
Norris, D. A., Sr., Fritz, K. A., Tan, E. M., Day, J. S., and Weston, W. L. (1982). *Clin. Res.* **50,** 264A.
O'Loughlin, S., Schroeter, A. L., and Jordon, R. (1978). *Arch. Dermatol.* **114,** 879–883.
Orr, W., Varani, J., and Ward, P. A. (1978). *Am. J. Pathol.* **93,** 405–422.
Park, M. S., Terasaki, P. I., Ahmed, A. R., and Tiwari, J. L. (1979). *Lancet Sept. 1,* pp. 441–442.
Patrucco, A., Rothfield, N. A., and Hirschhorn, K. (1967). *Arthritis Rheum.* **10,** 32–37.
Person, J. R., Rogers, R. S., III, and Perry, H. O. (1976). *Br. J. Dermatol.* **95,** 531–534.
Provost, T. T. (1979a). *Am. J. of Dermpathol.* **1,** 181–184.
Provost, T. T. (1979b). *J. Invest. Dermatol.* **72,** 110–113.
Provost, T. T., and Tomasi, T. B., Jr. (1973). *J. Clin. Invest.* **52,** 1779–1787.
Provost, T. T., and Tomasi, T. B., Jr. (1974). *Clin. Exp. Immunol.* **18,** 193–200.
Provost, T. T., Ahmed, A. R., Maddison, P., and Reichlin, M. (1977). *Arthritis Rheum.* **20,** 1457–1463.
Provost, T. T., Ahmed, R., Maize, J. C., Strauss, J., and Dobson, R. L. (1979). *Arch. Dermatol.* **115,** 156–160.
Provost, T. T., Andres, G., Maddison, P. J., and Reichlin, M. (1980). *J. Invest. Dermatol.* **74,** 407–412.

Provost, T. T., Zone, J. J., Synkowski, D., Maddison, P. J., and Reichlin, M. (1980). *J. Invest. Dermatol.* **75,** 495–499.
Provost, T. T., Arnett, F. C., and Reichlin, M. (1982). *Arthritis Rheum.* **25,** 541A.
Robinson, J. K., Baughman, R. D., and Provost, T. T. (1978). *Br. J. Dermatol.* **99,** 709–713.
Roenigk, H. H., and Pearson, R. W. (1981). *Arch. Dermatol.* **117,** 383.
Roenigk, H. H., Ryan, J. G., and Bergfeld, W. F. (1971). *Arch. Dermatol.* **103,** 1–10.
Rosenberg, F. R., Sanders, S., and Nelson, C. T. (1976). *Arch. Dermatol.* **112,** 962–970.
Rothman, S. (1957). *Arch. Dermatol.* **76,** 277–281.
Rucco, V. (1977). *Br. J. Dermatol.* **98,** 237–241.
Schlitz, J. R., and Michel, B. (1976). *J. Invest. Dermatol.* **67,** 254–260.
Schrager, M. A., and Rothfield, N. F. (1976). *J. Clin. Invest.* **57,** 212–221.
Sontheimer, R. D., Thomas, J. R., and Gilliam, J. N. (1979). *Arch. Dermatol.* **115,** 1409–1415.
Standberg, K., and Hagermark, O. (1977). *Acta Derm.-Venereol.* **57,** 487–492.
Storer, J. S., Galen, W. K., Nesbitt, L. T., and DeLeo, V. A. (1982). *J. Am. Acad. Dermatol.* **6,** 929–932.
Swanson, D. L., and Dahl, M. V. (1981). *J. Am. Acad. Dermatol.* **4,** 325–328.
Synkowski, D. R., and Provost, T. T. (1983). *J. Rheumatol.* **10,** 920–924.
Tappeiner, G., Helmut, H., Scholz, S., Albert, E., Linert, J., and Wolff, K. (1982). *J. Am. Acad. Dermatol.* **7,** 66–79.
Tan, E. M., Schur, P. H., Carr, R. I., and Kunkel, H. G. (1966). *J. Clin. Invest.* **45,** 1732–1740.
Tan, E. M., Cohen, A. S., Fries, J., Masi, A. T., McShane, D., Rothfield, N. F., Schaller, J., Talal, N., and Winchester, R. (1982). *Arthritis Rheum.* **25,** 53A.
Tan, S. G., and Powell, N. R. (1976). *Br. J. Dermatol.* **95,** 99–100.
Usher, B., and Senear, F. E. (1976). *Arch. Dermatol. Syphilol.* **13,** 761–781.
Van Joost, T., Cormane, R. H., and Pondman, K. W. (1972). *Br. J. Dermatol.* **87,** 466–474.
Ward, P. A. (1967). *J. Exp. Med.* **126,** 189–206.
Wertheimer, D., and Barland, P. (1976). *Arthritis Rheum.* **19,** 1249–1255.
Wintroub, B. U., Mihm, M. C., Jr., Goetze, E. J., Soter, N. A., and Austen, K. F. (1976). *N. Engl. J. Med.* **298,** 417–421.
Woo, T. Y., Barouski-Miller, P. A., Gelehrter, T. D., Anhalt, G. J., Hogan, V., Patel, H., and Diaz, L. A. (1982). *Clin. Res.* **30,** 719A.
Woo, T. Y., Hogan, V., Patel, H., Anhalt, G. J., Labib, R. S., Voorhees, J. J., and Diaz, L. A. (1983). *J. Invest. Derm.* **81,** 115–121.
Yaoita, H., Gullino, M., and Katz, S. I. (1976). *J. Invest. Dermatol.* **66,** 383–388.
Yaoita, H., Briggaman, R. A., Lawley, T. J., Provost, T. T., and Katz, S. I. (1981). *J. Invest. Dermatol.* **76,** 288–292.

CHAPTER 18

Autoimmune Hemolytic Anemia

BERNARD PIROFSKY
Division of Immunology, Allergy, and Rheumatology
Department of Medicine
The Oregon Health Sciences University
Portland, Oregon

The autoimmune hemolytic anemias have several unique features that make them pivotal syndromes in analyzing the autoimmune diseases. Historically, the concept of autoimmunization was first developed in an attempt to explain these diseases; in fact, the word "autoimmune" was coined to describe the phenomenon as originally observed with the erythrocyte. In addition, the pathological effects of the autoantibody on the target antigen cell (erythrocyte) can be precisely quantitated, in contrast to other tissues. Finally, various sensitive, quantitative methods have been developed to demonstrate and define the relevant autoantibodies.

I. HISTORICAL BACKGROUND AND PATHOPHYSIOLOGY

Hemolytic anemia, by definition, results from a shortening of the erythrocyte lifespan. Multiple etiologies have now been identified as inducing this effect. The processes can be broadly grouped in two main categories. In one form, an intracellular defect of the erythrocyte exists, leading to premature red cell death. Abnormalities in the membrane structure, enzyme makeup, and hemoglobin are the usual findings. In general, the defect and resulting hemolytic anemia are genetically inherited, leading to the term *congenital hemolytic anemia.* In contrast, another major group of hemolytic anemias exists with an abnormal extracellular environment that reduces the survival of normal erythrocytes. These have been termed *acquired hemolytic anemias,* and a variety of physical, chemical and toxic agents can induce the phenomena. Immunological mechanisms occupy a prominent etiological role in this grouping and have been termed *immune hemolytic anemia.* Figure 1 outlines this classification.

The immune hemolytic anemias can be further differentiated, depending on the status of the immune apparatus in inducing the hemolytic process. Early in the history of immunology, it was recognized that immunological reactions initiated by a host were directed against materials foreign to that host. Ehrlich's (1906) concept of "horror autotoxicus" formalized this basic observation. When a normal immune apparatus is stimulated to react against foreign erythrocytes, leading to their destruction, the resulting condition is termed *isoimmune hemolytic anemia.* For this process to occur, either the antigen (erythrocyte) or the antibody

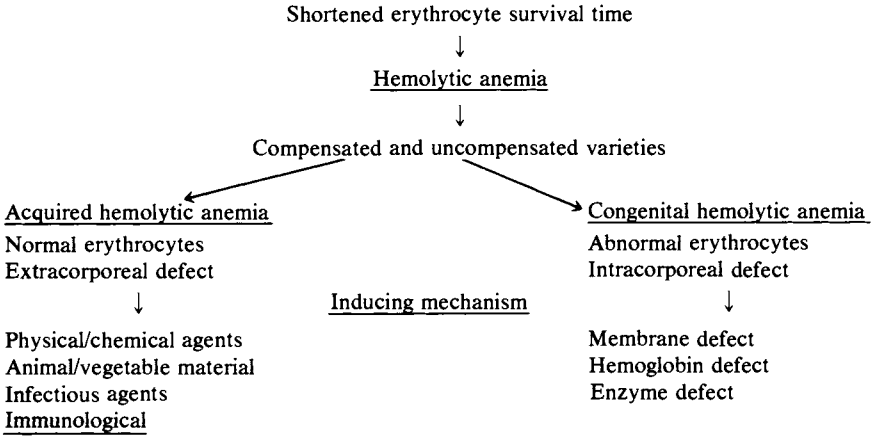


FIG. 1. Classification and Terminology of Hemolytic Anemia.

TABLE I
The Immunohemolytic Anemias: Classification Based on Etiological Mechanisms

Isoantibody immunohemolytic anemia
Passive antibody transfer type
Erythroblastosis fetalis
Minor-type transfusion reaction
Passive antigen transfer type
Major type transfusion reaction
Antiglobulin-positive immunohemolytic anemia
Nonimmunological protein fixation
Exogenous antigen
Modified autologous antigen
Autoimmune hemolytic anemia

must be foreign and introduced into the host to create the foreign state, permitting immune recognition and reaction as nonself. Clinically, this is seen in major and minor blood transfusion reactions and in hemolytic disease of the newborn (erythroblastosis fetalis). It should be emphasized that, although the clinical consequences may be catastrophic, the immune apparatus is functioning normally and performing its critical homeostatic role.

With the development of precise laboratory tests designed to demonstrate the role of antibodies in such immune hemolytic anemias, it rapidly became apparent that not all immune hemolytic anemias could be explained by classical isoimmune mechanisms. Dameshek and Schwartz (1940) called attention to this phenomenon and clearly demonstrated that hemolytic anemia could be induced by an immunological reaction initiated by a host and directed against the host's own erythrocytes. This process afflicting the erythrocyte was eventually called autoimmune, and the clinical state was named *autoimmune hemolytic anemia*. Table I outlines this classification.

Three profound consequences resulted from this major clinical and theoretical advance (1) Horror autotoxicus was no longer an absolute guideline for understanding immunological activity. This is not surprising. Ehrlich (1906), in his original development of the horror autotoxicus concept, predicted that this principle could be violated with great pathological consequences. (2) The capacity to violate horror autotoxicus and develop autoimmune reactions was incompatible with the reigning *instruction theory* of antibody formation. The inability of the instruction theory to explain or incorporate autoimmunization within its theoretical framework was a profound challenge to develop alternative antibody-

forming theories, and this eventually stimulated derivation of the *clonal selection* theory (Burnet, 1959). (3) Autoimmunization introduced the concept that abnormal immunological function could induce systemic disease. Accordingly, the functional capacity of the immune apparatus as a critical homeostatic mechanism recognizing self from nonself in a biological context assumed a major clinical and theoretical role. Every organ system suddenly became vulnerable as a target antigen against which immune reactions could be directed when the immune apparatus failed to function in a normal manner. This chapter will be specifically devoted to such abnormal immunological reactions directed against the erythrocyte, the autoimmune hemolytic anemias.

II. DEFINITION OF AUTOIMMUNE HEMOLYTIC ANEMIA

Any fruitful discussion of the autoimmune hemolytic anemias is dependent in great part on establishing a clear and precise definition of the abnormal process. Early in the development of the autoimmune concept, any hemolytic anemia that was not congenital or inherited (i.e., an acquired hemolytic anemia) was automatically termed autoimmune. This oversimplification was partially corrected by studies defining the nature of immune reactivity. However, such theoretical serological knowledge was distorted when applied to clinical situations. The problem was again simplified by using the phrase autoimmune to explain any clinical circumstance in which a host produced an erythrocyte antibody that resulted in premature death of the erythrocyte. The critical questions as to a normal or abnormal response of the immune apparatus and the nature of the antigen were not addressed with this broad terminology. In 1963, MacKay and Burnet provisionally defined autoimmune disease as follows: "Any condition in which clinical symptoms or functional changes result from immunologic reactions of immunologically competent cells or antibodies, produced by the individual, with normal components of the body."

This definition was excellent, and adherence to its postulates clarified many previous problems and clearly pointed to required additional investigations. Several critical aspects of autoimmunization were emphasized. The immunological nature of the process was unequivocally stated. The definition permitted immune mechanisms other than circulating antibody in the establishment of an autoimmune state. The application of cellular immunity in autoimmunization was implicit. The presence of tissue damage and functional impairment of the target organ system became essential for the diagnosis of an autoimmune disease. The simple presence of an

autoantibody or a sensitized T lymphocyte was not in itself necessarily assumed to be abnormal, and the phrase "disease" required the additional provision of clinical/pathological manifestations. Finally, the definition stated the dependence of autoimmunization on the host production of agents capable of reacting with normal host self-constituents. This provision suggested a specific pathological process in which biological self-recognition is abnormal, and horror autotoxicus as a controlling homeostatic state is abrogated. Extrapolation of this concept further implies that an autoimmune disease, irrespective of the target organ, has as its fundamental abnormality a malfunctioning immune apparatus. It further implies a basic unity of such diseases, in spite of the vast clinical variety which reflects the involvement of multiple and different target cellular antigens, either singularly or in various combinations.

Application of these concepts specifically to the erythrocyte allows us to define the autoimmune hemolytic anemias as follows: A clinical state characterized by an individual initiating an immunological reaction directed against his own normal unmodified erythrocytes, resulting in a shortening of the erythrocyte lifespan.

This definition suggests that the basic pathological defect lies in the immune apparatus, and that the erythrocyte represents a normal target for the aberrant immune reaction. It also implies that the erythrocyte may not be the only or even the major tissue against which the abnormal immune reactivity is directed. We should anticipate disease complexes in which multiple tissues will be involved, either at the time of diagnosis or subsequently throughout the natural history of the disease (Pirofsky, 1968). Immunological reactions directed against nonerythrocyte antigens, such as toxic materials, drugs, and bacteria and viruses and their products, etc., which fix to the erythrocyte surface and lead to erythrocyte destruction, are not truly autoimmune and do not carry the implication of a defective immune system or a primary immunologically medicated disease.

III. CLASSIFICATIONS OF AUTOIMMUNE HEMOLYTIC ANEMIA

The definition given above and its implications suggest that the autoimmune hemolytic anemias will be highly variable in their clinical presentations. This has been found to be true, and attempts have therefore been made to further define and classify clinically the disease entities. The obvious preferred method would be an etiological definition, documenting the specific immune defect and its pathogenesis. As yet, this idealized

situation has not been attained. Accordingly, two major clinical and serological subclassifications have been employed and found useful.

A. Cold- and Warm-Acting Autoimmune Hemolytic Anemia

The antibodies inducing autoimmune hemolytic anemia are of two major immunoglobulin classes, IgG and IgM varieties. The clinical presentations resulting from these two immunoglobulin classes differ greatly. In addition, the prognostic significance and type of therapy required are dissimilar. Accordingly, the clinician faced with autoimmune hemolytic anemia is obligated to define the process serologically, at least to the extent of the concerned immunoglobulin class. As a general rule, an IgM autoantibody inducing significant hemolytic anemia results in a *cryopathic* or *cold-acting autoimmune hemolytic anemia*. Hemolytic anemia caused by an IgG autoantibody generally leads to a *warm-acting autoimmune hemolytic anemia*. The one obvious exception to this general observation is *paroxysmal cold hemoglobinuria*, where an IgG autoantibody induces a cold-dependent hemolytic process.

B. Clinical Classification

The presence or absence of significant associated diseases plays a dominant role in the clinical presentation, prognosis, and required therapy of autoimmune hemolytic anemia. Accordingly, a clinical classification is a useful tool in the medical management of these conditions. When the hemolytic anemia appears as an isolated process in an otherwise healthy subject, the term *primary* or *idiopathic autoimmune hemolytic anemia* is employed. If there is a significant associated disease in a patient developing autoimmune hemolytic anemia, the process has been called *secondary* or *symptomatic autoimmune hemolytic anemia*. The relative frequency of these two varieties depends to a great extent on the length of time of follow-up examinations and the assumption of "significance" of the associated disease.

The early studies of autoimmune hemolytic anemia indicated that ~70% of cases were of the idiopathic form (Dausset and Colombani, 1959; Dacie, 1962). With our increased knowledge of autoimmunization, this frequency has dramatically changed. The formulation of a theoretical framework ascribing the disease to aberrant immune function rather than as a primary defect of the target cell (erythrocyte), and the concomitant demonstration of an immunological pathogenesis of a wide variety of clinical entities, have greatly increased the number of "significant" dis-

eases. In a large study we performed, only 18% of cases were found to be of an idiopathic variety (Pirofsky, 1969). We accordingly suggested that idiopathic autoimmune hemolytic anemia is illusory, and that all autoimmune hemolytic anemia is part of a multisystem, immunologically mediated abnormality. The apparent idiopathic or primary state is simply the "tip of an iceberg" syndrome and should be used as a herald sign alerting the clinician to an underlying multisystem disease that could become clinically apparent at any time (Pirofsky, 1968).

IV. CLINICAL ASPECTS OF AUTOIMMUNE HEMOLYTIC ANEMIA

Cold-acting (cryopathic) and warm-acting forms of autoimmune hemolytic anemia may present with a variety of different clinical patterns. Much of this variation represents the presence of different associated diseases and their influence on the general medical status. It is apparent that a warm-acting autoimmune hemolytic anemia occurring as the only disease or in association with active Hodgkin's disease or systemic lupus erythematosus will present entirely different symptomatic and physical findings. The problem is further complicated by the interrelationship between the hemolytic process and the underlying disease. In some cases, activity of the associated disease will dramatically affect the hemolytic anemia. In other patients, the hemolytic anemia and associated disease coexist, with their activities independent of each other.

It is beyond the scope of this chapter to review the myriad clinical variations that may occur with secondary autoimmune hemolytic anemia. Table II outlines several of the patterns seen in the cryopathic variety and Table III lists the clinical associations seen in 234 cases of warm-acting autoimmune hemolytic anemia. Clinical details may be found in the reviews of Dacie (1962), Pirofsky (1969, 1976), and Schubothe (1958). The following discussion summarizes many basic general clinical observations.

A. INCIDENCE

The autoimmune hemolytic anemias are relatively common diseases. At the Oregon Health Sciences University, an annual incidence of 1 case per 80,000 population was encountered over an 8-year period. The frequency of discovery of cases reflects in great part the alertness of the clinician and the laboratory staff, and the technical skill of the blood transfusion service.

TABLE II
Classification of the Cryopathic Autoimmune Hemolytic Anemias

Idiopathic or primary cold hemagglutinin disease
Secondary or symptomatic cold hemagglutinin disease
Associated with virus infection
Mycoplasma pneumoniae
Influenza
Infectious mononucleosis
Other viruses
Associated with reticuloendothelial neoplasia
Paroxysmal cold hemoglobulinuria
Luetic type
Nonluetic type

B. Age Distribution

Subjects with the cryopathic and warm-acting autoimmune hemolytic anemias differ in their age distribution. Cold-acting varieties typically involve the middle-aged and elderly; Schubothe (1958) noted a peak incidence of onset from 51 to 60 years and Dacie (1962) a peak at 70 to 80 years. In contrast, the warm-acting form occurs at any age. Dacie's (1962) patients ranged from 5 months to 78 years, and those of Pirofsky (1969) from 1 month to 87 years. The underlying associated disease appears to influence the age distribution rather than the hemolytic anemia. The occurrence of warm-acting autoimmune hemolytic anemia in the young is considered by some to be rare, but this is erroneous. O'Connor *et al.* (1956) reviewed the accumulated literature and added 18 additional cases in children under age 5. In our experience, the majority of cases occurring in the pediatric age group are associated with bacterial or viral infectious disease.

C. Race and Sex

There does not appear to be a racial predisposition for, or protection from, autoimmune hemolytic anemia. Blacks frequently develop these diseases. They also occur in Asians, and the syndromes are not rare in the tropics.

The sex distribution is influenced by the underlying pathological state rather than the autoimmune hemolytic anemia. For example, the majority of patients with SLE and autoimmune hemolytic anemia are female. In contrast, patients developing these syndromes with the lymphomas and

TABLE III

Primary and Secondary Types of Autoimmune Hemolytic Anemia[a]

Associated diseases	Number	Percent
Idiopathic (primary) type	44	18.2
Secondary (symptomatic) type	190	81.8
Reticuloendothelial neoplasia	114	48.7
The leukemia	77	32.9
Chronic lymphocytic	48	20.5
Others	29	12.4
The lymphomas	25	10.7
Plasma cell myeloma	9	3.8
Thymona	3	1.3
Other malignancies	20	8.5
Benign cyst and tumors	16	6.8
Collagen–vascular disease	35	15.0
Thyroid disease	25	10.7
Infectious disease		
Bacterial	53	22.6
Viral	21	9.0
Fungal	3	1.3
GI Tract disease	29	12.4
Miscellaneous group		
Trauma, erythrocytes	9	3.8
Drugs	7	3.0
Iron deficiency	3	1.3
Myeloproliferative disease	1	0.4
Heterozygous P-K deficiency	1	0.4

[a] There were 234 patients in the study.

leukemia show a predominance of males. The idiopathic varieties are more frequently seen in females. This suggests that many of these patients may eventually develop overt SLE. This is particularly true in subjects with an Evans' syndrome.

D. Genetic Predisposition

Although the familial occurrence of autoimmune hemolytic anemia is rare, 20 kindreds with 45 afflicted subjects have been reported (Lippman *et al.*, 1982). In spite of this uncommon overt genetic association, a genetic predisposition is apparent. If a careful family history is taken, a distinct pattern may appear. Multiple family members will frequently be found to have a variety of disease states generally considered to be immunologically mediated. Rheumatoid arthritis, nephritis, SLE, the leuke-

mias and lymphomas, ulcerative colitis, thyroid disease, pernicious anemia, multiple sclerosis, etc., are typical entities present in other family members. This distribution suggests that a hereditary defect exists, but that it does not specifically involve immune homeostasis of the erythrocyte. Rather, a more fundamental inheritable aberration of the immune apparatus may be present in which multiple target organs are eventually involved (Pirofsky, 1968, 1969; Pirofsky and Vaughn, 1968; Bardana and Pirofsky, 1970). A detailed review of this problem by Lippman *et al.* (1982) indicates the presence of a Mendelian dominant abnormality with impressive odds against linkage to HLA.

E. Presenting Symptoms

As mentioned above, a large number of symptoms is manifested by patients with autoimmune hemolytic anemia. The diversity reflects the variability of significant associated diseases present. Symptoms directly referable to the autoimmune hemolytic anemia are more limited and may be conveniently grouped into three areas. Symptoms resulting from the anemic state are most common, with weakness and dizziness frequently found. Evidence of cardiovascular complications is also common, with symptoms of congestive heart failure and coronary insufficiency particularly ominous.

The second major group of symptoms is a reflection of the hemolytic process. Jaundice is noted by about one-quarter of the patients, and dark urine or hemoglobinuria is occasionally encountered. Miscellaneous complaints of a lower frequency and less specificity are also reported. Fever is the most outstanding in this group. It is frequently difficult to date the exact onset of disease. The symptoms generally develop slowly and insidiously. Rarely, the disease will present in an explosive fashion with an acute hemolytic crisis. Table IV lists the relative frequency of presenting symptoms in 230 patients seen at our institution.

F. Presenting Signs

Physical findings in autoimmune hemolytic anemia generally reflect the hemolysis, resulting anemia and sequelae, and reticuloendothelial abnormalities. Jaundice is common and is a valuable physical sign. The clinical diagnosis of anemia is remarkably difficult, and the classical observation of pallor is unreliable. The effect of anemia, however, is easily documented in the cardiovascular system. Edema and evidence of congestive heart failure may be found. Enlargement of the liver, spleen, and lymph nodes is the most common abnormality. The hepatomegaly and spleno-

TABLE IV
Presenting Symptoms of 230 Patients with Autoimmune Hemolytic Anemia

Symptoms	Number	Percent
Anemia		
Weakness	203	88
Dizziness	115	50
Dyspnea	21	9
Heart failure	11	5
Edema	10	4
Pallor	9	4
Angina	5	2
Confusion	4	2
Syncope	1	.4
Hemolysis		
Jaundice	48	21
Dark urine	6	3
Hemoglobinuria	3	1
Miscellaneous		
Fever	85	37
Bleeding	23	10
GI Complaints	13	5
Cough	12	6
Weight loss	11	5
Anorexia	9	4
Arthritis	4	2

megaly are not massive, and the lymph node enlargement is slight to moderate. Approximately 75% of patients with autoimmune hemolytic anemia will have one or several of these enlarged reticuloendothelial structures. Table V summarizes the presenting physical signs in 230 cases of autoimmune hemolytic anemia.

V. LABORATORY TESTING IN AUTOIMMUNE HEMOLYTIC ANEMIA

Laboratory examinations are critical for the diagnosis of autoimmune hemolytic anemia, for efficient monitoring of therapeutic effects, and for evidence of either relapse or remission. A detailed analysis of all the laboratory findings is beyond the scope of this review. The reader should consult several excellent texts for many of the details (Dacie, 1962; Pirofsky, 1969; Petz and Garratty, 1980). The present discussion will be

Table V
Presenting Physical Signs in 230 Patients with Autoimmune Hemolytic Anemia

Physical sign	Number	Percent
Splenomegaly	120	52
Hepatomegaly	103	45
Lymphadenopathy	79	34
Jaundice	48	21
Thyromegaly	24	10
Edema	13	6
Heart failure	12	5
Pallor	10	4
Pleural effusion	4	2.2
Ascitis	4	2.2
Pneumonia	4	2.2
Abdominal mass	4	2.2
Skin lesions	4	2.2
Arthritis	4	2.2
Mediastinal mass	2	1.1
Splenic infarct	2	1.1

limited to a brief review of several critical laboratory parameters; the peripheral blood, the bone marrow, and serological testing.

A. Peripheral Blood

Examination of the peripheral blood is the most important initial test in autoimmune hemolytic anemia. In addition, repeated examinations are critical to evaluate the severity of disease and the response to therapy or progression of the abnormality. A complete blood count, including platelets and reticulocytes, is required. In 218 patients studied in our laboratory, the initial hematocrit ranged from 9.0 to 44.5%, with a median of 24%. Although the hematocrit may be only moderately depressed, meticulous follow-up is required. When these patients were sequentially followed, it became apparent that severe anemia is the rule. Lowest hematocrits ranged from 7.5 to 41.5%, with a median of 19%. A hematocrit $\leq$15% was observed in 49 patients.

The white blood cell count is highly variable and may reflect the presence of an associated significant disease, that is, the leukemias. In 38 patients with an idiopathic autoimmune hemolytic anemia, more than half the patients had normal WBC counts, although there was a range of 1400 to 27000/mm^3. Six patients had leukocyte counts $\leq$2000/mm^3, and these

occurred in a pattern of peripheral pancytopenia with hypoplastic or normal bone marrows.

Analysis of 77 patients with autoimmune hemolytic anemia in the absence of lymphoid malignancy revealed normal thrombocyte values in 60%, with a range of <1,000 to >1 × 10^6/mm^3. Of 31 patients with the idiopathic variety, 45% were found to be thrombocytopenic, with 7 subjects having values <20,000 mm^3. Thrombocytopenia occurring in association with autoimmune hemolytic anemia has been well recognized under the term Evans' syndrome (Evans and Duane, 1949). It should be emphasized that the anemia and thrombocytopenia do not necessarily appear at the same time. Follow-up is critical in such patients, and the majority will eventually appear with a full-blown SLE.

It is generally assumed that chronic and persistent reticulocytosis is inevitable in autoimmune hemolytic anemia. This was examined in 195 patients on the initial examination prior to diagnosis and therapy. Ninety-six patients, 49%, had normal reticulocyte values <2%. This was more common in secondary varieties (54.4%) than in the idiopathic form (25.7%). It might be assumed that either myelophthisic or chemotherapeutic states restricted the capacity of the bone marrow to respond to anemic stress. This could not be confirmed. The majority of patients responded to therapy with an adequate reticulocytosis. Five patients with the idiopathic variety had reticulocyte values of 0–0.3%. These cases were initially misdiagnosed as either aplastic anemia or erythrocyte aplasia. The true diagnosis was not suspected until positive antiglobulin tests were obtained during routine cross-matching for transfusions. Reticulocytosis generally occurred in these patients with the start of therapy. It should be emphasized that a lack of reticulocytosis should not sway the clinician from a diagnosis of autoimmune hemolytic anemia.

B. Bone Marrow Examination

The peripheral blood values found in autoimmune hemolytic anemia represent a summation of the immunologically mediated destructive process and the productive capacity of the bone marrow. Accordingly, bone marrow examination is critical for a complete evaluation of the disease process. It is perfectly possible for extensive peripheral blood cell destruction to occur and be sufficiently compensated by increased cell production to avoid anemia. The normal erythrocyte lifespan of 120 days could be reduced to 15 days, and a fourfold increase in erythrocyte production would correct this abnormality sufficiently to prevent anemia from occurring.

It is generally assumed that bone marrow hyperplasia, and more specifically erythrocyte hyperplasia, is universally found in hemolytic anemia. We have examined this assumption by reviewing bone marrow preparations obtained from 162 patients at the approximate time of onset of the hemolytic stage. A normal bone marrow was found in only 15% of cases, with 53% having erythrocytic hyperplasia and 43% generalized increased marrow cellularity. In the idiopathic variety, only 1 of 35 patients had a normal marrow; 71% had erythrocyte hyperplasia. Of particular interest was the observation that 8 of 35 had depressed numbers of erythrocyte precursors, with 3 patients presenting an erythrocytic aplasia pattern.

This is surprising to some. It suggests that the erythrocyte autoantibody is directed against antigens present on both mature and precursor erythrocytes. The resulting bone marrow, reticulocyte, and peripheral erythrocyte findings accordingly must represent the outcome of antigenic competition for antibody by mature and precursor erythrocytes. In general, the huge peripheral erythrocyte mass and their antigens must bind the overwhelming majority of available autoantibody. The comparatively limited destruction of precursor cells is obscured by intense proliferation of the remaining unafflicted precursor cells, resulting in the pattern of erythrocytic hyperplasia. In exceptional circumstances, however, large amounts of circulating antibody or antibody with a high avidity may be present, resulting in the destruction of large numbers of precursor cells and even mimicking erythrocytic aplasia. A detailed consideration of this possibility was presented previously (Pirofsky, 1969).

Morphological abnormalities of erythrocytic precursors resembling megaloblastic changes are commonly observed. Approximately 15% of patients we have studied had bone marrow examinations classified as megaloblastic (Pirofsky, 1969). Such changes may represent a relative lack of folic acid of Vitamin B_{12} produced by intense and chronic erythropoietic activity associated with autoimmune hemolytic anemia. However, measurements of serum folate and Vitamin B_{12} are generally normal, and supplementary therapy with folic acid or Vitamin B_{12} usually does not induce a therapeutic response. It should also be emphasized that positive antiglobulin tests are not infrequently found in classic pernicious anemia (Pirofsky and Vaughn, 1968) and have also been described in folic acid deficiency states (Pirofsky, 1969).

C. Serological Tests

The diagnosis of autoimmune hemolytic anemia cannot be made without the employment of appropriate serological tests to document the presence of erythrocyte autoantibodies. Current procedures are primarily de-

signed to demonstrate the fixation of antibodies to the erythrocyte surface. Such an interaction can be visualized as a two-step process. In the first phase, the antibody interacts with its erythrocyte surface antigen. In the second phase, erythrocytes coated with antibody interact with each other, eventually resulting in visible agglutination and/or hemolysis. Serologically, two general forms of erythrocyte antibodies exist in the autoimmune hemolytic anemias. A "complete" antibody is defined as one undergoing phase 1 and 2 of the hemagglutination reaction. An "incomplete" erythrocyte antibody initiates phase 1 but does not progress to phase 2 or visible agglutination. Additional procedures must be performed to demonstrate its fixation to the erythrocyte.

In the warm-acting autoimmune hemolytic anemias, the vast majority of pathological antibodies are of an "incomplete" variety. They are maximally reactive at 37°C and do not induce hemagglutination of erythrocytes suspended in saline. They are usually of an IgG immunoglobulin class. Because they do not spontaneously induce hemagglutination, specialized techniques must be employed to demonstrate the presence of such antibody, both free in the sera and fixed to the surface of the erythrocyte. Three major techniques are commonly employed. (a) Erythrocytes coated with pathological antibody will undergo agglutination when suspended in media containing large numbers of colloid anisometric molecules. Albumin, dextran, polyvinylpyrrolidine, carboxymethyl cellulose, gum acacia, etc., have all been successfully employed. (b) Treatment of erythrocytes with enzymes. Trypsin, papain, ficin, and bromelin are some of the most commonly used materials. (c) The antiglobulin test.

The antiglobulin test developed by Coombs *et al.* (1945) is the procedure most frequently used. Its rationale was based on the assumption that incomplete erythrocyte autoantibody fixed to the red cell surface was a gamma globulin molecule. Accordingly, an antibody directed against human gamma globulin (Coombs' sera) would react with such erythrocyte antibodies, thereby inducing hemagglutination. The process was visualized as supplying a functional bivalent antibody molecule capable of reacting with the gamma globulin antigen (erythrocyte antibody) on two different erythrocytes. This bridging phenomenon would expand in a fashion similar to the lattice framework seen in precipitation, leading to agglutination. Two basic test procedures were formulated. A *direct antiglobulin test* was used to document the presence of antibody on the erythrocyte surface. An *indirect antiglobulin test* was employed to demonstrate free antibody circulating in the serum. This was accomplished by incubating the test serum with appropriate donor erythrocytes, washing these cells free of sera, and then testing the washed erythrocytes in a direct antiglobulin test.

The choosing of appropriate donor erythrocytes requires some explanation. In order for an incomplete antibody to fix to an erythrocyte, the antigen against which the antibody is directed must be present on the erythrocyte surface. Accordingly, the use of an indirect antiglobulin test in an isoimmune hemolytic anemia (i.e., hemolytic disease of the newborn) would require the use of donor erythrocytes containing a blood group substance against which the isoantibody was directed. For an anti-D (Rh_o) form, D (Rh_o) (+) erythrocytes are needed. If the antibody had anti-K specificity, the donor erythrocytes must be K(+). Application of this technique to the warm-acting autoimmune hemolytic anemias gave different results. Positive indirect antiglobulin tests were obtained with all human red cells tested. Such data suggested to some that a lack of antigenic specificity existed and a true antibody was not involved.

Wiener *et al.* (1953) presented a brilliant set of speculations to explain the apparent lack of specificity of erythrocyte autoantibodies. It was postulated that antigenic material present on all human erythrocytes created the appearance of nonspecificity. After a series of studies, they concluded that the warm-acting autoantibodies were directed against the nucleus of the Rh–hr substance. Support for this concept has accumulated from two main sources. Occasional warm-acting erythrocyte autoantibodies directed against specific erythrocyte antigens have been found. The vast majority of these have been directed against rhesus group antigens. Several families have been shown to have erythrocytes that do not contain rhesus antigens. Use of such Rh (null) erythrocytes in the various test systems revealed that the majority of warm-acting autoantibodies have specificity within the rhesus system (Weiner and Vos, 1963).

Throughout the years the antiglobulin test has proven to be a surprisingly sensitive procedure and forms the cornerstone in the diagnosis of warm-acting autoimmune hemolytic anemia. Although false-negative and false-positive reactions can occur, it is a reliable and simple test with wide applicability. In a sense, its reliability has been so good as to introduce difficulties. We are currently faced with a circular form of reasoning in the diagnosis of warm-acting autoimmune hemolytic anemia. A positive antiglobulin test is so often found in these diseases that we now assume that all cases must have a positive antiglobulin reaction. In spite of this, all clinicians dealing with these diseases have encountered patients with classical presentations and negative antiglobulin reactions. In order to truly investigate such patients, elution procedures, the use of colloidal anisometric molecule solutions and enzyme treatment, and red cell survival studies with and without cross-transfusion analysis are required.

Serological analysis in the cold-acting autoimmune hemolytic anemias differs greatly from the studies outlined above. In these states the vast

majority of pathological antibodies are of a "complete" variety. They are maximally reactive at 4°C and induce agglutination of red cells suspended in saline and are usually of an IgM immunoglobulin class. In contrast to the warm-acting autoantibodies, they are efficient activators of the compliment system and lead to hemolysis. In general they are easily identified, and specialized ancillary techniques are not required to supplement simple agglutination procedures.

Pathological cold-acting autoantibodies are very similar to the normal cold-acting antibodies (cold agglutinins) present in all humans. They differ in two areas that are significant for diagnosis. (1) Agglutination titers at 4°C are high, generally >1:2000, in contrast to normal cold agglutinins, which have titers of 1:256 or less. (2) The thermal level of reactivity with erythrocytes is high. Cryopathic antibody capable of inducing hemolytic anemia generally will react with erythrocytes at temperatures >22°C. The normal cold agglutinins generally are inactive at temperatures >15°C.

The thermal level of reactivity is of major significance in the development and magnitude of cryopathic autoimmune hemolytic anemia. A cold-acting autoantibody with a titer $>10 \times 10^6$ at 0°C and a titer of 0 at 20°C is clinically insignificant. In contrast, a relatively low-titer antibody at 4°C with a titer of 1:2 at 30°C may lead to severe hemolytic anemia. The critical question is not the antibody titer, but rather whether the antibody is capable of reacting with erythrocytes at body temperatures. It should be emphasized that temperatures in the capillary circulation of the extremities can drop as low as 20–22°C.

Demonstration of a clinically significant cold-acting autoantibody can be accomplished with the following simple procedure. A whole blood sample should be collected in EDTA and warmed to 37°C prior to removing serum for testing. The rationale for these provisions is outlined below. A cold agglutinin titer study should be set up in duplicate. One rack of dilutions is to be incubated at 4°C and one at room temperature (22–25°C). A potentially significant autoantibody may be present if the titer at 4°C is >1:2000 or if agglutination occurs at any titer at room temperature. If this occurs, the titer should then be determined, at stepwise higher temperatures, until either 37°C is attained or agglutination disappears. This thermal level is a helpful guide in predicting the severity of the clinical problem.

Cold agglutinins generally induce overt agglutination of erythrocytes suspended in saline. Accordingly, antiglobulin testing is not required and cannot be done in the usual fashion. If it is desirable to search for an IgG warm-acting autoantibody in the presence of a cold agglutinin, it is necessary to eliminate the IgM-induced agglutination. This can be accomplished by disulfide reduction of the IgM autoantibody with dithiotriotol

at a concentration that will not inactivate the IgG autoantibody, followed by standard antiglobulin testing (Pirofsky and Rosner, 1974).

The usual commercial antiglobulin sera have antibodies directed against human IgG, with very little reactivity against IgM immunoglobulins. In view of this it was surprising and frequently confusing to observe that direct antiglobulin testing is generally positive in cold-acting autoimmune hemolytic anemia in the absence of warm-acting IgG autoantibodies. This apparent discrepancy is now easily explained. The usual commercial antiglobulin sera has antibodies against C3 and C4 as well as against IgG. When a cold hemagglutinin reacts with an erythrocyte antigen, complement is bound to the erythrocyte surface. In the usual laboratory manipulations, washing and incubation occurs at temperatures sufficiently warm to remove the antibody from the erythrocyte surface; complement, however, is irreversibly bound. Subsequent testing with antiglobulin sera will then lead to positive tests induced by anti-C3, -C4 reactions to the bound complement.

It is crucial that this reaction be understood in order to avoid misdiagnosis. As mentioned, all human sera have cold-acting autoantibodies that are not clinically significant. However, if a heparinized blood sample is obtained and stored in the refrigerator, these normal cold agglutinins will fix to the erythrocytes and bind complement. In subsequent testing the cold agglutinin is washed free, but the complement will remain on the erythrocyte and lead to a positive antiglobulin test due to anti-C3, -C4 antibodies in the antiglobulin sera. Collection of the blood sample in EDTA will prevent this complement activation. In addition, warming the blood sample prior to removal of sera for testing will prevent adsorbtion of the free cryopathic autoantibody to the erythrocytes. Clarification of this problem in clinical testing can be accomplished by using monospecific anti-IgG, anti-C3, and anti-C4 antiglobulin reagents.

The cold-acting autoantibodies were the first autoantibodies discovered. In 1903 Landsteiner accurately described and purified the normal human cold hemagglutinins. These antibodies had a remarkable broad spectrum of activity. They were found in the sera of many animal species (human, guinea pig, chicken, horse, dog, rabbit, cattle, cat, sheep, monkey, donkey, and goat), and they cross-reacted extensively. Because of this apparent nonspecificity, the term "panhemagglutinins" was applied to the entire group. Extensive studies by Wiener *et al.* (1956) clarified these observations. It was shown that the normal cold hemagglutinins and pathological cryopathic autoantibodies reacted with a blood group substance present on the erythrocytes of most humans and other animal species. The antigen was named I, with I chosen to emphasize the high degree of individuality of blood samples failing to react with anti-I. These

observations have been amply confirmed. In addition to anti-I, however, other specificities of cold-acting autoantibodies have been described. These include anti-H, anti-P, anti-O, anti-B, anti-M, anti-"not-I" and anti-i. This latter autoantibody is particularly common in infectious mononucleosis. Cord blood erythrocytes, in contrast to adult erythrocytes, have large amounts of i antigen and relatively small amounts of I. This finding is employed in specificity studies of cold-acting autoantibodies, with differential titers obtained using adult and cord blood erythrocytes.

VI. THERAPY OF AUTOIMMUNE HEMOLYTIC ANEMIA

The current review has emphasized that the autoimmune hemolytic anemias are not a homogeneous clinical entity. Variations exist in the magnitude and rapidity of hemolysis, in the association of additional pathological states, in the comparative influences of these two factors, and in the serological nature of the autoantibody. The effect of this heterogeneity is particularly apparent in divising efficient therapeutic procedures. This discussion will be limited to a general approach applicable to the usual case encountered.

A. Warm-Acting Autoimmune Hemolytic Anemia

1. Corticosteroids

Corticosteroids are the initial treatment of choice for the warm-acting autoimmune hemolytic anemias. Oral prednisone (60 mg as a single morning dose) should be given daily. Comparable doses of other corticosteroids may be utilized instead of prednisone. There is no evidence that higher doses or parenteral routes offer any advantage. A therapeutic response is frequently anticipated by stabilization of a dropping hematocrit or an increase in the reticulocyte response. Fifty percent of responding patients will do so within 7 days of initiating therapy. In our experience, a lack of response in 3 weeks generally indicates a therapeutic failure and the need for another therapeutic modality.

High-dose corticosteroid therapy may produce significant detrimental side effects. A major goal of this therapy is to maintain a hematological remission at the lowest possible dose of corticosteroids. To accomplish this, a slow but progressive reduction of corticosteroids should be carried out after a hematocrit of 30% is attained. A prednisone requirement >15 mg daily to maintain a hematocrit of 30% or higher is generally unaccepta-

ble for long periods of time and is an indication for alternative therapy. A therapeutic failure with steroids, or an unacceptable dosage, requires either splenectomy or immunosuppressive therapy. Clinicians differ as to the preferential procedure after steroids. In either case, if a therapeutic failure occurs with the chosen approach, the other therapy should be utilized. My personal preference is to use immunosuppressive therapy if prednisone fails, with splenectomy to be used if immunosuppression is unsuccessful or cannot be employed.

2. Immunosuppressive Therapy

Daily azathioprine, 2.0–2.5 mg/kg body weight in divided oral doses, is the therapy of choice. Some clinicians employ cyclophosphamide. The multiple system toxicity of this drug and the other alkylating agents and their oncogenic potential suggest that these agents should only be used if failure occurs with azathioprine or other analogs such as methotrexate. Azathioprine is a long-acting drug. Accordingly, therapeutic responses may not be seen for 10–14 days. Similarly, changes in dosage should not be initiated before a 2-week period, in order to evaluate the effect of the previous dose level. If no response is seen after 4 weeks of the initial therapy, dose levels may be increased by 25 mgs every 2 weeks, until either a response occurs or evidence of marrow depression is obtained. Low-dose corticosteroids, 10–20 mg daily, should be given concomitantly with azathioprine. When remission occurs it is frequently possible to reduce or eliminate prednisone without initiating relapse. At that point a slow reduction of azathioprine (i.e., 25 mg/month) should be attempted. Approximately 50% of patients unresponsive to corticosteroid therapy may respond to the above procedure.

3. Splenectomy

This is a well-established therapeutic approach. There is a surprisingly low morbidity and mortality when splenectomy is performed under ideal conditions. These include avoidance of the "desperation splenectomy," the availability of a team of surgeons experienced in splenectomy, and a close collaboration between the surgeon and immunohematologist. The long-term dangers of potentially fatal septicemia in adult asplenic subjects must be considered and appropriate immunizations carried out. We anticipate a response rate of ~50% in subjects failing corticosteroid therapy.

4. Other Therapies

A wide variety of other therapies has been employed. These are of dubious value; they have not been confirmed as therapeutically sound and

should be considered as investigational approaches. They should only be employed if the other therapies have been unsuccessful. The literature should be consulted for details of their use. Included in this group are heparin administration, plasmapheresis, lymphopheresis, vinblastine-laden platelets, antithymocyte antiserum, and thymectomy.

B. Cold-Acting Autoimmune Hemolytic Anemia

Frequently the degree of anemia is only mild to moderate, and vigorous therapy may not be required. If a secondary variety is present, the hemolytic process may be only transient. Hemolytic anemia following mycoplasmal infections and infectious mononucleosis falls in this category. When the anemia persists, therapy should be first directed to control environmental temperatures. Cold should be avoided, but frequently a simple change in temperature may be sufficient to precipitate a hemolytic crisis. The patient should be instructed to avoid cold temperatures and shifts of temperature by drafts. Emphasis should be placed on keeping the extremities protected from temperature changes with stockings and boots, gloves, earmuffs, etc. Relocation of the patient's permanent home to a warmer and more constant climate may be required.

If anemia is sufficiently severe, active therapy may be indicated but is frequently difficult. As a general rule, splenectomy is of no benefit and corticosteroid therapy is unsuccessful. The use of penicillamine has been reported, but its efficacy could not be confirmed. The therapy of choice is cytotoxic immunosuppression, similar to that employed in macroglobulinemia. The bulk of experience has been with chlorambucil. This drug must be given daily, and frequently months elapse before a therapeutic effect is apparent. Dosages of 2–6 mg daily are employed, and the patient is carefully followed to avoid bone marrow depression.

C. Transfusion Therapy for Autoimmune Hemolytic Anemia

The use of transfusions in the therapy of autoimmune hemolytic anemia involves the physician in a therapeutic dilemma that must be resolved by clinical judgment. Profound anemia may result in death. However, blood for transfusion in these conditions will generally be incompatible. Considerable skill is needed to decide when the risk of death overrides the risk of transfusing incompatible blood. The following should be considered in the decision:

1. Accept the fact that compatible blood will not be available. A frantic and prolonged search for compatible blood is self-defeating. In addi-

tion, the use of "least-incompatible" blood is not rational, does not resolve the dilemma, and offers no solution to the problem.

2. Avoid treating *your* apprehensions and evaluate the patient's requirements.

3. Do not treat the hematocrit. The danger level must be individualized for each patient.

4. Transfusions should not be utilized unless a life-threatening situation develops. This will usually involve the development of angina, pulmonary edema, congestive heart failure, and/or decreasing sensorium.

5. The risk of transfusion should not be minimized. However, severe reactions in a crisis or renal shutdown pattern are rare. In general, the transfused incompatible blood is destroyed at the same rate as the patient's own erythrocytes.

6. Avoid transfusions that may supply complement components, thereby intensifying hemolysis. Packed cells, washed cells, or older blood is indicated.

7. The benefits of transfusion are generally temporary. The hematocrit may rise but will generally drop to pretransfusion levels in 2–4 days.

REFERENCES

Bardana, E. J. B., and Pirofsky, B. (1970). *Int. Arch. Allergy Appl. Immunol.* **37,** 325–336.

Burnet, F. M. (1959). "The Clonal Selection Theory of Acquired Immunity." Vanderbilt Univ. Press, Nashville, Tennessee.

Coombs, R. R. A., Mourant, A. E., and Race, R. R. (1945). *Br. J. Exp. Pathol.* **26,** 255.

Dacie, J. V. (1962). "The Haemolytic Anaemias," 2nd ed., Part II. Grune & Stratton, New York.

Dameshek, W., and Schwartz, S. O. (1940). *Medicine (Baltimore)* **19,** 231.

Dausset, J., and Colombani, J. (1959). *Blood* **14,** 1280.

Ehrlich, P. (1906). "Collected Studies on Immunity." Wiley, New York. (Transl., C. Bolduan.)

Evans, R. S., and Duane, R. T. (1949). *Blood* **4,** 1196.

Landsteiner, K. (1903). *Munchen. Med. Wschr.* **50,** 1812.

Lippman, S. M., Arnett, F. C., Conley, C. L., Ness, P. M., Meyers, D. A., and Bias, W. B. (1982). *Am. J. Med.* **73,** 827.

MacKay, I. R., and Burnet, F. M. (1963). "Autoimmune Diseases." Thomas, Springfield, Illinois.

O'Connor, W. J., Vakiener, J. M., and Watson, R. J. (1956). *Pediatrics* **17,** 732.

Petz, L. D., and Garratty, G. (1980). "Acquired Immune Hemolytic Anemias." Churchill, New York.

Pirofsky, B. (1968). *Vox Sang.* **14,** 334.

Pirofsky, B. (1969). "Autoimmunization and the Autoimmune Hemolytic Anemias." Williams & Wilkins, Baltimore, Maryland.

Pirofsky, B. (1976). *Semin. Hematol.* **13,** 251.

Pirofsky, B., and Rosner, E. R. (1974). *Vox Sang.* **27,** 480.

Pirofsky, B., and Vaughn, M. (1968). *Am. J. Clin. Pathol.* **50,** 459.
Schubothe, H. (1958). "Serologie und Klinische Bedeutung der Autohamantikorper." Karger, Basel.
Weiner, W., and Vos, G. H. (1963). *Blood* **22,** 606.
Wiener, A. S., Gordon, E. B., and Gallop, C. (1953). *J. Immunol.* **71,** 58.
Wiener, A. S., Unger, L. J., Cohen, L., and Feldman, J. (1956). *Ann. Intern. Med.* **44,** 221.

CHAPTER 19

Idiopathic Thrombocytopenic Purpura

RICHARD H. ASTER
The Blood Center of Southeastern Wisconsin, Inc.
and Medical College of Wisconsin
Milwaukee, Wisconsin

I. INTRODUCTION

Idiopathic thrombocytopenic purpura (ITP) is a term used by convention to describe patients who have thrombocytopenia resulting from increased platelet destruction not attributable to known causes. Thus, ITP has until recently been largely a diagnosis of exclusion. Evidence that ITP is immunologically mediated has been accumulating for many years, and recent laboratory advances offer the hope that specific diagnostic tools will be available in the near future. The lack of such methods has almost certainly caused a diverse group of disorders to be labeled "ITP" in the past.

From the clinical findings alone, it has long been apparent that "ITP" consists of at least two distinct entities, usually designated "acute ITP" and "chronic ITP." These two disorders differ greatly in their clinical

TABLE I
Characteristics of Acute and Chronic Idiopathic Thrombocytopenic Purpura (ITP)

Acute ITP	Chronic ITP
Age of onset: 2–5 years	Age of onset: adolescence to middle age
Equally common in males and females	Females affected three times as often as males
Preceding infection in two-thirds of cases	Preceding infection rare
Acute onset of symptoms, often very severe	Insidious onset of symptoms, often mild
Spontaneous recovery in 80–90%	Spontaneous recovery rare

features (Table I) and almost certainly in pathogensis, and they will therefore be treated separately.

II. ACUTE IDIOPATHIC THROMBOCYTOPENIC PURPURA

A. General Characteristics

Acute ITP is a disease of childhood characterized by sudden onset of thrombocytopenia, usually associated with purpura and other hemorrhagic manifestations. The peak incidence is at 1 to 5 years of age (Fig. 1). It is generally accepted that an identical disease affects adults on occa-

FIG. 1. Age distribution of 305 children diagnosed as having idiopathic thrombocytopenic purpura. (From Lusher and Iyer, 1977.)

sion. In about two-thirds of patients, symptoms appear 1–3 weeks after an infection, usually viral. Eventual complete recovery is the rule in 80–90% of cases, regardless of treatment. The disorder is quite common, and many series consisting of hundreds of cases have been reported (McClure, 1975; Lusher and Iyer, 1977; McWilliams and Mauer, 1979; Karpatkin and Karpatkin, 1981; Burstein and Burns, 1982). A similar disorder common among blacks of central Africa has been designated *Onyalai*. Acute ITP, which often occurs during convalescence from an infection, should be distinguished from thrombocytopenia seen during active infection in which factors such as disseminated intravascular coagulation or suppression of platelet production may be operative (Yeager and Zinkham, 1980).

B. Pathogenesis

1. Platelet Kinetics; Sites of Platelet Destruction

The sudden onset of severe thrombocytopenia, the almost universal finding of increased numbers of megakaryocytes in the bone marrow, the presence of megathrombocytes (giant platelets) in the peripheral blood, and the rapid destruction of isologous platelets when transfused provide convincing evidence that thrombocytopenia of acute ITP is caused by accelerated destruction of platelets in the peripheral blood. The sites of platelet destruction have not been clearly established by experiment, but by analogy with chronic ITP it is likely that platelets are cleared from the circulation by the phagocytic cells of the spleen and the liver.

2. Role of Viruses

The relationship between viral infection and the subsequent, acute thrombocytopenia characteristic of this disorder is presently unexplained. Viruses are known to be capable of proliferating in animal megakaryocytes *in vivo* (Osborn and Shahidi, 1973; Brown and Axelrad, 1976), and changes in megakaryocytes suggestive of viral infestation have been noted during viral infections in humans (Oski and Naiman, 1966). Moreover, relatively tight binding of certain viruses to human platelets has been demonstrated *in vitro* (Terada *et al.*, 1966; Turpie *et al.*, 1973), and normal platelets pretreated with influenza A virus can be lysed in the presence of complement by antibodies specific for viral hemagglutinin (Kazatchkine *et al.*, 1984). It seems unlikely, however, that these mechanisms operate to reduce platelet levels in acute ITP, because the onset of symptoms usually occurs after viremia has subsided. Nor has evidence

been obtained to support two alternative possibilities: that viruses somehow alter the structure of the platelet membrane to induce neoantigens capable of provoking an immune response, or that antibodies formed against viruses occasionally cross-react with some constituent of the platelet membrane. Some of the suggestions advanced to explain acute ITP are depicted in Fig. 2.

3. Immune Complexes and Platelet Destruction

Lack of evidence for a direct action of viruses on platelets in acute ITP has led to the suggestion that platelet destruction is somehow related to the immune response mounted against the infecting organism. Immune complexes have been shown in numerous studies to be capable of binding to platelet membranes (Pfueller and Luscher, 1972; Israels *et al.*, 1973; Moore *et al.*, 1978), probably to an Fc receptor. Moreover, two disorders are recognized in which immune complexes formed in response to exogenous antigens are thought to provoke severe, even life-threatening, but transient thrombocytopenia. These are posttransfusion purpura and drug-induced immunological thrombocytopenia, in which the presumptive exogenous antigens postulated to provoke the syndrome are, respectively, a platelet alloantigen (Shulman *et al.*, 1961; Lau *et al.*, 1980) and a drug such as quinidine or quinine that may be complexed to a cell membrane or plasma protein (Shulman, 1958; Christie and Aster, 1982). In one study, it was found that antibodies from children who developed acute ITP follow-

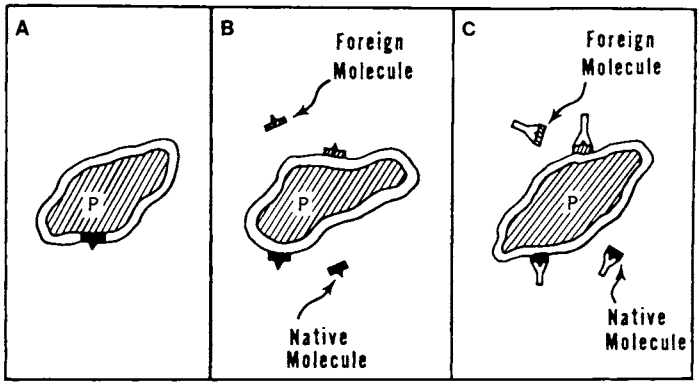


FIG. 2. Antigens of possible importance in the pathogenesis of ITP. (A) Natural constituent of the platelet membrane (platelet membrane autoantigen). (B) Autoantigen or foreign antigen, not of platelet origin, tightly bound (adsorbed) to the platelet surface. (C) Absorbed immune complexes consisting of exogenous (foreign) or autologous material complexed with antibody and presumably bound to the platelet Fc receptor. (From McMillan, 1977.)

ing rubella infection were capable of aggregating human platelets in the presence of rubella antigen, whereas serum obtained postrecovery from other children lacked this property (Myllyla *et al.*, 1969). However, no direct evidence is yet available to link immune complexes to platelet destruction.

4. Evidence for Platelet Autoantibodies

Owing perhaps to shortcomings of available serological techniques (discussed under chronic ITP, Section III), platelet-specific antibodies have been difficult to demonstrate in the serum of patients with acute ITP. In recent years, it has become possible to measure immunoglobulins on the surface of autologous platelets, using methods to be discussed. However, assays of this type are especially difficult to perform in acute ITP, because platelet levels are often extremely low and sometimes only small quantities of blood can be obtained from children. In every report to date, elevated levels of platelet-associated IgG (PA-IgG) were found as is true also of chronic ITP. These studies have demonstrated two interesting and possibly significant differences between the acute and chronic forms of ITP: Children with acute ITP seem to have higher levels of PA-IgG (Lightsey *et al.*, 1979; Sugiura *et al.*, 1980; Ware *et al.*, 1982) and more often have elevated levels of platelet-associated IgM (PA-IgM) (van Leeuwen *et al.*, 1981a) than adults or children with chronic ITP. As discussed below, recent studies have called into question the specificity of measurements of PA-IgG and PA-IgM. Until methodological questions are resolved, such data must be interpreted with caution.

5. Cellular Immunity; Abnormalities of Immunoregulation

Little has been done to test the possibility that acute ITP is related to abnormal immunoregulation, in part because of the young age group in which the disorder occurs and its self-limited nature. In one study, a variety of immunological abnormalities was found in patients and their families, but the analysis included children with both acute and chronic ITP (McIntosh *et al.*, 1981). In another, children with acute ITP were found to have normal lymphocyte responsiveness to phytohemagglutinin, normal basal levels of tritiated thymidine incorporation into lymphocytes, and normal numbers of T and B cells (Tomar and Stuart, 1981). The antigen HLA-Aw32 was significantly increased in one group of patients (Evers *et al.*, 1978), but the significance of this finding has been challenged (Westphal and Mueller-Ruchholtz, 1979). As will be noted subsequently, more definite evidence of defective immunoregulation has been obtained in patients with chronic ITP.

C. Clinical Presentation

Children between the ages of 1 and 10 years are most often affected (Fig. 1), but, as already noted, a similar disorder can occur in adults. Males and females are affected with equal frequency. Typically, the first signs are petechial hemorrhages on the trunk, extremities, and buccal mucosa. In severe cases, frank gastrointestinal and urinary tract bleeding occurs. The most severe hemorrhagic symptoms usually persist only for a few days, after which bleeding manifestations are often limited to sporadic petechiae. The association of the disorder with a preceding viral infection, which may be a specific (rubella, varicella, and rubeola) or more commonly a nonspecific infection of the upper respiratory tract, has already been mentioned. Acute ITP may also occur after immunization with live vaccines for mumps, chickenpox, measles, and smallpox (McClure, 1975; Carpentieri and Haggard, 1975), and in adults appears to have been induced in a few instances by injection of BCG vaccine for the treatment of cancer (Norton *et al.*, 1977). Symptoms usually occur within 1–3 weeks of the acute viral manifestations, but in a few instances the latent period has been as short as 2 days or as long as 6 weeks. Examination discloses multiple petechial hemorrhages, which may involve any area of the skin or mucous membranes. In severe cases, hemorrhagic bullae are found in the buccal mucosa. The liver and spleen are typically of normal size. When present, lymphadenopathy is usually attributable to the preceding infection.

D. Laboratory Features

Thrombocytopenia is usually very severe, more so than in most cases of chronic ITP of adulthood. Marrow aspiration almost without exception demonstrates hypercellularity and increased numbers of megakaryocytes, many of which are immature, reflecting an increased rate of megakaryocytopoiesis. Serum IgG levels were slightly reduced in one series of patients (Khalifa *et al.*, 1976). Levels of C3 were normal in one group (Khalifa *et al.*, 1976) but reduced in another (McIntosh *et al.*, 1981). Measurements of platelet-associated immunoglobulins have been discussed.

E. Therapy and Prognosis

1. General

Although the initial hemorrhagic symptoms in acute ITP are dramatic and may appear life threatening, ~85% of children improve spontane-

ously, usually within 4–6 weeks after the onset of symptoms. Recovery is generally permanent, although cases have been reported in which recurrences were triggered by subsequent infections or vaccinations (Lusher and Iyer, 1977). About 1% of the cases reported in the literature have ended fatally, primarily because of intracerebral hemorrhage. However, reporting is biased toward the most severe cases and the overall frequency of cranial hemorrhage is probably <0.1%. During the thrombocytopenic period, aspirin and other drugs that inhibit platelet function should be avoided (Woerner *et al.*, 1981). Patients with severe hemorrhagic symptoms should be hospitalized for observation during the acute stage.

2. Adrenal Corticosteroids

Although the usefulness of corticosteroids in the treatment of chronic ITP (below) has been amply demonstrated, whether these agents are beneficial in acute ITP is still controversial (McClure, 1975; Lusher and Iyer, 1977). Some groups have found that the duration of thrombocytopenia is unaffected (Lammi and Lovric, 1973; McClure, 1975) or even prolonged by corticosteroid therapy (Lusher and Iyer, 1977; Lusher *et al.*, 1984). In other studies, platelet levels seem to have been restored to normal more rapidly in children given corticosteroid therapy (Simons *et al.*, 1975; McWilliams and Mauer, 1979; Sartorius, 1984).

3. Platelet Transfusions

Platelet transfusions are generally ineffective in the treatment of acute ITP because transfused platelets are rapidly destroyed, presumably by the same mechanisms that act on autologous platelets. Transfusions should, however, be given to children with life-threatening bleeding, especially when signs of intracerebral hemorrhage are present.

4. Treatment of Intracranial Hemorrhage

The complication of intracranial hemorrhage in acute ITP is, fortunately, very rare. In addition to platelet transfusions, emergency splenectomy and, where indicated, neurosurgery have been recommended (Zerella *et al.*, 1978; Woerner *et al.*, 1981).

5. Plasma and Plasma Fractions

In studies performed nearly 20 years ago, it was found that infusion of large quantities of normal, fresh-frozen plasma was followed by rapid elevation of platelet levels in >50% of a small series (Reiquam and Pros-

per, 1966). Recent reports have also demonstrated possible benefit in some patients following fresh-frozen plasma transfusion (Lehoczky and Keleman, 1982). Treatment with whole plasma or plasma fractions may be advisable in selected cases, although the rationale for this form of therapy is not yet established. Exchange transfusion with fresh-frozen plasma also appears to have been beneficial in several severely affected children (Lightsey *et al.*, 1975; Novak and Wilimas, 1978). In a very recent report, intravenous injections of large quantities of gamma globulin (400 mg/kg/day for 5 days) was followed by an elevation of platelets to normal in ~1 week in each of six cases (Imbach *et al.*, 1981). Experience with this form of treatment was recently reviewed (Bussel and Hilgartner, 1984). The possible rationale for treatment with plasma and intravenous gamma globulin will be considered in the discussion of chronic ITP (see Section III). In a self-limited and nearly always benign disorder such as acute ITP, it seem advisable to limit the use of plasma and gamma globulin until their effectiveness has been fully established.

6. Development of Chronicity

As noted earlier, ~10–15% of children with apparent acute ITP fail to recover spontaneously. After 6 months, it is generally assumed that the disorder has become "chronic," but exceptions to this rule have been documented. No combination of clinical findings and laboratory measurements yet allows a prediction to be made as to which children are destined to enter a chronic phase. However, recent reports are consistent with the possibility that a very high initial level of PA-IgG (Lightsey *et al.*, 1979, Ware *et al.*, 1982), an elevated level of PA-IgM (van Leeuwen *et al.*, 1981a), and a normal basal level of tritiated thymidine incorporation by peripheral blood lymphocytes (Tomar and Stuart, 1981) are associated with a good prognosis.

III. CHRONIC IDIOPATHIC (AUTOIMMUNE?) THROMBOCYTOPENIC PURPURA

A. General Characteristics

A disorder probably identical to that now designated "chronic idiopathic thrombocytopenic purpura" was first described by Werlhof in 1735, but discovery of the relationship between hemorrhagic symptoms and lack of circulating blood platelets appears to have been made by Krauss ~100 years ago. In contrast to acute ITP, chronic ITP is primarily a disorder of adults, is at least twice as common in females as in males, is

relatively insidious in its onset, is not associated with preceding viral infections, and resolves spontaneously only on rare occasions. No precise statistics on prevalence are available, but chronic ITP, like acute ITP, is a common disorder. Indeed, it seems possible that it may be the most common of all autoimmune disorders in which the cell targeted by the immune system for autoreaction has been clearly identified. A disease similar to chronic ITP has been described in many animal species (Dodds and Wilkins, 1977). Numerous reviews of chronic ITP have been published (Mueller-Eckhardt, 1977; DiFino *et al.*, 1980; Karpatkin, 1980; McMillan, 1981; Kelton and Gibbons, 1982).

B. Pathogenesis

1. Background

For many years after chronic ITP had been recognized as a specific disease entity, a controversy raged as to whether thrombocytopenia was caused by accelerated destruction of platelets in the peripheral blood or by inadequate production of platelets in the bone marrow. Increased numbers of megakaryocytes typically found in the marrow were considered to favor the first viewpoint, but the smooth contour and immature appearance of many of these cells led some to suggest that platelet release was defective. The somewhat fortuitous discovery that dramatic improvement can occur after splenectomy (Kaznelson, 1919) supported the possibility that the spleen itself was the primary site of premature platelet destruction. However, advocates of the proposal that platelet production was inadequate cited the fact that splenomegaly of diverse etiologies is associated with low platelet levels and argued that removal of the spleen acted to eliminate splenic suppression of megakaryocytopoiesis.

With the advent of platelet transfusion therapy and the development of methods for measuring platelet life span with radioisotopic tracers, it was recognized that platelet survival is nearly always shortened in ITP (Hirsch and Gardner, 1952; Cohen *et al.*, 1961; Aster and Keene, 1969; Harker, 1970). The plasma factor that provokes destruction of platelets in the peripheral blood may, to some extent, act on megakaryocytes to impair platelet production (Baldini, 1978). However, calculations of platelet turnover have shown that the effective rate of platelet production is at least normal and often supernormal (Harker, 1970; Branehog *et al.*, 1974).

2. Demonstration of Factors Active against Platelets

Not long after ITP was first described, it was noted that offspring of women suffering from the disorder were sometimes thrombocytopenic at

birth, suggesting the presence of an antiplatelet factor, possibly an antibody, capable of crossing the placenta and affecting fetal platelets. Direct evidence for the existence of such a factor was provided by Harrington *et al.* (1956), who transfused plasma from patients with ITP to normal subjects and found that varying degrees of thrombocytopenia developed in about one-half of the recipients (Fig. 3). A similar effect was observed by Watkins *et al.* (1967) in one-third of plasma recipients. The latter workers localized the antiplatelet activity to the IgG fraction of plasma and found that it became bound *in vitro* to human but not animal platelets. They further demonstrated that reinfusion of autologous plasma to a patient postrecovery caused a significant reduction of platelet levels (Shulman *et al.*, 1965).

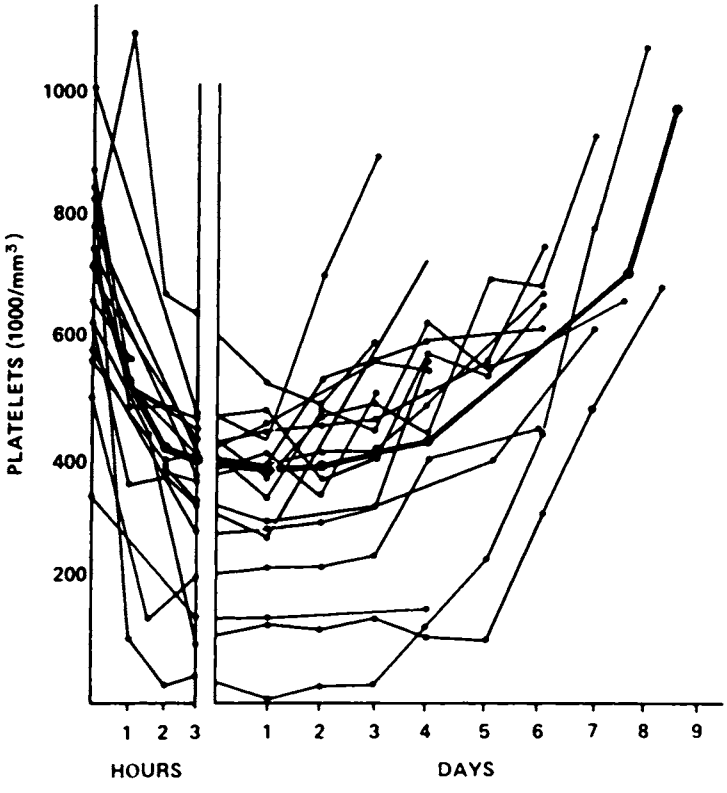


FIG. 3. Effect of infusing 500 ml of blood or its equivalent in plasma from patients with ITP to normal recipients. (From Harrington *et al.*, 1956.)

3. Sites of Platelet Destruction

Using radioisotopic platelet labels, it was found that tagged normal platelets are destroyed in the spleen in most patients, but that a significant degree of hepatic destruction can occur in some instances (Aster and Keene, 1969; Ries and Price, 1974; Heyns *et al.*, 1982). In a patient who had relapsed after splenectomy, hepatic destruction of platelets was demonstrated (Aster and Keene, 1969). Direct evidence for splenic destruction of platelets was obtained by electron microscopic studies of spleens removed at surgery (Tavassoli and McMillan, 1975; Luk *et al.*, 1980).

4. Evidence for Autoantibodies

Together, these findings were consistent with the possibility that platelet destruction in chronic ITP was provoked by autoantibodies capable of sensitizing platelets and promoting their destruction in the reticuloendothelial system, especially in the spleen (Fig. 4). Thus, except for the target cell involved, ITP appeared to be analogous to autoimmune hemolytic anemia in its pathogenesis. Attempts to demonstrate autoantibodies against platelets experimentally date back to the studies of Minot (1916). However, platelets lent themselves less readily than erythrocytes to serological studies, and there followed over the next 50 years a series of sometimes conflicting reports in which platelet autoantibodies were thought to have been demonstrated by agglutination (Harrington *et al.*, 1953; Stefanini and Mele, 1958), platelet lysis (Sauer and von Loghem, 1954), complement fixation (Stefanini and Mele, 1958), and numerous other techniques (reviewed by Baldini, 1966). Others, however, were often unsuccessful in autoantibody identification using these and other techniques (Jackson *et al.*, 1963; Shulman *et al.*, 1964; Mueller-Eckhardt and Boehm, 1968). Typical of the frustrations experienced by platelet serologists during these years was a cooperative study in which serum samples were exchanged among six different laboratories with wholly different results (Sefanini and Mele, 1958).

In recent years, more satisfactory tools for characterization of immunoglobulins reactive with platelets have become available. These have provided important new information but are still somewhat controversial, as will be discussed. Dixon and his colleagues (1975) appear to have been the first to measure specific quantities of IgG immunoglobulins on platelets of patients with ITP. This was accomplished by using sheep red blood cells coated with human IgG as targets for complement-dependent lysis by rabbit anti-IgG antibodies. The IgG in material to be tested, in this case autologous platelets, was assayed by incubating it with the anti-IgG, thus

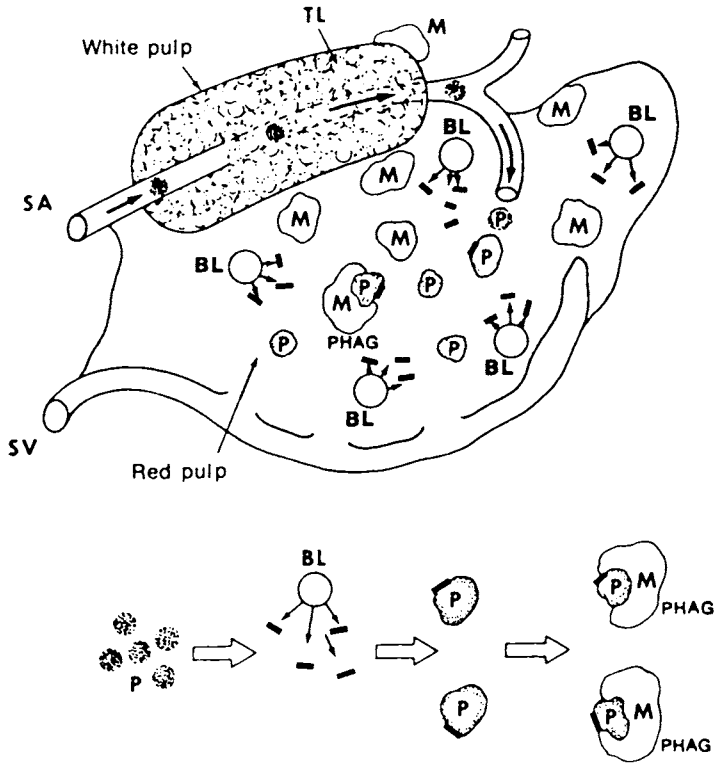


FIG. 4. The splenic microenvironment in idiopathic thrombocytopenic purpura (ITP). After traveling through the white pulp, platelets move slowly through the sinusoidal and extrasinusoidal spaces of the red pulp where they are exposed to a relatively high concentration of antibody produced by splenic B cells and come in close contact with splenic macrophages. Abbreviations: BL, B lymphocyte; M, macrophage; P, platelet; PHAG, phagocytosis of "sensitized" platelet; TL, T lymphocyte; ■, antiplatelet antibody. (From McMillan *et al.*, 1974.)

inhibiting lysis of the subsequently added sheep cells. This antiglobulin consumption test served, in essence, the same purpose as the Coombs' test used to detect autoantibodies on red cells, but had the added advantage of being quantitative. Elevated levels of platelet-associated IgG (PA-IgG) ranging from 2 to 100 times the normal value were found in each of 17 patients with chronic ITP.

These important observations led to a profusion of studies in which PA-IgG was measured by modification of the antiglobulin consumption test and by other techniques both qualitative and quantitative. Methods em-

ployed for this purpose included immunofluorescence (von dem Borne *et al.*, 1978; Sugiura *et al.*, 1980), binding of radioactive antiimmunoglobulin (Soulier *et al.*, 1975; Mueller-Eckhardt *et al.*, 1978; Cines and Schreiber, 1979), radial immunodiffusion (Morse *et al.*, 1981), electroimmunoassay (Kunicki *et al.*, 1982a), nephelometry (Morse *et al.*, 1982), enzyme-linked immunoassay (ELISA) (Leporrier *et al.*, 1979), and binding of radiolabeled staphylococcal protein A (Kekomaki, 1977). This work led to the interesting finding that each platelet normally carries ~4,000 molecules of IgG on its surface (1 fg) and an additional 12,000 molecules (3 fg) in its cytoplasm. Normal PA-IgG appears to be in equilibrium with plasma IgG (Sugiura *et al.*, 1981). Subsequently, platelet-bound complement components were also detected in chronic ITP (Hauch and Rosse, 1977; Cines and Schreiber, 1979; McMillan and Martin, 1981), and measurement of platelet-associated IgM (PA-IgM) has recently been described (Follea *et al.*, 1982; Kunicki *et al.*, 1982b).

As of 1981, studies in >300 patients had demonstrated elevated levels of PA-IgG in >90% (McMillan, 1981). Elevation of platelet-bound C3 was found in more than half these patients (Hauch and Rosse, 1977; Cines and Schreiber, 1979). Occasional patients demonstrated elevated platelet C3 but normal PA-IgG (Cines and Schreiber, 1979). Elevated levels of PA-IgM were found in more than half the patients studied; usually but not always this was in association with elevated PA-IgG (Kunicki *et al.*, 1982b; Follea *et al.*, 1982).

Application of these techniques to measurement of serum autoantibody levels in ITP has been less successful (reviewed by Mueller-Eckhardt, 1977; and by McMillan, 1981). This is perhaps explained by technical difficulties related to the extra manipulation of platelets necessary for indirect assays and the fact that in many patients most antibody is platelet bound. The platelet-bound IgG appears to include all four IgG subclasses, but mainly IgG-I (Rosse *et al.*, 1980; von dem Borne *et al.*, 1980).

From the foregoing review, it might be assumed that the serological complexities of chronic ITP will soon be resolved. However, several recent reports have questioned the specificity of PA-IgG measurements and by inference those of PA-IgM and platelet-associated C3. Pfueller *et al.* (1981) measured both PA-IgG and total platelet protein in chronic ITP, using a radioimmunoassay. In more than one-half of the patients, total platelet protein was elevated in proportion to PA-IgG, suggesting that the binding of IgG to these platelets was not immunologically specific. In another study, elevated PA-IgG was found in a number of patients in whom thrombocytopenia was thought not to be immunologically mediated (Kelton *et al.*, 1982). The picture has been further complicated by the

report of Kelton and Denomme (1982) that normal platelets vary greatly in IgG content when fractionated on the basis of density.

A possible explanation for the apparent nonspecificity of PA-IgG determinations in some patients is provided by recent findings that even small amounts of hemolysis can cause a spurious elevation of PA-IgG, possibly because of sedimentation of red cell microvesicles with the platelet button (Kunicki and Aster, 1981) and the fact that plasma of some patients with *in vivo* platelet destruction contains circulating microvesicles and membranous debris rich in IgG that may be isolated with platelets using conventional separation techniques (Shulman *et al.*, 1982). Circulating microvesicles and platelet fragments were reported earlier in chronic ITP by Zucker-Franklin and Karpatkin (1977).

The aforementioned reports dictate that measurements of PA-IgG, PA-IgM, and platelet-associated C3 should be interpreted with caution but do not totally rule out their validity, since, as noted above, frequently only one or two of these key indices are increased in individual patients, and it is unlikely that this could be a consequence of nonspecific plasma trapping. Moreover, normal values for all three indices are not uncommonly found in patients with thrombocytopenia of nonimmune etiology. A possible means of circumventing the problem of circulating microvesicles and debris was suggested by LoBuglio *et al.* (1983), who obtained uniformly low values for PA-IgG in nonimmune thrombocytopenias when platelets were centrifuged through a density gradient prior to assay. Interestingly, platelets from patients with idiopathic thrombocytopenia purpura were found to carry 500–6000 molecules of autoantibody IgG by this method. These values are much lower than those found by others using different methods but are comparable to the amounts found on red cells in autoimmune hemolytic anemia. Similar findings were made by Shaw *et al.* (1984) using radiolabeled staphyloccal protein A for measurement of cell-bound IgG.

5. Immune Complexes

The propensity of platelets to bind certain classes of immune complexes, which has already been mentioned, raises the question of whether platelet destruction in some patients with ITP is mediated by immune complexes rather than by autoantibodies. Although immune complexes have been identified in plasma of some patients with chronic ITP (Lurhuma *et al.*, 1977; Wautier *et al.*, 1980), no evidence linking these to platelet destruction is yet available. It is of interest that a possible role for immune complexes in the destruction of platelets transfused to thrombocytopenic patients has been proposed (Saffai-Kutti *et al.*, 1980).

6. Relationship of Autoantibodies to Platelet Destruction

Despite the apparent shortcomings of techniques available for assay of platelet-associated IgG, IgM, and C3, the reports cited above provide abundant evidence that autoantibodies, at least those of the IgG class, provoke platelet destruction in chronic ITP. Many of these appear to be capable of activating complement at the platelet surface. In a surprisingly high percentage of patients, autoantibodies of the IgM class appear to be active as well.

In nearly all studies to date, an inverse relationship between PA-IgG and platelet levels has been observed (Fig. 5). Kernoff and his colleagues (1980) also demonstrated a direct relationship between PA-IgG and the

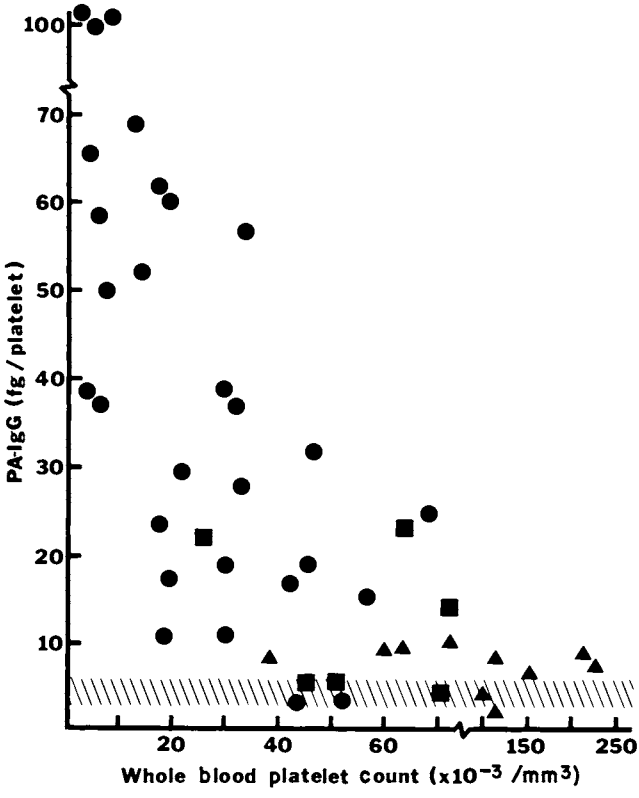


FIG. 5. Relationship between platelet-associated IgG (PA-IgG) and whole blood platelet count in patients with acute or chronic ITP (●), ITP in remission or partial remission (■), and nonimmune thrombocytopenia (▲). Dashed area indicates normal range for PA-IgG. (From Kunicki *et al.*, 1982a.)

degree to which platelet life span is shortened. It seems probable that platelets are destroyed primarily in the spleen when they are minimally sensitized with antibody (Aster and Keene, 1969; Shulman *et al.*, 1965) and in the liver when they are heavily sensitized (Aster and Keene, 1969; Heyns *et al.*, 1982). Synthesis of autoantibody has been demonstrated in cultures of splenic tissue (Karpatkin *et al.*, 1972; McMillan *et al.*, 1974). Splenic destruction of platelets may, therefore, be favored by a high concentration of antibody within that organ. The failure of splenectomy to be curative in all cases of ITP, and the neonatal thrombocytopenia sometimes seen in offspring of women who have responded successfully to splenectomy, make it clear that significant amounts of antibody can be synthesized in extrasplenic sites.

7. Nature of the Autoantigens

To date, little has been learned about the nature of the target antigens against which autoantibodies are directed in ITP. No convincing evidence has been obtained to suggest that the antibodies have alloantigenic specificity. However, it was found that they often fail to bind to platelets from patients with type I Glanzmann's thrombasthenia, suggesting that the target antigen may be expressed on glycoproteins IIb and/or IIIa that are known to be lacking in patients with that disorder (van Leeuwen *et al.*, 1982). It is of great interest that two of the recognized platelet alloantigens, $P1^{A1}$ (Zw^a) (Kunicki and Aster, 1979) and Bak^a (van Leeuwen *et al.*, 1981b), also appear to be expressed on this glycoprotein complex. These findings suggest that the IIb–IIIa complex is quite polymorphic and offer an interesting analogy with the Rh complex of red cells against which antierythrocyte autoantibodies are often directed. Two recent reports indicate that the major sialoglycoprotein of platelets, GPIb, can also be a target for autoantibody in ITP (Szatkowski *et al.*, 1984; Woods *et al.*, 1984).

8. Effect of Autoantibodies on Platelet Function

Since the IIb–IIIa complex is essential for binding of activated platelets to each other and for the formation of an effective hemostatic plug (reviewed by Shattil and Bennett, 1981), it seems possible that binding of autoantibodies to this molecular complex may explain the platelet dysfunction manifested by some patients with ITP (Clancy *et al.*, 1972; Heyns *et al.*, 1978). Indeed, the provocative suggestion has been made that, in some patients, autoantibodies induce platelet dysfunction without causing thrombocytopenia (Karpatkin, 1980; Weiss *et al.*, 1980; Hymes *et al.*, 1981).

9. Hereditary Factors

Little evidence is available to implicate genetic factors in the pathogenesis of chronic ITP. However, *in vitro* immunological abnormalities detected in close relatives of patients were interpreted as being indicative of a hereditary tendency to develop the disorder (Stuart *et al.*, 1978), and other forms of autoimmune disease were found in greater than expected frequencies in relatives of another group of patients (Conley, 1981). Increased frequencies of the antigens HLA-B8 and HLA-B12 (Goebel *et al.*, 1977) and of HLA-DRw2 (Karpatkin *et al.*, 1979) were observed in two series of patients. In other studies, no association was found between ITP and HLA-A, -B (Mueller-Eckhardt *et al.*, 1979; Veenhoven *et al.*, 1979) or HLA-DR (Mayr *et al.*, 1981).

10. Cellular Immunity; Abnormalities of Immunoregulation

In vitro stimulation of lymphocytes by platelets from patients with ITP (Piessens *et al.*, 1970; Wybran and Fudenberg, 1972) and production of migration-inhibition factor by lymphocytes in the presence of platelets (Clancy, 1972; Morimoto *et al.*, 1977) have been reported.

As in other autoimmune diseases, abnormalities of immunoregulation have been sought in chronic ITP. In several recent studies, subnormal numbers of circulating suppressor T-lymphocytes were found in small series of patients (Trent *et al.*, 1981; Lauria *et al.*, 1981). Induction of this change by immune complexes has been suggested (Trent *et al.*, 1981). Abnormal transformation and capping of autologous lymphocytes in response to various stimuli (Quagliata and Karpatkin, 1979) and abnormal reactions in autologous mixed lymphocyte cultures (Zinberg *et al.*, 1982) have also been described. The relationship of these findings to the pathogenesis of ITP remains unclear. Abnormalities of immunoregulation are characteristic of some disorders with which ITP is known to be associated (see Section III,B,12).

11. Hormonal Influences

A possible role for hormonal influences in chronic ITP is suggested by the high frequency with which the disorder occurs in women during the postpubertal, premenopausal period by the relapses that sometimes occur during pregnancy (Laros and Sweet, 1975) and by the response of some patients to treatment with a synthetic androgen (Ahn *et al.*, 1983). The interesting suggestion has been made that the increased prevalence of the disorder in women may be related to increased expression of Fc receptors on female platelets (Moore *et al.*, 1981).

12. Association with Other Diseases

A significant relationship between chronic ITP and other disorders of immunological and nonimmunological etiology has long been recognized. A syndrome indistinguishable from chronic ITP may occur in association with thyrotoxicosis (Hymes *et al.*, 1981; Herman *et al.*, 1978), systemic lupus erythematosus (reviewed by Budmann and Steinberg, 1977), Hashimoto's thyroiditis (Hymes *et al.*, 1981), myasthenia gravis (Veenhoven *et al.*, 1979), sarcoidosis (Knodel and Beekman, 1980), and a variety of malignant disorders including carcinoma (Schwartz *et al.*, 1982) and lymphoma (chronic lymphocytic leukemia, Hodgkin's disease, and non-Hodgkin's lymphoma) (Kaden *et al.*, 1979; Kirshner *et al.*, 1980). The association of autoimmune hemolytic anemia and chronic ITP is designated "Evans' syndrome" (Pui *et al.*, 1980; DiFino *et al.*, 1980). Recent reports indicate that at least two different autoantibodies, rather than one autoantibody reactive with both red cells and platelets, are present in such patients (Monch *et al.*, 1981; Pegels *et al.*, 1982).

Recently, ITP has been described in homosexual men having decreased numbers of helper T cells in their circulation without overt acquired immunodeficiency syndrome (AIDS) (Morris *et al.*, 1982) and following successful allogeneic bone marrow transplantation (Minchinton *et al.*, 1982), and in patients with classical hemophilia treated with large quantities of commercial factor VIII concentrates (Suffredini and Qureschi, 1982; Ratnoff *et al.*, 1983). The basis for these associations is not clear, but it seems possible that abnormalities of immunoregulation and/or circulating immune complexes reactive with platelets may be of importance. Hyperreactivity of the reticuloendothelial system has been suggested as a contributing factor in patients with coexistent ITP and thyrotoxicosis (Kurata *et al.*, 1980).

Chronic ITP frequently occurs in women of childbearing age, but there is no direct evidence that this apparent relationship is statistically significant.

13. Summary

Some of the mechanisms thought to be active in the pathogenesis of chronic ITP are summarized in Fig. 6.

C. Clinical Presentation

Chronic ITP occurs most often between puberty and 50 years of age, but, as noted above, it can occur in children and in the aged. In children, the disorder is similar to the adult form in its response to therapy and,

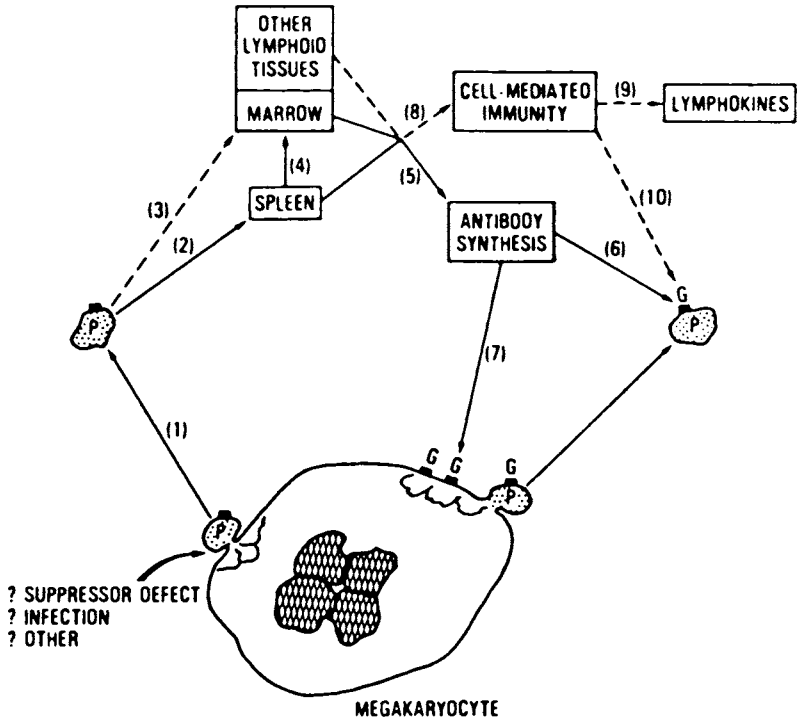


FIG. 6. Immunological mechanisms that may operate to promote platelet destruction in chronic ITP. Platelet-associated autoantigen or tightly adherent nonplatelet antigen is expressed on the platelet surface (1). Antibody production against this antigen is stimulated, probably first in the spleen (2), but later in the marrow and other extra splenic sites (3). Subsequently, the immune response against this antigen becomes more generalized (4). Antiplatelet antibody synthesis (5) results in levels of autoantibody (G) sufficient to sensitize platelets (6) and promote platelet destruction. Autoantibody may also bind to the same antigenic determinant expressed on the megakaryocyte membrane (7), suppressing platelet production in some instances. The degree to which cell-mediated immunity is activated (8) with consequential lymphokine production (9) and possible cell-mediated effects on platelets (10) is unknown. Dashed lines denote mechanisms as yet unconfirmed by experimental evidence. As discussed in the text, an abnormality of immunoregulation may predispose a patient to initial autoantibody production. (From McMillan, 1981.)

presumably, in its pathogenesis. Females are affected about three times as often as males. The disease may be less common in blacks than in Caucasians (DiFino *et al.*, 1980). Initial bleeding manifestations are usually less dramatic than in children with acute ITP, and, in some patients, the disorder may be present in subclinical form for many months before the diagnosis is made. Hemorrhagic lesions may occur anywhere on the

skin or in the gastrointestinal and urinary tracts. In the most severely affected patients, hemorrhagic bullae are commonly found in the buccal mucosa. Bleeding into joints and the retina is, fortunately, rare. Intracranial hemorrhage is also unusual, but is a potential risk in every patient. The spleen and liver are typically not enlarged.

D. Laboratory Findings

Platelet levels may range from a few hundred per ml in severely affected patients to as high as 100,000 per ml. On blood smears, large bizarrely shaped platelets characteristic of stress thrombopoiesis are often seen. Several cases have been described in which platelets were adherent to circulating neutrophils (White *et al.*, 1978). Serological tests for disseminated lupus erythematosus are positive in the small percentage of patients who have associated SLE (Budman and Steinberg, 1977). The bleeding time is often prolonged in proportion to the severity of the thrombocytopenia, but there are many exceptions to this rule, especially after institution of corticosteroid therapy. As noted above, platelet dysfunction can be demonstrated by *in vitro* testing in some patients, suggesting that autoantibodies in such cases inhibit platelet function.

Platelet-associated IgG (PA-IgG) has been found to be elevated in >90% of the cases studied, as noted above in the discussion of pathogenesis. Precautions to be taken in interpreting such results have also been mentioned. In more than half the cases, PA-IgM and C3 are elevated. There is reason to expect that methodological problems inherent in the assay of platelet-associated plasma factors will soon be resolved and that such measurements will soon become valuable for diagnosis and for the assessment of prognosis and/or response to therapy.

Marrow aspiration usually reveals normal or increased numbers of megakaryocytes, many of which are small and immature. As noted above, the "nonbudding" appearance of these cells is probably not indicative of any abnormality. IgG has been demonstrated on the surface of megakaryocytes by immunofluorescent techniques (McMillan *et al.*, 1978).

Increased numbers of reactive lymphoid nodules and plasma cells peripheral to small vessels of the marginal zone and suggestive of active antibody production have been described in spleens removed at surgery (Tavassoli and McMillan, 1975), and phagocytosis of platelets by splenic macrophages has been demonstrated on electron microscopy of splenic sections (Tavassoli and McMillan, 1975; Luk *et al.*, 1980). Large numbers of lipid-laden macrophages are found in the spleens of perhaps 10% of patients (King and Harsock, 1968; Takahashi *et al.*, 1977).

E. Treatment and Prognosis

1. General Considerations

Treatment of chronic ITP has been discussed in a number of recent publications (DiFino *et al.*, 1980; Karpatkin, 1980; McMillan, 1981; Lacey and Penner, 1977; Picozzi *et al.*, 1980; Kelton and Gibbons, 1982). Patients with the most severe thrombocytopenia and hemorrhagic symptoms sometimes recover spontaneously in 1 or 2 weeks. These individuals may have acute ITP of the type seen in children or sensitivity to an unrecognized exogenous antigen or drug.

2. Adrenal Corticosteroids

The effectiveness of adrenal corticosteroids in chronic ITP was discovered empirically soon after these drugs became available. Their mechanism of action is still not fully understood, but it probably involves inhibition of the phagocytic activity of macrophages, especially in the spleen (Shulman *et al.*, 1965; Handin and Stossel, 1978). Antibody synthesis is probably inhibited as well. Because the bleeding time is sometimes normalized after institution of corticosteroid therapy before the platelet count changes, it has been suggested that these drugs exert a beneficial effect on capillary integrity, but this has been difficult to establish experimentally. Most patients respond with a rise in platelet count and a decrease in the severity of symptoms when treated with prednisone in the range of 0.5–2.0 mg/kg body weight. Much larger doses are sometimes required in severe cases.

3. Platelet Transfusions

As in acute ITP, transfused platelets are likely to be rapidly destroyed by the same mechanisms that act on autologous platelets. Moreover, hemorrhagic symptoms usually respond readily to corticosteroid therapy. Platelet transfusion should, however, be given to patients with potentially life-threatening symptoms.

4. Splenectomy

Removal of the spleen appears to benefit patients with chronic ITP by at least two mechanisms: removal of the organ that has the greatest ability to extract antibody-coated platelets from the circulation, and elimination of a major site of autoantibody production. Splenectomy is usually indicated in patients who failed to recover spontaneously within a few weeks

of diagnosis and who require maintenance doses of prednisone that are sufficiently high to induce adverse side effects. About 75–85% of patients develop a normal platelet level within 1 or 2 weeks of splenectomy. Even those in whom platelet counts fail to rise may experience a reduction in doses of corticosteroids needed to prevent hemorrhage. In most patients, restoration of platelet counts to normal is followed by the disappearance of all traces of autoantibody. However, persistent elevation of PA-IgG (Luiken *et al.*, 1977), and a slight but persistent reduction of platelet life span (Branehog, 1975; Burger *et al.*, 1978), have been described following splenectomy-induced remission.

There is as yet no definite answer to the intriguing question of how to predict which patients will be benefited by removal of the spleen. In different studies, a good prognosis has been thought to correlate with response to corticosteroid therapy (Brennan *et al.*, 1975; Thompson *et al.*, 1972), early restoration of platelet levels to normal after surgery (MacPherson and Richmond, 1975; DiFino *et al.*, 1980), and the demonstration before surgery that transfused, radiolabeled platelets are destroyed preferentially in the spleen (Najean and Ardaillou, 1971; Viala *et al.*, 1975; Burger *et al.*, 1978). The latter suggestion has been especially controversial, but the prevailing view is that study of the sites of platelet destruction is not useful in predicting the response to splenectomy (Aster and Keene, 1969; Ries, 1977; Richards and Thompson, 1979; Heyns *et al.*, 1982). Older patients are probably less likely than are younger patients to respond to removal of the spleen (DiFino *et al.*, 1980).

The early suggestion that systemic lupus erythematosus is likely to become manifest after splenectomy has not been borne out by more recent studies (Best and Darling, 1962; Homan and Dineen, 1978). Although splenectomy is usually permanently curative, relapses occur in some patients many years after surgery. In most, this is a consequence of exacerbation of the disease and production of autoantibody in quantities sufficient to provoke destruction of platelets in extrasplenic sites. Hypertrophy of accessory spleens has been shown to be responsible in a number of well-documented cases (DiFino *et al.*, 1980; Verheyden *et al.*, 1978; Davis *et al.*, 1980), and these should be looked for using modern imaging techniques (Davis *et al.*, 1980). Even in adults, the danger of increased susceptibility to infection post splenectomy should be carefully considered (Karpatkin, 1980).

5. Immunosuppression

Patients who fail to respond to splenectomy or are not candidates for operation sometimes improve when treated with immunosuppressive agents. Recommended regimens include azothioprine, cyclophos-

phamide, and vincristine together with prednisone (these have recently been reviewed by McMillan, 1981). The highest rate of response, in excess of 50%, appears to have been achieved with vincristine (Ahn *et al.*, 1974; Ries, 1976; Lacey and Penner, 1977). Improvement following vinca alkaloid therapy is sometimes temporary, however (DiFino *et al.*, 1980; McMillan, 1981). Risks of immunosuppressive therapy, including susceptibility to infection and malignancy, should be carefully considered in selecting patients for immunosuppressive therapy.

6. Vinblastine-Treated Platelets

Significant elevations of platelet levels have been achieved in patients refractory to conventional therapy by infusion of normal platelets treated with vinblastine, which binds to microfilamentous platelet proteins (Ahn *et al.*, 1978; Nenci *et al.*, 1981). It has been suggested that effectiveness of vinblastine- or vincristine-treated platelets may be limited to patients in whom platelet survival time is quite short because of dissociation of the alkaloid from platelets *in vivo* (Kelton *et al.*, 1981). Vincristine may be superior to vinblastine for this type of therapy (Agnelli *et al.*, 1982).

7. Plasma Exchange

Intensive exchange of patient plasma for normal, fresh-frozen plasma by exchange pheresis appears to have been beneficial in a few instances (Branda *et al.*, 1978; Patten and Reuter, 1980; Marder *et al.*, 1981). Benefits are nearly always transient at best, and this form of therapy should probably be limited to patients with life-threatening symptoms.

8. Intravenous Gamma Globulin

Several reports have documented dramatic but sometimes temporary elevations of platelet levels in patients with chronic ITP refractory to other forms of treatment (Imbach *et al.*, 1981; Schmidt *et al.*, 1981; Fehr *et al.*, 1982; Bussel *et al.*, 1982; Bussel and Hilgartner, 1984). Experimental studies suggest that the transfused IgG acts by inducing temporary blockade of the reticuloendothelial system, thus inhibiting ingestion of antibody-coated platelets by phagocytes (Fehr *et al.*, 1982). In one series of children, the duration of benefits sometimes exceeded the presumed life span of the transfused IgG for unknown reasons (Bussel *et al.*, 1982).

9. Other approaches

Recent reports suggest a beneficial effect from danazol, an impeded androgen (Harrington *et al.*, 1982; Ahn *et al.*, 1983), and colchicine

(Strother *et al.*, 1982). The latter report is of particular interest because of the minimal side effects associated with use of that drug.

IV. CONCLUSIONS

There is reason to hope that the designation "idiopathic thrombocytopenic purpura" will soon become obsolete because of advances in laboratory diagnosis and in the understanding of the physiology of immunoregulation. At the very least, it should soon be possible to establish the diagnosis of autoimmune thrombocytopenic purpura with certainty in the laboratory, and, perhaps, to distinguish between the "acute" and "chronic" forms of the disease prospectively. Even if these objectives are reached, important unresolved questions will remain to challenge basic and clinical investigators.

One such question is the unknown mechanism by which fulminating destruction of platelets is somehow provoked by what to all appearances is a normal immunological response to an infecting organism. Another is the role that abnormalities of immunoregulation, arising *de novo* or associated with a coexisting disorder, play in provoking the production of autoantibodies reactive with the platelet membrane. Also intriguing are the questions, "Why are platelets so often the target for an autologous immune response?" and "What is the site (or sites) on the platelet membrane toward which this response is directed?" Answers to these and related questions are likely to be forthcoming in the next decade and should lead to improved forms of therapy, especially for the unfortunate minority of patients who fail to respond to presently available modes of treatment.

REFERENCES

Agnelli, G., DeCunto, M., Grisele, P., and Nenci, G. P. (1982). *Blood* **60,** 1235.

Ahn, Y. S., Harrington, W. J., Seelman, R. C., and Eytel, C. S. (1974). *N. Engl. J. Med.* **291,** 376–380.

Ahn, Y. S., Byrnes, J. J., Harrington, W. J., Cayer, M. L., Smith, D. S., Brunskill, D. E., and Pall, L. M. (1978). *N. Engl. J. Med.* **298,** 1101–1107.

Ahn, Y. S., Harrington, W. J., Simon, S. R., Mylvaganam, R., Pall, L. M., and So, A. G. (1983). *N. Engl. J. Med.* **308,** 1396–1399.

Aster, R. H., and Keene, W. R. (1969). *Br. J. Haematol.* **16,** 61–73.

Baldini, M. G. (1966). *N. Engl. J. Med.* **274,** 1245–1251, 1302–1306, 1360–1367.

Baldini, M. G. (1978). *JAMA, J. Am. Med. Assoc.* **239,** 2477–2479.

Best, W. R., and Darling, D. R. (1962). *Med. Clin. North Am.* **46,** 19–44.

Branda, R. F., Tate, D. Y., McCullough, J. J., and Jacob, H. S. (1978). *Lancet* **1,** 688–690.

Branehog, I. (1975). *Br. J. Haematol.* **29,** 413–426.

Branehog, I., Kutti, J., and Weinfeld, A. (1974). *Br. J. Haematol.* **27,** 127–143.
Brennan, M. F., Rappeport, J. M., Maloney, W. C., and Wilson, R. E. (1975). *Am. J. Surg.* **129,** 490–492.
Brown, W. M., and Axelrad, A. A. (1976). *Int. J. Cancer* **18,** 764–773.
Budman, D. R., and Steinberg, A. D. (1977). *Ann Intern. Med.* **86,** 220–229.
Burger, T., Schmelczer, M., Kett, R., and Kutas, J. (1978). *Acta Med. Acad. Sci. Hung.* **35,** 213–224.
Burstein, Y., and Berns, L. (1982). *Pediatr. Ann.* **11,** 323–331.
Bussel, J., Kimberly, R., Hilgartner, M., Cheung, N., O'Malley, J., Salmon, J., and Barandum, S. (1982). *Blood* **60,** (Suppl. 1), 185a.
Bussel, J. B., and Hilgartner, M. W. (1984). *Brit. J. Haematol.* **56,** 1–17.
Carpentieri, U., and Haggard, M. E. (1975). *Tex. Med.* **71,** 81–83.
Christie, D. J., and Aster, R. H. (1982). *J. Clin. Invest.* **70,** 989–998.
Cines, D. B., and Schreiber, A. D. (1979). *N. Engl. J. Med.* **300,** 106–111.
Clancy, R. (1972). *Lancet* **1,** 6–9.
Clancy, R., Jenkins, E., and Firkin, B. (1972). *N. Engl. J. Med.* **286,** 622–626.
Cohen, P., Gardner, R. H., and Barnett, G. O. (1961). *N. Engl. J. Med.* **264,** 1294–1299, 1350–1355.
Conley, C. L. (1981). *Johns Hopkins Med. J.* **149,** 101–109.
Davis, H. H., Varki, A., Heaton, A., and Siegel, B. A. (1980). *Am. J. Hematol.* **8,** 81–86.
DiFino, S. M., Lachant, N. A., Kirshner, J. J., and Goettlieb, A. J. (1980). *Am. J. Med.* **69,** 431–442.
Dixon, R., Rosse, W., and Ebbert, L. (1975). *N. Engl. J. Med.* **292,** 230–236.
Dodds, W. J., and Wilkins, R. J. (1977). *Am. J. Pathol.* **86,** 489–491.
Evers, K. G., Thouet, R., Haase, W., and Kruger, J. (1978). *Eur. J. Pediatr.* **129,** 267–272.
Fehr, J., Hofmann, V., and Urs Kappler, C. M. (1982). *N. Engl. J. Med.* **306,** 1254–1258.
Follea, G., Mandrand, B., and Dechavanne, M. (1982). *Thromb. Res.* **26,** 249–258.
Goebel, K. M., Hahn, E., and Havemann, K. (1977). *Br. J. Haematol.* **35,** 341–342.
Handin, R. I., and Stossel, T. P. (1978). *Blood* **51,** 771–779.
Harker, L. A. (1970). *Br. J. Haematol.* **19,** 95–104.
Harrington, W. J., Sprague, C. C., Minnich, V., Morre, C. V., Aulvin, R. C., and Dubach, R. (1953). *Ann. Intern. Med.* **38,** 433–469.
Harrington, W. J., Minnich, V., and Arimura, G. (1956). *Prog. Hematol.* **1,** 166–192.
Harrington, W. J., Ahn, Y. S., So, A. G., Ayub, J., Simon, S., Pall, L. M., and Mylvaganam, R. (1982). *Blood* **60** (Suppl. 1), 186a.
Hauch, T. W., and Rosse, W. F. (1977). *Blood* **50,** 1129–1136.
Herman, J., Resnitzky, P., and Fink, A. (1978). *Isr. J. Med. Sci.* **14,** 469–475.
Heyns, A. D. P., Fraser, J., and Retief, F. P. (1978). *J. Clin. Pathol.* **31,** 1239–1243.
Heyns, A. P., Lotter, M. G., Badenhorst, P. N., DeKock, F., Pieters, H., Herbst, C. van Reenen, O. R., Kotze, H., and Minnaar, P. C. (1982). *Am. J. Hematol.* **12,** 167–177.
Hirsch, E. O., and Gardner, F. H. (1952). *J. Lab. Clin. Med.* **39,** 556–569.
Homan, W. P., and Dineen, P. (1978). *Ann. Surg.* **187,** 52–56.
Hymes, K., Blum, M., Lackner, H., and Karpatkin, S. (1981). *Ann. Intern. Med.* **94,** 27–30.
Imbach, P., D'Apuzzo, V., Hirt, A., Russi, E., Vest, M., Barandun, S., Baumgartner, C., Morell, A., Schoni, M., and Wagner, H. B. (1981). *Lancet* **1,** 1228–1231.
Israels, E. D., Nisli, J., Paraskevas, F., and Israels, L. G. (1973). *Thromb. Diath. Haemorrh.* **29,** 434–444.
Jackson, D. P., Schmid, H. J., Zieve, P. D., Levin, J., and Conley, C. L. (1963). *J. Clin. Invest.* **42,** 383–390.
Kaden, B. R., Rosse, W. F., and Hauch, T. W. (1979). *Blood* **53,** 545–551.

Karpatkin, M., and Karpatkin, S. (1981). *Am. J. Pediat. Hematol. Oncol.* **3,** 213–219.
Karpatkin, S. (1980). *Blood* **56,** 329–343.
Karpatkin, S., Struck, N., and Siskind, G. W. (1972). *Br. J. Haematol.* **23,** 167–176.
Karpatkin, S., Fotino, M., Gibofsky, A., and Winchester, R. J. (1979). *J. Clin. Invest.* **63,** 1085–1088.
Kasnelson, P. (1919). *Wein. Klin. Wochensch.* **29,** 145–150.
Kazatchkine, M. D., Lambrey, C. R., Kieffer, N., Maillet, F., and Nurden, A. T. (1984). *J. Clin. Invest.* **74,** 976–984.
Kekomaki, R. (1977). *Med. Biol.* **54,** 112–114.
Kelton, J. G., and Denomme, G. (1982). *Blood* **60,** 136–139.
Kelton, J. G., and Gibbons, S. (1982). *Sem. Thromb. Hemostasis.* **8,** 83–104.
Kelton, J. G., McDonald, J. W. D., Barr, R. M., Walker, I., Nicholson, W., Neame, P. B., Hamid, C., Wong, T. Y., and Hirsh, J. (1981). *Blood* **57,** 431–436.
Kelton, J. G., Powers, P. J., and Carter, C. J. (1982). *Blood* **60,** 1050–1053.
Kernoff, L. M., Blake, C. H., and Shackleton, D. (1980). *Blood* **55,** 730–733.
Khalifa, A. S., Lusher, J. M., Cejka, J., and Zuelzer, W. W. (1976). *Acta Haematol.* **56,** 205–211.
King, F. M., and Harsock, R. J. (1968). *Am. J. Clin. Pathol.* **49,** 250–255.
Kirshner, J. J., Zamkoff, K. W., and Gottlieb, A. J. (1980). *Am. J. Med. Sci.* **280,** 21–28.
Knodel, A. R., and Beekman, J. F. (1980). *JAMA, J. Am. Med. Assoc.* **243,** 258–259.
Kunicki, T. J., and Aster, R. H. (1979). *Mol. Immunol.* **16,** 353–360.
Kunicki, T. J., and Aster, R. H. (1981). *Transfusion* **21,** 639–640.
Kunicki, T. J., Koenig, M. B., Kristopeit, S. M., and Aster, R. H. (1982a). *Blood* **60,** 54–58.
Kunicki, T. J., Koenig, M. B., and Aster, R. H. (1982b). *Blood* **60** (Suppl.1), 188a.
Kurata, Y., Nishioed, Y., Tsubakio, T., and Kitani, T. (1980). *Acta Haematol.* **63,** 185–190.
Lacey, J. V., and Penner, J. A. (1977). *Sem. Thromb. Hemostasis* **3,** 160–173.
Lammi, A. T. and Lovric, V. A. (1973). *J. Pediatr. (St. Louis)* **83,** 31–36.
Laros, R. K., and Sweet, R. L. (1975). *Am. J. Obstet. Gynecol.* **122,** 182–191.
Lau, P., Sholtis, M., and Aster, R. H. (1980). *Am. J. Hematol.* **9,** 331–336.
Lauria, F., Mantovani, V., Catovsky, D., Guarini, A., Gobbi, M., Gugliotta, L., Mirone, E., and Tura, S. (1981). *Scand. J. Haematol.* **26,** 156–160.
Lehoczky, D., and Kelemen, E. (1982). *N. Engl. J. Med.* **307,** 1150–1151.
Leporrier, M., Dighiero, G., Auzemery, M., and Binet, J. L. (1979). *Br. J. Haematol.* **42,** 605–611.
Lightsey, A. L., McMillan, R., and Koenig, H. M. (1975). *JAMA, J. Am. Med. Assoc.* **232,** 734–737.
Lightsey, A. L., Koenig, H. M., and McMillan, R. (1979). *J. Pediatr. (St. Louis)* **94,** 201–204.
LoBuglio, A. F., Court, W. S., Vincour, L., Maglott, J. G., and Shaw, G. M. (1983). *N. Engl. J. Med.* **309,** 459–465.
Luiken, G. A., McMillan, R., Lightsey, A. L., Gordon, P., Zevely, S., Schulman, I., Gribble, T. J., and Longmire, R. L. (1977). *Blood* **50,** 317–323.
Luk, S. C., Musclow, E., and Simon, G. T. (1980). *Histopathology* **4,** 127–136.
Lurhuma, A. Z., Riccomi, H., and Masson, P. L. (1977). *Clin. Exp. Immunol.* **28,** 49–55.
Lusher, J. M., and Iyer, R. (1977). *Sem. Thromb. Hemostasis* **3,** 175–199.
Lusher, J. M., Emami, A., Ravindranath, Y., and Warrier, A. I. (1984). *Am. J. Pediatr. Hematol. Oncol.* **6,** 149–158.
McClure, P. D. (1975). *Pediatrics* **55,** 68–74.
McIntosh, S., Johnson, C., Hartigan, P., Baumgarten, A., and Dwyer, J. M. (1981). *J. Pediatr. (St. Louis)* **99,** 525–530.

McMillan, R. (1977). *CRC Crit. Rev. Clin. Lab. Sci.* **8,** 303–332.
McMillan, R. (1981). *N. Engl. J. Med.* **304,** 1135–1147.
McMillan, R., and Martin, M. (1981). *Br. J. Haematol.* **47,** 251–256.
McWilliams, N. B., and Mauer, H. M. (1979). *Am. J. Hematol.* **7,** 87–96.
McMillan, R., Longmire, R. L., Yelenosky, R., Donnel, R. L., and Armstrong, S. (1974). *N. Engl. J. Med.* **291,** 812–817.
McMillan, R., Luiken, G. A., Levy, R., Yelenosky, R., and Longmire, R. L. (1978). *JAMA, J. Am. Med. Assoc.* **239,** 2460–2462.
MacPherson, A. I. S., and Richmond, J. (1975). *Br. Med. J.* **1,** 64–66.
Marder, V. J., Nubacher, J., and Anderson, F. W. (1981). *Transfusion* **21,** 291–297.
Mayr, W. R., Mueller-Eckhardt, G., Kruger, M., Mueller-Eckhardt, C., Lechner, K., and Niessner, H. (1981). *Tissue Antigens* **18,** 56–57.
Minchinton, R. M., Waters, A. H., Kendra, J. R., and Barrett, A. J. (1982). *Lancet* **2,** 627–629.
Minot, G. R. (1916). *Am. J. Med. Sci.* **152,** 48–55.
Monch, H., Briethaupt, H., and Mueller-Eckhardt, C. (1981). *Blut* **42,** 27–32.
Moore, A., Ross, G. T., and Nachman, R. L. (1978). *J. Clin. Invest.* **62,** 1053–1060.
Moore, A., Weksler, B. B., and Nachman, R. L. (1981). *Thromb. Res.* **21,** 469–474.
Morimoto, C., Abe, T., Hara, M., and Homma, M. (1977). *Clin. Immunol. Immunopathol.* **8,** 181–189.
Morris, L., Distenfeld, A., Amorosi, E., and Karpatkin, S. (1982). *Ann. Intern. Med.* **96,** 714–717.
Morse, B. S., Giuliani, D., and Nussbaum, M. (1981). *Blood* **57,** 809–811.
Morse, B. S., Guiliani, D., and Nussbaum, M. (1982). *Am. J. Hematol.* **12,** 271–275.
Mueller-Eckhardt, C. (1977). *Sem. Thromb. Hemostasis* **3,** 125–159.
Mueller-Eckhardt, C., and Boehm, H. (1968). *Klin. Wochenschr.* **46,** 986–998.
Mueller-Eckhardt, C., Mahn, I., Schulz, G., and Mueller-Eckhardt, G. (1978). *Vox Sang.* **35,** 357–365.
Mueller-Eckhardt, C., Mayr, W., Lechner, K., Mueller-Eckhardt, G., Niessner, H., and Pralle, A. (1979). *Scand. J. Haematol.* **23,** 348–352.
Myllyla, G., Vaheri, A., Vesikari, T., and Penttinen, K. (1969). *Clin. Exp. Immunol.* **4,** 323–332.
Najean, Y., and Ardaillou, N. (1971). *Br. J. Haematol.* **21,** 153–164.
Nenci, G. G., Agnelli, G., Decunto, M., and Gresele, P. (1981). *Acta Haematol.* **66,** 117–121.
Norton, J. H., Shulman, N. R., Corash, L., Smith, R. L., Au, F., and Rosenberg, S. A. (1977). *Cancer* **41,** 820–826.
Novak, R., and Wilimas, J. (1978). *Pediatrics* **92,** 434–436.
Osborn, J. E., and Shahidi, N. T. (1973). *J. Lab. Med.* **81,** 53–63.
Oski, F. A., and Naiman, J. L. (1966). *N. Engl. J. Med.* **275,** 352–356.
Patten, E., and Reuter, F. P. (1980). *Transfusion* **20,** 589–593.
Pegels, J. G., Helmerhorst, F. M., van Leeuwen, E. F., van de Plas-Van Dalen, C., Engelfriet, C. P., and von dem Borne, A. E. G. Kr. (1982). *Br. J. Haematol.* **51,** 445–450.
Pfueller, S. L., and Luscher, E. F. (1972). *Immunochemistry* **9,** 1151–1165.
Pfueller, S. L., Cosgrov, L., Firkin, B. G., and Tew, D. (1981). *Br. J. Haematol.* **49,** 293–302.
Picozzi, V. J., Roeske, W. R., and Creger, W. P. (1980). *Am. J. Med.* **69,** 690–694.
Piessens, W. F., Wybran, J., Manaster, J., and Strijckmans, P. A. (1970). *Blood* **36,** 421–427.

Pui, C. H., Wilimas, J., and Wang, W. (1980). *J. Pediatr. (St. Louis)* **97,** 754–758.
Quagliata, F., and Karpatkin, S. (1979). *Blood* **53,** 341–349.
Ratnoff, O., Menitove, J. E., Aster, R. H., and Lederman, M. (1983). *N. Engl. J. Med.* **308,** 439–442.
Reiquam, C. W., and Prosper, J. C. (1966). *J. Pediatr. (St. Louis)* **68,** 880–885.
Richards, J. D. M., and Thompson, D. S. (1979). *J. Clin. Pathol.* **32,** 1248–1252.
Ries, C. A. (1976). *N. Engl. J. Med.* **295,** 1136–1137.
Ries, C. A. (1977). *Ann. Intern. Med.* **86,** 194–195.
Ries, C. A., and Price, D. C. (1974). *Ann. Intern. Med.* **80,** 702–707.
Rosse, W. F., Adams, J. P., and Yount, W. J. (1980). *Br. J. Haematol.* **46,** 109–114.
Saffai-Kutti, S., Zaroulis, G. C., Day, N. K., Good, R. A., and Kutti, J. (1980). *Vox Sang.* **39,** 22–27.
Sartorius, J. A. (1984). *Am. J. Pediatr. Hematol. Oncol.* **6,** 165–174.
Sauer, A. J., and van Loghem, J. J. (1954). *Vox Sang.* **4,** 120–125.
Schmidt, R. E., Budde, U., Schafer, G., and Stroehmann, I. (1981). *Lancet* **2,** 475–476.
Schwartz, K. A., Slichter, S. J., and Harker, L. A. (1982). *Br. J. Haematol.* **51,** 17–24.
Shattil, S. J., and Bennett, J. S. (1981). *Ann. Intern. Med.* **94,** 108–118.
Shaw, G. M., Axelson, J., Maglott, J. G., and LoBuglio, A. F. (1984). *Blood* **63,** 154–161.
Shulman, N. R. (1958). *J. Exp. Med.* **107,** 665–690.
Shulman, N. R., Aster, R. H., Leitner, A., and Hiller, M. C. (1961). *J. Clin. Invest.* **40,** 1597–1620.
Shulman, N. R., Marder, V. J., Hiller, M. C., and Collier, E. M. (1964). *Prog. Hematol.* **4,** 222–304.
Shulman, N. R., Marder, V. J., and Weinrach, R. S. (1965). *Ann. N. Y. Acad. Sci.* **124,** 499–542.
Shulman, N. R., Leissinger, C. A., Hotchkiss, A., and Kautz, C. (1982). *Blood* **60** (Suppl. 1), 191a.
Simons, S. M., Main, C. A., Yaish, H. M., and Rutzky, J. (1975). *J. Pediatr.* **87,** 16–22.
Soulier, J. C., Paterrau, C., and Brouet, J. (1975). *Vox Sang.* **29,** 253–268.
Stefanini, M., and Mele, R. H. (1958). *Acta Haematol.* **20,** 195–202.
Strother, S. V., Zuckerman, K. S., and LoBuglio, A. F. (1982). *Blood* **60** (Suppl. 1), 192a.
Stuart, M. J., Tomar, R. H., Miller, M. L., and Davey, F. R. (1978). *JAMA, J. Am. Med. Assoc.* **239,** 939–942.
Suffredini, A., and Qureshi, G. D. (1982). *JAMA, J. Am. Med. Assoc.* **247,** 2497–2498.
Sugiura, K., Steiner, M., and Baldini, M. (1980). *J. Lab. Clin. Med.* **96,** 640–653.
Sugiura, K., Steiner, M., and Baldini, M. (1981). *Thromb. Haemostasis* **45,** 27–33.
Szatkowski, N. S., Kunicki, T. J., and Aster, R. H. (1984). *Circulation* **70,** Suppl. II, 357.
Takahashi, K., Hakozaki, H., Terashima, K., and Kogima, M. (1977). *Acta Pathol. Jpn.* **27,** 447–462.
Tavassoli, M., and McMillan, R. (1975). *Am. J. Clin. Pathol.* **64,** 180–191.
Terada, H., Baldini, M., Ebbe, S., and Maddoff, M. A. (1966). *Blood* **28,** 231–228.
Thompson, R. L., Moore, R. A., Hess, C. E., Whevy, M. S., and Leavell, B. S. (1972). *Arch. Intern. Med.* **130,** 730–734.
Tomar, R. H., and Stuart, M. J. (1981). *Am. J. Dis. Child.* **135,** 446–449.
Trent, R. J., Clancy, R. L., Danis, V., and Basten. (1981). *Clin. Exp. Immunol.* **45,** 9–17.
Turpie, A. G., Chernesky, M. A., Larke, R. P. B., Packham, M., and Mustard, J. F. (1973). *Lab. Invest.* **28,** 575–583.
van Leeuwen, E. F., von dem Borne, A. E. G., van der Plas-Van Dalen, C., and Engelfriet, C. P. (1981a). *Scand. J. Haematol.* **26,** 285–291.

van Leeuwen, E. F., von dem Borne, A. E. G. Kr., von Riesz, L. E., Nijenhuis, L. E., and Engelfriet, C. P. (1981b). *Blood* **57,** 49–54.

van Leeuwen, E. F., van der Ven, J. T. M., Engelfriet, C. P., and von dem Borne, A. E. G. Kr. (1982). *Blood* **59,** 23–26.

Veenhoven, W. A., Kaars Sijpesteijn, J. A., and van der Schans, G. S. (1979). *Acta Haematol.* **62,** 153–158.

Verheyden, C. N., Beart, R. W., Clifton, M. D., and Phyliky, R. L. (1978). *Mayo Clin. Proc.* **53,** 442–446.

Viala, J. J., Dechevanne, M., and Ville, D. (1975). *Lyon Med.* **234,** 419–425.

von dem Borne, A. E. G. Kr., ver Heugt, F. W. A., Oosterhof, F., von Riesz, E., Brutel de la Rivigre, A., and Engelfriet, C. P. (1978). *Br. J. Haematol.* **39,** 195–207.

von dem Borne, A. E. G. Kr., Helmerhorst, F. M., van Leeuwen, E. F., Pegels, H. G., von Ries, Z. E., and Engelfriet, C. P. (1980). *Br. J. Haematol.* **45,** 319–327.

Ware, R., Kinney, T. R., Friedman, H., Falleta, J., and Rosse, W. F. (1982). *Blood* **60** (Suppl. 1), 192A.

Watkins, S. P., Cowan, D. H., and Shulman, N. R. (1967). *J. Clin. Invest.* **46,** 1129–1130.

Wautier, J. L., Boizard, B., Wantier, M. P., Kadeva, H., and Caen, J. (1980). *Nouv. Rev. Fr. Hematol.* **22,** 29–36.

Weiss, H. J., Rosove, M. H., Lages, B. A., and Kaplin, K. L. (1980). *Am. J. Med.* **69,** 711–717.

Westphal, E., and Muller-Ruchholtz, W. (1979). *Eur. J. Pediatr.* **131,** 81–83.

White, L. A., Brubaker, L. H., Aster, R. H., Henry, P. H., and Adelstein, E. H. (1978). *Am. J. Hematol.* **4,** 313–323.

Woerner, S. J., Abildgaard, C. F., and French, B. N. (1981). *Pediatrics* **67,** 453–460.

Woods, V. L., Kurata, Y., Montgomery, R. R., Tani, P., Mason, D., Oh, E. H., and McMillan, R. (1984). *Blood* **64,** 156–163.

Wybran, J., and Fudenberg, H. H. (1972). *Blood* **40,** 856–861.

Yeager, A. M., and Zinkham, W. H. (1980). *Johns Hopkins Med. J.* **146,** 270–274.

Zerella, J. T., Martin, L. W., and Lampkin, B. C. (1978). *J. Pediatr. Surg.* **13,** 243–246.

Zinberg, M., Francus, T., Weksler, M. E., Siskind, G. W., and Karpatkin, S. (1982). *Blood* **59,** 148–151.

Zucker-Franklin, D., and Karpatkin, S. (1977). *N. Engl. J. Med.* **297,** 517–523.

CHAPTER 20

Autoimmune Neutropenia

PARVIZ LALEZARI
Division of Immunohematology
Department of Medicine
Montefiore Medical Center
and Albert Einstein College of Medicine
New York, New York

I. INTRODUCTION

Immunologically induced neutropenias are probably as common as the more familiar autoimmune disorders of red cells and platelets. Investigation and characterization of these disorders were delayed because of difficulties in the separation of neutrophils, technical problems in the detection of their antibodies, and inability to discriminate between autoantibodies and those produced by alloimmunization.

The blood level of leukocytes, similar to that of red cells and platelets, represents a balance between the rate of production and the rate at which the cells are destroyed or consumed. Leukocytes, however, are heterogeneous cells each with separate proliferative characteristics, functions, and life spans. These differences result in distinguishable pathophysiological features in each cell type and call for the use of a proper terminology.

Leukopenia, granulocytopenia, and neutropenia should no longer be used interchangeably. Autoimmune leukopenia would probably be an appropriate term for instances in which granulocytes, lymphocytes, and monocytes are affected collectively. Granulocytopenia should be reserved for cases with documented reduction of all granulocytes, and neutropenia, the subject on which we concentrate, should be used when neutrophils are the only affected cells.

II. PATHOPHYSIOLOGY AND MECHANISM OF IMMUNE DESTRUCTION OF NEUTROPHILS

A. Neutrophil Antibodies and Their Targets

Association of neutropenia with "leukocyte" antibodies is long recognized (Dausset, 1956), and the ability of leukocyte antibodies in man and xenogeneic leukoagglutinins in experimental animals to cause neutropenia is well documented. Among the early studies reviewed by Walford (1960), those performed by Kissmeyer-Nielsen (1954), Miescher (1954), and Butler (1958) are of considerable importance as they show that the infusion of sera of neutropenic patients into normal subjects reduces the neutrophil counts. The work of Moeschlin and Wagner (1952) called attention to the immunological mechanism in the aminopyrine-induced leukopenia; later, observation in this laboratory that fetal–maternal neutrophil incompatibility could lead to neonatal neutropenia contributed to the general acceptance of the immune neutropenias as true clinical entities. Further studies in neonatal neutropenia led to discovery of several antigens specific for blood neutrophils (Lalezari, 1977).

An intriguing observation was that alloantigens involved in neonatal neutropenia could also be the target of autoantibodies found in some neutropenic children and adults (Lalezari, 1977), and in fact new neutrophil "alloantigens" were identified by sera of patients who had autoimmune neutropenia (Verheugt *et al.*, 1978; Claas *et al.*, 1979). These observations indicated that the neutrophil antigens recognized in the allogeneic reactions are also heeded by the individual's own immunoregulatory system. It is the alteration of this regulatory mechanism that probably is the main cause of autoimmunity against the neutrophils. In this respect, neutrophil antigens appear to behave like the Rh determinants that often are the targets in autoimmune hemolytic anemias. A brief review of neutrophil antigens further demonstrates their intimate involvement in autoimmune neutropenias and illustrates that hematological features can be pre-

TABLE I
Distribution Patterns of Various Neutrophil Antigens

Antigens expressed only on myeloid cells (Tissue-specific antigens)	
Neutrophil-specific antigens (N groups)	NA, NB, NC, ND, NE, HGA-3
Granulocyte antigens (G Groups expressed on neutrophils, eosinophils, and basophils)	Antigens of cold-reactive antibodies
Granulocyte–monocyte antigens (GM Groups)	HGA-1
Granulocyte–Endothelial–monocyte (GEM Groups)	
Antigens with wide tissue distribution (Systemic antigens)	HLA, Ii, Group 5

dicted in various forms of the disease according to the properties of the target antigens. For example, autoimmune neutropenia of infancy and alloimmune neonatal neutropenia, two etiologically separate entities in which the target antigens appear to be the same, have identical clinical and hematological manifestations.

Neutrophil antigens thus far identified have been classified (Lalezari, 1983) on the basis of the patterns of their distribution on various cell types, on the basis of their expression on various stages of cell maturation, and according to the optimal temperatures required for their *in vitro* (and presumably *in vivo*) reactivity. A summary of neutrophil antigens characterized based on their distribution pattern is given in Table I. These antigens are divided into those found only on the myeloid cells (tissue-specific) and those with wide tissue distribution (systemic). Among the myeloid cell antigens, the neutrophil (N) groups have been shown to be the target of both alloimmune and autoimmune neutropenias. These neutrophil-specific antigens are distinct from the G groups expressed on granulocytes (neutrophils, eosinophils, and basophils), from the GM antigens shared between granulocytes and monocytes, and from GEM antigens found on granulocytes, monocytes, and endothelial cells.

In Table II, leukocyte antigens are divided on the basis of the relationship between the expression of antigens and the stage of cell maturation. Some antigens are expressed on both the mature and immature cells, and their concentration appears to remain constant in the course of differentiation. These antigens may represent the structures that constitute the "building blocks" of the cell membrane. The HLA belong to the group in which the antigen concentration decreases as the cells differentiate. Curi-

TABLE II
Relationship between Antigen Expression on Leukocytes and Cell Maturation

Antigen types	Examples
Expressed on both mature and immature cells	Group 5, HGA-1
Expressed mainly on immature cells capable of division	HLA System
Develops on cell maturation	NA,NB,NC,HGA-3

ously, it appears that HLA antigens, particularly those controlled by the *HLA-D* region, completely disappear from the cell surfaces when the cells lose capacity to divide. These observations suggest that the HLA system probably is involved in the regulation of cell division and differentiation and is lost when these functions are completed. Antigens developed exclusively on mature cells are probably related to the structures involved in the specialized functions of these cells. Table III calls attention to the fact that some leukocyte antibodies react best at physiological temperatures, some are cold reactive, and some have temperature requirements that are intermediary.

The class of antigens exclusively expressed on neutrophils is likely to be the target of autoimmunity when neutropenia is the sole hematological abnormality. In these cases, the eosinophils and basophils are spared and even may be increased, together with monocytes, because of a myeloid proliferative response. The exclusive expression of these antigens on the mature cells is the reason for the characteristic bone marrow findings, namely, the overabundance of the precursors and the selective decrease in the mature polymorphonuclear neutrophils. Antibodies to neutrophil-specific antigens are detected best at physiological temperatures. As these antigens are presumed to be associated with specialized neutrophil functions, their involvement in autoimmunity may result in a dysfunction of the cells that may escape cell destruction.

The true agranulocytosis occurs when the target antigens are expressed on all granulocytes. Unfortunately, in many reported cases the serological and hematological evaluations have not been carried out with the specific question of the possible effects of antibodies on various granulocytes in mind. The bone marrow myeloid hypoplasia seen in some patients with agranulocytosis suggests that in these cases the target antigens are either expressed on precursors alone, preventing generation of the immature cells or belong to the category of the structural antigens that are ex-

TABLE III
Optimal Temperatures for the Reaction of Various Leukocyte Antibodies

Warm (30°–37°C)	NA, NB, NC, HGA-3
Intermediary (18°–25°C)	ND1
Cold-reactive (4°C)	Ii; Cold granulocyte agglutinins, cold granulocytotoxins

pressed on both the mature and immature cells (Table II). The potential role of the HLA-type antigens as a target for autoimmune disorders has been suggested by Roodman *et al.* (1980) but has not been adequately explored. Because of the expression of HLA primarily on the dividing cells, their autoimmune disorder if it occurs would probably not only be manifested by a reduction in the bone marrow cells, but may also result in diminished lymphoid cell proliferation. The patient studied by Cline *et al.* (1976) may represent such an example.

The toxicity and systemic manifestations observed in some severely leukopenic patients have been generally attributed to infection. The possibility should be considered, however, that in some of these patients the target antigens may not be limited to the myeloid cell lines or to the hematopoietic tissue but may have a wide tissue distribution. In such patients the disease would affect multiple organs. Supportive data for these speculations may become available when investigators attempt to identify the target antigens in individual clinical examples and try to relate the clinical manifestations with the profiles of the involved antigens.

In characterization of the specificities of autoantibodies, a word of caution is in order: Studies carried out on the specificities of red cell autoantibodies by the sensitive AutoAnalyzer technology have revealed subtle differences between the Rh-related antigens recognized by the alloantibodies and those that react with autoantibodies (Lalezari and Berens, 1977), the autoantibodies reacting with determinants that are not as restricted as those defined by alloantibodies. Similar differences may exist in neutrophil specifities that may not be recognized by the techniques currently employed.

B. Neutrophils and Immune Complexes

Receptors are known to exist on human polymorphonuclear neutrophils for IgG, IgA, C3b, C3d, C4, and chemotactic factors. Neutrophils

readily react with immune complexes *in vitro*, and many disorders attributed to the immune complexes are associated with neutropenia (Starkebaum *et al.*, 1980; Camussi *et al.*, 1981). The possibility has been considered that in some cases neutropenia may be the only obvious manifestation of an immune complex disease. Protracted infection, often associated with chronic neutropenias, is known to generate immune complexes. In such complicated cases, therefore, it may become difficult to ascertain the secondary nature of the immune complexes.

C. Mechanism of Immune Destruction of Neutrophils

The mechanism of autoimmune destruction of the myeloid cells is not fully understood. Based on the available information on other cells, and in accordance with the conclusions drawn from cases studied by various investigators, multiple mechanisms appear to be involved. These mechanisms may operate alone or in concert and include destruction by humoral factors; phagocytosis, antibody-dependent lymphocyte-mediated cytotoxicity (ADCC), and antibody-independent direct cell-mediated cytotoxicity (or suppression). An additional mechanism may be entrapment of the agglutinated cells in the capillaries. Finally, immunological injury to the myeloid cell may cause a functional alteration rather than the destruction of the cell itself. At the precursor level, such a possible effect would lead to the blocking of cell proliferation or maturation, and at the peripheral level, neutrophils would be numerically normal but functionally impaired.

1. Humoral Mechanisms of Neutrophil Destruction

In an analogy to intravascular lysis in hemolytic anemia, this mechanism would involve the cooperation of antineutrophil antibodies and complement system, leading to formation of the "membrane attack complex." Complement-mediated granulocytotoxicity can be shown by the various modifications of the basic dye exclusion technique or by the release of a ^{51}Cr radiolabel. Autoantibodies with cytotoxicity have been described in the sera of neutropenic patients (Drew and Terasaki, 1978; Blaschke *et al.*, 1979). The actual intravascular lysis of neutrophils, however, has not been documented in any of these cases. In one example of chronic neutropenia (Markenson *et al.*, 1975) that proved to be fatal, we detected an IgM autoantibody with a cytotoxicity titer in excess of 1×10^{-6}. Although IgG autoantibodies were also present and could explain the neutropenia, the consideration of intravascular lysis of neutrophils by

this unusual cytotoxic antibody appeared inevitable. Unfortunately, methods for direct demonstration of *in vivo* neutrophil lysis are not available, and measurement of the release of neutrophil products that may reflect *in vivo* cell lysis (like hemoglobin in erythrocytes) has been limited to the studies on lysozymes (Boxer *et al.*, 1975). Preliminary studies with Drs. Herbert and Jacob suggested that serum transcobalamin I and III, largely products of specific neutrophil granules (Herbert and Colman, 1980), are either normal or elevated in autoimmune neutropenias, but are low in neutropenias due to inadequate cell production (unpublished data).

2. Neutrophil Phagocytosis

Opsonized neutrophils are more likely to be cleared by the phagocytic system, as illustrated by the case reported by Blaschke *et al.* (1979). Leukophagocytosis has been considered a major means of cell destruction, even in xenogeneic systems when rabbit antineutrophil antibodies were infused into guinea pigs (Simpson and Ross, 1971). In *in vitro* experiments, Boxer and Stossel (1974) showed that neutrophils coated by allogeneic or autologous antineutrophil antibodies can stimulate the phagocytic cells and provoke a measurable metabolic response in the absence of complement. The leukophagocytosis seen in the spleens of some cases of autoimmune neutropenia indicates that this organ is one location for the clearance of opsonized neutrophils. Data similar to those available for the determination of red cell survival time and the site of their clearance can now be obtained for neutrophils by the use of ^{111}Indium-radiolabeling techniques (Weiblen *et al.*, 1979). It may be shown by this technology that in most cases of autoimmune neutropenia, a selective site for the clearance of neutrophils cannot be demonstrated. Rather, neutrophils are likely to be removed by phagocytic cells widely distributed in many tissues.

3. Antibody-Dependent Lymphocyte-Mediated Cytotoxicity (ADCC)

In this mechanism, a group of lymphocytes equipped with Fc-γ-receptors (killer cells) recognize antibody-coated cells and cause their destruction. Logue *et al.* (1978) examined this mechanism and showed ADCC in two neutropenic patients who also had a connective tissue disorder. In two neutropenic patients reported by Bom-van Noorloos *et al.* (1980), the patient had a T-cell lymphoproliferative disease, and the abnormal cells had the killer but not the suppressor activities. Neutrophil-bound Ig were shown by immunofluorescence test in both cases.

4. Cell-Mediated Cytotoxicity, Bone Marrow Suppression, and Neutropenia

Immune suppression of hematopoiesis has been considered a cause of aplastic anemia (Cline and Golde, 1978), and inhibition of the growth of CFU-C colonies by lymphocytes of neutropenic patients has been demonstrated in Felty's syndrome (Abdou *et al.*, 1978), rheumatic disorders (Bagby and Gabourel, 1979), systemic lupus (Kagan, 1979), and autoimmune nephropathy (Roodman *et al.*, 1980). The mechanism of this suppression is not entirely clear. Although frequent association with SLE and other autoimmune disorders is suggestive of an immunologically induced effect, direct cytotoxicity and bone marrow cell destruction by the patients lymphocytes have not been demonstrated. A possibility remains that at least in some of these patients, the bone marrow inhibition is not caused by cytotoxicity but rather by an imbalance in the production of lymphocyte products that appear to modulate granulopoiesis (Barr and Stevens, 1982).

5. Antibody-Induced Leukoembolization and Leukostasis, a Possible Mechanism for Neutrophil Destruction

Neutrophils are agglutinated by activated complement ($C5^a$) *in vitro* (Craddock *et al.*, 1977) and *in vivo* (Hammerschmidt *et al.*, 1978). Moreover, neutrophils exposed to activated complement have been shown to adhere to endothelial cells and cause a vascular injury (Jacob *et al.*, 1980). Microleukoembolization resulting from *in vivo* complement-dependent leukoagglutination has been considered to be the cause of postdialysis neutropenia, and a contributing factor in shock–lung syndrome, in retinopathy associated with acute pancreatitis (Jacob *et al.*, 1980), and in the pulmonary transfusion reaction that occurs following accidental infusion of antileukocyte antibodies (Thompson *et al.*, 1971).

Agglutination is a common response when neutrophils react with their specific antibodies in vitro. This reaction has been found to be an active process akin to neutrophils response to chemotactic stimuli (Lalezari, 1977). Although *in vivo* antibody-mediated neutrophil agglutination has not been directly demonstrated, the occurrence of a pulmonary transfusion reaction due to anti-NA2, a neutrophil-specific antibody (Yomtubian *et al.*, 1982), clearly indicates that this reaction can occur *in vivo*. In autoimmune neutropenia, however, the reaction would probably be a continuous and slow process without discernible clinical manifestations. Some neutrophil antibodies are known to activate the complement system and thereby may initiate a complement-mediated neutrophil aggregation

as well. A similar mechanism may be involved in the immune complex disease-related neutropenias in which antibodies are not directed against any specific antigens on neutrophils.

6. Antibody-Induced Neutrophil Dysfunction

Neutrophil functions include motility, response to chemotactic stimuli, adherence, phagocytosis, and bacteriocidal effects by means of generation of toxic radicals and the release of various granular enzymes. Many of these functions are initiated at the cell surface through specific receptor–ligand interactions. Data obtained in this laboratory have shown that the Fc and complement receptors on neutrophils can be blocked by the use of xenogeneic antineutrophil antibodies. Monoclonal neutrophil antibodies have been produced that inhibit lysosomal enzyme release by chemotactic factors without blocking their binding to the cell membrane (Cotter *et al.*, 1981). Although similar effects have not as yet been demonstrated for the allogeneic neutrophil antibodies, clinical observations point to the possibility that autoantibodies may compromise selective neutrophil functions without causing their destruction. The report by Kramer *et al.* (1980) exemplifies the inhibition of neutrophil motility by an IgG autoantibody, causing a nonneutropenic clinical condition indistinguishable from chronic neutropenic states. Autoimmune neutrophil dysfunction and neutropenia may coexist. This is indicated indirectly by the frequent observation in the neutropenic patients that the severity of clinical manifestations is not proportional to the number of neutrophils.

III. LABORATORY DIAGNOSIS

Techniques available for demonstration of autoimmunity against the myeloid cells include leukocyte agglutination (Lalezari, 1977), antiglobulin consumption (Engelfriet and van Loghem, 1961), cytotoxicity (Blaschke *et al.*, 1979; Drew and Terasaki, 1978), opsonization and inhibition of phagocytosis (Boxer and Stossel, 1974), immunofluorescence (Verheught *et al.*, 1978), quantitation of neutrophil-bound Ig by radiolabeling techniques (Lightsey *et al.*, 1977; Logue and Shimm, 1980), and by staph protein A (Stossel, 1981). Methods recently developed for the testing of cell-mediated immunity should now be included in these procedures.

In autoimmune neutropenia, as in other autoimmune diseases, examination of the target cells (direct tests) should be most informative. The reduced number of neutrophils in neutropenic patients, however, often

limits application of most of these techniques to the less optimal indirect methods. The exception is in the immunofluorescence test (Verheugt *et al.*, 1978), with which antibodies can be demonstrated on the single cells that are often possible to isolate from most neutropenic patients. In addition, by the use of appropriate FITC-labeled antiglobulin reagents it is possible to determine the class (IgG, IgM, IgA) and subclasses of the neutrophil-bound immunoglobulins.

Despite technical advances, the tests developed for the study of autoimmune neutropenias have many shortcomings. In the immunofluorescence test, neutrophils should be differentiated from the frequently present normal monocytes that regularly give a strong reaction with FITC-labeled anti-human IgG. Antigen modulation, experienced by other investigators (Weitzman *et al.*, 1979), has not been an obstacle in this test. Among the indirect tests, leukocyte agglutination in the absence of EDTA does not give reliable and reproducible results (Walford, 1960; Stossel, 1981). It is not generally recognized, however, that the addition of EDTA to the reaction mixture inhibits the nonspecific clumping of neutrophils and at the same time enhances antibody-dependent leukoagglutination. Unfortunately, EDTA-dependent leukoagglutination is more useful for the detection of alloantibodies and gives a positive reaction only in more severe cases of autoimmune neutropenias.

Many normal, nonneutropenic patients have been found in this laboratory to have circulating cytotoxic antibodies. Therefore, we consider these antibodies clinically significant only when they are detected in high titers. Enthusiasm for the tests based on opsonization and inhibition of phagocytosis is tempered by their complexity. Moreover, examples of strong agglutinating anibodies have been found in severe cases of alloimmune neonatal neutropenia that have failed to cause neutrophil opsonization. The quantitative Ig assays also have limitations created by the nonspecific presence of a large amount of Ig molecules on normal neutrophils and by potential difficulties in differentiating between specific antibodies and immune complexes. These technical problems indicate the need for the use of multiple techniques and further methodological developments.

IV. MANIFESTATIONS AND CLINICAL VARIATIONS

Autoimmune neutropenia in general has a chronic course and is manifested by various types of persistent infectious complications. Acute "agranulocytosis," by contrast, is often drug induced. Following are the most frequent variations of autoimmune neutropenia recognized.

A. Idiopathic Autoimmune Neutropenia

Clinical manifestations of idiopathic autoimmune neutropenias in adolescent and adult patients encompass a wide range, from absence of symptoms to devastating forms complicated by uncontrollable sepsis. In the majority of these cases, the patients develop intermittent cutaneous infections and stomatitis. Splenomegaly is an exception rather than the rule. The hematological profile reveals selective absence of peripheral blood neutrophils; monocytes, eosinophils, and basophils are present or increased. The bone marrow typically shows increased myeloid cellularity with diminished number of mature cells. Diagnosis if often made by demonstrating cell-bound or circulating antibodies.

B. Autoimmune Neutropenia of Infancy

This syndrome appears to be an unusually common disorder among young infants. More than 150 examples with basically identical clinical features have been studied in our laboratory since the first report (Lalezari *et al.*, 1975). The disease is characterized by a severe neutropenia that is diagnosed when the child is ~5–7 months old. Neutropenia is often associated with various forms of mild bacterial or fungal infections, and spontaneous recovery within 1 to 4 years is the rule. Diagnosis is made by demonstrating neutrophil-bound and less frequently circulating antibodies in the patients' blood. The absence of antineutrophil antibodies in the maternal sera reveals the autoimmune nature of the disease. The antibodies may be demonstrated to have specificities, similar to those found in alloimmune neonatal neutropenia in the younger infants. The cause of this disorder remains unclear. The disease may represent a "physiological autoimmune" state in which the autologous antigens (or those orginating from the mother but also present on fetal cells) provoke an immune response in an immune system in which the normal regulatory mechanisms have not yet become fully operative. The autoimmune process ceases, however, when this regulatory mechanism reaches maturity. The validity of this hypothesis is now being tested in this laboratory.

C. Secondary Autoimmune Neutropenias

1. Autoimmune Neutropenia Associated with Autoimmune Hemolytic Anemia and Thrombocytopenia

Evans' syndrome, a combination of autoimmune hemolytic anemia and thrombocytopenia, is well recognized (Evans *et al.*, 1951). A special com-

bination designated "alternating autoimmune hemocytopenia" has been observed in this laboratory on several occasions. The disease is characterized by its multiphasic manifestation. In the initial phase, the disease is limited to one of the hematological cells and responds to treatment. After a free interval, cell destruction recurs but now involves another cell type. Alternation between hemolysis, thrombocytopenia, and neutropenia continues over several years, and eventually a pancytopenia that is resistent to treatment prevails.

2. Autoimmune Neutropenias Associated with Other Autoimmune Diseases

Association of neutropenia with systemic lupus, Felty's syndrome, Sjogren's syndrome, Graves' disease, and lymphoproliferation disorders have been well documented, and we have observed cases associated with hairy cell leukemia and sarcoidosis.

V. MANAGEMENT OF AUTOIMMUNE NEUTROPENIAS

Many neutropenic patients, especially children, tolerate the disease well and require antibiotics only intermittently for the management of complicating infections. Splenectomy and steroid therapy are not as successful as in autoimmune hemolytic anemia and thrombocytopenia. The reason for these differences is not understood and may reflect pathophysiological features unique to neutrophils, especially differences in the mechanism of cell destruction. It may be reasoned that patients in whom neutrophil destruction is primarily mediated by leukoembolization, ADCC, or intravascular neutrophil lysis are not likely to benefit from splenectomy or steroids. Moreover, monocyte production is almost invariably increased in autoimmune neutropenias, and chronic infection probably serves as a continuous stimulus for their activation.

The activated macrophages are known to be more destructive and less susceptible to the inhibition by steroids. In practice, many adult patients with symptomatic neutropenias are splenectomized and receive multiple courses of steroids, which should be given only intermittenly. In several cases, plasmapheresis has been attempted with doubtful results, but transfusion of compatible neutrophils appears to be helpful in the control of sepsis. For example, transfusion of NA2-negative neutrophils was effective in the management of pulmonary and brain abscesses in a 12-year-old boy who had an autoimmune neutropenia due to anti-NA2. Unfortunately, after multiple transfusions the patient developed resistance and

eventually succumbed to infection. Large doses of gamma globulin administered intravenously have been shown to reverse the neutropenia temporarily (Pollack *et al.*, 1982; Bussel *et al.*, 1983). This treatment should be considered in the complicated cases, particularly in autoimmune neutropenia of infancy.

VI. SUMMARY

Autoimmune neutropenia (AIN), selective destruction of blood neutrophils by immunological mechanism, is being recognized with increasing frequency. In a large number of these patients, the antibody targets are antigens expressed only on the mature cells. This explains the increased number of myeloid precursors, and the absence or the reduction of the more mature cells in the bone marrow. In some cases, autoimmunity may be directed against antigens with wider cell distribution, or against antigens also expressed on precursor cells. In such cases other cell types (lymphocytes, monocytes, and eosinophils) would be reduced, and the bone marrow may be hypoplastic.

The mechanism of cell destruction in these disorders has not been fully clarified. Leukophagocytosis, cell lysis, and ADCC- and cell-mediated cytotoxicity appear to occur. An additional factor may prove to be microembolization of aggregated mature neutrophils. Antibody-induced neutrophil dysfunction such as blocking of the Fc and complement receptors and inactivation of surface enzymes may be an explanation for disproportionately severe clinical manifestations occasionally observed.

The techniques available for diagnosis of AIN need improvement. Nevertheless, diagnosis can be established by a combination of the agglutination and the direct immunofluorescence tests in most cases, especially in a newly recognized form that has been termed autoimmune neutropenia of infancy. Most patients with AIN require treatment only for intercurrent infections and complications. Steroids should not be used indiscriminately. Splenectomy is helpful only in individual cases. Neutrophil transfusion and intravenous infusion of gamma globulin are indicated in selected patients.

REFERENCES

Abdou, N. I. *et al.* (1978). *J. Clin. Invest.* **61,** 738–743.
Bagby, G. G., and Gabourel, J. D. (1979). *J. Clin. Invest.* **64,** 72–82.
Barr, R. D., and Stevens, C. A. (1982). *Am. J. Hematol.* **12,** 323–326.
Blaschke, J. *et al.* (1979). *Am. J. Med.* **66,** 862–866.
Bom-van Noorloos, A. A. *et al.* (1980). *N. Engl. J. Med.* **302,** 933–937.

Boxer, L. A., and Stossel, T. P. (1974). *J. Clin. Invest.* **53,** 1534–1545.
Boxer, L. A. *et al.* (1975). *N. Engl. J. Med.* **293,** 748–753.
Bussel, J. *et al.* (1983). *Blood* **62,** 398–405.
Butler, J. J. (1958). *Am. J. Med.* **24,** 145–152.
Camussi, G. *et al.* (1981). *Afr. J. Clin. Exp. Immunol.* **2,** 311–317.
Claas, F. H. J. *et al.* (1979). *Tissue Antigens* **13,** 129–134.
Cline, M. J., and Golde, D. W. (1978). *Am. J. Med.* **64,** 301–310.
Cline, M. J. *et al.* (1976). *N. Engl. J. Med.* **295,** 1489–1493.
Cotter, T. G. *et al.* (1981). *J. Immunol.* **127,** 1355–1360.
Craddock, P. R. *et al.* (1977). *J. Clin. Invest* **60,** 260–264.
Dausset, J. (1956). *In* "Immuno-Hematologie Biologique et Clinique" pp. 641–670. Editions Medicales Fammarion, Paris.
Drew, S. I., and Terasaki, P. I. (1978). *Blood* **52,** 941–952.
Engelfriet, C. P., and van Loghem, J. J. (1961). *Br. J. Haematol.* **7,** 223–238.
Evans, R. S. *et al.* (1951). *Arch. Intern. Med.* **87,** 48–65.
Hammerschmidt, D. E. *et al.* (1978). *Blood* **52** (Suppl. 1), 125.
Herbert, V., and Colman, N. (1980). *In* "Lithium Effects on Granulopoiesis and Immune Function" (A. H. Rossoff and W. A. Robinson, eds.), pp. 61–78. Plenum, New York.
Jacob, H. S. *et al.* (1980). *N. Engl. J. Med.* **302,** 789–794.
Kagan, W. A. (1979). *Blood* **54,** (Suppl. 1), 155a.
Kissmeyer-Nielsen, F. (1954). *Acta Med. Scand.* **150,** 349–353.
Kramer, N. *et al.* (1980). *N. Engl. J. Med.* **303,** 1253–1258.
Lalezari, P. (1977). *In* "The Granulocyte: Function and Clinical Utilization" (T. J. Greenwalt and G. A. Jamieson, eds.), pp. 209–225. Alan R. Liss, New York.
Lalezari, P. (1983). *In* "Immunohaematology" (C. P. Engelfriet, A. E. G. Kr. von dem Borne, and J. J. van Loghem, eds.), pp. 178–187. Elsevier Sci. Publ., Amsterdam.
Lalezari, P., and Berens, J. A. (1977). *Hum. Blood Groups, Proc. Int. Convoc. Immunol., 5th, 1976* pp. 44–55.
Lalezari, P. *et al.* (1975). *N. Engl. J. Med.* **293,** 744–747.
Lightsey, A. L. *et al.* (1977). *Annu. Intern. Med.* **86,** 60–62.
Logue, G. L., and Shimm, D. S. (1980). *Annu. Rev. Med.* **31,** 191–200.
Logue, G. L. *et al.* (1978). *Blood* **51,** 97–108.
Markenson, A. L. *et al.* (1975). *Proc. 18th Congr. Am. Soc. Hematol., Dallas,* p. 174.
Miescher, P. A. (1954). *Acta Haematol.* **11,** 152–167.
Moeschlin, S., and Wagner, K. (1952). *Acta Haematol.* **8,** 29–41.
Pollack, S. (1982). *N. Engl. J. Med.* **307,** 253.
Roodman, G. D. (1980). *Am. J. Med.* **69,** 325–328.
Simpson, D. M., and Ross, R. (1971). *Am. J. Pathol.* **66,** 79–102.
Starkebaum, G. (1980). *J. Lab. Clin. Med.* **96,** 238–251.
Stossel, T. P. (1981). *In* "A Seminar on Immune-Mediated Cell Destruction" C. A. Bell, ed.), pp. 199–208. Am. Assoc. Blood Banks.
Thompson, J. S. *et al.* (1971). *N. Engl. J. Med.* **284,** 1120–1125.
Verheugt, F. W. A. *et al.* (1978). *Vox Sang.* **35,** 13–17.
Walford, R. L. (1960). *In* "Leukocyte Antigens and Antibodies" pp. 1–182. Grune & Stratton, New York.
Weiblen, B. J. *et al.* (1979). *J. Lab. Clin. Med.* **94,** 246–255.
Weitzman, S. A. (1979). *J. Clin. Invest.* **64,** 321–325.
Yomtubian, R. *et al.* (1982). *Transfusion (Philadelphia)* **22,** 426.

CHAPTER 21

Sperm and Testicular Autoimmunity

KENNETH S. K. TUNG
Department of Pathology
University of New Mexico School of Medicine
Albuquerque, New Mexico

ALAN C. MENGE
Department of Obstetrics and Gynecology
The University of Michigan Medical Center
Ann Arbor, Michigan

I. INTRODUCTION

Spermatozoa (abbreviated "sperm") was one of the first tissue antigens known to be immunogenic to the autologous host (Landsteiner, 1899; Metchnikoff, 1899; Metalnikoff, 1900). The several decades following this discovery encompassed efforts aimed at utilizing sperm antigens in antifertility vaccines (see review by Katsh, 1959), a revitalized subject (Anderson and Alexander, 1983). The study of testicular autoimmunity was launched in the early 1950s when an autoimmune disease of the testis, experimental allergic orchitis (EAO), was induced by immunization of

guinea pigs with homologous sperm or testicular antigens in complete Freund's adjuvant (CFA) (Voisin *et al.*, 1951; Freund *et al.*, 1953, 1955). More recently, increasing circumstantial evidence has indicated that autoimmunity to sperm may be a cause of infertility in men, and that immunity to sperm also results in infertility in women (reviewed in Jones *et al.*, 1975; Shulman, 1975; Menge, 1980). Despite this early beginning, the possible pathological consequences and clinical relevance of sperm and testicular autoimmunity remain poorly defined and the pathogenetic mechanisms of EAO incompletely understood. In this chapter, we will review in detail the models of autoimmune orchitis as examples of organ-specific autoimmune disease. In addition, we will examine the clinical and experimental evidence for sperm autoimmunity and alloimmunity as possible causes of human infertility.

II. AUTOIMMUNE DISEASES OF THE TESTIS

A classification of epididymo-orchitis and aspermatogenesis of proven or probable immunological etiology is presented in Table I.

Elucidation of the pathogenesis of testicular autoimmune diseases has been based largely on studies in EAO in the guinea pig. Since murine EAO can now be induced consistently, incisive dissection of this disease will be forthcoming, as examplified by recent immunogenetic analysis of murine EAO. The studies on naturally occurring models of testicular autoimmune diseases have provided evidence for coinheritance of autoimmunity of the gonads and that of other endocrine organs. In this chapter, we will analyze in detail the pathology and pathogenesis of EAO. We will then describe postvasectomy autoimmune orchitis and the naturally occurring models of testicular autoimmune disease. As an introduction to these topics, a consideration of the testis as an immunological target organ and the nature of aspermatogenic antigens, as well as the immunoregulatory mechanisms against testicular autoimmunity, will be presented.

A. The Testis as an Immunological Target Organ

It has been estimated that only one-fifth of the theoretical number of progeny of spermatogonia eventually develop into sperm (Huckins, 1978). During spermiation, numerous residual bodies, rich in surface autoantigens, are retained by the germinal epithelium (Fawcett, 1975). The redundant autoantigenic by-products of spermatogenesis and spermiogenesis are removed by the remarkable Sertoli cells, which are able to phagocytose and degrade rapidly. While the blood–testis barrier is well developed

Classification of Epididymo-Orchitis and Aspermatogensis with Proven or Probable Immunological Etiology[a]

I. Experimental allergic orchitis
   A. Active immunization
      1. Sperm or testis antigens with[b,c] or without adjuvant[d]
      2. Chemically modified soluble aspermatogenic antigens[e]
      3. Parotid gland antigens in CFA[f]
   B. Passive immunization
      1. Immune serum to untreated[g] or CFA-injected recipients[h]
      2. Immune lymphoid cells[i]
      3. Cellular extracts from lymphoid cells[j]
II. Naturally occurring orchitis
   A. Beagle dog (A line) with or without autoimmune thyroiditis[k]
   B. T-Locus mouse with t^{w18} haplotype[l]
   C. Infertile dark mink[m]
   D. Rhesus monkey[n]
   E. Man (?)[o]
III. Postvasectomy orchitis
   Guinea pig[p], rabbits,[q], monkey[n]
IV. Physical trauma to the contralateral testis.
   A. Cryoinjury[r].
   B. Heat injury[s].
   C. Physical injury[t].
V. Postinfective orchitis
   Mumps orchitis[u]
VI. Epididymo-orchitis following neonatal thymectomy[v]

[a] Reprinted from Tung (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.
[b] Voisin *et al.* (1951).
[c] Freund *et al.* (1953).
[d] Bishop (1961).
[e] Pokorna and Vojtiskova (1964).
[f] Nagano and Okumura (1973a).
[g] Tung *et al.* (1971a).
[h] Pokorna (1970a).
[i] Tung *et al.* (1971b).
[j] Pokorna (1970b).
[k] Fritz *et al.* (1976).
[l] Dooher *et al.* (1981).
[m] Tung *et al.* (1981b).
[n] Tung and Alexander (1980).
[o] Salomon *et al.* (1982).
[p] Tung, 1978.
[q] Bigazzi *et al.* (1976).
[r] Zappi *et al.* (1974).
[s] Rappaport *et al.* (1969).
[t] Raitsina and Nilovsky (1967).
[u] Mostofi (1977).
[v] Taguchi and Nishizuka (1981).

at the level of the seminiferous tubules, other parts of the sperm compartment are not as impenetrable to macromolecules; there are regional differences in the immunological blood–testis barrier (Tung *et al.*, 1971a; Johnson, 1973).

1. Blood Vessels, Lymphatics, and Interstitial Space

Although the testicular capillaries and venules are not fenestrated, they allow diffusion of IgG (Mancini *et al.*, 1965). The interstitial space is composed of an extensive labyrinth of lymphatics (Fawcett *et al.*, 1969) and islands of blood vessels surrounded by Leydig cells, macrophages, mast cells, and plasma cells (Hermo and Lalli, 1978). Resident macrophages have been isolated from perfused testis (Tung *et al.*, 1979), and it has been stated that the Leydig cell/macrophage ratio is 4 : 1 (Christensen, 1977). Ia molecules are detected on ~40% of resident testicular macrophages (Miller *et al.*, 1984). Circulating lymphocytes reach the testicular interstitium, and from there enter the lymphatics to be recovered by cannula placed in the afferent testicular lymph vessels (Morris, 1968). Thus, humoral antibody and recirculating lymphocytes attendant on a systemic immune response can readily reach the interstitial space of the testis.

2. The Blood–Testis Barrier at the Seminiferous Tubules

By analysis of testicular fluids from rams, Setchell (1967) first directed attention to the existence of a permeability barrier surrounding the seminiferous tubules of the mammalian testis. These studies found that proteins were abundant in plasma and testicular lymph but existed in very low concentration in the rete testis fluid. Soon after this important discovery, the morphological basis of the blood–testis barrier was defined.

The seminiferous tubules in most species are surrounded by a layer of smooth muscle-like myoid cells. These cells act as a partial barrier for large molecules in the rat and guinea pig, but not in primates. Antibody IgG to tubular basement membrane antigens, when injected into guinea pig testis, becomes bound to the basement membrane around the rete where myoid cells are absent but not to the basement membrane surrounding the seminiferous tubules (Tung *et al.*, 1971a). The most significant barrier at the level of the seminiferous tubules is provided by the Sertoli cells and the multiple specialized tight junctions between them. Dym and Fawcett (1970) have divided the germinal epithelium of the seminiferous tubules into two compartments separated by the Sertoli cell barrier: (1) a basal compartment containing the preleptotene spermatocytes and spermatogonia, and (2) an adluminal compartment containing the more mature germ cells. The tight junctions between Sertoli cells have

been revealed by freeze-fracture technique as unique intramembranous strands of particles located in grooves on the E face (Gilula *et al.*, 1976). Forty to fifty parallel strands, separated from one another by ~0.1 μm, are located on the entire circumference of the basal portion of Sertoli cells and provide what is probably the strongest tissue barrier in the body (Neaves, 1977).

3. Differential Permeability of Tissue Barriers Along the Sperm-Containing Compartment

The barrier of the rete testis is more permeable to serum proteins, including IgG, than is the barrier at the seminiferous tubules. Apart from the difference in the tight junctional complexes of the two regions, higher concentration of IgG exists in fluid obtained from the rete as compared with fluid from the seminiferous tubules (Johnson and Setchell, 1968; Koshimies *et al.*, 1971). After antisperm antiserum was injected intravenously into normal guinea pigs, IgG (presumably antibody) was observed to bind to the acrosome of sperm in the rete but not to sperm in the seminiferous tubules (Tung *et al.*, 1971a). The tight junctions of the ductus efferentes are scanty and do not completely surround the circumference of the cell (Suzuki and Nagano, 1978). Rete testis and ductus efferentes are therefore the most likely sites for the entrance of immune reactants into the sperm compartment and for the exit of any soluble sperm antigens. Although lymphoid cells are frequently detected between epithelial cells in the rete testis, ductus efferentes, and the vas (Dym and Romrell, 1975), it is not known whether leukocytes can penetrate the intact blood–testis barrier.

B. Testis-Specific Autoantigens and Aspermatogenic Antigens

Testis-specific autoantigens are antigens in the testis that readily elicit antibodies and/or cell-mediated immune responses in the autologous host, whereas aspermatogenic antigens are testis-specific antigens that have the capacity to induce EAO. Leydig cells, Sertoli cells, and germ cells are the three likely sources of testis-specific autoantigens. Although serum autoantibodies to Leydig cells and Sertoli cells have been detected in human autoimmune diseases (including polyendocrine autoimmunity, amenorrhea, and unexplained infertility) (Irvine *et al.*, 1969), antigens associated with these cells from prepubertal animals do not elicit an autoimmune response in normal experimental animals even when they are incorporated in CFA. In contrast, sperm and their progenitor cells readily induce an autoantibody response and EAO.

1. Testis-Specific Autoantigens Readily Elicit Autoimmune Responses in the Absence of Adjuvant

Testis-specific autoantigens develop at puberty, and they are absent during the developmental stages of the immune system. Immunological unresponsiveness to these antigens might be incomplete, since acquired immunological unresponsiveness is readily induced in the neonatal but not the adult animals. That immunological tolerance to testis-specific autoantigens is indeed incomplete can be demonstrated by the ease with which antibody to sperm and spermatid surface antigens, as well as acrosomal antigens, are elicited in guinea pigs immunized with autologous sperm or testicular cells. Autoantibodies to sperm often follow testicular injuries resulting from obstructed vas deferens (Phadke and Padukone, 1964), vasectomy (Zappi *et al.*, 1974), and testicular biopsy (Hjort *et al.*, 1974).

Immunological foreignness of testis-specific autoantigens is further supported by the study on the age-related incidence of naturally occurring serum antibody to sperm antigens. The age-related incidences of antibodies to self and foreign antibodies are quite different. Autoantibodies such as antinuclear and antithyroid antibodies occur rarely in children and young adults but increase gradually with age (Goodman *et al.*, 1963; Doniach and Roitt, 1964; Cannat and Seligmann, 1965). On the other hand, antibodies to foreign antigens tend to have an early onset, reach a peak before puberty, and gradually decline (Friedberger and Furstenheim, 1929; Thomsen and Kettel, 1929; Lewi, 1968). When a large number of sera from different age groups of both sexes were studied for the age-related prevalence of antisperm antibodies, the pattern matched best with that of antibodies to foreign antigens; prevalence of antinuclear antibodies in the same group of individuals fit the pattern for autoantibodies (Tung *et al.*, 1976). The high incidence of antisperm antibodies in children probably results from stimulation by foreign antigens that cross-react with sperm.

The number of testis-specific autoantigens capable of autoantibody induction is more numerous than those present in other organs. In sera of vasectomized men, antibodies to as many as eight internal antigens in sperm (Tung, 1975, 1976), including nuclear protamine (Samuel *et al.*, 1975), were detectable (Fig. 1). Thus unlike most autoantigens, the immune responses of the body to sperm resemble the responses to foreign antigens.

2. Cells of the Germ Cell Lineage Are the Main or Only Source of Testis-Specific Autoantigens

While adult testis incorporated in CFA readily induced EAO in the guinea pig, Katsh (1960) found that testes from prepubertal animals, defi-

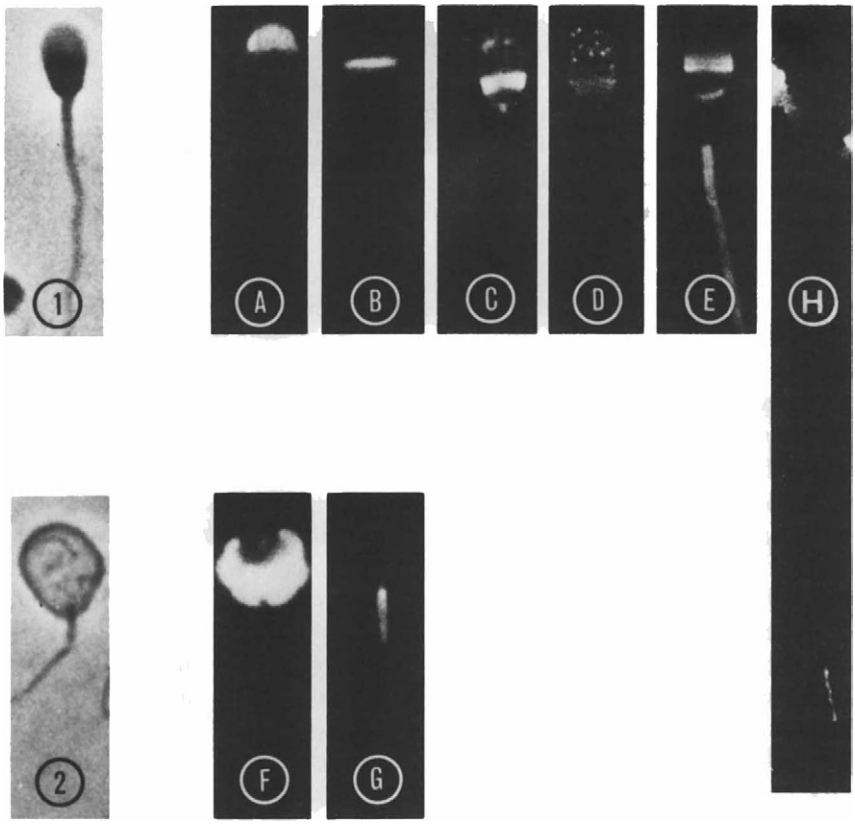


FIG. 1. Antisperm antibodies from vasectomized men detectable by indirect immunofluorescence using (1) methanol-fixed sperm smear, and (2) sperm that have been treated with dithiotreitol and trypsin. Included are antibodies to diffuse acrosomal antigen (A), equatorial antigen (B), postacrosomal antigen (C), speckled acrosomal antigen (D), tail mainpiece antigen (E), tail endpiece antigen (F), nuclear protamine (F), and midpiece antigen (G). Reproduced from Tung (1975), with permission of the publisher.

cient in germ cells beyond the primary spermatocytes, failed to induce testicular disease. Similarly, guinea pigs immunized with testes rendered atrophic from EAO did not develop testicular disease. Prepubertal guinea pigs immunized with adult testis in CFA did not develop EAO until spermatogenesis of the immunized guinea pigs was fully developed (Bishop *et al.*, 1961; Johnson, 1970a). Furthermore, EAO can be readily induced by immunization with epididymal sperm in CFA. When autoantisera from guinea pigs immunized with testis homogenates or epididymal sperm were used to localize testis-specific autoantigen in the guinea pigs, they reacted with the surface of sperm, residual body, and late spermatids (Tung *et al.*,

1979), and with the acrosomal antigens of secondary spermatocytes, spermatids, and sperm. The precise ontogeny of cell-surface testis-specific autoantigen may differ among species. For example, rats autoimmunized with testicular cells or sperm developed antibody that reacted with pachytene spermatocytes as well (Tung and Fritz, 1978). Specifically, Lewis rats, immunized with purified pachytene spermatocytes in CFA, developed aspermatogenesis, orchitis, and antibody to spermatocytes, spermatids, and sperm.

The possible existence of autoantigens on Sertoli cells capable of eliciting antibody and cytotoxic T-lymphocyte response was suggested by the finding that female mice immunized with Sertoli cells from H-2-histocompatible males produced antibodies and cytotoxic T lymphocytes to H-Y antigen and to a second, cell-lineage-specific antigen shared between Sertoli cells and ovarian follicular cells (Ciccareses and Ohno, 1978). Wekerle and his collaborators (Wekerle and Begemann, 1976; Wekerle, 1977, 1978a,b) found that when suspensions of rat lymph node cells or splenic cells were mixed with syngeneic testicular cells, the lymphocytes formed rosettes around a centrally located cell that was described as having the morphology of Sertoli cells. When the rosettes were isolated on a hypaque-ficoll gradient, the rosetting lymphocytes transformed into lymphoblasts and became cytotoxic to monolayer of testicular cells. However, these findings cannot yet be regarded as evidence for the existence of autoantigens on Sertoli cells. The criteria of Sertoli cell identity were insufficiently rigorous. Whether they represented spontaneous rosettes between T cells and macrophages (Seeger and Oppenheim, 1970), or between Leydig cells and lymphocytes (Rivenzon *et al.*, 1974), was not considered. Even if the cells under consideration were Sertoli cells, the surface antigens recognized by the autologous lymphocytes might nevertheless have come from sperm or spermatids that were phagocytosed, processed by the Sertoli cells, and subsequently presented on the Sertoli cell surface. It will be important to determine whether Sertoli cells, isolated free of sperm and spermatids, also rosette with and stimulate autoreactive lymphocytes.

3. Purified Aspermatogenic Antigens of Guinea Pig Testis and Sperm

A number of aspermatogenic proteins and glycoproteins have been purified to apparent homogeneity from the guinea pig testis and caudal epididymal sperm (Table II). These antigens differ from one another in amino acid composition, carbohydrate content, molecular weight, and

TABLE II
Purified Aspermatogenic Antigens from Guinea Pig Testis and Sperm

Antigen	Source	Characteristics	Reference
AP1	Testis	Protein, MW 15,000, pI 8.6; activity disappears after reduction–alkylation	Jackson *et al.* (1975)
GP1	Testis	Glycoprotein, MW 29,000, PI 3.9	Hagopian *et al.* (1975)
GP4	Testis	Glycoprotein, MW 13,000	Hagopian *et al.* (1975)
AP2	Sperm	Protein, MW 9,500, pI 5.5	Teuscher *et al.* (1983a)
AP3	Testis	Protein, MW 12,500, pI 8.9; activity not affected by reduction–alkylation	Teuscher *et al.* (1983b)

isoelectric point. The first purified aspermatogenic antigen of sperm (AP2) was isolated from soluble acrosomal contents released from guinea pig sperm during calcium-dependent, ionophore-induced acrosome reaction (Teuscher *et al.*, 1983a). While AP1 and AP3 have similar molecular weights and pI, they differ in amino acid composition. Moreover, AP1 is inactive after it has been reduced and alkylated (Jackson *et al.*, 1976), whereas similarly treated AP3 remains active (Teuscher *et al.*, 1983b). That the aspermatogenic capacity of AP3 is preserved after reduction–alkylation indicates that the disease-inducing determinant(s) in the AP3 molecule is sequential rather than conformational. If true, it should be possible to identify and synthesize an aspermatogenic peptide.

Although the purified aspermatogenic proteins are highly active, each being capable of EAO induction at or below 1 μg, the antigen yield is uniformly low. Furthermore, the isolation procedures are technically demanding. Hence, immunological studies based on these antigens have not yet been carried out. In the studies to be described herein, several crude but biologically active aspermatogenic antigenic preparations have been employed. Their isolation schemes are outlined in Table III.

C. Immunological Unresponsiveness to Testis-Specific Autoantigens

Having considered the blood–testis barrier and the unusual autoimmunogenicity of testicular-specific autoantigens, we will evaluate the current concept of the mechanisms that may prevent EAO.

TABLE III
Crude Aspermatogenic Preparations Described in This Chapter[a]

Antigenic preparations	Isolation procedures
ASPM (Freund *et al.*, 1955)	GP Testis homogenate extracted in acetic acid; supernatant in 30% ammonium sulfate (AS); precipitate in 70% AS
T (Toullet *et al.*, 1973)	GP Sperm in water; centrifuge (500 *g*); centrifuge the supernatant (80,000 *g*); sonicate and wash the precipitate
S (Toullet *et al.*, 1973)	GP Sperm in water; centrifuge (500 *g*); centrifuge the supernatant after 5% trichloracetic acid (TCA); excluded volumn in Sephadex G-100
P (Toullet *et al.*, 1973)	GP Sperm in water; centrifuge (500 *g*); centrifuge the supernatant (80,000 *g*); precipitate after 5% TCA; aqueous phase after chloroform–butanol; "included" volumn in Sephadex G-100
TCAsup (Jackson *et al.*, 1975)	GP Testis homogenized in chloroform–methanol–acetone; extracted with HCl, pH 3, precipitated in 80% AS; supernatant in 5%
G75m (Teuscher *et al.*, 1983b)	TCAsup; the second peak on Sephadex G 75 (proteins with MW between 13,000 and 50,000)

[a] Reprinted from Tung *et al.* (1981a), by permission of the publisher, Munksgaard, Copenhagen.

1. Antigen Sequestration Theory

This has been the prevailing but unproven theory of immunological unresponsiveness to testis-specific autoantigens, which has been discussed (Shulman, 1975; Beer and Billingham, 1976; Tung, 1977). According to this view, the blood-testis barrier completely separates testis-specific autoantigens from the immune system: The antigens do not reach the lymphoid tissues to stimulate an immune response, nor do immune reactants reach the antigens in the testis to initiate EAO. It follows that EAO occurs only if and when the blood–testis barrier becomes incomplete. This theory probably requires modification for the following reasons.

First, there is no proof that the testis-specific autoantigens are completely sequestered. Second, if testis-specific autoantigens are indeed totally sequestered, it would be difficult to envisage how the disease EAO could adoptively be transferred by systemic injection of sensitized lymphocytes or antisperm antiserum. Under these circumstances, either the immune reactants can cross the blood–testis barrier or testis-specific autoangiens can reach outside the barrier. In this regard, we know now that the rete testis and the ductus efferentes do not provide complete barriers for IgG. It will be important to determine whether soluble testis-specific

autoantigens exist in the lumen of the sperm compartment and whether traces of the antigens are detectable outside the ductus efferentes or in testicular lymph. It will also be important to further define autoantigens on Sertoli cells. Third, the complete antigen sequestration theory, as it stands, excludes other known active control mechanisms against immune response to testis-specific autoantigens; that is, the barrier would not permit any antigen to reach the lymphoid system to stimulate and maintain a negative or suppressive immunoregulatory response.

2. Evidence in Support of Active Immunoregulation against Testicular Autoimmunity

a. The Presence in the Rat of T Lymphocytes That React with Syngeneic Testicular Cells and Can Adoptively Transfer EAO. Incubation of lymph nodes or spleen but not thymus lymphocytes of untreated rats with syngeneic dissociated testicular cells led to the formation of lymphocyte rosettes around a central testicular cell (Wekerle and Begemann, 1976; Wekerle, 1978a,b). After 4–5 days, the lymphocytes in the rosette transformed to lymphoblasts. The lymphoblasts, defined as T cells on the basis of undetectable C3b receptors, were cytotoxic to monolayers of testicular cells, and when adoptively transferred to the testis of syngeneic recipients they induced an inflammatory lesion. Experiments using congenic strains of rats demonstrated that the cytotoxic T lymphocytes and the target testicular cells had to share the major histocompatibility antigens in order for cytotoxicity to occur. As a control for the specificity of the cytotoxicity, lymphocytes that rosetted with testicular cells were found to respond poorly against allogeneic lymphocytes *in vitro* when compared with lymphocytes that did not form rosettes. This study provided evidence for the existence in normal unimmunized animals of precursor effector T lymphocytes against testicular autoantigens.

b. Neonatal Thymectomy Is Associated with Testicular Disease Resembling EAO. Rats thymectomized within the first week after birth were often infertile when they became adults (Hattori and Brandon, 1977; Lipscomb *et al.*, 1979). Infertility was associated with testicular atrophy, interstitial edema, multinucleated giant spermatids within aspermatogenic seminiferous tubules, and empty epididymides. Interstitial cells appeared normal. This finding has now been confirmed and extended in a similar study in the mouse (Taguchi and Nishizuka, 1981). Mice of the (C57B1/6 × A/J)F_1 strain subjected to thymectomy between days 2 and 4 after birth developed severe epididymitis (70–80%) and orchitis (20%) several weeks later. The lesions were accompanied by heavy granular deposits of mouse IgG and C3 that resembled immune complexes along ductal and tubular

basement membrane. The disease could be adoptively transferred to syngeneic, untreated recipients by spleen cells from thymectomized mice. On the other hand, epididymo-orchitis was prevented if the thymectomized animals received splenic cells from normal untreated syngeneic adults within 3–4 weeks of thymectomy.

This interesting phenomenon can best be interpreted as the consequence of removal of the thymus as a central lymphoid organ. Thymectomy at a critical period in life depletes subpopulations of T cells that subserve the immunosuppressive function (i.e., suppressor T cells or their precursors). Unopposed effector T cells that have immigrated from the thymus earlier may spontaneously induce epididymo-orchitis and aspermatogenesis. Further evidence for the existence of suppressor T lymphocytes in the prevention of testicular autoimmunity comes from studies on antigen-mediated prevention of EAO induction.

c. *Antigen-Mediated Suppression of Immune Response to Testis-Specific Autoantigens and of EAO Induction.* That treatment of guinea pigs with testis antigen can lead to unresponsiveness to EAO induction is best illustrated by the so-called immune deviation phenomenon (Asherson, 1967). Guinea pigs that were immunized with testis antigens in saline or in incomplete Freund's adjuvant became unresponsive to subsequent development of EAO when challenged with testis antigens in CFA (Chutna and Rychlikova, 1964; Brown *et al.*, 1967). Chutna (1970) later demonstrated that the first injection of testis antigen in saline abrogated both the complement-dependent, cytotoxic antibody response and the delayed hypersensitivity reaction to sperm antigens but not the anaphylactic antibody response. Furthermore, it was shown that the state of unresponsiveness could be reversed if guinea pigs were given cyclophosphamide at the time of the first injection.

d. *Immunoregulatory Mechanisms within Testis.* The blood–testis barrier at the level of the seminiferous tubule is clearly an important immunological barrier (Johnson, 1973). A study by Hurtenbach *et al.* (1980) suggested that cells within the seminiferous tubules can also impose an immunosuppressive influence through the induction of suppressor T lymphocytes. However, the physiological significance of this mechanism remains to be determined.

D. Experimental Allergic Orchitis

Although EAO has been subjected to more detailed analysis in the guinea pig, some recent studies on EAO in the rat, mouse, and rabbit have added a new dimension to our understanding of this disease (summarized

in Table IV). Certain aspects of EAO in the four species of animals will therefore be described separately, with major emphasis on guinea pig studies.

1. Guinea Pig

a. Disease Induction. Experimental autoimmune orchitis is most reproducibly induced by a single intradermal injection of sperm, testicular homogenate, or their extracts, in CFA. Although CFA is not absolutely required, in its absence multiple large doses of testes homogenate were needed (Bishop, 1961), the animals were deficient in some vitamins (Brown and Glynn, 1973), or the soluble antigen was chemically modified (Pokorna and Vojtkiskova, 1964). The resultant disease is milder and its occurrence less predictable. *Mycobacterium tuberculosis* in the CFA can be replaced by the water-soluble adjuvant fraction isolated from the cell wall of *M. smegmatis* (Toullet *et al.*, 1974), by its synthetic analogs, MDP_1 and MDP_2, or by polynucleotides (Toullet *et al.*, 1977).

b. Immunopathology. The histopathology of EAO is apparent 8 to 9 days after immunization, although earlier lesions have been reported (by day 6) (Brown *et al.*, 1972; Nagano and Okumura, 1973b). The histopathological findings of EAO are complex and consist of any combinations of the four lesions to be described. The temporal and causal relationship between these lesions and their pathogenetic significance remain unclear.

Degenerative Changes of Spermatids and Exfoliation of Germinal Epithelium. Deformation in the acrosomal apparatus and the developing tail complex appear as early as day 6 or 7 after immunization (Brown *et al.*, 1972; Nagano and Okumura, 1973b). Later, exfoliation of spermatids, spermatocytes, and sperm is apparent. Some spermatids fuse and become multinucleated giant cells. Desquamated cells accumulate in the lumen of seminiferous tubules, the rete, and the epididymis, which often are dilated. Perivascular or peritubular infiltration of inflammatory cells is not a consistent feature (Brown *et al.*, 1963). This morphological pattern represents a common testicular response to a multitude of testicular injuries, including irradiation (Vilar, 1971), cytotoxic drugs (Kierszenbaum, 1970), and antifertility drugs (Ericsson, 1971). Therefore, this finding is not diagnostic of EAO.

Mononuclear Infiltrative Lesions in the Testis. An early diagnostic lesion of EAO consists of focal clusters of lymphocytes and macrophages abutting the basement membrane of the seminiferous tubules (Waksman, 1959; Tung *et al.*, 1970). These cells then enter the lumen of the seminiferous tubules through holes in the tubular basement membrane and between

TABLE IV

Studies and Some Findings on Experimental Allergic Orchitis Based on Different Species[a]

Studies	Guinea pig	Rat	Mouse	Rabbit
Disease induction				
Dose (number of antigen)	Small (1)	Multiple	Large (1)	Large (3)
CFA requirement	+	+	+	+
Pertussis requirement	−	+	+	−
In vitro induction of EAO	NS	Yes	NS	NS
Homologous versus heterologous antigen	Homologous	Homologous	Homologous	Homologous and heterologous
Thymic dependency	NS	NS	Yes	NS
By chemically modified antigen	Yes	NS	NS	NS
Immunopathological analysis				
Sequential studies	Yes	No	No	No
Extrapolation to pathogenesis	Yes	Yes	Yes	Yes
Antigen-related pathology	Yes	NS	NS	Yes
Immune complexes in testis	No	No	Yes	Yes
Epididymitis	Yes	No	Yes	No
Vasitis	Yes	No	Yes	No
In vitro study of pathogenesis				
MHC-restricted T-cell cytotoxicity	NS	Yes	NS	NS
Biological activities of antibody	Yes	NS	Yes	NS
Cell-mediated immune response	Yes	NS	Yes	NS
In vivo study of pathogenesis				
Passive transfer with serum	Yes	NS	Yes	NS
Passive transfer with LNC	Yes	Yes	Yes	NS
Passive transfer with T cells	Yes	NS	Yes	NS
Passive transfer with non-T cells	Yes	NS	Yes	NS

TABLE IV (*Continued*)

Studies	Guinea pig	Rat	Mouse	Rabbit
Passive transfer with cell eluates	NS	NS	Yes	NS
Influence of the blood–testis barrier	Yes	NS	NS	NS
Immunogenetic analysis	NS	NS	Yes	NS
Antigen-mediated Suppression of EAO	Yes	NS	NS	NS

[a] Reprinted from Tung (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc. Abbreviations: NS, not studied; LNC, lymph node cells; CFA, complete Freund's adjuvant; MHC, major histocompatibility complex.

Sertoli cells. Germinal epithelial cells adjacent to the lesions desquamate, resulting in focal aspermatogenesis. The fully developed orchitic lesion consists of heavy infiltrates of macrophages, lymphocytes, and eosinophils in the interstitium and surrounding blood vessels and of polymorphonuclear neutrophils in the lumen of blood vessels and seminiferous tubules. "Invasive" macrophagic lesions are associated with a decrease in the blood–testis barrier to systematically injected acroflavine (Johnson, 1970b).

Neutrophil-Rich Lesions in the Efferent Duct, the Caput, Body, and Cauda Epididymis, and the Vas Deferens. Numerous neutrophils, often containing periodic acid Schiff-positive granules, are seen between epithelial cells lining these regions of the sperm compartment (Tung *et al.*, 1970). Heavy or sparse mononuclear cellular infiltrates are found outside the ducts. Late lesions include large abscesses, granulomata, and periductal fibrosis.

Complete Aspermatogenesis with Severe Atrophy of Seminiferous Tubules and Interstitial Fibrosis. In this end-stage lesion of EAO, one may find abundant mitotic activity of spermatogenic cells, representing regeneration (Tung *et al.*, 1970)

c. Pathogenesis

Extrapolation from Immunopathological Findings of EAO. Degeneration and exfoliation of sperm, spermatids, and spermatocytes are believed to represent a lesion mediated by humoral antibody that has reached the lumen of the seminiferous tubules (Brown *et al.*, 1972; Nagano and Okumura, 1973b; Toullet and Voisin, 1974). Further support for this concept has been provided by the findings that sperm, with higher

concentration of surface autoantigens than spermatids, were more susceptible to complement-dependent cytotoxic antibody and that seminiferous tubules containing sperm and spermatids were most severely damaged in EAO (Johnson, 1970b). The macrophagic "invasive" lesion in the testis resembles a delayed hypersensitivity type reaction mediated by T lymphocytes (Waksman, 1959; Tung *et al.*, 1970). Except for traces of IgG adjacent to the infiltrating mononuclear cells (Johnson, 1970b), there is no specific binding of immunoglobulin or complement to spermatids or sperm. The neutrophil-rich lesions of the ductus efferentes and the epididymis are similar to an Arthus-type reaction involving insoluble antigen–antibody complexes and complement (Toullet *et al.*, 1973; Tung *et al.*, 1970). In support of this possibility was the finding of IgG and the complement component C3 bound to sperm acrosome in the rete, adjacent to the seminiferous tubules and cauda epididymis where neutrophil-rich lesions are also most commonly detected (Tung *et al.*, 1970; Nagano and Okumura, 1973b). Toullet *et al.* (1973) postulated that this type of lesion may be induced by sperm P antigen, which appears selectively to induce precipitating and complement-activating antisperm antibody.

Initiation of EAO. This requires the interaction between adequate amounts of immune reactants (humoral antibodies and/or sensitized lymphocytes) with aspermatogenic antigens in the testis; this event can be modulated by the blood–testis barrier. On immunizing young guinea pigs with adult testicular homogenate in CFA, it was found that EAO did not occur in guinea pigs 74–80 days of age when the seminiferous tubules were already fully developed, but the sperm had not reached the rete or the excurrent duct consistently (Johnson, 1970a). However, when the blood–testis barrier of these testes was weakened by cadmium sulfate, by needling of the testis, or by intratesticular injection of *Mycobacterium tuberculosis*, orchitis at the level of seminiferous tubules appeared within 24 h. To explain how EAO is initiated in guinea pigs immunized with testicular or sperm antigens in CFA, several possibilities are being considered. They are not mutually exclusive.

First, immune reactants may react with soluble aspermatogenic antigens present outside the blood–testis barrier. Soluble antigens may leak out of the rete and ductus efferentes where the blood–testis barrier is weak (Brown and Glynn, 1969). The reaction of sensitized lymphocytes with aspermatogenic antigen, presented by Ia-positive interstitial macrophages, outside the ductus efferentes provides the most likely mechanism in adoptive transfer of EAO by systemic injection of lymph node cells and peritoneal exudate cells.

Second, immune reactants may enter the ductus efferentes or the rete,

by retrograde flow (Tuck *et al.*, 1970) to reach the antigen-containing cells in the seminiferous tubules. Inside the seminiferous tubules, both antibody and sensitized lymphocytes can react with target antigens to initiate different types of immune reactions. Exfoliated sperm and spermatids that reach the rete and ductus efferentes will greatly increase the concentration of target antigens in these vulnerable regions of the sperm compartments. This mechanism may explain why the rete testis and the ductus efferentes are most commonly and severely affected in EAO induced by active or passive immunization (Waksman, 1959; Johnson, 1970b; Kantor and Dixon, 1972). Alternatively, sensitized T lymphocytes, on reaction with aspermatogenic antigens in the seminiferous tubules, may release lymphokines that increase the permeability of the blood–testis barrier and attract macrophages from the interstitium. The focal nature of these lesions may be determined by the unpredictable locations where sensitized T cells and aspermatogenic antigens meet. Although Sertoli cells may not contain aspermatogenic antigens, they are damaged as innocent bystanders to lymphocytoxins that have been detected following the reaction between sensitized lymphocytes and sperm *in vitro* (Toder *et al.*, 1975). Thus, different immunopathological pictures may be found depending on the dominant immune response in a given animal: antibody response leads mainly to desquamative lesions with minimal orchitis, and cell-mediated immune response leads to severe orchitis.

Third, immune reactants may enter the ductus efferentes and proceed to flow only toward the epididymis where inflammatory lesions are elicited. In occasional animals, lesions are found exclusively in the ductus efferentes, the epididymis, and the vas.

Fourth, the entrance of immune reactants to the sperm compartment may be conditioned by a second factor that weakens the blood–testis barrier. Pertinent to this possibility is the observation that CFA enhances the passive transfer of EAO by immune sera (Pokorna, 1970a; Willson *et al.*, 1973; Toullet and Voisin, 1976) and may do so by weakening the blood–testis barrier (Brown and Glynn, 1969). Intravenously administered horseradish peroxidase was found to enter the cytoplasm of spermatogonia and Sertoli cells but failed to penetrate the tight junctions between Sertoli cells (Willson *et al.*, 1973). Others (Voisin and Toullet, 1968; Brown and Glynn, 1969; Johnson, 1973) have postulated that the delayed hypersensitivity resulting from the reaction between sensitized T lymphocytes and aspermatogenic antigens weakens the blood–testis barrier and permits antibody to reach aspermatogenic antigens.

Finally, in addition to the sequestered aspermatogenic antigens, there may be a second set of autoantigens that normally exist outside the blood–testis barrier. Immunological tolerance to the nonsequestered anti-

gen, such as molecules on Sertoli cell surface, normally prevents the induction of testicular autoimmunity. However, autoimmune disease ensues when the tolerance state is terminated by immunization with appropriate adjuvants.

Although the above-mentioned possibilities for initiation of EAO are speculative, they provide useful working hypotheses for future experimentation.

The Role of T-Lymphocyte-Mediated Mechanisms. There is evidence for T-lymphocyte response to testicular-specific antigens in guinea pigs with EAO, because delayed-type hypersensitivity to sperm or sperm extracts has been demonstrated (Baum *et al.*, 1961; Chutna and Pokorna, 1967; Voisin and Toullet, 1968; Brown and Glynn, 1969; Mazzolli, 1971; Muir *et al.*, 1976) and because peritoneal cells from sensitized guinea pigs produce macrophage migration-inhibitory factor in the presence of sperm antigens (Mazzolli, 1971; Tung *et al.*, 1971b; Marcus *et al.*, 1975; Meng and Tung, 1983). Cell-mediated immune response to aspermatogenic antigen precedes the development of EAO. In guinea pigs immunized with G75m antigens in CFA, the proliferative response of lymph node cells in the presence of G75m was first detected on day 4, reached maximal activity by day 7, and declined after day 15 (Meng and Tung, 1983). The lymphoproliferative response correlated with severity of orchitis. Peritoneal exudate cells capable of secreting macrophage inhibitory factor in the presence of G75m were detected on day 7 and reached plateau activity on day 10. On the other hand, EAO was not detected until day 10 and reached a plateau incidence on day 14. Using delipidated acid extract of guinea pig testis, Hojo *et al.* (1980) found a similar temporal relation between lymph node cell proliferative response, delayed-type hypersensitivity reaction, and EAO occurrence in the guinea pig. Furthermore, in both studies cell-mediated immune response was found to decline by day 18, only to be followed by a second peak of activity on day 30 (Hojo *et al.*, 1980). These findings provide indirect evidence for involvement of cell-mediated immunity in EAO and point to the possible existence of suppressive or negative immunoregulation once EAO is fully developed.

More direct evidence for a T-cell-mediated mechanism of EAO has come from studies based on the adoptive transfer of the disease. Systemic transfer of EAO was first reported by Stone *et al.* (1969). When lymph node cells from two donors that had been immunized 5 days previously with testicular homogenate in CFA were injected intraperitoneally into 1 recipient, testicular pathology became apparent 9 days later. When Tung *et al.* (1971c) attempted to repeat this experiment, only 6 of 31 (20%) of the recipients developed orchitis. The lesions were diffuse in 1 and focal

in 5 recipients. More recently, Carlo *et al.* (1976) succeeded in the transfer of severe EAO to 100% of strain XIII syngeneic recipients by intraperitoneal injection of mixtures of peritoneal exudate cells and lymph node cells from donors immunized with the purified GP1 aspermatogenic antigen in CFA. This study demonstrated one useful application of purified aspermatogenic antigen in the study of EAO.

Since all studies on systemic adoptive transfer of EAO have utilized mixtures of T and B lymphocytes, the results did not distinguish T lymphocytes or humoral antibodies secreted by B cells as the cause of disease in the recipients. To specifically address the questions of T-cell requirement in EAO induction, Tung *et al.* (1971c) and Kantor and Dixon (1972) based their studies on the technique of local adoptive transfer in inbred strain XIII guinea pigs. Lymph node cells or peritoneal exudate cells and their subpopulations were injected immediately beneath the tunica albuginea, using aseptic technique.

Lesions indistinguishable from EAO were transferable with this technique; >90% of the recipients of lymph node cells and all of the recipients of peritoneal exudate cells developed lesions in the seminiferous tubules and the rete testis. The adoptive transfer of EAO apparently requires that aspermatogenic antigens be present in the testis of the recipient, since sexually immature 10-day-old recipients did not develop any evidence of pathology. In a successful transfer, the injected lymph node cells were first distributed in the interstitial spaces throughout the testis, but tended to concentrate in the rete. During the first 2 days, the infiltrating cells were predominantly small lymphocytes, lymphoblasts, and cells in mitosis. From day 3 on, monocytes and macrophages dominated and abutted the boundary tissue of the seminiferous tubules. On day 4, numerous clusters of mononuclear cells invaded the tubules and the rete. From day 4 to day 8, spermatids and spermatocytes adjacent to the clustered macrophages began to desquamate and aspermatogenesis ensued. When sufficient numbers of cells were injected, lesions were as severe as EAO induced by active immunization. Lesions produced by local adoptive transfer were antigen specific, since cells from donors immunized with CFA alone or CFA with several irrelevant antigens did not transfer disease to susceptible animals.

The requirement of T lymphocytes in the adoptive transfer of EAO was determined by comparing the efficiency of transfer between highly enriched T cells and non-T cells from peritoneal exudate cells (Tung *et al.*, 1977) and lymph node cells (Tung *et al.*, 1981a). In these experiments, guinea pig T cells were isolated, either by rosette formation of these cells with rabbit erythrocytes and separation from the nonrosetted cells in a hypaque-ficoll gradient or by depletion of macrophages and neutrophils

on a nylon column. The enriched T-cell fraction transferred "invasive" macrophagic lesions, and the number of T cells correlated with the prevalence of adoptive transfer. These studies thus show that sensitized T lymphocytes can mediate an antigen-specific, macrophagic "invasive" lesion like that in EAO.

The Role of Non-T-Cell-Mediated Mechanisms. In adoptive transfer of EAO by lymph node cells, the non-T-cell fraction, which contained mainly B lymphocytes, macrophages, and presumably K and NK cells, also transferred EAO with the same efficiency as the cell population enriched in T cells (Tung *et al.*, 1981a). Furthermore, the lesions transferred by non-T cells were indistinguishable from those induced by T cells.

The Role of Humoral Antibody. Guinea pigs with EAO produce antibody of different immunoglobulin classes: (1) IgG_1 antibody, detected by immunofluorescence and anaphylactic-type reactions (Katsh and Bishop, 1958; Baum *et al.*, 1961; Chutna and Pokorna, 1967; Voisin and Toullet, 1968; Tung *et al.*, 1971b); (2) IgG_2 antibody, detected by immunofluorescence (Tung *et al.*, 1971a), complement fixation (Tung *et al.*, 1971b), Arthus reaction (Baum *et al.*, 1961; Voisin and Toullet, 1968), and cytophilic activity to macrophages (Barrera *et al.*, 1976), and (3) IgM antibody, characterized by its elution characteristic in G-200 chromatography and by its sensitivity to 2-mercaptoethanol treatment (Chutna and Pokorna, 1967). These antibodies react with the acrosomal apparatus (Tung *et al.*, 1971a; Toullet *et al.*, 1973), as well as with the cell-surface antigens of sperm, residual body, and later spermatids (Tung *et al.*, 1979), and are cytotoxic to these cells (Chutna and Rychlikova, 1964; Mancini *et al.*, 1969; Johnson, 1970c; Toullet and Voisin, 1974) by rapidly inducing typical ultrastructural lesions on the plasma membrane and acrosomal membrane (LeBouteiller *et al.*, 1975). Antisperm antibodies are detected 6 days after immunization. However, in animals immunized with the better-defined aspermatogenic antigen G75m, maximal antibody to G75m was detected at a time when active orchitis had subsided (Meng and Tung, 1983).

Passive transfer of EAO by immune sera has been reported by several investigators (Pokorna, 1970b; Tung *et al.*, 1971a; Willson *et al.*, 1973; Nagano and Okumura, 1973b; Toullet and Voisin, 1976). The results of these studies indicate that after transfer, IgG is localized on acrosomes of sperm in the rete and cauda epididymis and less consistently in the seminiferous tubules; that infiltrations of neutrophils become apparent in the rete and the cauda epididymis; and that mild to severe aspermatogenesis develops in the seminiferous tubules without mononuclear infiltrates. Frequent and severe lesions, however, are found only in recipients that have been injected with CFA. The transfer is antigen specific, since normal

guinea pig serum or serum from guinea pigs immunized with CFA does not produce similar changes. These results indicate that humoral antibodies can reach sperm in the seminiferous tubules, rete, and cauda epididymis, and can induce two lesions that are found in EAO. The need to inject recipients with CFA to produce severe lesions may be related to the apparent increase in permeability of the blood–testis barrier from treatment with CFA. The success of serum transfer of EAO may depend on the antigen specificity of the antiserum used. Toullet and Voisin (1976) found that antiserum to P antigen transferred most efficiently (64%); anti-T antiserum was less efficient (40%), while anti-S serum was inefficient. This finding correlated well with their earlier data on the decreasing cytotoxic efficiency of antisera to T, P, and S antigens against testicular cells (Toullet and Voisin, 1974).

d. Summary. The immunopathological and pathogenetic studies on guinea pigs with EAO suggest that EAO in this species is mediated by complex immunological mechanisms. The rete testis and ductus efferentes are clearly the weak spots in the blood–testis barrier. Humoral antibody can enter these sites, and from there reach the sperm and spermatids in the seminiferous tubules and the epididymis. It is possible that soluble aspermatogenic antigens may also leak out at this site. Whether antibody lymphocytes can traverse the barrier at the seminiferous tubules under influence of CFA needs study. Antibodies appear to be involved in eliciting neutrophil-rich lesions in the rete and epididymis and aspermatogenesis in the seminiferous tubules. Adoptive transfer experiments, on the other hand, show that both T lymphocytes and cell populations rich in B cells, macrophage, and presumably K and NK cells can cause a mononuclear infiltrative lesion, which also leads to aspermatogenesis. Whether these lymphocytes reach the aspermatogenic antigens by a retrograde route from the rete or outside the seminiferous tubules, or both, requires further study. Thus, multiple immunological mechanisms are involved in producing the complete picture of EAO, but each mechanism itself can initiate its own characteristic lesions.

2. Mouse

A highly reproducible model of murine EAO has greatly facilitated our understanding of this autoimmune disease, since advantage can be taken of the numerous inbred strains, congenic strains, and hypothymic mice for precise immunogenetic and pathogenetic analyses.

Although several investigators reported induction of EAO in the mouse (Pokorna *et al.*, 1963; Bohme, 1965; Hargis *et al.*, 1968), these studies were not followed up by any incisive immunological analysis. Recently, based on a modified protocol of Pokorna *et al.* (1963), severe EAO has

TABLE V
Influence of H-2 on the Relative Severity of Orchitis and Vasitis in Mice

Study	Mouse	Non-*H-2* genes	*H-2* haplotype	*n*	Orchitis[a]	Vasitis[a]
A	A.SW/SnJ	—	s	31	4.4 ± 2.8	0.4 ± 0.8
	NFS/N	—	s	53	2.6 ± 3.6	0.4 ± 0.8
	SJL/J	—	s	31	3.6 ± 2.6	1.7 ± 1.5
	BALB.K	—	k	22	0.3 ± 0.8	2.7 ± 3.7
	AKR	—	k	25	0.3 ± 0.6	2.3 ± 1.7
	B6/C3H	—	k	22	2.0 ± 2.5	3.3 ± 2.3
B	A/J	A/J	a	25	4.4 ± 3.7	3.3 ± 2.5
	A.SW	A/J	s	31	4.4 ± 2.8	0.4 ± 0.8
	BALB/cBy	BALB/c	d	29	5.6 ± 4.1	2.1 ± 2.2
	BALB.K	BALB/c	k	22	0.3 ± 0.8	2.7 ± 3.7

[a] Disease severity of orchitis and vasitis is expressed as disease indices (means ± SEM) derived from a range of 1 to 10.

been induced in mouse (Bernard *et al.*, 1978; Sato *et al.*, 1981; Kohno *et al.*, 1983). Pertussis organisms or their extracts (Munoz and Arai, 1982) must be included as adjuvant. Induction of EAO in the mouse was thymus dependent. Hypothymic BALB/c nu/nu mice failed to develop EAO. After they were reconstituted with thymic cells from litter mate nu/+ mice, hypothymic mice regained capacity to develop EAO.

Inbred and congenic mice have been used to investigate the genetic influence on orchitis and vasitis development (Teuscher *et al.*, 1985). Mice with the $H\text{-}2^s$ haplotype tended to develop severe orchitis with little or no vasitis, whereas mice with the $H\text{-}2^k$ haplotype developed mainly vasitis (Table V). In study that employed appropriate *H-2* congenic mice, the $H\text{-}2^s$ haplotype appeared to code for absence of vasitis, and the $H\text{-}2^k$ for absence of orchitis.

Influence of the major histocompatibility complex (*H-2*) genes on orchitis development was evaluated in *H-2* congenic mice derived from BALB/cBy and from C57BL/10 mice. It was clear that severe orchitis was associated with the $H\text{-}2^d$ haplotype (Table VI). Moreover, results on BALB/cBy and C57BL/10 mice with *H-2* gene recombinance suggested that gene(s) coding for severe orchitis could be mapped to the *D* end of $H\text{-}2^d$.

The influence of non-*H-2* genes on orchitis development was evaluated in the F_1 progeny of high and low responders and in the backcross progeny of the F_1 with the appropriate parental strain. While BALB/cBy was a high responder and DBA/2 a low responder, all of the F_1 mice were low responders (Table VII). Furthermore, the ratio of low responders to high responders among the backcross mice between (BALB/cBy × DBA/2) F_1

TABLE VI

Orchitis Severity Can Be Mapped to the *D End* of *H-2*^{d} in the Major Histocompatibility Complex (MHC)[a,b]

Study	Mouse	Non-*H-2* Genes	Major histocompatibility complex					*n*	Orchitis[c]
			K	I-A	I-E	S	D		
A	B10.D2	C57BL/10	d	d	d	d	d	46	4.2 ± 3.6
	B10.BR	C57BL/10	k	k	k	k	k	33	0.4 ± 0.4
	C57BL/10	C57BL/10	b	b	b	b	b	17	1.7 ± 1.4
	BALB/cBy	BALB/c	d	d	d	d	d	29	5.6 ± 4.1
	BALB.K	BALB/c	k	k	k	k	k	28	0.3 ± 0.8
	BALB.B10	BALB/c	b	b	b	b	b	29	0.4 ± 0.7
B	B10.BR	C57BL/10	k	k	k	k	k	28	0.3 ± 0.8
	B10.A	C57BL/10	k	k	k	*d*	*d*	11	5.3 ± 3.3
	BALB.HTG	BALB/c	d	d	d	d	b	25	1.5 ± 2.3
	BALB/cBy	BALB/c	d	d	d	d	*d*	29	5.6 ± 4.1

[a] Shown by H-2 congenic strains of C57BL/10 and BALB/cBy mice; A studies.
[b] Studies with mice with gene recombinance within the MHC; B studies.
[c] Severity of orchitis is expressed as disease indices (means ± SEM) ranging from 1 to 10.

and BALB/cBy was such that orchitis resistance could be accounted for by the effect of a single gene locus. Since BALB/cBy and DBA/2 are both *H-2*^{d}, the phenotype observed is controlled by gene(s) located outside *H-2*.

Also of interest was the finding that severity and prevalence of orchitis differed between substrains of BALB/c mice; the difference was as marked as that noted between responder and nonresponder inbred mice (Teuscher *et al.*, 1985b). EAO is the first example of an organ-specific autoimmune disease with different susceptibility between mouse substrains that became diverged from one another for less than 40 years. This finding underlines the importance of selecting and defining sublines of inbred mice for biologic research, and offers additional support for rapid divergence of phenotypes between inbred mouse strains (Fitch and Atchley, 1985).

The genetic control of murine EAO induced by crude testis homogenate is complex, with dicotomy in the control of orchitis and vasitis. Moreover, both *H-2* and non-*H-2* genes are involved in the control of orchitis development.

3. Rabbit

Severe orchitis and aspermatogenesis can be induced in rabbits with an active aspermatogenic antigen preparation (TCAsup) (Tung and Woo-

TABLE VII

Orchitis Severity is Controlled by Gene Locus Mapped outside the Major Histocompatibility Complex

Study	Mouse	H-2	*n*	Orchitis indices[a]
A	BALB/cBy	d	26	5.6 ± 4.1
	DBA/2	d	29	0.1 ± 0.2
	BALB/cBy × DBA/2	d/d	20	0.5 ± 1.7
	DBA/2 × BALB/cBy	d/d	17	0.9 ± 2.0
B	(BALB/cBy × DBA/2)	dd/d	11	7.7 ± 1.5[b]
	× BALB/cBy		8	2.8 ± 1.3[b]

[a] Severity of orchitis is expressed as disease indices (means ± SEM) ranging from 1 to 10.

[b] The pathology indices between the two groups are significantly distinguishable as high and low responders ($p < .001$). The number of gene loci = log (number of low responders/number of animals studied/log 0.5, = 1.25.

droffe, 1978). While both heterologous and homologous TCAsup in CFA elicited EAO, the immune response and the attendant testicular immunopathology of the two groups differ in several respects. Guinea pig TCAsup elicited in all animals a high titer of antibody that reacted with antigenic determinants common to guinea pig and rabbit sperm. Multiple injections were required to induce severe disease, which consisted mainly of aspermatogenic seminiferous tubules and mild orchitis. Immunofluorescence revealed massive granular deposits in the basement membrane of rabbit IgG, IgM, complement component C3, and sperm acrosomal antigen. When eluted from the testis by citrate buffer at pH 3.2, the testicular-bound IgG was found to contain 13 times more antisperm antibody per unit IgG than serum IgG. The immunopathological findings in this group of animals resembled the findings of orchitis in vasectomized rabbits (Bigazzi *et al.*, 1976) and added further support for the immunological nature of postvasectomy orchitis in rabbits.

In contrast, rabbits immunized with homologous TCAsup antigen developed serum antisperm antibody of low level and infrequently. However, a single injection of the antigen in CFA is sufficient to induce severe orchitis and aspermatogenesis. Unlike rabbits injected with heterologous antigens, immune complexes were not detected.

4. Rat

As in the mouse, induction of EAO in the rat required multiple injections of testicular antigens in CFA, and this was potentiated by pertussis

vaccine. In the rat, EAO was studied only rarely until it was selected by Wekerle and his co-workers as a model to investigate autoreactive T lymphocytes. Results of their work have already been described in Section II,C,2a.

E. Postvasectomy Autoimmune Orchitis

Vasectomy leads to autoimmune responses to testis-specific autoantigens (Reviewed in Teuscher *et al.*, 1981) and to autoimmune orchitis.

1. Rabbit

Rabbits that had been vasectomized 8 to 10 months previously developed evidence of immune complexes in the testis (Bigazzi *et al.*, 1976). Granular IgG and C3 were present along the basement membrane of the seminiferous tubules. Ultrastructurally, electron-dense deposits were found between the tubular basement membrane and Sertoli cells. In areas that corresponded to immune complex deposition, mononuclear cellular infiltration was evident. The IgG dissociated from the testis by phosphate-buffered saline at 56°C reacted with sperm acrosome. It can be envisaged that sufficient soluble sperm antigen leaks outside the blood–testis barrier, reacts with interstitial antisperm antibody, and forms immune complexes *in situ*. Such immune complexes then lead to inflammation and initiate or amplify tissue destruction.

2. Guinea Pig

Fifty percent of vasectomized inbred strain XIII guinea pigs operated on 16 months previously developed typical lesions of EAO, which consisted of macrophage invasion of seminiferous tubules (Tung, 1978). This finding coincides temporally with the development of positive cell-mediated immune response to the aspermatogenic antigen G75m, as detected by the macrophage migration-inhibition assay (Tung *et al.*, 1981a). When peritoneal exudate cells from the vasectomized guinea pigs were adoptively transferred to syngeneic recipients, similar testicular lesions developed in 30% of the animals. The passive transfer of testicular disease was specific, since peritoneal cells from sham-vasectomized guinea pigs did not transfer lesions.

3. Rhesus Monkey

Tung and Alexander (1980) compared the immunopathological findings in 23 long-term (5–15 years) vasectomized and 11 age-matched control rhesus monkeys. While orchitis and/or aspermatogenesis were detected in

27% of the control animals, they were found in 87% of vasectomized monkeys. As reported previously by Alexander (1972), granular IgA, IgM, and/or C3, presumptive evidence of immune complexes, were found along the basement membrane of different parts of the sperm compartment. Immune complexes were most commonly detected in the regions of the ductus efferentes and caput epididymis in both vasectomized and control animals. After vasectomy, immune complexes became detectable in the cauda epididymis and the rete testis. While the immunological nature of the testicular disease remains to be proven, the finding of immune complexes suggests an increase in permeability to sperm antigens in certain regions of the sperm compartment following vasectomy.

F. Spontaneous Testicular Autoimmune Diseases

1. Aspermatogenesis and Testicular Immune Complexes in Infertile Men

Salomon and co-workers (1982) described the presence of electron-dense deposits and discrete clusters of membranovesicular particles in thickened basal lamina surrounding seminiferous tubules in testicular biopsies of infertile men. Of 70 patients, 35% were positive. When the testes were studied by immunoperoxidase technique at the ultrastructural level, human IgG (Fig. 2) and/or C3 were clearly demonstrable in 70% of 30 cases. The findings provided strong evidence for involvement of immunological mechanism in human testicular disease leading to infertility. It is interesting that aspermatogenesis with sperm antigen–antibody complexes is also the immunopathological feature of rabbit EAO induced by heterologous testis antigen (Tung and Woodroffe, 1978), and is found in a subset of the dark mink with secondary infertility (Tung *et al.*, 1981b) (see below).

2. The Infertile "A" Line Beagle Dog

The purebred beagle colony at the Division of Biological and Medical Research, Argonne National Laboratory, has been a closed breeding unit since 1960. The ancestral relationships of the animals in this colony include two partially inbred lines (A and B). The A line beagle dogs, derived from three sibling progenitors, have a high incidence of lymphocytic thyroiditis (Fritz *et al.*, 1976) and increasing evidence of male infertility. Of 69 dogs analyzed, 22 had different degrees of nodular lymphocytic orchitis involving the interstitium and/or the seminiferous tubules. The pres-

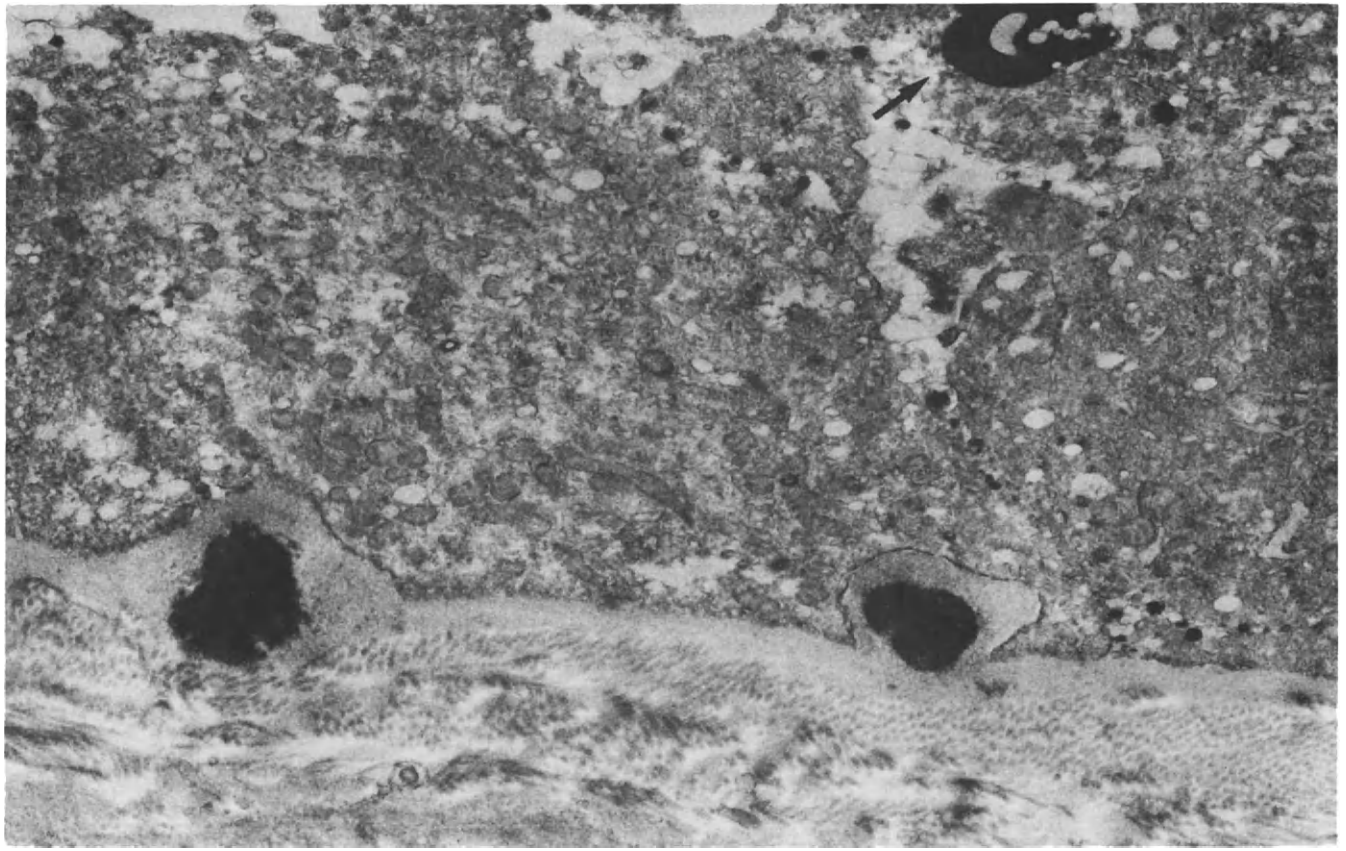


FIG. 2. Ultrastructural detection of human IgG by immunoperoxidase technique of human IgG in the testis of an infertile man. IgG appears as discrete deposits in thickened basement membrane of seminiferous tubules with hypospermatogenic epithelium. Reproduced from Salomon *et al.* (1982), with permission of the publisher.

ence of orchitis was associated with an increased frequency of aspermatogenesis or degenerative changes, *A* ancestral gene frequency, small testicles, and hypospermia or azospermia.

Lymphocytic thyroiditis and orchitis tended to coexist. Statistical analysis indicated that the occurrence of lesions in the two organs was related and both diseases were in turn related to the A ancestral composition. The A line beagle dogs represent one of the best examples of genetically determined autoimmune disease involving the gonads and other endocrine organs (Sotsiou *et al.*, 1980).

3. The T/t^{w18} Backcross Mouse

In the course of raising a congenic mouse strain with the t^{w18} haplotype on the inbred BTBRTF/NeV (genotype +/+) background, repeated backcross matings were carried out. Although the original T^{j}/t^{w18} breeders were fertile, it was evident that the male (6%) and female (14%) backcross progeny with the tailless phenotype (T/t^{w18} genotype) were infertile (Dooher *et al.*, 1981). Because of the nature of the breeding conditions by which these mice were generated, the infertility was considered to be a consequence of genetic factors located on chromosome 17. Of 20 infertile T/t^{w18} backcross males, 8 (40%) had bilateral or unilateral orchitis. The testicular lesions consisted of focal infiltrations of lymphocytes and plasma cells that were present mainly in the interstitial space. Electron microscopic study revealed separation of peritubular adventitial cells from the Sertoli cells by lymphocytes. Inside the seminiferous tubules, there were degenerative germ cells and phagocytosis of germ cells by Sertoli cells.

4. The Infertile Black Mink

The process of breeding mink in Utah for a fine black fur has selected for the undesirable phenotype of male infertility. Some mink were infertile soon after puberty (primary infertility), and others became infertile after a period of proven fertility (secondary infertility) (Tung *et al.*, 1981b).

Mink with secondary infertility had high levels of antisperm antibodies, detectable by indirect immunofluorescence. Many testes had severe orchitis and/or aspermatogenesis and epididymitis. In many of these testes, there were peritubular granular deposits of mink IgG and C3, presumptive evidence of immune complexes (Fig. 3). Although sperm antigen was not detected in the immune complexes, the acid eluate from testes with immune complexes recovered IgG that was enriched (×10) in antisperm

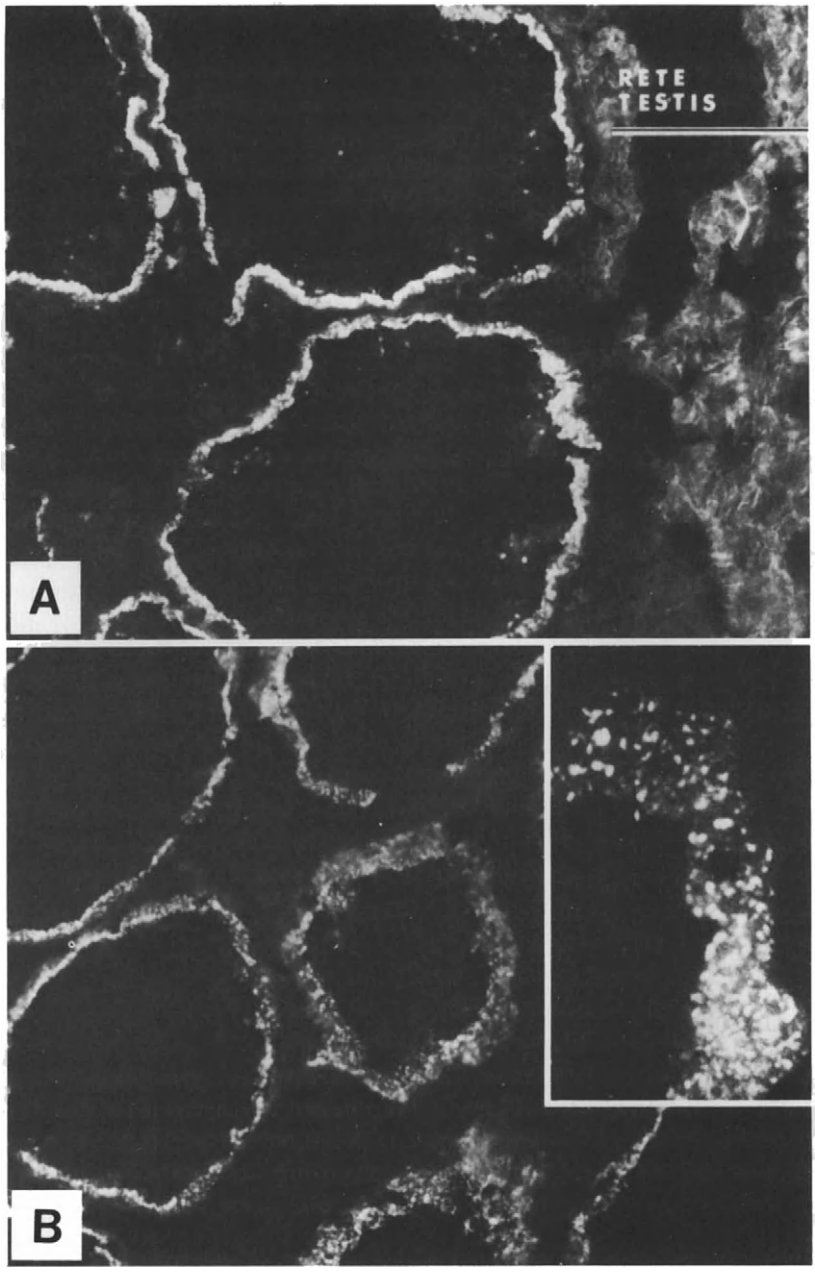


FIG. 3. Deposits of mink IgG (A) and C3 (B) in the testis of a mink with secondary infertility. Granular deposits (evidenced in insert of B) are detectable in the tubular basement membrane throughout the testis, and are absent in the rete testis. Reproduced from Tung *et al.* (1981b), with permission of the publisher.

antibody activity. Furthermore, phosphate-buffered saline failed to elute antisperm antibody from the same pool of testicular tissues. These findings are thus consistent with *in situ* deposition of sperm antigen–antibody complexes.

While autoimmune testicular disease in the secondary infertile mink is clearly an important feature of the infertile state, it is unclear how the autoimmune disease evolves. The primary defect may lie with some nonimmunological mechanisms that determine the development and maintenance of normal testicular function. In this context, recent findings in mink with primary infertility are pertinent. These animals, unlike mink with secondary infertility, exhibited minimal evidence of testicular autoimmunity. The major abnormality is defective production of gonadotropin-releasing hormone (GnRH) (Tung *et al.*, 1984), which may be the result of abnormal hypothalamic function or of defective neuroendocrinological control mechanism(s) of the hypothalamus. Thus, the serum levels of luteinizing hormone (LH) and testosterone in primary infertile mink at the time of breeding were significantly below that of the fertile dark mink or mink of other fur colors. However, they responded to exogenous GnRH by prompt elevation of LH, and they responded to exogenous hCG by prompt elevation of testosterone. If a similar endocrinological defect is also found in mink with secondary infertility, this disease model may point to an association between abnormal hormone regulation of the testis and testicular autoimmune disease.

III. ANTISPERM IMMUNITY AND INFERTILITY

A. Clinical Evidence

Some clinical observations provide strong but indirect support for existence of immunological infertility. Antisperm antibodies are detected at a higher frequency in infertile than fertile couples (Jones, 1976; Menge, 1980). Antisperm antibodies are detected in the local secretion of the genital tracts, particularly in the cervical mucus (Soffer *et al.*, 1976; Menge *et al.*, 1977). Immunoglobulins have been identified on surface of ejaculated sperm (Jager *et al.*, 1978). Some antisperm antibody can activate the complement pathway, resulting in cytotoxicity or immobilization of sperm (Fjallbrant, 1968; Jager *et al.*, 1972). In rare patients, spontaneous sperm autoagglutination can be detected, and this is most likely due to agglutinating antibodies (Rumke and Hellinga, 1959; Fjallbrant, 1968). In the presence of antisperm antibody, sperm fail to penetrate cervical mucus, or they exhibit a peculiar "shaking" appearance on the edge of the

cervical mucus (Fjallbrant, 1968; Kremer and Jager, 1976). Antisperm antibodies have been shown to inhibit penetration of human sperm to zona-free hamster ova (Menge and Black, 1979). Vasectomy leads to antisperm antibodies, and some patients who have vasovasostomy, and subsequently show an adequate sperm count, remain infertile.

Since antisperm antibody detection is increasingly used in the clinical evaluation of infertile patients, the methods and clinical significance of these procedures will be detailed.

1. Methods of Antisperm Antibody Detection

Classically, the two methods that have been used to detect sperm antibodies are agglutination and cytotoxicity. Early methodology for assessing sperm agglutination varies with each report. Kibrick and co-workers (1952) published a reliable agglutination procedure that subsequently has been widely used clinically, especially in infertile men. This method, which allows macroscopic observation, has become known as the gelatin agglutination test (GAT). A simpler method designed originally for determining sperm agglutination with human sera was published by Franklin and Dukes (1964). This is a microscopic procedure and has been most widely used with sera from infertile women. A microagglutination procedure was subsequently developed by Friberg (1974) utilizing tissue-typing trays (TAT). Mammalian sperm in the presence of cytotoxic antibody and complement lose their flagellar movement concomitantly with the occurrence of cell death, thus the term sperm-immobilizing antibody. The immobilization method most commonly used was standardized by Isojima and co-workers (1972).

Presented in brief are the procedures standardized by a WHO Workshop on Sperm Immunology (Rose *et al.*, 1976) and the more recent procedures. These methods are discussed to ensure that differences among methods are understood; they are summarized in Table VIII.

a. Sperm Agglutination Assays

Gelatin Agglutination Test (GAT). Motile sperm cells are suspended in a 5% gelatin solution in physiological saline or in a buffer (e.g., Baker's solution) and mixed with an equal volume of serum serially diluted in small-bore culture tubes. The suspension in gelatin allows free swimming sperm to remain evenly dispersed, but formation of medium-to-large sperm agglutinates becomes visible macroscopically; these tend to settle down, clearing the medium. Tail-to-tail or mixed type of agglutination is detected by the GAT.

Tube-Slide Agglutination Test (TSAT). Formerly known as the Franklin–Dukes method, the procedure receives its present name in that

TABLE VIII

Methods for Detecting Sperm Antibodies

	Agglutination tests							Antiglobulin reactions			
				Immobilization						Radio-immunoassay	
	Gelatin	Tube-slide	Tray	Tube	Micro-tray	Indirect immuno-fluores-cence	Cervical mucus: penetration/ contact	Mixed anti-globulin reaction (coated RBCs particles)	Anti-globulin-coated particles	Serum	Sperm
Sperm concentration (10^6)	40	50	40	40	10	Wash and dilute for readable	Ejaculate	Neat	10	6.7	50
Sperm volume (ml)	0.1–0.15	0.05	1 μl	0.025	1 μl	smear	0.1–0.2	1 drop	5 μl	0.3	0.1
Serum volume (ml)	0.2–0.3	0.5	5 μl	0.25	10 μl	.025–0.1		1 drop	50 μl	0.3(1.0)	1.0
Initial dilution (1:)	4	1.5	4	1,4	4	16	1	(1 drop)	4	2	—
Serial dilution	Yes	No	Yes	Yes	Yes	Limited	No	No	Yes	Yes	No
Incubation (h)	1	2–4	2 or 4	1–2	1–2	0.5	0.25–1.0	10 min	10 min	0.75	0.75
Incubation (°C)	37	37	37 or 22	37	37	33	22 or 37	22–35	22	37	37
Used with											
Cervical mucus	–	+	+	+	+	+	—	—			
Seminal plasma	+	+	+	–	+	+	—	—			
Observation	Macro	Micro	Micro	Micro		Micro	Micro	Micro		Radioactivity	
Reaction	T agg	Ht agg	Ht agg	Motility		Fluores-cence	Penetration	RBCs/particle binding to sperm		[^{125}I]anti-Ig binding to sperm	
Ig Class detected	AG	AGM	AGM	GM		AGM	AG	AGM		AGM	

the sperm and serum mixture is incubated in a culture tube, after which a sample is carefully withdrawn and transferred to a microslide for examination. Serum samples are considered to be positive if they agglutinate 10% or more of the motile sperm cells or if the number of agglutinates exceed 6 per 100 motile free sperm (Shulman, 1975). The TSAT method detects mainly head-to-head agglutination and to a lesser extent tail-to-tail agglutination. The method has yielded considerable variation among studies and has given some high values for sera from women who are pregnant, taking oral contraceptives, or possibly in the luteal phase of the menstrual cycle. A nonimmunoglobulin serum component with β mobility in electrophoresis and associated with steroids causes a head-to-head agglutination that may be observed in serum at low dilution (Boettcher *et al.*, 1970). The nonimmunoglobulin spermagglutinating components, referred to as β-sperm agglutinin, has a molecular weight exceeding 600,000, and is not a β lipoprotein as first reported. The β-sperm agglutinin activity apparently is directed against seminal plasma components absorbed loosely onto the sperm head, whereas agglutinating antibody activity is against antigens exposed on the plasma membrane of the sperm head, tail, or both (Ingerslev and Hjort, 1979; Kay and Boettcher, 1980).

Tray Agglutination Test (TAT). This procedure, although based on principles similar to those of the TSAT, has some definite advantages over the TSAT and GAT methods. Microliter quantities of reagents are utilized on a disposable tissue-typing tray, serial dilutions are easily made, and observation, performed with an inverted microscope, is rapid and does not cause disruption of agglutinates. Both the head and tail types of agglutination are readily detected by the TAT. A sample is considered to be positive if two or more consecutive dilutions show significant agglutination. The TAT will also detect the small head-to-head agglutinates caused by the β-sperm agglutinin, but this is a problem only at serum dilutions of 1:16 or less (Ingerslev and Hjort, 1979).

b. Sperm Immobilization Test (SIT). This is a complement-dependent assay that requires a suitable source of complements. Fresh and fresh-frozen serum samples from rabbit, guinea pig, and AB Rh-positive human volunteers have been used. Serum with natural sperm toxic factors, which are found commonly in guinea pig serum, should be looked for and if found excluded. Evaluation is made by determining the ratio of the percentage of motile sperm in control serum to the percentage in the unknown sample. This ratio is referred to as the sperm-immobilizing value (SIV) (Isojima *et al.*, 1972). Sera with SIV of 2 or greater are considered to be positive. This value combined with serial dilution of the sera allows determination of titers.

The micro-SIT assay has become useful in combination with a method yielding sperm samples with a high percentage of motility by overlayering semen with a buffer containing protein and an energy source (Hellema and Rumke, 1978). This provides samples with lower sperm concentration but a higher degree of motility and forward progression. Contamination by seminal plasma, which possesses anticomplement activity, is also reduced. The micro-SIT described is used in one of our laboratories (ACM) to detect antibodies in serum and cervical mucus samples. After the reagents are added, the tissue-typing tray is flooded with light mineral oil. The micro-SIT allows for a greater number of samples to be run and analyzed in duplicate or triplicate. Several studies have shown that sperm-immobilizing activity is highly correlated with sperm cytotoxic activity (Boettcher *et al.*, 1977). Cytotoxic tests using a double-fluorescent vital staining technique (Mathur *et al.*, 1981a) or a single fluorescent dye (Shepherd *et al.*, 1982) are alternatives to visible vital dyes. They require, however, an inverted fluorescent microscope.

c. Indirect Immunofluorescence Method. In this procedure sperm cells are usually prepared as smears are air dried and methanol fixed (Hjort and Hansen, 1971). Although this is a sensitive method, indirect immunofluorescence on fixed sperm detects submembranous antigens that most likely are not involved in the antifertility effects of antibodies. Indicative of this observation is the general lack of association of this method with agglutination and immobilization results. Antisperm antibodies generally react with the acrosome and main tail piece, whereas nonspecific staining often occurs on the equatorial segment, neck, and midpiece. IgG and IgM antibodies stain the acrosome, and IgG antibodies, the tail.

d. Paraimmunological Techniques. These techniques involving sperm–cervical mucus interactions are not highly specific for the detection of sperm antibodies. Described as *in vitro* postcoital tests, the methods are receiving more attention and use in both clinical and experimental studies (Kremer, 1965; Kremer and Jager, 1976).

Sperm-Cervical Mucus Penetration Test (S-CMPT). The procedure can be completed using midcycle periovulatory cervical mucus, either drawn up into a capillary tube or placed onto a microslide and covered. Semen is added, and after an incubation period the distance and degree of penetration by sperm are evaluated. Bovine estrous cervical mucus has been shown useful as a control in antisperm antibody studies with human sperm (Mangione *et al.*, 1981).

Sperm-Cervical Mucus Contact Test (S-CMCT). This procedure may be used in conjunction with or independent of the S-CMPT. It was

occasionally observed that sperm on contact with or shortly after entering cervical mucus would lose their forward progression, but maintained, often for hours, a jerking/vibrating motion that suggested that they were bound to the mucous micelles. This type of flagellation-in-place was termed "shaking phenomenon." The reaction was found to be associated with high titers of agglutinating antibodies in either the male or the female partner. The results are based on the percentage of motile sperm exhibiting the shaking phenomenon.

e. Antiglobulin Methods. These methods can be used to detect immunoglobulin (presumed antisperm antibody) on sperm surface from men suspected of having immunological infertility, or to analyze antibody in serum, seminal plasma, or cervical mucus following reaction of the antibody with sperm from a fertile individual. This category includes the mixed antiglobulin reaction (MAR), utilizing immunoglobulin-coated erythrocytes (Jager *et al.*, 1978; Hendry *et al.*, 1982) or plastic beads (Bronson *et al.*, 1982). The location of bound antibodies can be visualized microscopically by the presence of attached erythrocytes or beads on the sperm. Advantages of the test are the rapidity of the assay, the identification of Ig class of antisperm antibody, and the assurance that antigens being studied are located on the cell surface. The disadvantages are that viable sperm with high motility must be used, and the semen must be free of debris and foreign cells.

The radiolabeled antiglobulin method is the most objective method listed. Potentially, it can quantify antibody and identify the Ig class of the antibody (Hass *et al.*, 1980, 1982). However, the method as described requires multiple centrifugation and washings and handling of radioactive reagents.

f. Comments of Methodology. All methods need to be well controlled, with the inclusion of known positive and negative serum samples and the use of high-quality donor semen being required. To avoid nonspecific effects on sperm, serum should be diluted at least 1:4, and for screening purposes at least two dilutions should be utilized. Based on our experience with the TAT, SIT, and S-CMPT methods, the following observations can be made: Microtrays must be clean and free of foreign substances; semen donors need to be screened and their sperm cells examined in the systems before routine use; semen should exceed 2 ml in volume, 80×10^6 sperm/ml, and 70% in initial motility, and 75% of the motile sperm should show good forward progression and absence of spontaneous agglutination. To ensure that a semen sample is free of agglutinates, it should routinely be passed through a glass-wool column (Paulson and Polakoski, 1977). Despite these precautions, the TAT will occasion-

ally need to be repeated for some semen samples. When samples from a new sperm donor are used, it is advisable to compare the results based on the new sperm sample with those of a known sperm sample. It is important to be able to dilute the donor semen sample at least twice and preferably more for use in the SIT, as seminal plasm contains complement-inhibitory factors that may interfere with the system. Use of a micro-SIT with highly motile sperm that swim up to a layer of buffer in the semen-overlay method eliminates this problem.

Serum, if properly obtained and stored, should present few problems in terms of nonspecific sperm agglutination, except for the β-sperm agglutinins. Genital tract secretions, seminal plasma, and cervical mucus from individuals with local infections of the reproductive tract may cause non-immunological sperm agglutination, sperm immobilization, or both. Different organisms or their products have been shown to produce these effects directly or possible even indirectly through induction of leukocytosis. Neutrophils may be activated by seminal plasma or by infectious agents that may be present. Increased complement levels are associated with an inflammatory response that is frequently observed in the cervices of infertile women. It is conceivable that sperm immobilization seen is postcoital tests and *in vitro* assays is the result of antigen–antibody reaction against microorganism, with sperm an innocent bystander. It has been shown that microorganisms from seminal plasma or of cervical/vaginal origin bind to the sperm membrane and are transported through cervical mucus (Toth *et al.*, 1982).

The ideal assay for detection and quantification of sperm antibodies should be easy, rapid, repeatable, inexpensive, specific, quantitative, objective, and capable of detecting the immunoglobulin class of the antibody. Also, as individual sperm antigens are being identified, the assay may need to quantify antibodies against the different antigens. To date, none of the available methods possesses all these characteristics. The radiolabeled antiglobulin method is probably the most objective, and it is capable of detecting titers and class of antisperm antibodies. The method currently has several drawbacks: lack of stability of reagents, use of radiolabeled materials, the need for an expensive counter, and the need for centrifugation of sperm cells with the attendant pitfalls.

2. Clinical Applications of Antisperm Antibodies in the Diagnosis of Immunological Infertility

a. Antisperm Antibody Prevalence

Women. Among reports on infertile women, there exists a wide variation in the incidence of positive cases within each method (Table

IX). This may be due to the use of undiluted serum, the serum dilution considered positive by the different investigators, the criteria used in evaluating positive results, and the experimental design. Concerning the latter, evidently some studies have disregarded the fertility status of the partner when the couple was not being evaluated as a whole unit, while others have not included control groups. Likewise, racial background may influence results. High incidence of immobilizing antibody has been reported in infertile Japanese women (Isojima *et al.*, 1978). After the different variations have been taken into consideration, there remains a real difference in antibody prevalence between the infertile and the control groups. As mentioned above, a nonimmunoglobulin component in sera during pregnancy probably accounts for the high values in the pregnant group detected by agglutination method. Thus, Ingerslev and Hjort (1979) found that the majority of positive sera with H-H agglutinins detected by TAT from pregnant women and presumably fertile women was caused by a nonantibody component that could be removed from serum by absorption with seminal plasma.

Of all the techniques described, the SIT appears to be the most specific, with positive cases being limited to women from infertile couples. Usefulness of the RIA method has still to be confirmed. The least clinically useful method is the immunofluorescence technique; the results do not correlate with fertility status or with the other methods of antisperm antibody detection.

The emphasis has shifted from analysis only of serum samples to include analysis of cervical mucus for antisperm antibodies (Table X). With only one exception, the incidence of sperm-immobilizing activity was higher in mucus samples than in sera from infertile women. The results support the contention that the cervix is capable of antibody secretion independent of the serum. Detection of both serum and cervical mucus antibodies correlates with poor sperm penetration of the mucus. Antibodies that immobilize sperm or cause the shaking sperm phenomenon also correlate with poor mucus penetration (Jager *et al.*, 1978; Telang *et al.*, 1978; Menge *et al.*, 1982).

Men. Studies that evaluate antisperm antibodies in the serum and seminal plasma of men from infertile couples are summarized in Table XI. The GAT has been the most widely used method in analyzing sera of men. With this method, some fertile men were positive at low dilutions. Some of the variations in positive cases among reports are due to the selection of serum dilution considered significant. Moreover, data on control subjects have been omitted in several studies. Spermagglutinins in serum and seminal plasma samples of men are also readily detectable by the TAT. A direct comparison of GAT and TAT results revealed a high degree of

TABLE IX

Incidence of Women Positive for Serum Antisperm Antibodies[a]

Control % (number)	Unexplained infertility % (number)	Known cause of infertility % (number)	Pregnant % (number)	Reference
Tube-slide agglutination test				
9 (44)	67 (67)	16 (277)		Dukes and Franklin (1968)
20 (25)	38 (64)		34 (44)	Schwimmer *et al.* (1967)
0 (65)	14 (78)			Kolodny *et al.* (1971)
3 (113)	15 (389)[b]			Shulman *et al.* (1978)
8 (83)	38 (72)	19 (93)	46 (83)	Isojima *et al.* (1972)
	26 (310)[c]	17 (99)	18 (50)	Jones (1976)
	17 (304)	2 (550)	0 (221)	Mettler (1977)
	20 (51)	16 (37)	10 (48)	Petrunia *et al.* (1976)
	9 (70)	7 (154)		Hanafia *et al.* (1972)
	6 (487)	6 (489)		Lehmann *et al.* (1977)
Tray agglutination test				
	9 (58)			Mettler (1977)
	17 (134)[d]			Friberg (1974)
1 (150)	9 (326)[e]		6 (116)	Ingerslev and Hjort (1979)
	20 (698)[f]			Menge *et al.* (1982)
0 (150)	6 (95)	6 (231)	2 (116)	Ingerslev and Ingerslev (1980)
Gelatin agglutination test[b]				
3 (112)	18 (409)			Shulman *et al.* (1978)
11 (28)	34 (29)	25 (32)	38 (32)	Isojima *et al.* (1972)
	3 (153)			Schoenfeld *et al.* (1976)
0 (66)	1 (158)			Harrison (1968)
	8 (487)	6 (489)		Lehmann *et al.* (1977)
	2 (326)			Ingerslev and Ingerslev (1980)
	6 (1754)			Rumke *et al.* (1980)
	9 (172)			Moghissi *et al.* (1980)
Gelatin Agglutination and/or Immobilization Test				
	11 (377)			Ansbacher *et al.* (1971)
Sperm immobilization test				
0 (56)	16 (154)	1 (248)	0 (100)	Isojima *et al.* (1978)
	9 (310)	5 (99)	0 (50)	Jones (1976)
	6 (698)[b]			Menge *et al.* (1982)
0 (231)	2 (304)	0 (550)		Mettler (1977)
	10 (51)	5 (37)	0 (48)	Petrunia *et al.* (1976)
	3 (153)			Schoenfeld *et al.* (1976)
0 (66)	2 (158)			Harrison (1968)
	2 (487)	1 (489)		Lehmann *et al.* (1977)
	1 (95)	0 (231)		Ingerslev and Ingerslev (1980)
	2 (172)			Moghissi *et al.* (1980)

TABLE IX (*Continued*)

Control % (number)	Unexplained infertility % (number)	Known cause of infertility % (number)	Pregnant % (number)	Reference
Indirect immunofluorescence				
79 (80)	76 (100)		62 (66)	Hjort and Hansen (1971)
9 (80)	22 (100)[c]		5 (66)	
	46 (51)[b]	51 (37)	56 (48)	Petrunia *et al.* (1976)
30 (66)	28 (158)[g]			Harrison (1968)
	30 (64)[g]	51 (51)	20 (61)	Wall *et al.* (1975)
Radiolabeled antiglobulin test				
0 (53)	13 (318)			Haas *et al.* (1980)

[a] Serum was tested undiluted, except as noted. Reprinted from Menge (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.
[b] 1:4 dilution.
[c] 1:5 dilution.
[d] 1:8 dilution.
[e] 1:32 dilution.
[f] 1:16 dilution.
[g] 1:25 dilution.

association for sera with T-T agglutination, but as expected the GAT missed the sera with H-H agglutination (Hellema and Rumke, 1976). Two of the three studies using 1:16 dilution reported similar incidences of positive cases, whereas the third study had twice the incidence. As in women, the SIT gave rather consistent results with sera from infertile men and was negative for control men in one study, whereas indirect immunofluorescence is of limited value.

Two interesting points are also presented in Table XI: (1) Results of two studies indicated persistence of a high incidence of antisperm antibodies after vasovasostomy; and (2) the existence of spermagglutin antibody in seminal plasma is generally dependent on its occurrence in serum. Penetration of cervical mucus was significantly reduced by sperm from men demonstrating serum or seminal plasma antisperm antibodies.

b. Prognostic Significance of Antisperm Antibody Detection. An important aspect in evaluating methodology is the possible association of antibody titer with subsequent fertility. Only a limited number of such studies have been reported.

Women. Six studies utilizing four different methods for detection of sperm antibodies are reported in infertile women (Table XII). In four of the studies, a significant reduction in subsequent fertility was associated

TABLE X

Incidence of Sperm Antibodies in Serum and Cervical Mucus Samples of Infertile Women[a]

Method	Control women (% positive)			Infertile women (% positive)			Reference
	Number	Serum	Mucus	Number	Serum	Mucus	
Gelatin and/or tube-slide agglutination				148	26	5	Shulman *et al.* (1978)
Modified sperm immobilization					ND[b]	35	
Tray agglutination				172	9	14	Moghissi *et al.* (1980)
Sperm immobilization	36	0	0	39[c]	26	21	Canturia (1977)
				47	11	28	Isojima *et al.* (1978)
				24	8	21	Soffer *et al.* (1976)
				459	13	30	Menge *et al.* (1982)
				26[c]	ND	38	Dor *et al.* (1979)
				16[d]	ND	6	
				172	2	9	Moghissi *et al.* (1980)
	33	ND	3	58	7	17	Chen and Jones (1981)
Sperm cytotoxicity				59	8	20	Eyquem and D'Almeida (1973)
Indirect immunofluorescence					20	12	
Sperm immobilization		0	0		0	24	Menge *et al.* (1977)
Gelatin agglutination	15	0	ND	100	8	ND	
Indirect immunofluorescence (IgA)		7	7		22	32[e]	
Indirect immunofluorescence (IgG)		27	13		35	34[f]	

[a] Reprinted from Menge (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.

[b] ND, not done.

[c] Women selected on the basis of poor postcoital tests.

[d] Women with normal postcoital tests.

[e] 66% of positive samples from women with negative sera.

[f] 32% of positive samples from women with negative sera.

TABLE XI

Incidence of Sperm Antibodies in Sera and Seminal Plasma of Infertile Men[a]

Method	Serum titer-positive	Antibody-positive			Reference
		Fertile men % (number)	Infertile men % (number)	Seminal plasma % (number)[b]	
Gelatin agglutination	32	0 (416)	3 (2015)	69 (16)	Rumke and Hellinga (1959)
	64	1 (500)	4 (400)	71 (42)	Fjallbrant (1968)
	4	4 (207)	7 (657)	32 (44)	Husted (1975)
	4		9 (381)		Shulman *et al.* (1978)
	4		20 (600)		Schoenfeld *et al.* (1977)
	4		88 (16)[d]	64 (14)	Friedman (1978)
	25		3 (113)		Harrison (1968)
	4		4 (172)		Moghissi *et al.* (1980)
	4		16 (213)	85 (34)[e]	Hendry *et al.* (1982)
Tray agglutination	16		19 (134)	47 (85)	Friberg (1975)
	32		36 (80)	90 (29)	Kremer *et al.* (1977)
	16		15 (698)		Menge *et al.* (1982)
	16		69 (29)[d]	28 (29)	Linnet and Fogh-Andersen (1979)
	4		7 (180)	17 (12)	Marmar *et al.* (1980)
	4		70 (23)[d]	12 (16)	Marmar *et al.* (1980)
Tube-slide agglutination	4		5 (347)		Shulman *et al.* (1978)

(Continued)

TABLE XI (*Continued*)

Method	Serum titer-positive	Antibody-positive: Fertile men % (number)	Antibody-positive: Infertile men % (number)	Antibody-positive: Seminal plasma % (number)[b]	Reference
Sperm immobilization	4		6 (300)		Schoenfeld *et al.* (1977)
	4		7 (429)		Menge *et al.* (1982)
	25		3 (113)		Harrison (1968)
	4		2 (172)		Moghissi *et al.* (1980)
	4	0 (65)	3 (409)		Jones (1979)
Gelatin agglutination and/or immobilization	4		6 (377)		Ansbacher *et al.* (1971)
Indirect immunofluorescence	10	18 (207)[c]	22 (657)		Husted (1975)
	10	5 (100)	21 (63)		Wall *et al.* (1975)
	25		26 (113)		Harrison (1968)
Radiolabeled antiglobulin	1	0 (10)	7 (296)	72 (18)[f]	Haas *et al.* (1982)

[a] Reprinted from Menge (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.
[b] Samples from men with serum antibodies.
[c] Male blood donors.
[d] Men after vasovasostomy.
[e] Mixed antiglobulin reaction on sperm.
[f] Direct test on sperm.

TABLE XII

Incidence of Pregnancy in Women of Infertile Couples Subsequent to Antisperm Antibody Analysis[a]

Antibody method	Serum dilution-positive	Pregnancy incidence: Negative test % (number)	Pregnancy incidence: Positive test % (number)	Reference
Tube-slide agglutination	≥5	56 (232)	44 (55)	Jones (1979)
	≥5	43 (105)[b]	39 (31)[b]	
	≥1	27 (460)	22 (27)	Lehmann *et al.* (1977)
Tray agglutination	≥16	46 (301)	4 (75)	Menge *et al.* (1982)
Gelatin agglutination	≥4	27 (449)	24 (38)	Lehmann *et al.* (1977)
Immobilization	≥1	56 (265)	36 (22)	Jones (1979)
	≥1	45 (122)[b]	14 (14)[b]	
	≥1	27 (476)	18 (11)	Lehmann *et al.* (1977)
	≥4	40 (344)	12 (32)	Menge *et al.* (1982)
	≥4[c]	36 (146)	12 (65)	
	≥1[c]	29 (48)	10 (10)	Chen and Jones (1981)
Gelatin agglutination and immobilization	≥4 ≥1	52 (159)	42 (26)	Ansbacher *et al.* (1973)
Tube-slide agglutination and immobilization	≥1	28 (61)	15 (13)	Mettler (1977)
	≥1	35 (37)[b]	0 (11)[b]	
Passive hemagglutination and cytotoxicity[d]	≥8	16 (70)	0 (58)	Mathur *et al.* (1981a)

[a] Reprinted from Menge (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.

[b] Women from couples with 3 or more years of infertility.

[c] Immobilizing activity in cervical mucus.

[d] Serum plus cervical mucus result.

with positive antibody analysis. Jones (1979) and Mettler (1977) noted a significant influence of positive SIT results only in those women from infertile couples who had been followed for 3 or more years. The SIT results suggested a significantly lower incidence of pregnancy in women who have antibody in either serum or cervical mucus. An increased incidence of spontaneous abortion was also found in women with positive

TABLE XIII
Incidence of Pregnancy in Couples Subsequent to Antisperm Antibody Analysis in Men[a]

Antibody method	Serum dilution-positive	Pregnancy incidence: Negative test % (number)	Pregnancy incidence: Positive test % (number)	Reference
Gelatin agglutination	≥16		48 (31)	
	32–128		16 (58)	Rumke *et al.* (1974)
	≥256		12 (48)	
Tray agglutination	≥16	43 (302)	16 (74)	Menge *et al.* (1982)
	≥64	41 (334)	7 (42)	
Immobilization	≥4	39 (350)	19 (26)	Menge *et al.* (1982)
Gelatin agglutination and immobilization	≥4	52 (159)	31 (13)	Ansbacher *et al.* (1973)
Passive hemagglutination and cytotoxicity[b]	≥8	22 (51)	0 (77)	Mathur *et al.* (1981b)

[a] Reprinted from Menge (1980), by permission of the publisher, Elsevier Science Publishing Co., Inc.
[b] Serum plus seminal plasma results.

serum antibodies against sperm (Jones, 1976; Mathur *et al.*, 1981a; Menge *et al.*, 1982).

Men. The overall results in men appear clearer than in the women (Table XIII). Serum levels of antisperm antibodies were negatively associated with subsequent incidence of pregnancy in the couples. Sperm agglutinins of high titers were especially related to persistent infertility. Subsequent fertility was also significantly lower in situations where the man's sperm or the women's cervical mucus was found to give poor results in the S-CMPT (Menge *et al.*, 1982).

c. Response to Treatment with Corticosteroids. Additional support for the negative effect of antisperm antibodies on fertility is the improved fertility rate in men after corticosteroid treatment (Table XIV). While the average pregnancy rate for couples in which the men had antisperm antibodies was 17%, the rate for men after treatment with corticosteroids was 40% (Tables XIII and XIV). Treatment generally reduced antibody titers, but a convincing correlation between antibody titers and pregnancy rate was not observed. Other factors such as semen quality and amount of antibody entering the ejaculate may also play a role.

TABLE XIV

Fertility after Corticosteroid Treatment of Patients with Antisperm Antibodies

Reference	Treatment	Number treated[a]	(%)	Pregnancies	Follow-up period (months)
Alexander and Sampson (1981)	Prednisone 60 mg/d/week	14 M 6 W	8 dropped, 6 partial drop, 6 no drop	9 (45)	4
	None	20 C	Not determined	3 (15)	
Hendry *et al.* (1979)	Prednisone 15 mg/d/to 1 year	29 M		4 (14)	
Mathur *et al.* (1981a)	Prednisone 15 mg/d, 3–6 months	25 M	Antibody titers fell	9 (36)	
Hendry *et al.* (1981)	Methylprednisdone 96 mg/d/week, alternate months for 6 months	45 M	Little difference in titers between men successfully or unsuccessfully fathering children	14 (31)	
Shulman and Shulman (1982)	Methylprednisolone 96 mg/d/week with repeated treatment cycles	71 M	Pregnancy-positive group had greater incidence of decreased titers	31 (44)	12
Katz and Newill (1980)	Methylprednisolone 96 mg/d/week, repeated in alternate	7 M		4 (57)	
Hargreave and Elton (1982)	Methylprednisolone 96 mg/d/week/month for 6 months and/or betamethasone 0.5–2.0 mg/d/week repeated in alternate weeks for 6 months	33 M	Antibody titers fell but no significant association with pregnancy occurrence	6/33 (18%) 11/27 (41%)	6 12
	No treatment (antibody-negative)	304 M		61/304 (20%)	12

[a] M, men; W, women; C, couples.

B. Experimental Evidence

1. Effects of Immunization with Sperm, Testis, or Their Antigens on Female Reproductive Performance *in vivo*

Among 37 reports by 18 different authors (or laboratories) published since 1964, 33 (90%) reported reduced infertility following immunization with sperm, testis, or teratocarcinoma antigens. Infertility has been induced in many species, including rabbit, mouse, guinea pig, heifer, sheep, and baboon (reviewed in Tung, 1983).

Three different approaches have been adopted. First, females that had been immunized with sperm, testis, or teratocarcinoma antigens were mated or artificially inseminated following ovulation induction or were transplanted with embryo after being made pseudopregnant. Control animals were immunized with adjuvant, with saline, or with seminal plasma and adjuvant. Second, ejaculated or epididymal sperm were incubated with serum, cervical mucus, or their IgG, Fab (of IgG), or secretory IgA components from homologous or heterologous animals immunized with sperm or testis. The mixture was then artificially inseminated into nonimmunized animals in which ovulation had been induced. Third, antiserum to sperm antigens was transferred systematically into recipients at defined stages of pregnancy. This approach provides evidence that humoral antibodies are important pathogenetic mechanisms of infertility. For the last two approaches, control animals received preimmune sera, sera from animals immunized with adjuvant alone, or antisperm antisera that had been absorbed with sperm or testis.

It is clear that female infertility can be induced in rabbit, mouse, guinea pig, and heifer. In rabbit and mouse, which have been adequately evaluated, the three parameters that were most consistently reduced are fertilization rate, embryo survival, and litter size.

Three purified sperm-specific enzymes have been carefully evaluated for their antifertility properties. The enzyme LDH-C4 induced infertility in rabbit, mouse, and baboon. As one might expect, immunization with heterologous LDH-C4 resulted in a greater antibody response, and hence infertility, than immunization with homologous LDH-C4 (Goldberg *et al.*, 1981). It was found that the antifertility effects influenced both preimplantation and postimplantation events (Goldberg and Lerum, 1972). Although hyaluronidase readily induced antibody responses in homologous females, fertility reduction was not observed (Morton and McAnulty, 1979). Immunization with purified homologous acrosin in the sheep demonstrated at best a marginal effect on fertility rate (Morton and McAnulty, 1979). Recently, a cell-surface sialoglycoprotein (RSA-1) with a molecular weight of 15,000 was isolated from rabbit sperm by O'Rand and Porter

(1979). Monovalent (Fab) antibodies to this antigen prevented fertilization *in vitro* and reduced fertility *in vivo* (O'Rand, 1981).

2. *In Vitro* Effect of Autoantibodies, Isoantibodies, or Monoclonal Antibodies to Sperm Antigens on Prefertilization and Fertilization Events.

Bivalent or monovalent antibody to sperm can specifically inhibit fertilization in rabbit, guinea pig, hamster, and mouse. In studies where different steps of prefertilization and fertilization events were evaluated, all steps were inhibited (reviewed in Tung *et al.*, 1981a). The data strongly suggest that molecules involved in the cellular recognition during fertilization are sperm specific, and that many are autoantigens. Monoclonal antibodies to sperm antigens have been produced by many laboratories, and several of these interfered with fertilization *in vitro* (Bellve and Moss, 1983).

IV. SUMMARY AND CONCLUSIONS

The anatomy of the blood–testis barrier and genetic control of immune responses to aspermatogenic antigens are important factors that determine the development of autoimmune disease of the testis. There are regional differences in the barrier, so that the seminiferous tubules are not permeable to macromolecules whereas the rete testis and the ductus efferentes are permeable. Although there is ample evidence for the "foreign" immunogenicity of testis-specific autoantigens, these antigens need not be completely sequestered within the blood–testis barrier. If antigens in soluble form exist in the lumen of the spermatogenic compartment, trace amounts can reach the immune system, where they may elicit an antigen-mediated suppression of immune response to testis-specific autoantigens and prevent EAO development. On the other hand, autoimmune disease of the testis develops when animals are deliberately immunized with large doses of antigens, with chemically modified antigens, or with antigens in appropriate adjuvants. In vasectomy, autoimmune orchitis occurs when aspermatogenic antigens leak outside the barrier. In naturally occurring testicular autoimmunity, orchitis and autoimmune disease of other endocrine organs may coexist, and autoimmune orchitis can also be found in association with abnormal endocrinological control of the testis.

Studies on EAO in the guinea pig have revealed the complexity of the immunopathology and the multitude of pathogenic mechanisms operative in this disease. This is explicable in part by (1) the complex nature of genetic control of the disease, as evidenced in a recent study on murine

EAO, and (2) the unusually large number of aspermatogenic autoantigens that have been isolated from guinea pig testis and sperm. It is possible that the immune responses to different aspermatogenic antigens are under separate genetic controls.

Experimental autoimmune orchitis is not merely an experimental model, since naturally occurring autoimmune orchitis has been documented in several species. More importantly, immune complexes have recently been detected in the testes of infertile men. This finding suggests that aspermatogenesis associated with infertility in men may be the end stage of an earlier episode of immunological orchitis.

There is both clinical and experimental evidence that unexplained infertility may be associated with and caused by antisperm antibodies. Antibodies to surface antigens of sperm are found more frequently in infertile men and women than in fertile couples. Since antibodies are also detectable in the genital secretions, they may be produced by the secretory immune system and can reach the sites where prefertilization and fertilization events occur.

The clinical applications of antisperm antibodies have helped to identify a subgroup of men and women who have the clinical diagnosis of unexplained infertility. This group of patients appears to be less likely to become fertile spontaneously than infertile patients without antisperm antibodies. Preliminary studies have suggested that infertile men with antisperm antibodies may respond to treatment with corticosteroid.

Although the assays based on spermagglutination and complement-dependent sperm immobilization are subjective and not strictly quantitative, they are currently the most popular and probably the most reproducible among antisperm antibody assays. Nevertheless, continued efforts to improve these assays are clearly indicated.

ACKNOWLEDGMENTS

The study is supported by NIH grants HD 12247 and HD 14505 from the Institute of Child Health and Human Development, and by an NIH fellowship (HD 06515) awarded to KSKT.

REFERENCES

Alexander, N. J. (1972). *J. Reprod. Fertil.* **31,** 399–406.
Alexander, N. J., and Sampson, J. H. (1981). *J. Androl.* **2,** 33.
Anderson, D. J., and Alexander, N. J. (1983). *Fertil. Steril.* **40,** 557–571.
Ansbacher, R., Pangan, S. M., and Srivannaboon, S. (1971). *Fertil. Steril.* **22,** 298–302.
Ansbacher, R., Yeung, K. K., and Behrman, S. J. (1973). *Fertil. Steril.* **24,** 305–308.
Asherson, G. L. (1967). *Br. Med. Bull.* **23,** 24–29.

Barrera, C. N., Mazzolli, A. B., and Mancini, R. E. (1976). *Fertil. Steril.* **27,** 21–27.
Baum, J., Boughton, B., Mongar, J. L., and Schild, H. O. (1961). *Immunology* **4,** 95–110.
Beer, A. E., and Billingham, R. E. (1976). *In* "The Immunobiology of Mammalian Reproduction." Prentice-Hall, Englewood Cliffs, New Jersey.
Bellve, A. R., and Moss, S. B. (1983). *Biol. Reprod.* **28,** 1–26.
Bernard, C. C. A., Mitchell, G. F., Leydon, J., and Bargerbos, A. (1978). *Int. Arch. Allergy Appl. Immunol.* **56,** 256–263.
Bigazzi, P. E., Kosuda, L. L., Hsu, K. C., and Andres, G. A. (1976). *J. Exp. Med.* **143,** 382–404.
Bishop, D. W. (1961). *Proc. Soc. Exp. Biol. Med.* **107,** 116–120.
Bishop, D. W., Narbaitz, R., and Lessof, M. (1961). *Dev. Biol.* **3,** 444–485.
Boettcher, B., Hay, J., Kay, D. L., Baldo, B. A., and Roberts, T. K. (1970). *Int. J. Fertil.* **15,** 143–148.
Boettcher, B., Hjort, T., Rumke, P., Shulman, S., and Vyazov, O. E. (1977). *Acta Pathol. Microbiol. Scand., Suppl. No. 258,* pp. 1–69.
Bohme, D. (1965). *Immunol. Allerg. Forsch.* **128,** 31–51.
Bronson, R., Cooper, G. W., and Rosenfeld, D. L. (1982). *Am. J. Reprod Immunol.* **2,** 222–224.
Brown, P. C., and Glynn, L. E. (1969). *J. Pathol. Bacteriol.* **98,** 277–282.
Brown, P. C., and Glynn, L. E. (1973). *J. Pathol.* **111,** 77–83.
Brown, P. C., Glynn, L. E., and Holborow, E. J. (1963). *J. Pathol. Bacteriol.* **86,** 505–520.
Brown, P. C., Glynn, L. E., and Holborow, E. J. (1967). *Immunology* **13,** 307–314.
Brown, P. C., Dorling, J., and Glynn, L. E. (1972). *J. Pathol.* **106,** 229–233.
Cannat, A., and Seligmann, M. (1965). *Sem. Hop. Paris* **4,** 1090–1098.
Canturia, A. A. (1977). *Br. J. Obstet. Gynecol.* **84,** 865–870.
Carlo, D. J., Hagopian, A., Jackson, J. J., Limjuco, J. A., and Eylar, E. H. (1976). *J. Immunol.* **116,** 619–622.
Chen, C., and Jones, W. R. (1981). *Fertil. Steril.* **35,** 542–545.
Christensen, A. K. (1977). *In* "The Testis in Normal and Infertile Men" (P. Troen and H. R. Nankin, eds.), p. 444. Raven, New York.
Chutna, J., and Pokorna, Z. (1967). *Folia Biol. (Praha)* **13,** 68–78.
Chutna, J., and Rychlikova, M. (1964). *Folia Biol. (Praha)* **10,** 177–187.
Ciccareses, S., and Ohno, S. (1978). *Cell* **13,** 643–650.
Doniach, D., and Roitt, I. M. (1964). *Sem. Hematol.* **1,** 313–343.
Dooher, G. B., Artzt, R., Bennett, D., and Hurtenbach, U. (1981). *J. Reprod. Fertil.* **62,** 505–511.
Dor, J., Nebel, L. A., Soffer, Y., Mashiach, S., and Serr, D. M. (1979). *Int. J. Fertil.* **24,** 94–100.
Dukes, C. D., and Franklin, R. R. (1968). *Fertil Steril.* **19,** 263–271.
Dym, M., and Fawcett, D. W. (1970). *Biol. Reprod.* **3,** 308–326.
Dym, M., and Romrell, L. J. (1975). *J. Reprod. Fertil.* **42,** 1–7.
Ericsson, R. J. (1971). *Proc. Soc. Exp. Biol. Med.* **137,** 532–535.
Eyguem, A., and D'Almeida, M. (1973). *In* "Immunology of Reproduction" (K. Bratanov and R. G. Edwards, eds.), p. 344. Bulgarian Acad. of Sci., Sofia.
Fawcett, D. W. (1975). *In* "Male Reproductive System" (R. O. Greep and D. W. Hamilton, eds.), pp. 21–55. Am. Physiol. Soc., Bethesda, Maryland.
Fawcett, D. W., Heidger, P. J., and Leak, L. V. (1969). *J. Reprod. Fertil.* **19,** 109–119.
Fitch, W. M., and Atchley, W. R. (1985). *Science* **228,** 1169–1175.
Fjallbrant, B. (1968). *Acta Obstet. Gynecol. Scand.* **47** (Suppl. 4), 1–84.
Franklin, R. R., and Dukes, C. D. (1964). *Am. J. Obstet. Gynecol.* **89,** 6–9.

Freund, J., Lipton, M. M., and Thompson, G. E. (1953). *J. Exp. Med.* **97,** 711–725.
Freund, J., Thompson, G. E., and Lipton, M. M. (1955). *J. Exp. Med.* **101,** 591–603.
Friberg, J. (1975). *Acta Obstet. Gynecol. Scand., Suppl. No. 36,* pp. 1–76.
Friedberger, E., and Furstenheim, A. (1929). *Z. Immun. Forsch.* **64,** 294–319.
Friedman, S. (1978). *Andrologia* **10,** 251–252.
Fritz, T. E., Lombard, L. S., Tyler, S. A., and Norris, W. P. (1976). *Exp. Mol. Pathol.* **24,** 142–158.
Gilula, N. B., Fawcett, D. W., and Aoki, A. (1976). *Dev. Biol.* **50,** 142–168.
Goldberg, E., and Lerum, J. (1972). *Science* **176,** 686–687.
Goldberg, E., Wheat, T. E., Powell, J. E., and Stevens, V. C. (1981). *Fertil. Steril.* **35,** 214–217.
Goodman, M., Rosenbla, M., Gottleib, J. S., Miller, J., and Chen, C. H. (1963). *Arch. Gen. Psychiatry* **8,** 518–526.
Haas, C. G., Cines, D. B., and Schreiber, A. D. (1980). *N. Engl. J. Med.* **303,** 722–727.
Haas, C. G., Weiss-Wik, R., and Wolf, D. P. (1982). *Fertil. Steril.* **38,** 54–61.
Hagopian, A., Jackson, T., Carlo, D. J., Limjuco, G. A., and Eylar, E. H. (1975). *J. Immunol.* **115,** 1731–1743.
Hanafia, M. J., Epstein, J. A., and Sobrero, A. J. (1972). *Fertil. Steril.* **23,** 493–497.
Hargis, B. J., Malkiel, S., and Berkeha, J. (1968). *J. Immunol.* **101,** 374–376.
Hargreave, T. B., and Elton, R. A. (1982). *Fertil. Steril.* **38,** 586–590.
Harrison, R. F. (1968). *Int. J. Fertil.* **23,** 288–293.
Hattori, M., and Brandon, M. R. (1977). *In* "Immunological Influence on Human Fertility" (B. Boettcher, ed.), pp. 311–322. Academic Press, New York.
Hellema, H. W. J., and Rumke, P. (1976). *Fertil. Steril.* **27,** 284–292.
Hellema, H. W. J., and Rumke, P. (1978). *Clin. Exp. Immunol.* **31,** 1–11.
Hendry, W. S., Stendronska, J., Hughes, L., Cameron, K. M., and Pugh, R. C. B. (1979). *Lancet* **2,** 498–500.
Hendry, W. S., Stedronska, J., Parslow, J., and Hughes, L. (1981). *Fertil. Steril.* **36,** 351–355.
Hendry, W. F., Stedronska, J., and Lake, R. A. (1982). *Fertil. Steril.* **37,** 108–112.
Hermo, L., and Lalli, M. (1978). *Biol. Reprod.* **19,** 92–100.
Hjort, T., and Hansen, K. B. (1971). *Clin. Exp. Immunol.* **8,** 9–23.
Hjort, T., Husted, S., and Linnet-Jepsen, P. (1974). *Clin. Exp. Immunol.* **18,** 201–212.
Hojo, K. C., Hiramine, C., and Ishitaki, M. (1980). *J. Reprod. Fertil.* **59,** 113–123.
Huckins, C. (1978). *Anat. Rec.* **190,** 905–926.
Hurtenbach, U., Morgenstern, F., and Bennett, D. (1980). *J. Exp. Med.* **151,** 827–838.
Husted, S. (1975). *Int. J. Fertil.* **20,** 113–120.
Ingerslev, H. J., and Hjort, T. (1979). *Fertil. Steril.* **31,** 496–502.
Ingerslev, H. J., and Ingerslev, M. (1980). *Fertil. Steril.* **33,** 514–520.
Irvine, W. J., Chan, M. M., and Scarth, L. (1969). *Clin. Exp. Immunol.* **4,** 489–503.
Isojima, S., Tsuchiya, K., Tanaka, C., Naka, O., and Adachi, H. (1972). *Am. J. Obstet. Gynecol.* **112,** 199–207.
Isojima, S., Koyama, K., and Kamata, R. (1978). *In* "Immunology of Reproduction" (K. Bratanov, ed.), p. 104. Bulgarian Acad. of Sci., Sofia.
Jackson, J. J., Hagopian, A., Carlo, D. J., Limjuco, G. A., and Eylar, E. H. (1975). *J. Biol. Chem.* **250,** 6141–6150.
Jackson, J. J., Hagopian, A., Carlo, D. J., Limjuco, G. A., and Eylar, E. H. (1976). *Biochim. Biophys. Acta* **427,** 251–261.
Jager, S., Kremer, J., and Stochteren-Draaisma, T. (1978). *Int. J. Fertil.* **23,** 12–21.
Johnson, M. H. (1970a). *J. Pathol.* **101,** 129–139.

Johnson, M. H. (1970b). *J. Reprod. Fertil.* **22,** 119–127.
Johnson, M. H. (1970c). *J. Pathol.* **102,** 131–138.
Johnson, M. H. (1973). *Adv. Reprod. Physiol.* **6,** 279–324.
Johnson, M. H., and Setchell, B. P. (1968). *J. Reprod. Fertil.* **17,** 403–406.
Jones, W. R. (1976). *In* "Immunology of Human Reproduction" (J. S. Scott and W. R. Jones, eds.), pp. 375–400. Academic Press, New York.
Jones, W. R. (1979). *Med. J. Aust.* **2,** 188–192.
Jones, W. R., Ing, R. M. Y., and Hobbin, E. R. (1975). *In* "Development of Vaccine for Fertility Regulation" pp. 17–35. Scriptor, Copenhagen.
Kantor, G. L., and Dixon, F. J. (1972). *J. Immunol.* **108,** 329–338.
Katsh, S. (1959). *Am. J. Obstet. Gynecol.* **77,** 946–956.
Katsh, S. (1960). *Int. Arch. Allergy* **16,** 241–275.
Katsh, S., and Bishop, D. W. (1958). *J. Embryol. Exp. Morphol.* **6,** 94–104.
Katz, M., and Newill, R. (1980). *Lancet* **1,** 1306.
Kay, D. J., and Boettcher, B. (1980). *Arch. Androl.* **4,** 37–43.
Kierszenbaum, A. L. (1970). *Virchows Arch. B.* **5,** 1–2.
Kilbrick, S., Belding, D. L., and Merrill, B. (1952). *Fertil. Steril.* **3,** 430–438.
Kohno, S., Munoz, J. A., Williams, T. S., Teuscher, C., Bernard, C. C. A., and Tung, K. S. K. (1983). *J. Immunol.* **130,** 2675–2682.
Kolodny, R. C., Koehler, B. C., Toro, G., and Masters, W. H. (1971). *Obstet. Gynecol.* **38,** 576–582.
Koshimies, A. I., Kormano, M., and Lahti, A. (1971). *J. Reprod. Fertil.* **27,** 463–465.
Kremer, J. (1965). *Int. J. Fertil.* **10,** 209–215.
Kremer, J., and Jager, S. (1976). *Fertil. Steril.* **27,** 335–340.
Kremer, J., Jager, S., and Kuiken, J. (1977). *In* "The Uterine Cervix in Reproduction" (V. Insler and G. Bettendorf, eds.), pp. 181–186. Thieme, Stuttgart.
Landsteiner, K. (1899). *Zbe. Bakt.* **25,** 546–549.
LeBouteiller, P., Toullet, F., and Voissin, G. A. (1975). *Immunology* **28,** 983–999.
Lehmann, F., Stripling, K., Budel, B., Kreb, D., and Masson, D. (1977). *In* "The Uterine Cervix in Reproduction" (V. Insler and G. Bettendorf, eds.), pp. 204–211. Thieme, Stuttgart.
Lewi, S. (1968). *Geront.* **14,** 160–173.
Linnett, L., and Fogh-Andersen, P. (1979). *J. Clin. Lab. Immunol.* **2,** 245–248.
Lipscomb, H. L., Gardner, P. J., and Sharp, J. G. (1979). *J. Reprod. Immunol.* **1,** 209–217.
Mancini, R. E., Vilar, O., Alvarez, B., and Seigure, A. C. (1965). *J. Histochem.* **13,** 376–385.
Mancini, R. E., Monastirsky, R., Fernandez Collazo, E., Seiguer, A. C., and Alonso, A. (1969). *Fertil. Steril.* **20,** 779–798.
Mangione, C. M., Medley, N. E., and Menge, A. C. (1981). *Int. J. Fertil.* **26,** 20–24.
Marcus, Z. H., Nebel, L., Soffer, Y., Stahl, Y., and Toullet, F. (1975). *Fertil. Steril.* **26,** 1024–1034.
Marmar, J. L., Praiss, D. E., and DeBaneductis, T. J. (1980). *Fertil. Steril.* **34,** 365–368.
Mathur, S., Williamson, H. O., Derrick, F. C., Madyostha, P. R., Melchers, J. T., Holtz, G. L., Baker, E. R., Smith, C. L., and Fudenberg, H. H. (1981a). *J. Immunol.* **126,** 905–907.
Mathur, S., Baker, E. R., Williamson, H. O., Derrick, F. C., Teague, K. J., and Fudenberg, H. H. (1981b). *Fertil. Steril.* **36,** 486–495.
Mazzolli, A. B. (1971). *J. Reprod. Fertil.* **26,** 161–166.
Meng, A. L., and Tung, K. S. K. (1983). *J. Reprod. Fertil.* **69,** 279–288.
Menge, A. C. (1980). *In* "Immunological Aspects of Infertility and Fertility Regulation" (D.

S. Dhindsa and G. F. B. Schumacher), pp. 205–224. Elsevier-North-Holland, New York.
Menge, A. C., and Black, C. E. (1979). *Fertil. Steril.* **32,** 214–218.
Menge, A. C., Schwanz, M. L., Riolo, R. L., Greenberg, V. N., and Neda, T. (1977). *In* "The Uterine Cervix in Reproduction" (V. Insler and G. Bettendorf, eds.), pp. 221–230. Thieme, Stuttgart.
Menge, A. C., Medley, N. E., Mangione, C. M., and Dietrich, J. W. (1982). *Fertil. Steril.* **38,** 439–446.
Metalnikoff, S. (1900). *Ann. Inst. Pasteur* **14,** 577–589.
Metchnikoff, E. (1899). *Ann. Inst. Pasteur* **13,** 734–769.
Mettler, L. (1977). *In* "Fortschritte der Fertilitätsforschung" Vol. IV, pp. 1–121. Grosse Verlag, Berlin.
Miller, S. C., Bowman, B. M., and Roberts, L. K. (1984). *J. Leuk. Biol.* **36,** 679–687.
Moghissi, K. S., Sacco, A. G., and Borin, K. (1980). *Am. J. Obstet. Gynecol.* **136,** 941–950.
Morris, B. (1968). *Rev. Fr. Hematol.* **8,** 525–534.
Morton, D. B., and McAnulty, P. A. (1979). *J. Reprod. Immunol.* **1,** 61–73.
Mostofi, F. J. (1977). *In* "Pathology" (W. A. D. Anderson and J. M. Kissane, eds.), 7th ed., pp. 1013–1037. Mosby, St. Louis, Missouri.
Muir, V. Y., Turk, J. L., and Hanley, H. G. (1976). *Clin. Exp. Immunol.* **24,** 72–80.
Munoz, J. J., and Arai, H. (1982). *In* "Seminars in Infectious Diseases" (J. B. Robbins, J. C. Hill, and J. C. Sadoff), Vol. IV, pp. 395–400. Thieme-Stratton, New York.
Nagano, T., and Okumura, K. (1973a). *Virchows, Arch. B* **14,** 237–246.
Nagano, T., and Okumura, K. (1973b). *Virchows. Arch. B* **14,** 223–235.
Neaves, W. B. (1977). *In* "The Testis" (A. D. Johnson, W. R. Gomes, and N. L. Van Demark, eds.), Vol. IV, pp. 125–162. Academic Press, New York.
O'Rand, M. G. (1981). *Biol. Reprod.* **25,** 621–628.
O'Rand, M. G., and Porter, J. P. (1979). *J. Immunol.* **122,** 1248–1254.
Paulson, J. D., and Polakoski, K. L. (1977). *Fertil. Steril.* **28,** 179–181.
Petrunia, D. M., Taylor, P. J., and Watson, J. I. (1976). *Fertil. Steril.* **27,** 655–661.
Phadke, A. M., and Padukone, K. (1964). *J. Reprod. Fertil.* **7,** 164–170.
Pokorna, Z. (1970a). *Folia Biol. (Praha)* **16,** 320–329.
Pokorna, Z. (1970b). *Folia Biol. (Praha)* **16,** 222–224.
Pokorna, Z., and Vojtiskova, M. (1964). *Folia Biol. (Praha)* **10,** 261–267.
Pokorna, Z., Vojtiskova, M., Rychlikova, M., and Chutna, J. (1963). *Folia Biol. (Praha)* **9,** 203–209.
Raitsina, S. S., and Nilovsky, M. N. (1967). *Folia Biol. (Praha)* **13,** 450–456.
Rappaport, F. T., Sampath, A., Kano, K., McCluskey, R. T., and Milgrom, F. (1969). *J. Exp. Med.* **130,** 1411–1425.
Rivenzon, A., Rivenzon, M., and Madden, R. E. (1974). *Cell. Immunol.* **14,** 411–416.
Rose, N. R., Hjort, T., Rumke, P., Harper, M. J. K., and Vyazov, O. (1976). *Clin. Exp. Immunol.* **23,** 175–199.
Rumke, P., and Hellinga, G. (1959). *Am. J. Clin. Pathol.* **32,** 357–363.
Rumke, P., Van Amstal, N., Messer, R. N., and Begemer, P. D. (1974). *Fertil. Steril.* **25,** 393–398.
Rumke, P., Renckens, C. N. M., Van Amstel, N., and Bezemer, P. D. (1980). *Proc. 1st Congr. Int. Soc. Immunol. Reprod.*, Abstr. 5.
Salomon, F., Saremaslani, P., Jakob, M., and Hedinger, C. E. (1982). *Lab. Invest.* **47,** 555–567.
Samuel, T., Kolk, A. H. J., Rumke, P., and Van Lis, J. M. (1975). *J. Clin Exp. Immunol.* **21,** 65–74.

Sato, K., Kirokawa, K., and Hatakeyama. (1981). *Virch. Arch. (Pathol. Anat.)* **392,** 147–158.
Schoenfeld, C., Amelar, R. D., and Dublin, L. (1976). *Fertil. Steril.* **27,** 1199–1203.
Schoenfeld, C., Amelar, R. D., and Dublin, L. (1977). *Arch. Androl.* **1,** 111–114.
Schwimmer, W. B., Ustay, K. A., and Behrman, S. J. (1967). *Fertil. Steril.* **18,** 167–180.
Seeger, R. C., and Oppenheim, J. J. (1970). *J. Exp. Med.* **132,** 44–65.
Setchell, B. (1967). *J. Physiol. (London)* **189,** 63–65.
Shephard, R. W., Anderson, D. J., and DeWolf, W. C. (1982). *Surg. Forum* **32,** 600–602.
Shulman, J. F., and Shulman, S. (1982). *Fertil. Steril.* **38,** 591–599.
Shulman, S. (1975). "Reproduction and Antibody Responses," CRC Press, Cleveland, Ohio.
Shulman, S., Guerrero, M., and Sudo, N. (1978). *In* "Proceedings of the Fourth European Sterility Congress" (J. Cortex-Priete and K. Semm, eds.), pp. 529–535. DeProduccion, Madrid.
Soffer, Y., Marcus, Z. H., Bukowsky, I., and Caspi, E. (1976). *Int. J. Fertil.* **21,** 89–95.
Sotsiou, F., Bottazzo, G. F., and Doniach, D. (1980). *Clin. Exp. Immunol.* **39,** 97–111.
Stone, S. H., Lerner, E. M., II, Goode, J. H., Jr. (1969). *In* "International Convocation on Immunology" (N. R. Rose and F. Milgrom, eds.), pp. 339–341. Karger, Basel.
Suzuki, F., and Nagano, T. (1978). *Anat. Rec.* **191,** 503–520.
Taguchi, O., and Nishizuka, Y. (1981). *Clin. Exp. Immunol.* **46,** 425–434.
Telang, M., Reyniak, J. V., and Shulman, S. (1978). *Int. J. Fertil.* **23,** 200–206.
Teuscher, C., Wild, G. C., Johnson, E., and Tung, K. S. K. (1981). *Ricerca Clin. Lab.* **11,** 313–329.
Teuscher, C., Wild, G. C., and Tung, K. S. K. (1983a). *J. Immunol.* **130,** 317–322.
Teuscher, C., Wild, G. C., and Tung, K. S. K. (1983b). *J. Immunol.* **130,** 2683–2688.
Teuscher, C., Smith, S. M., Goldberg, E. H., Shearer, G. M., and Tung, K. S. K. (1985a). *Immunogenetic.* (In press).
Teuscher, C., Potter, M., and Tung, K. S. K. (1985b). (In press).
Thomsen, O., and Kettel, K. (1929). *Z. Immun. Forsch.* **63,** 67–93.
Toder, V., Marcus, Z. H., and Nebel, L. (1975). *Eur. J. Immunol.* **5,** 680–683.
Toth, A., O'Leary, W. M., and Ledger, W. (1982). *Obstet. Gynecol.* **59,** 556–559.
Toullet, F., and Voisin, G. A. (1974). *J. Reprod. Fertil.* **37,** 299–313.
Toullet, F., and Voisin, G. A. (1976). *Clin. Exp. Immunol.* **26,** 549–562.
Toullet, F., Voisin, G. A., and Nemirovsky, G. (1973). *Immunology* **24,** 635–753.
Toullet, F., Audibert, L., Chedid, L., and Voisin, G. A. (1974). *Ann. Immunol. Pasteur, Paris* **125C,** 901–910.
Toullet, F., Audibert, F., Voisin, G. A., and Chedid, L. (1977). *Ann. Immunol. Pasteur, Paris* **128C,** 267–269.
Tuck, R. R., Setchell, B. P., Waites, G. M. H., and Young, J. A. (1970). *Pfuegers Arch.* **138,** 225–243.
Tung, K. S. K. (1975). *Clin. Exp. Immunol.* **20,** 93–104.
Tung, K. S. K. (1976). *Clin. Exp. Immunol.* **24,** 292–299.
Tung, K. S. K. (1977). *In* "Immunobiology of Gametes" (M. Edidin and M. H. Johnson, eds.), pp. 157–180. Cambridge Press, Cambridge.
Tung, K. S. K. (1978). *Science* **201,** 833–835.
Tung, K. S. K. (1980). *In* "Immunological Aspects of Infertility and Fertility Regulation" (D. S. Dhindsa and G. F. B. Schumacher, eds.), pp. 33–91. Elsevier-North-Holland, New York.
Tung, K. S. K. (1983). *In* "Immunology of Reproduction" (T. G. Wegman and T. J. Gill III, eds.), pp. 389–423. Oxford Univ. Press, London and New York.

Tung, K. S. K., and Alexander, N. J. (1980). *Am. J. Pathol.* **101,** 17–29.
Tung, K. S. K., and Woodroffe, A. J. (1978). *J. Immunol.* **120,** 320–328.
Tung, K. S. K., Unanue, E. R., and Dixon, F. J. (1970). *Am. J. Pathol.* **60,** 313–324.
Tung, K. S. K., Unanue, E. R., and Dixon, F. J. (1971a). *J. Immunol.* **106,** 1463–1472.
Tung, K. S. K., Unanue, E. R., and Dixon, F. J. (1971b). *Int. Arch. Allergy Appl. Immunol.* **41,** 565–574.
Tung, K. S. K., Unanue, E. R., and Dixon, F. J. (1971c). *J. Immunol.* **106,** 1453–1462.
Tung, K. S. K., Cook, W. D., McCarty, T. A., and Robitaille, P. (1976). *Clin. Exp. Immunol.* **25,** 73–80.
Tung, K. S. K., Leong, C., and McCarty, T. A. (1977). *J. Immunol.* **188,** 1774–1779.
Tung, K. S. K., Han, B. L.-P., and Evan, A. P. (1979). *Dev. Biol.* **68,** 224–238.
Tung, K. S. K., Teuscher, C., and Meng, A. C. (1981a). *Immunol. Rev.* **55,** 217–255.
Tung, K. S. K., Ellis, L. E., Teuscher, C., Meng, A., Blaustein, J. C., Kohno, S., and Howell, R. (1981b). *J. Exp. Med.* **154,** 1016–1032.
Tung, K. S. K., Ellis, L. E., Child, G. V., and Dufau, M. (1984). *Endocrinology (Baltimore)* **114,** 922–929.
Tung, P. S., and Fritz, I. B. (1978). *Dev. Biol.* **64,** 297–315.
Vilar, O. (1971). *In* "Progresos en Endocrinologia" (L. Cardonet and L. Lencioni, eds.), p. 78. Editorial Panamaricana, Buenos Aires.
Voisin, G. A., and Toullet, F. (1968). *Am. Inst. Pasteur* **114,** 727–755.
Voisin, G. A., Delauney, A., and Barber, M. (1951). *Ann. Inst. Pasteur* **81,** 48–63.
Waksman, B. H. (1959). *J. Exp. Med.* **109,** 311–324.
Wall, J. R., Stedronska, J., David, R. D., Harrison, R. F., Goriup, D., and Lessof, M. H. (1975). *Fertil. Steril.* **26,** 1035–1041.
Wekerle, H. (1977). *Nature (London)* **267,** 357–358.
Wekerle, H. (1978a). *Eur. J. Immunol.* **8,** 294–302.
Wekerle, H. (1978b). *J. Exp. Med.* **147,** 233–250.
Wekerle, H., and Begemann, M. (1976). *J. Immunol.* **116,** 159–161.
Willson, J. T., Jones, N. A., and Katsh, S. (1973). *Anat. Rec.* **176,** 85–100.
Zappi, E., Nemirovsky, M., and Shulman, S. (1974). *Immunology* **26,** 477–488.

CHAPTER 22

Autoimmune Diseases of Muscle: Myasthenia Gravis and Myositis*

ROGER L. DAWKINS AND MICHAEL J. GARLEPP
Departments of Clinical Immunology
Royal Perth Hospital
and The Queen Elizabeth II Medical Centre
Perth, Western Australia, Australia

* Publication 8348 of the Departments of Clinical Immunology, Royal Perth Hospital, and The Queen Elizabeth II Medical Centre, Perth, Western Australia 6000.

INTRODUCTION

There are at least two major types of autoimmune disease of skeletal muscle. Myasthenia gravis (MG) is characterized by reversible fatiguability generally due to antiacetylcholine-receptor antibody (anti-AChR). By contrast, polymyositis refers to a group of inflammatory disorders associated with fixed weakness of proximal musculature. Whereas MG is due to a disturbance of neuromuscular transmission, polymyositis appears to be the consequence of necrotizing myositis resulting in loss of muscle fibers per se.

In this chapter we will compare and contrast these two different processes. It should be emphasized that the two can sometimes coexist, and that recent developments have suggested that there may be merit in considering autoimmune muscle disease as an extremely heterogeneous group of overlapping disorders. There may be many reactions that can affect skeletal muscle to produce rather similar clinical manifestations. Ultimately it should be possible to identify these different reactions and to treat the conditions accordingly.

II. CLASSIFICATION OF AUTOIMMUNE DISEASES OF MUSCLE

On purely clinical grounds it is possible to recognize MG because of the characteristic nature and distribution of weakness and because of its associated features. It is clear that there are several forms, and various clinical classifications have been proposed over the years. Restricted ocular MG (ROMG) appears to be different from the generalized form (GMG). The neonatal type (NMG) is clearly distinct from congenital MG in terms of age of onset, duration, distribution, and associations. Such clinical observations enabled Simpson (1960) to develop these classifications and to predict that GMG and NMG would prove to be due to an antibody that inhibited neuromuscular transmission. The antistriational autoantibody (AStr) discovered at the same time did not appear to be important in terms of inducing neuromuscular block, and in fact it proved to be associated with thymoma rather than MG per se. The major clarification came with the demonstration of anti-AChR experimentally (Patrick and Lindstrom, 1973). Subsequently it was shown that anti-AChR antibody was present in most patients with GMG and NMG, but not in patients with some of the other forms of myasthenia.

Similarly, several classifications of polymyositis and dermatomyositis have been proposed, and clinical observations make it clear that there are at least several subtypes. For example, juvenile dermatomyositis appears

to be quite different from the usual adult form. Again an experimental model was useful in providing insights. The demonstration that cell-mediated immunity (CMI) is important in experimental allergic myositis (EAM) paved the way for the demonstration of lymphocytic-mediated myotoxicity in some but not all adult patients with polymyositis (Dawkins and Mastaglia, 1973). A minority of patients with MG had similar findings. Characterization of the autoantibodies associated with polymyositis and mixed connective tissue disease (MCTD) (Sharp *et al.*, 1972) led to the identification of a subset of patients with polymyositis and antiribonucleoprotein (RNP). Accordingly, it is now possible to define the relationship between polymyositis and similar multisystem autoimmune diseases by the use of serology, and it is also possible to recognize some subtypes of polymyositis.

III. EXPERIMENTAL MODELS

A. Experimental Autoimmune Myasthenia Gravis

With the purification of AChR from *Torpedo* and *Electrophorus* spp., it became possible to produce antibodies to the AChR in experimental animals. In 1973 Patrick and Lindstrom reported the development of experimental autoimmune myasthenia gravis (EAMG) in rabbits immunized with purified electrophorus AChR. These animals developed a flaccid paralysis that could be temporarily reversed by anticholinesterase. Anti-AChR were demonstrated using an assay that took advantage of the exquisite specificity and high affinity of the α neurotoxin from the elapid snake *Bungarus multicinctus* when reacted with AChR. Radioactively labelled α-bungarotoxin was used to tag the detergent-solubilized AChR, and anti-AChR, bound to this complex was precipitated using antiimmunoglobulin.

Subsequently several groups have reported the induction of EAMG in a variety of animals including rabbits, mice, goats, and monkeys using AChR from either *Torpedo* or *Electrophorus* spp.; EAMG has also been induced in rats using purified syngeneic AChR. Anti-AChR have been shown to be primarily responsible for the muscle weakness in these animals. In rats at least there is a transient acute inflammatory phase associated with muscle weakness (for reviews, see Lindstrom, 1979; Vincent, 1980).

The induction of EAMG in mice has provided an opportunity to examine the effects of *H-2* and immunoglobulin allotype in this model. Berman *et al.* (1981) have shown that genes associated with both of these systems

help determine susceptibility to EAMG. Christadoss *et al.* (1981) have also demonstrated that genes in the *I* region of *H-2* influence the lymphoproliferative response to AChR. Interestingly, EAMG in mice can be treated, or its induction prevented, by administration of anti-*Ia* sera (Waldor *et al.*, 1983).

B. Experimental Autoallergic Myositis

Generalized myositis may be induced in experimental animals by immunization with skeletal muscle or fractions of skeletal muscle in complete Freund's adjuvant (Dawkins, 1975). These animals develop an uniphasic disease characterized by segmental necrosis of skeletal muscle. Lymphocyte-mediated cytotoxicity to cultured muscle has been demonstrated and is thought to be the pathogenic mechanism in EAM. Antibodies reactive with contractile proteins of skeletal muscle (e.g., actomyosin) are also demonstrable in sera (Dawkins, 1975). These antibodies resemble those found in patients with MG and a thymoma. Interestingly, they are not found in patients with isolated polymyositis.

Inflammatory muscle disease may also be induced experimentally by means of a number of viruses including Semliki Forest virus, encephalomyocarditis virus, Ross River virus, and Coxsackie viruses (Whitaker, 1982). Polymyositis induced by Coxsackie B1 virus in mice is strain dependent (Ray *et al.*, 1979) and may provide a model to explore the immunogenetic basis of some forms of polymyositis.

C. Spontaneous Myasthenia Gravis and Myositis in Animals

The breadth and complexity of autoimmune muscle disease is illustrated by the spontaneous disease that occurs in the African *Mastomys*. These animals appear to be predisposed to several autoimmune diseases and thymoma. In some respects their muscle disease resembles MG *plus* polymyositis, and in this regard they may be similar to the occasional patient with coexistent thymoma, MG, and polymyositis.

Idiopathic canine MG has the potential to be even more valuable than EAMG in studying the genetics and pathogenic mechanisms of MG. The affected dogs exhibit the same clinical, histopathological, and autoimmune features as seen in the human disease, and each of the various subgroups of disease is represented (i.e., congenital, late onset thymoma, etc.) (Garlepp *et al.*, 1979; Palmer, 1980).

IV. CLINICAL FEATURES

A. Presenting Features

As with other forms of myopathy, the presenting feature of autoimmune muscle disease will generally be weakness. The nature of the weakness will largely determine the clinician's approach.

If there is excessive fatiguability, that is, reversible weakness that improves after rest, the differential diagnosis will include disorders of the neuromuscular junction such as MG. Clinical testing is generally sufficient to demonstrate deterioration with exercise and improvement after rest, but it is often useful to show that the patient improves after anticholinesterase therapy. Electromyography may also assist in that it may be possible to demonstrate a decremental response and to distinguish this change from that which occurs in the Eaton–Lambert syndrome and other disorders associated with myasthenia.

In other patients, the weakness may be "fixed" in that it remains relatively constant throughout the day and from day to day. In such patients one can generally assume that there is an abnormality of the muscle fiber per se rather than of the neuromuscular junction. Other possibilities, such as denervation, need to be excluded. Evidence in favor of myofibril necrosis might be provided by demonstration of elevated serum creatine kinase activity and by biopsy.

The distribution of weakness is also very useful. In GMG, extraocular, facial, bulbar, and proximal musculature is generally involved. In polymyositis extraocular muscles are spared, but proximal and truncal muscles are usually affected.

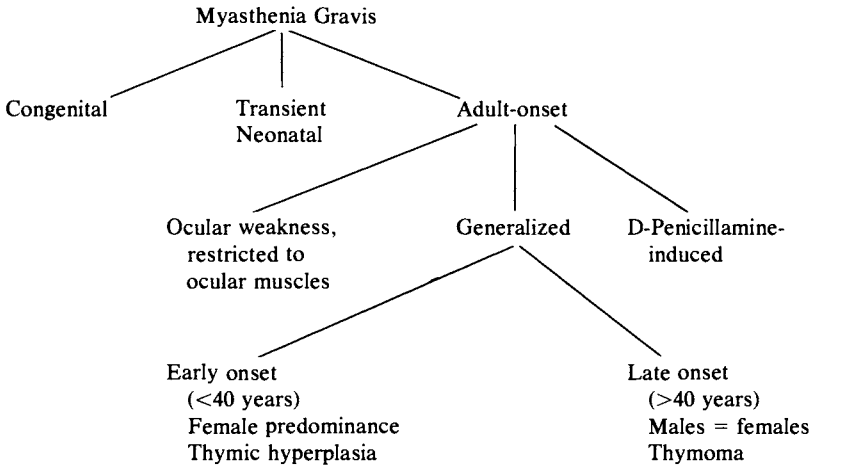
Other symptoms and signs of autoimmune muscle disease depend on the nature of the lesion. If there is an inflammatory component, the patient may complain of muscle pain and tenderness, but these features are generally more pronounced in the presence of synovitis, tenosynovitis, or interstitial inflammation.

Associated features are also important. For example the characteristic dermal involvement of dermatomyositis is almost pathognomonic. The diagnostic criteria for MG, polymyositis, and MCTD have been discussed elsewhere (Dawkins, 1983).

B. Clinical Subclassifications

Patients who fulfil the clinical criteria for MG may be subclassified in a number of ways. They may be classified according to the severity of their

TABLE I
Subgroups of Myasthenia Gravis[a]



[a] Patients suffering from D-Penicillamine-induced MG may present with ocular or generalized muscle involvement but are almost invariably anti-AChR positive. Several forms of congenital, familial, and juvenile myasthenia have been described.

disease (Osserman, 1958), but often pass from one group to another as exacerbations or remissions occur. Division may also be made according to the muscle groups affected, that is, ocular (OMG) versus generalized. Although in many patients the disease does progress from ocular to generalized skeletal muscle involvement, in some weakness remains confined to the extraocular muscles throughout the disease course (Drachman, 1978). The term restricted OMG is only appropriate after observation for ~2 years.

Another classification is based on the age of onset of disease (Table I). A number of syndromes of myasthenia has been described during infancy and childhood (Fenichel, 1978; Engel *et al.*, 1981a). Patients with adult-

TABLE II
Clinical Subgroups of Polymyositis

Adult polymyositis/dermatomyositis
Juvenile dermatomyositis
Polymyositis associated with C2 deficiency
D-Penicillamine-induced polymyositis
Mixed connective tissue disease

TABLE III
Differential Diagnosis of Inflammatory Muscle Disease

Polymyositis
Dermatomyositis
Mixed connective tissue disease
Systemic lupus erythematosus
Rheumatoid arthritis
Progressive systemic sclerosis
Myasthenia gravis and myasthenic myopathy
Infectious and parasitic myositis
Muscular dystrophy
Endocrine myopathy
Metabolic myopathy
Drug and toxic myopathies
Neurogenic atrophy
Polymyalgia rheumatica
Granulomatous myopathy
Carcinomatous myopathy

onset generalized MG (GMG) may then be grouped according to other characteristics. The characteristic patient in one group with idiopathic GMG is female, with disease onset prior to 40 years of age and an associated hyperplastic thymus. The other subgroup is characterized by later age of onset and the presence of a thymoma. D-Penicillamine (D-P)-induced MG has been included as a separate subgroup based upon its unique immunological and immunogenetic features (Garlepp *et al.*, 1983).

The clinical syndrome of polymyositis has been variously subclassified according to age of onset, skin involvement, and associated features such as C2 deficiency (Bohan and Peter, 1975; Mastaglia and Walton, 1982; Dawkins, 1983). The inclusion of D-P-induced polymyositis is again based on etiological and immunogenetic features (Table II). The diseases that should be considered in the differential diagnosis of polymyositis are shown in Table III. Polymyositis may occur together with several of these conditions, but inflammatory changes in muscle may also be seen that do not justify classification as polymyositis.

V. HISTOPATHOLOGY

A. Myasthenia Gravis

In MG, the organs that display notable histopathological features are thymus and skeletal muscle. Approximately 90% of patients have thymic

abnormalities. Germinal center formation is present in ~60% and thymoma in ~30%. Indirect immunofluorescent staining of frozen sections of thymus using AStr antibody usually reveals myoid cells in the thymus of those patients with hyperplasia, but these are rarely seen in those patients with thymoma. Histopathological examination of skeletal muscle sections in MG occasionally reveals accumulations of mononuclear cells, predominantly lymphocytes, in the interstitium. These have been referred to as "lymphorrhages," are most often found in those patients with thymoma, and can be shown to be associated with focal myofiber necrosis. On occasions these changes are of the same degree as seen in polymyositis.

Electron microscopic examination of the motor endplate in MG reveals simplification of the postjunctional folds and widening of the synaptic space which may contain fragments of degenerated postsynaptic membrane. Immunoglobulin and complement are said to be bound to the postsynaptic membrane, but the presynaptic membrane appears intact (Engel *et al.*, 1981b).

B. Polymyositis

The key features of polymyositis are segmental myofiber necrosis and inflammatory infiltration. The extent of these changes varies from patient to patient, and some muscle biopsies will prove to be negative, possibly reflecting a sampling problem.

VI. IMMUNOLOGY

A. Immunological Features

Anti-AChR are found in the serum of >90% of patients with active adult-onset GMG. The prevalence of anti-AChR varies in the subgroups of MG (see Chapter 24). Myasthenic syndromes with onset in early childhood have been shown to have diverse etiologies that only rarely are due to anti-AChR (Engel *et al.*, 1981a). By contrast, transient neonatal MG is due to the transfer of anti-AChR across the placenta.

The difficulty in detecting anti-AChR in ROMG might suggest qualitative differences from GMG. Comparison of the autoantibody profiles in ROMG and GMG shows an increase of thyroid autoantibodies and a decrease of antinuclear antibodies in the former (Table IV). Furthermore, the thyrotoxicosis found in MG is associated with ROMG rather than GMG (Garlepp *et al.*, 1981a). It is probable that ROMG is an autoimmune disease mediated by an antibody to extraocular muscle that is not readily

TABLE IV
Autoantibodies in Subgroups of Myasthenia Gravis (%)

Autoantibodies	Active adult-onset (generalized) $n = 72$	Restricted ocular $n = 61$	Juvenile/ congenital $n = 35$	Healthy controls $n = 200$
Anti-AChR	91	21	29	<1
ANF	26	10	3	4
Antithyroid	10	39	9	7

detectable in the current assay (see Chapter 24). Some published data suggest unique antigens on extraocular muscle (Vincent and Newsom-Davis, 1980; Kodama *et al.*, 1982).

Antistriational antibodies are found in some 30% of patients with MG and almost invariably in those patients with MG with a thymoma. The specificity of these AStr antibodies, as reflected by their immunofluorescent staining pattern, varies from patient to patient (Peers *et al.*, 1977). Efforts to identify the antigenic targets have shown that the antigen is rarely a single contractile protein but may result from the interaction of components of the sarcomere. It is notable that AStr antibodies are induced in ~20% of patients treated with D-P, although there is currently no evidence to indicate a thymic abnormality in these patients. Lymphocytotoxic antibodies may be seen in up to 50% of patients with MG, while anti-DNA may be elevated in the absence of active systemic lupus erythematosus (SLE).

In polymyositis, the most notable serological feature is the presence of one or more antibodies to nuclear antigens. Although antinuclear antibodies as detected by immunofluorescence are relatively infrequent (16–35%) and of low titer, many sera have antibodies reactive with one or more saline extractable nuclear antigens that have been detected by immunodiffusion in agarose (Table V). Antibodies to Jo-1 and PM-1 may be specific for polymyositis and dermatomyositis, while antibodies to Mi are said to be found only in dermatomyositis. Antibodies to the Ku antigen may identify the combination of polymyositis and progressive systemic sclerosis (Mimori *et al.*, 1981). High titers of antibodies to RNP have been used as part of the definition of MCTD (Sharp *et al.*, 1972), but there is little doubt that high-titer anti-RNP is associated with inflammatory muscle disease occurring in diverse clinical situations.

Nishikai and Homma (1977) detected antibodies to myoglobin in 70% of

TABLE V

Antinuclear Antibody–Antigen Systems in Inflammatory Muscle and Connective Tissue Disease[a]

Antigen	Enzyme sensitivity	Disease						
		DM	PM	PM/PSS	PSS	MCTD	MG	SLE
JO-1	Trypsin	±	++	−	−	−	−	−
JO-2	NA	−	+	NA	NA	NA	NA	NA
PM-1	Trypsin	+	+++	+++	−	−	NA	−
Mi	Trypsin	+	−	−	−	−	−	−
PA-1	Trypsin	NA	+	NA	NA	NA	NA	NA
Ku	Trypsin	NA	−	+++	−	−	NA	−
RNP/ENA	RNase, trypsin	NA	NA	+	−	+++	−	++
Scl-70	Trypsin	NA	NA	NA	+	NA	NA	−
n-DNA	DNase	−	−	−	−	+	−	+++
Sm	Trypsin	−	−	−	−	−	−	++
ANF	Various	+	+	+	+++	+++	+	+++

[a] From Garlepp and Dawkins (1984) with kind permission of the editor of *Clinics in Rheumatic Diseases*. Abbreviations: NA, not available; −, <5%; ±, 5–10%; +, 10–30%; ++, 30–50%; +++, >50%.

patients with polymyositis, but also found them with high frequency in other diseases (e.g., MG). Wada *et al.* (1983) reported the detection of antibodies to skeletal muscle myosin in 90% of patients with polymyositis as well as in a smaller percentage of patients with other diseases of muscle (e.g., MG, muscular dystrophy). Detection of these antibodies bore no relationship to the presence of AStr antibodies. By contrast, Carrano *et al.* (1983) were unable to demonstrate reactivity of polymyositis sera with crude actomyosin by an ELISA assay, although reactivity with this antigen correlated with the presence of AStr antibodies in sera from patients with MG. We conclude that antibodies to contractile proteins are rarely detectable in polymyositis, and in this respect the human disease contrasts with EAM.

B. Immunological Derangements

In both MG and polymyositis extensive efforts have been made to detect immunological derangements. There is no convincing evidence for a derangement in lymphocyte subsets, although a recent report (Berrih *et al.*, 1981) described decreased OKT8-bearing cells (suppressor/cytotoxic) in MG and decreased suppressor function in patients without thymoma

but normal activity in those with thymoma. Interestingly, Zilko *et al.* (1979) also showed decreased suppressor function in patients with HLA-B8, that is, in those who tend *not* to have a thymoma. In Japanese patients, although lymphopenia was seen there was no disturbance of the OKT4/OKT8 ratio.

Hypocomplementemia in MG was reported by Nastuk *et al.* (1960), and Christiansen *et al.* (1978) reported a marginal decrease in C4 concentration in patients with HLA-B8. This decrease may, at least in part, be due to the presence of a *C4* null allele as part of the HLA-A1,B8,C4AQ0,C4B-1,BfS,DR3 supratype found in these patients (Dawkins *et al.*, 1983).

Reduced complement concentrations are not a feature of polymyositis (Cumming *et al.*, 1977). If C3 and C4 concentrations are normal but hemolytic activity is low, then C2 deficiency might be suspected, particularly given a family history of relevant diseases (Dawkins *et al.*, 1982a). In MCTD, serum complement activity is reduced in some cases. If serum C4 is decreased and anti-DNA antibodies are elevated, then the possibility of coincident SLE should be considered.

Immunoglobulin-G concentrations in MG are generally normal, although antibody responses may be lower than in healthy controls (Dawkins *et al.*, 1976). Polymyositis/dermatomyositis has been reported in the presence of agammaglobulinemia and hypogammaglobulinemia, and some patients do have low or borderline low serum IgG concentrations and respond poorly to tetanus immunization. By contrast, patients classified as MCTD tend to have elevated serum IgG concentrations.

C. Mechanisms of Damage

1. Myasthenia Gravis

Several pieces of evidence indicate that anti-AChR is the primary pathogenic agent in MG (Table VI). The mechanism by which the autoantibody damages the motor endplate *in vivo* is less clear (Table VII). Experimental evidence in support of each mechanism has been obtained using human sera, sera from animals with EAMG, and monoclonal anti-AChR derived from experimental animals (Lennon and Lambert, 1981). It is possible that multiple mechanisms will prove to be operative in MG, and that one or other will be especially important in different patients.

Anti-AChR from different patients vary in specificity (Garlepp *et al.*, 1981b) and are often directed at different determinants on the AChR. Each determinant may have particular significance in terms of susceptibility to the various mechanisms shown in Table VII. The antigenic determinant(s) against which anti-AChR are directed may determine the thresh-

TABLE VI
Anti-AChR As the Pathogenic Agent in Myasthenia Gravis

Present in majority of patients with generalized MG
Titer fluctuates with disease activity in the individual patient
Therapy that reduces anti-AChR titer is effective treatment
Passive transfer (transplacental or to mice) produces disease signs
Presence of anti-AChR in D-P-induced MG
Presence of anti-AChR in idiopathic canine MG
Anti-AChR responsible for disease signs in EAMG
Monoclonal experimental anti-AChR produces signs of EAMG

old titer that must be reached in order to produce overt signs of MG. This would, in part, explain the poor relationship between titer and disease severity when different individuals are compared in contrast, to the strict relationship to disease activity in the individual patient (see Chapter 24). Antiidiotype antibodies to anti-AChR have recently been demonstrated in patients with MG (Dwyer *et al.*, 1983). These might further contribute to this anomaly. Antiidiotype antibodies could inhibit, or alternatively, enhance, anti-AChR binding.

A further consideration is that the current assay does not detect all antibodies acting at the motor end plate (either anti-AChR or others) (Table VIII). This might provide explanations for anti-AChR-negative MG and the apparent absence of anti-AChR in ROMG (see above). The relative rarity of transient neonatal MG (12–20%) despite the frequent transfer of anti-AChR across the placenta may also be dependent on anti-AChR specificity and/or the presence of antiidiotype antibodies.

2. Polymyositis

Despite the presence of several autoantibodies in polymyositis, none seems to play a role in producing the damage to skeletal muscle. The relevance of the recent observation that antibodies to Jo-1 react with histidine tRNA synthetase (Bernstein and Mathews, 1983) remains to be

TABLE VII
Possible Mechanisms of Action of Anti-AChR

Direct blocking of access to AChR
Complement activation and membrane destruction
Modulation of AChR and increased degradation
Alteration of AChR function by interference with conformational changes

TABLE VIII
Reactivity of Human Anti-AChR with Murine AChR *in vitro* and *in vivo*

MG Serum	Cross-reactivity with murine AChR (%)		Increase in degradation rate of AChR (%)[c]	Mean exercise time (min)[d]
	In vitro[a]	*in vivo*[b]		
Serum 1	58	54	80	6.4
Serum 2	2	2	0	11.5
Serum 3	4	60	55	4.5

[a] Determined by comparing the precipitation of murine and human detergent solubilized AChR.
[b] Determined by absorption of anti AChR *in vivo* after passive transfer to mice.
[c] The effect on turnover of AChR in cultured BC_3H-1 muscle cells compared to that of normal human serum.
[d] Mice were given a sublethal dose of D-tubocurarine and exercised on a rotating drum 24 h after passive transfer of test sera. Although anti-AChR in serum 3 is virtually unreactive with detergent-solubilized mouse AChR, it does bind to the AChR in the membrane and produce an effect at the neuromuscular junction.

determined. The presence of immunoglobulin and complement in the necrotic fibers in polymyositis seems to be a consequence of the damage rather than its cause (Fig. 1). In muscle biopsies from patients with MCTD, IgG can occasionally be demonstrated in apparently intact myofibers in the absence of complement components or fibrinogen (Dawkins *et al.*, 1982a). Again the significance of this apparent "*in vivo*" binding of IgG is unclear. Controversy also exists as to the significance of apparent *in vivo* binding of speckled antinuclear antibody seen in skin and muscle biopsies of patients with MCTD, and occasionally, dermatomyositis. Some suggest that this is an artifact of high-titer circulating antinuclear antibody, which binds to the nuclei during processing of the biopsy. Others (Alarcon-Segovia *et al.*, 1978) have demonstrated entry of anti-RNP into human monocytes via Fc receptors and speculated that the entry of autoantibody into viable cells in this way may be of pathological importance.

It has been known for some years that peripheral blood lymphocytes from patients with polymyositis may be directly cytotoxic for cultured muscle and produce lymphotoxin in response to autologous muscle (see Whitaker, 1982). Electron microscopic examination of inflammatory lesions of muscle has shown intimate contact between lymphocytes and muscle cells with processes of lymphoid cells invaginating the sarcolemma of the muscle cell (Mastaglia and Walton, 1982). Peripheral blood lymphocytes have also been shown to proliferate in response to allogeneic

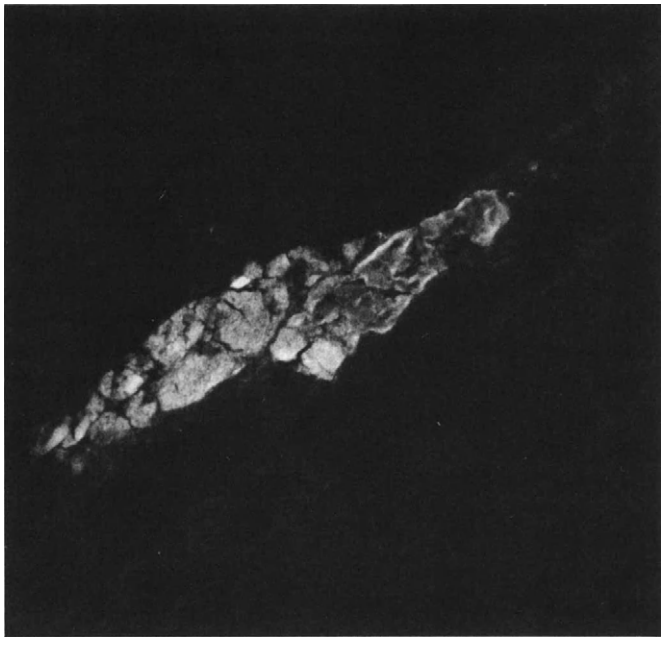


FIG. 1. Indirect immunofluorescent staining of immunoglobulin in a biopsy of skeletal muscle from a patient with polymyositis. Similar intrafiber staining was seen using antifibrinogen and anticomplement reagents (originally ×160). (Reproduced from Garlepp and Dawkins, 1984, with permission from the editor of *Clinics in Rheumatic Diseases*.)

or autologous muscle antigens. These findings when taken together suggest a major role for cell-mediated immunity in the pathogenesis of polymyositis. The demonstration that similar mechanisms are operational in EAM (Dawkins, 1975) provides further support.

A number of observations suggest that overlap can exist between the two mechanisms of damage to muscle. Patently, MG and polymyositis can coexist in the one patient. Moreover, many patients with MG, particularly those with a thymoma, will have lymphorrhages present in skeletal muscle that correspond to areas of myofiber necrosis (see above). Furthermore, anti-AChR have been detected in some patients with D-P-induced polymyositis, and in at least one case frank MG has developed (Carroll *et al.*, 1982; Essigman, 1982). The inflammatory changes in these patients may represent the acute phase of MG, analogous to that seen in EAMG. They might equally indicate the induction of both polymyositis and MG in the same patient.

Both diseases involve an immune assault on skeletal muscle. However, the mechanism that predominates and the clinical manifestations that

TABLE IX
HLA (%) in Myasthenia Gravis[a]

Race	Subgroup of MG	Antigen	Frequency	
			Disease	Controls
Caucasian	Early-onset GMG	A1	76	40
		B8	72	28
		DR3	60	23
		A1,B8,DR3	60	18
	Ocular	DR5	57	20
American black	GMG	A1	27	10
		B8	28	12
		DR3	23	24
		DR5	55	32
Japanese	GMG	Bw44	37	13
		Bw51	32	13
Asian Indian	GMG	B8	20	12
		Bw21	16	4
		Bw35	24	7

[a] Selected antigen frequencies. Data are from Dawkins *et al.* (1982b, 1983).

ensue appear to depend on the immunogenetic makeup of the individual (see Section VI,D). The factors (environmental or other) that trigger the abrogation of control of these responses may also be determined by genetic factors.

D. Immunogenetics

1. Myasthenia Gravis

Although it has been clear for some years that the HLA antigens B8 and DR3 are increased in frequency in Caucasians with MG (Dawkins, 1980), the association is by no means complete; that is, the p(T/D) does not approach unity. This might be due to the heterogeneity of the disease, but equally may be due to need for disease susceptibility genes linked to MHC markers. Both possibilities seem to be true. Myasthenia gravis is heterogeneous, and the HLA associations differ between subgroups of disease. Furthermore, if one considers races other than Caucasians, then different HLA associations are seen (Table IX).

Thus it appears that genes associated with the A1,Cw7,B8,C4AQ0,C4B1,BfS,DR3 supratype are important in determining the development of GMG in young Caucasian females. The nature of such genes

TABLE X
HLA (%) in D-Penicillamine-Induced Muscle Disease

Antigen	D-P MG	D-P Polymyositis	Idiopathic MG	RA	Healthy controls
A1	44	38	64	38	42
A2	48	54	40	60	42
B8	22	23	54	36	28
Bw35	33	31	14	15	11
B27	33	0	10	8	4
DR1	73	0	20	20	18
DR3	9	11	46	23	23
DR4	26	78	20	67	32
A1,B8,DR3	7	11	46	11	18
Bw35,DR1	27	0	3	5	5
A2,B27,DR1	27	0	0	1.5	0.5

is unknown. HLA-B8 has been associated with impaired nonspecific suppressor function in these patients (Zilko *et al.*, 1979) as well as with high-titer anti-AChR (Dawkins, 1980; Keesey *et al.*, 1982). Such a gene might also operate by conferring susceptibility to certain environmental trigger factors. This may be the case in D-P-induced MG, where the frequency of A1,B8,DR3 is reduced but DR1 is increased, probably by virtue of the supratypes Bw35,DR1 and A2,B27,DR1 (Table X). Such supratypes may allow excessive anti-AChR production after exposure to D-P. The effect seems to require the continued presence of D-P, since discontinuance of the drug usually results in a fall in anti-AChR titer and remission of MG (Fig. 2).

By contrast, in D-P-induced polymyositis, there is no increase in DR1 although DR4 is greatly increased (Table X). This may simply reflect the fact that most patients were being treated with D-P for rheumatoid arthritis (RA), but it serves to emphasise the unique association between the DR1-containing supratypes and D-P-induced MG.

Analysis of immunoglobulin allotypes in MG again illustrates the immunogenetic heterogeneity of the disease (Table XI). In Japanese MG there is an increased frequency of phenotypes containing G1m(2), whereas in Caucasian patients there is no overall disturbance of Gm phenotype frequencies (Nakao *et al.*, 1980; Garlepp *et al.*, 1984a). However, Gm does influence autoantibody development in Caucasian patients. The homozygous phenotype Gm(3;5) is associated with high-titer anti-AChR in females with early-onset GMG, but with AStr antibody (and thymoma) in

TABLE XI
Immunoglobulin Allotypes (%) and Autoantibodies in Myasthenia Gravis

Patient group	*n*	*Gm* phenotype				
		1	1,2	1,2,3;5	1,3;5	3;5
Controls	200	5	8	14.5	32.5	39.5
Myasthenia gravis	122	6	3	8	37	46
Females <40 years	41	2.5	5	7.5	39	46
Males and females >40 years	78	7.5	2.5	9	36	45
Females <40 years (high-titer anti-AChR)	23	4.5	0	4.5	22	70
Males and females >40 years (AStr)	45	7	0	7	22	64

the patients with later onset of GMG (Smith *et al.*, 1983; Garlepp *et al.*, 1984a).

2. Polymyositis

Immunogenetic analysis of polymyositis also requires breakdown of the disease into its various subgroups. Most studies have shown that HLA-B8 and -DR3 are increased in adult polymyositis, but Hirsch *et al.* (1981)

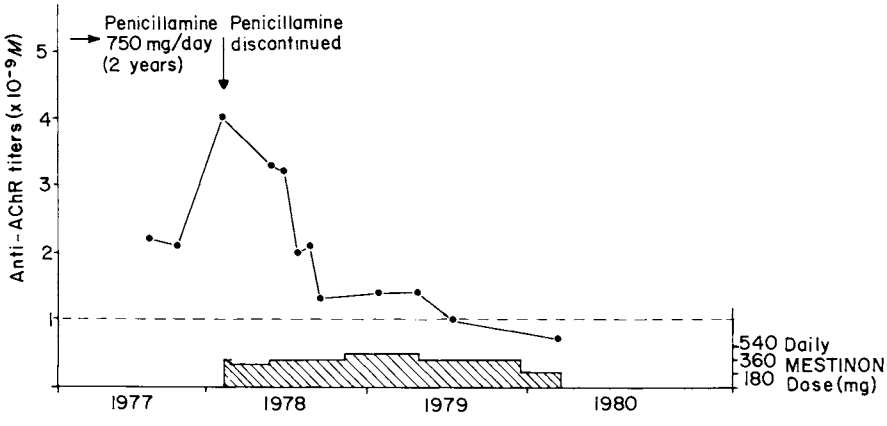


FIG. 2. Titer of anti-AChR in a patient with rheumatoid arthritis treated with D-penicillamine. The titer fell gradually after D-penicillamine was withdrawn. Clinical evidence of MG disappeared over the same period. (Reproduced from Dawkins *et al.*, 1981, with kind permission of the editor of *Journal of Rheumatology*.)

TABLE XII
HLA Associations in Polymyositis and Juvenile Dermatomyositis[a]

Disease	Antigen	Antigen frequency (%)	
		Disease	Controls
Adult PM	B8	44–61	21–28
	DR3	67	24
	B14	40	8
Juvenile DM	B8	35–43	21
	DR3	43–57	23–30
	A1,B8,DR3	21	5

[a] Summarized from Garlepp and Dawkins (1984).

were unable to show an increase in a small group of patients with adult dermatomyositis. In juvenile dermatomyositis it seems clear that B8 and DR3 are increased in frequency, most probably by virtue of the A1,B8,DR3 supratype (Table XII).

Adult dermatomyositis has been reported in association with C2 deficiency (Leddy *et al.*, 1975; Dawkins *et al.*, 1982a). The supratype in both cases was A25,B18,DR2,BfS. The role of C2 deficiency in the induction of autoimmune disease is unknown.

E. Laboratory Diagnosis

1. Myasthenia Gravis

In most laboratories, anti-AChR is detected by a modification of the original radioimmunoassay described by Lindstrom (1977). When human AChR is used as antigen, this assay is extremely sensitive. Assays that detect blocking anti-AChR or rely on immunofluorescence or blocking of binding to concanavalin A are less useful diagnostically. Provided that a suitable reference range for anti-AChR has been established (see Chapter 24), the assay becomes extremely valuable in the diagnosis of MG. The assay is most sensitive in active adult-onset GMG. Sensitivity is lower in ROMG and childhood-onset MG (see Figs. 11 and 12, Chapter 24). Most importantly, the specificity of the assay is extremely high. A titer above the reference range (i.e., >1.6 units in this laboratory) is confirmatory for MG, although a negative result does not exclude MG. False-positives rarely if ever occur. In patients with SLE, Graves' disease, or a thymoma, anti-AChR has been demonstrated in the absence of obvious MG

(Garlepp *et al.*, 1982). We believe that such patients may have latent, early, or subclinical MG (see Figs. 13 and 14, Chapter 24).

The assay is of value in monitoring the effect of therapy (e.g., plasmapheresis, alternate-day steroids) or the progress of patients in remission, since it is generally agreed that titer does correlate with disease activity in the individual patient (Vincent and Newsom-Davis, 1980; Dawkins *et al.*, 1982a). Fluctuations in titer that exceed the threshold will result in exacerbation of disease. Sequential titers must be obtained in order to provide an assessment of each individual's threshold (see Fig. 14, Chapter 24). The assay may be used as a confirmatory test in D-P-induced MG, although for the reasons discussed in Chapter 24 routine screening of patients with RA treated with D-P is not recommended.

The detection of AStr antibody by immunofluorescence (Peers *et al.*, 1977) is of value in relation to a thymoma in MG. A negative result argues against the presence of a thymoma. A positive result, however, has a low predictive value, as AStr antibody is occasionally found in the absence of a thymoma (e.g., after D-P).

2. Myositis

Most of the antibodies listed in Table V are detected by means of immunodiffusion in agarose and identified by the presence of a line of identity with a reference serum. The availability of quantitative and reference data is limited, so that the diagnostic utility of these antibodies is difficult to determine. The p(T/D) is low in all cases, so that none can be used as excluding tests. In polymyositis it seems that the $p(\bar{T}/\bar{D})$ (specificity) for antibodies to Jo-1 and PM-1 is high, so that these two antibodies have potential as confirmatory tests. Similarly, antibodies to Ku in the "overlap syndrome" of polymyositis and progressive systemic sclerosis might possibly be used in this way.

The presence of high-titer anti-RNP is a *sine qua non* for the diagnosis of MCTD. Anti-RNP (low- and high-titer) can be found in other multisystem autoimmune diseases including dermatomyositis. Its diagnostic utility is further complicated by technical difficulties in its detection and quantitation (Dawkins, 1983). This antibody may be detected initially by a speckled antinuclear immunofluorescent staining pattern. It should be remembered that antibodies to other nuclear antigens may also give rise to this pattern, and that treatment with RNAase can lead to loss of antigens other than RNA. Most laboratories use immunodiffusion as a confirmatory test for the presence of anti-RNP. The definition of high-titer anti-RNP is also subject to local interpretation. In this laboratory, a titer >1 in 512 by immunofluorescence is considered of clinical and diagnostic importance.

VII. TREATMENT

A. Approaches to Therapy

Before planning therapy, it is crucial to consider the differential diagnosis and to exclude certain conditions that require a different approach. For example, hypothyroidism can mimic polymyositis but clearly requires different therapy. Any deviation from the euthyroid state is said to make the control of MG more difficult, and it is usual to correct any abnormality at the outset. Systemic lupus erythematosus can present as an inflammatory myopathy and may respond to standard therapy, but it is preferable to determine the extent of involvement and especially the extent of glomerulonephritis before commencing corticosteroid therapy.

It is also important to consider the actual mechanisms of injury before commencing therapy. If anti-AChR appears to be responsible for the neuromuscular block, it is reasonable to consider approaches that will lead to a fall in titer. If D-P appears to be the trigger factor, it may be possible to discontinue the drug and to observe in the hope that there may be spontaneous recovery.

B. Myasthenia Gravis

In adult-onset GMG, it is usual to commence therapy with anticholinesterase drugs. If symptoms can be controlled with dosages which do not lead to complications, it may not be necessary to proceed with any additional therapy.

The role of thymectomy remains somewhat controversial. If there is evidence of a thymoma and a suspicion that the tumor may be locally invasive, thymectomy should be performed for this indication irrespective of the severity of the associated MG. If however thymic hyperplasia is suspected, the indications are somewhat less clear. Some patients in this group do appear to respond well, but an adequate controlled study has not been undertaken and it remains unclear whether thymectomy should be performed soon after diagnosis or only after a trial of other forms of therapy. It is also unclear as to whether thymectomy reduces anti-AChR titer.

Alternate-day corticosteroid therapy is undoubtedly effective and can lead to a reduction in the titer of anti-AChR. The mechanism of this effect remains unclear; it is not necessary to induce immunosuppression as indicated by negative responses to ubiquitous antigens used for delayed-type hypersensitivity testing. The fall in anti-AChR after 100 mg of prednisolone on alternate days is illustrated in Fig. 3.

Plasmapheresis is useful, especially in patients with severe progressive

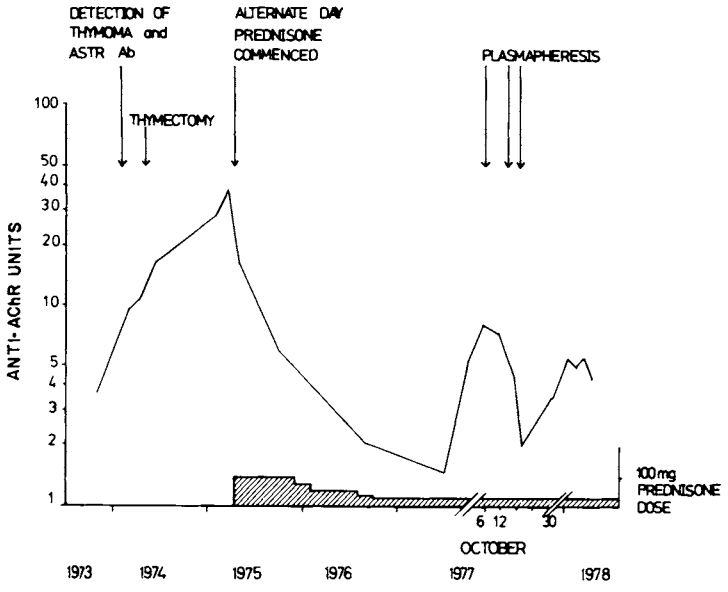


FIG. 3. Sequential titers of anti-AChR in a patient identified by incidental discovery of AStr antibody on routine immunofluorescence screening for autoantibodies. Myasthenia gravis was mild initially but progressed until after the anti-AChR titer fell. As prednisone therapy was reduced, the myasthenia pairs became more severe. Plasmapheresis was effective but anti-AChR titer rose again in parallel with further clinical deterioration. (Reproduced from Dawkins *et al.*, 1982a, with permission from Churchill Livingstone.)

disease. It may be necessary to undertake multiple exchanges before the anti-AChR titer falls substantially, but such an approach is appropriate, at least as a means of tiding the patient over the initial period of alternate-day corticosteroid therapy. It may also be appropriate in preparing the patient for thymectomy.

Intravenous administration of high-dose 7 S IgG has been reported to produce a fall in anti-AChR titer and clinical improvement in a small group of patients (Fateh-Moghadam *et al.*, 1984). It has been suggested as a possible mode of therapy for emergency or refractory cases. Confirmation and further clinical evaluation is required.

C. Polymyositis

Corticosteroid therapy is the mainstay of treatment for inflammatory myopathy leading to weakness. It is usual to commence therapy with 40–80 mg prednisone each day. A sustained remission can be induced, but this may require at least months of therapy and relatively high doses. In those who fail to respond, methotrexate and other forms of immunosup-

pression may be useful. We use intravenous methotrexate initially, but the oral form may also be effective.

Patients have been reported occasionally to respond to plasmapheresis, but the rational basis for this therapy has not been established. Conceivably, titers of the associated antibodies will be reduced.

VIII. CONCLUDING REMARKS

In this chapter we have taken the view that autoimmune muscle disease is heterogeneous but that some mechanisms of injury can be identified.

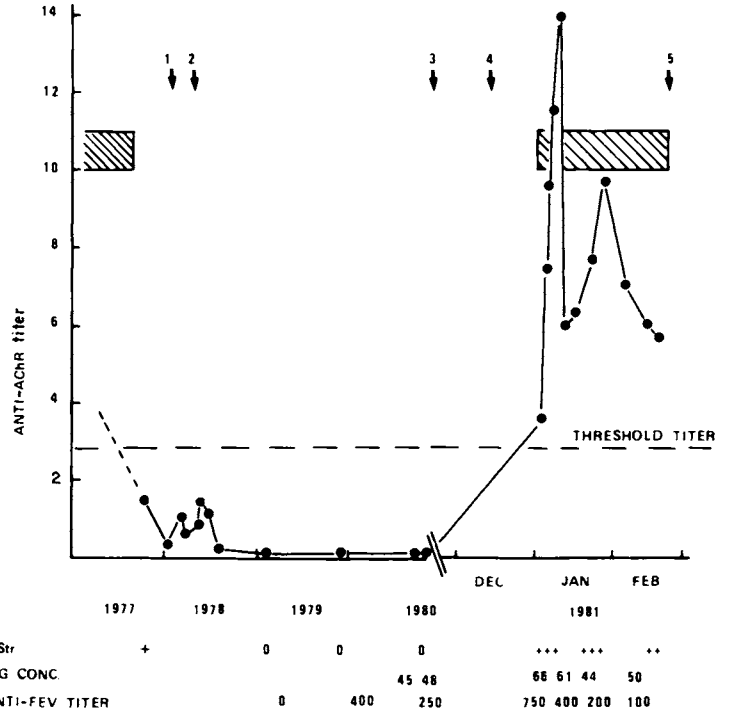


FIG. 4. Sequential serological data from a dog that presented with MG in 1977. It underwent spontaneous remission and suffered an exacerbation with concomitant increases in anti-AChR and AStr titers following vaccination for feline enteritis virus (FEV) and an eye infection. Shaded areas indicate periods of active MG. 1, administration of rat AChR; 2, commencement of D-P (stopped October, 1978); 3, FEV vaccination; 4, eye infection; 5, death. Relative IgG concentrations are expressed in mm^2 and FEV antibody concentrations as the reciprocal of the titer. (Reproduced from Garlepp *et al.*, 1984b, with kind permission of the editor of *Clinical Immunology and Immunopathology*.)

There are already suggestions that this approach has practical value in relation to diagnosis and therapy, and further progress can be anticipated. It seems highly likely that at least most forms of autoimmune muscle disease can be considered to be the consequence of aberrant immunoregulation in the genetically predisposed subject. Various inducing factors appear important, and it should be possible to identify at least some of these. Further study of the drug-induced syndromes should be particularly useful in this regard. We also see substantial merit in the study of spontaneous diseases occurring in animals and have illustrated the value of this approach by reference to a dog studied sequentially (Fig. 4).

IX. SUMMARY

Myasthenia gravis is heterogeneous, but at least some forms are due to the action of anti-AChR. Individual subjects vary in the threshold that must be exceeded if clinical features are to develop. Diagnostic assays for MG depend on the development of appropriate reference data and some understanding of the threshold phenomenon. Monitoring the anti-AChR titer is useful and provides the clinician with a rational approach to therapy.

Polymyositis is probably even more heterogeneous. In at least some patients it appears that cell-mediated mechanisms are of predominant importance, and for at least these patients corticosteroid therapy may be effective. The distinction between polymyositis as a clinical entity and inflammatory muscle disease associated with other clinical entities remains difficult, but there are suggestions that the identification of particular antinuclear antibodies will assist in the better classification of these disorders.

REFERENCES

Alarcon-Segovia, D., Ruis-Arguelles, A., and Fishbein, E. (1978). *Nature (London)* **271,** 67–69.

Berman, P., Patrick, J., Heinemann, S., Klier, F. G., and Steinbach, J. H. (1981). *Ann. N. Y. Acad. Sci.* **377,** 237–257.

Bernstein, R. M., and Mathews, M. B. (1983). *Arthritis Rheum.* **26,** S9.

Berrih, S., Gaud, C., Back, M. A., Le Brigand, H., Binet, J. P., and Bach, J. F. (1981). *Clin. Exp. Immunol.* **45,** 1–8.

Bohan, A., and Peter, J. B. (1975). *N. Engl. J. Med.* **292,** 344–347.

Carrano, J. A., Swanson, N. R., and Dawkins, R. L. (1983). *J. Immunol. Methods* **59,** 301–314.

Carroll, G. J., Peter, J. B., and Garlepp, M. J. (1982). *In* "Immunogenetics in Rheumatol-

ogy" (R. L. Dawkins, F. T. Christiansen, and P. J. Zilko, eds.), pp. 326–330. Excerpta Medica, Amsterdam.
Christadoss, P., Lennon, V. A., Krco, C. J., Lambert, E. N., and David, C. (1981). *Ann. N. Y. Acad. Sci.* **377,** 258–277.
Christiansen, F. T., Houliston, J. B., and Dawkins, R. L. (1978). *Muscle Nerve* **1,** 467–470.
Cumming, W. J. K., Hudgson, P., Lattimer, D., Sussman, M., and Wilcox, C. B. (1977). *Lancet* **2,** 978.
Dawkins, R. L. (1975). *Clin. Exp. Immunol.* **21,** 185–201.
Dawkins, R. L. (1980). *In* "Histocompatibility Testing 1980" (P. I. Terasaki, ed.), pp. 662–667. UCLA Tissue Typing Laboratory, Los Angeles, California.
Dawkins, R. L. (1983). *In* "Immunology in Medicine" (E. J. Holborow and W. G. Reeves, eds.), pp. 439–465. Academic Press, London.
Dawkins, R. L., and Mastaglia, F. L. (1973). *N. Engl. J. Med.* **288,** 434–438.
Dawkins, R. L., O'Reilly, C., Grimsley, G., and Zilko, P. J. (1976). *Ann. N. Y. Acad. Sci.* **274,** 461–467.
Dawkins, R. L., Garlepp, M. J., McDonald, B. L., Williamson, J., Zilko, P. J., and Carrano, J. (1981). *J. Rheumatol.* **8** (Suppl. 7), 169–172.
Dawkins, R. L., Garlepp, M. J., and McDonald, B. L. (1982a). *In* "Skeletal Muscle Pathology" (F. L. Mastaglia and J. Walton, eds.), pp. 461–482. Churchill Livingstone, London.
Dawkins, R. L., Christiansen, F. T., and Zilko, P. J. (1982b). "Immunogenetics in Rheumatology: Disease and D-Penicillamine." Excerpta Medica, Amsterdam.
Dawkins, R. L., Christiansen, F. T., Kay, P. H., Garlepp, M. J., McCluskey, J., Hollingsworth, P. N., and Zilko, P. J. (1983). *Immunol. Rev.* **70,** 5–22.
Drachman, D. B. (1978). *N. Engl. J. Med.* **298,** 136–142.
Dwyer, D. S., Bradley, R. J., Urquhart, C. K., and Kearney, J. F. (1983). *Nature (London)* **301,** 611–614.
Engel, A. G., Lambert, E. H., Mulder, D. M., Gomez, M. R., Whitaker, J. N., Hart, Z., and Sahashi, K. (1981a). *Ann. N. Y. Acad. Sci.* **377,** 614–639.
Engel, A. G., Sahashi, K., and Fumagalli, G. (1981b). *Ann. N. Y. Acad. Sci.* **377,** 158–174.
Essigman, W. K. (1982). *Ann. Rheum. Dis.* **41,** 617–620.
Fateh-Moghadam, A., Wick, M., Besinger, U., and Geursen, R. G. (1984). *Lancet* **1,** 848–849.
Fenichel, G. M. (1978). *Arch. Neurol.* **35,** 97–103.
Garlepp, M. J., and Dawkins, R. L. (1984). *Clin. Rheum. Dis.* **10,** 35–51.
Garlepp, M. J., Farrow, B., Kay, P., and Dawkins, R. L. (1979). *Immunology* **37,** 807–810.
Garlepp, M. J., Dawkins, R. L., Christiansen, F. T., Lawton, J., Luciani, G., McLeod, J., Bradley, J., Genkins, G., and Teng, C. S. (1981a). *J. Neuroimmunol.* **1,** 325–332.
Garlepp, M. J., Kay, P. H., Dawkins, R. L., Bucknall, R. C., and Kemp, A. (1981b). *Muscle Nerve* **4,** 282–288.
Garlepp, M. J., Kay, P. H., and Dawkins, R. L. (1982). *J. Neuroimmunol.* **3,** 337–350.
Garlepp, M. J., Dawkins, R. L., and Christiansen, F. T. (1983). *Br. Med. J.* **286,** 338–340.
Garlepp, M. J., Kay, P. H., and Dawkins, R. L. (1984a). *Disease Markers* **2,** 429–436.
Garlepp, M. J., Kay, P. H., Farrow, B. R., and Dawkins, R. L. (1984b). *Clin. Immunol. Immunopath.* **31,** 301–306.
Hirsch, T. J., Enlow, R. W., Bias, W. B., and Arnett, F. C. (1981). *Hum. Immunol.* **3,** 181–186.
Keesey, J., Naiem, F., Lindstrom, J., Roe, D., Herrmann, C., and Walford, R. (1982). *Arch. Neurol.* **39,** 73–77.

Kodama, K., Sikorska, H., Bandy-Dafoe, P., Bayly, R., and Wall, J. R. (1982). *Lancet* **2,** 1353–1356.

Leddy, J. P., Griggs, R. C., Klemperer, M. R., and Frank, M. M. (1957). *Am. J. Med.* **58,** 83–91.

Lennon, V. A., and Lambert, E. H. (1981). *Ann. N. Y. Acad. Sci.* **377,** 77–96.

Lindstrom, J. (1977). *Clin. Immunol. Immunopathol.* **7,** 36–43.

Lindstrom, J. (1979). *Adv. Immunol.* **27,** 1–50.

Mastaglia, F. L., and Walton, J. (1982). *In* "Skeletal Muscle Pathology" (F. L. Mastaglia and J. Walton, eds.), pp. 360–392. Churchill Livingstone, Edinburgh.

Mimori, T. A., Akizuki, M., Yagamata, H., Inada, S., Hoshida, S., and Homma, M. (1981). *J. Clin. Invest.* **68,** 611–620.

Nakao, Y., Matsumoto, H., Miyazaki, T., Nishitani, H., Ohta, K., Fujita, T., and Tsuji, K. (1980). *Lancet* **1,** 677–680.

Nastuk, W. L., Plescia, O. J., and Osserman, K. E. (1960). *Proc. Soc. Exp. Biol. Med.* **105,** 177–184.

Nishikai, M., and Homma, M. (1977). *JAMA, J. Am. Med. Assoc.* **237,** 1842–1844.

Osserman, K. E. (1958). "Myasthenia Gravis." Grune & Stratton, New York.

Palmer, A. C. (1980). *Vet. Clin. N. Am.Small Anim. Pract.* **10,** 213–218.

Patrick, J., and Lindstrom, J. (1973). *Science* **180,** 871–872.

Peers, J., McDonald, B. L., and Dawkins, R. L. (1977). *Clin. Exp. Immunol.* **27,** 66–73.

Ray, C. G., Minnich, L. L., and Johnson, P. C. (1979). *J. Infect. Dis.* **140,** 239–243.

Sharp, G. C., Irwin, W. A., Tan, E. M., and Holman, H. (1972). *Am. J. Med.* **52,** 148–149.

Simpson, J. A. (1960). *Scott. Med. J.* **5,** 419–436.

Smith, C. I. E., Grubb, R., Hammarstrom, L., and Matell, G. (1983). *J. Immunogenet.* **10,** 1–9.

Vincent, A. (1980). *Physiol. Rev.* **60,** 756–823.

Vincent, A., and Newsom-Davis, J. (1980). *J. Neurol. Neurosurg. Psychiatry* **43,** 590–600.

Wada, K., Yeno, S., Hazama, T., Ogasahara, S., Kang, T., Takahashi, M., and Tarui, S. (1983). *Clin. Exp. Immunol.* **52,** 297–304.

Waldor, M. K., Sriram, S., McDevitt, H. O., and Steinman, L. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80,** 2713–2717.

Whitaker, J. N. (1982). *Muscle Nerve* **5,** 573–592.

Zilko, P. J., Dawkins, R. L., Holmes, K., and Witt, C. (1979). *Clin. Immunol. Immunopath.* **14,** 222–230.

CHAPTER 23

Antireceptor Antibodies

LEONARD C. HARRISON
Department of Diabetes and Endocrinology
The Royal Melbourne Hospital
Melbourne, Victoria, Australia

I. INTRODUCTION

A. The Receptor Concept

Biological signals such as hormones, neurotransmitters, and other small molecules initiate their effects by binding to specific cellular receptors. The receptor-based concept of signal recognition and transduction applies to all levels of biological communication: to single- and multicelled organisms, to autocrine, paracrine, and endocrine systems, and to the transmission of "hardwired" nervous system information across synaptic junctions. Furthermore, the efficacy of most drugs can be attributed to their ability to mimic, antagonize, or modify the actions of endogenous biological signals at the level of receptors.

The prime function of a receptor is selective recognition, as evidenced by stereospecific and high-affinity binding of its respective ligand, from among the millions of other molecules present in the extracellular milieu. While recognitive specificity is the cardinal property of receptors, it is by no means absolute. Many hormones, for example, belong to structurally related families and engage in "cross-talk" with each other's receptors, particularly when present at supraphysiological concentrations. In some cases, the specificity of ligand–receptor interactions is not only inherent but is also determined by local geography, as exemplified by the compartmentalization of neurotransmitters to synapses.

The binding of a ligand is followed by conformational and/or covalent modifications in the receptor, which lead to activation of postbinding (postreceptor) pathways and bioeffects characteristic of the ligand. For many of the hormones and neurotransmitters that act initially at the cell surface, the ligand–receptor complex is coupled by a GTP binding protein to either the activation (e.g., β catecholamines, glucagon, ACTH) or the inhibition (e.g., α-2 catecholamines, opiates, prostaglandins) of adenyl cyclase located on the inner face of the plasma membrane, thereby regulating the production of cAMP, the intracellular second messenger, and the activity of cAMP-dependent protein kinases. However, there is a major group of anabolic, growth-promoting peptide hormones including insulin and insulinlike growth factors, growth hormone, prolactin, and epidermal growth factor for which the postbinding events in hormone action remain largely undefined, despite intensive research.

Receptors are large glycoproteins with molecular weights in the range 10^4–10^6 (i.e., orders of magnitude greater than their ligands). Generally, they have a subunit structure and/or are composed of functional domains and a hydrophobic domain necessary for membrane anchorage. Receptor concentrations range from a few hundred to 10^5 per cell and reflect a dynamic equilibrium between rates of synthesis and degradation, and for cell-surface receptors, rates of recycling and membrane insertion.

B. Historical Background

It is now accepted as axiomatic that biological signals require receptors for their expression. Indeed, a receptor can be regarded as a biochip containing all the programmed information unique for its ligand. The function of the ligand or biological signal is simply to "throw the switch" to activate the program. This concept is not modern and was espoused, in principle, by Emile Fischer in his lock-and-key hypothesis for enzyme–substrate interactions. It formed the basis of the receptor concept introduced separately by Newport Langley (1852–1925) and Paul Ehrlich

(1854–1915). Langley (1878) explained the opposing actions of atropine and pilocarpine on saliva flow in cats by the assumption that these drugs competed for a specific substance:

> We may, I think, without much rashness, assume that there is a substance or substances in the nerve endings or gland cells with which both atropin and pilocarpin are capable of forming compounds. On this assumption then, the atropin or pilocarpin compounds are formed according to some law of which their relative mass and chemical affinity for the substance are factors.

This deduction was the germ of the receptor theory. Later, Langley (1905) in seeking to explain the antagonism between nicotine and curare on muscle, proposed that the two drugs competed for the same "receptive substance."

Ehrlich (1906), after studying bacterial toxin–antitoxin reactions, came to the conclusion that cells possess side chains, or receptors (Fig. 1). He believed that antitoxin made by cells normally resides on cell surfaces as side chains and that the circulating antitoxins result from excessive production of "side-chains," triggered by toxin. He suggested that side-chain specificity is based on steric complementarity. Furthermore, he distinguished between toxin and toxoid, the latter being harmless but still antigenic, and he proposed that toxin molecules contain two distinct chemical groupings, a haptophore responsible for binding and a toxophore responsible for biological effects; toxoid molecules contained only haptophore (Fig. 1). Thus, Ehrlich first addressed the question of agonism versus antagonism, the molecular basis of which remains a challenge to this day. His view of bacterial toxins as bifunctional molecules has been essentially validated.

We now appreciate that Ehrlich's side chains are the antibodies on the surface of B lymphocytes that act as receptors for antigen, whose expression is "up-regulated" and whose secretion is stimulated by antigen, but progress in defining the molecular biology of receptors in the immune system has been relatively slow. The evolution of receptor theory and its validation has occurred predominantly in pharmacology and endocrinology, disciplines that were founded on the basis of ligands rather than receptors. This is understandable in view of the relative ease with which many drugs and classic hormones can be identified, isolated, and labeled. Hormone receptor techniques now provide not only paradigms for immunology, but, ironically, have allowed the discovery of naturally occurring autoantibodies against cell-surface receptors of nonimmune cells. It is noteworthy that Ehrlich proposed that side chains existed as *Normalantikorper*, whereas his contemporary Karl Landsteiner held that antibody production was a new event, *Andersleistung*, that only followed exposure

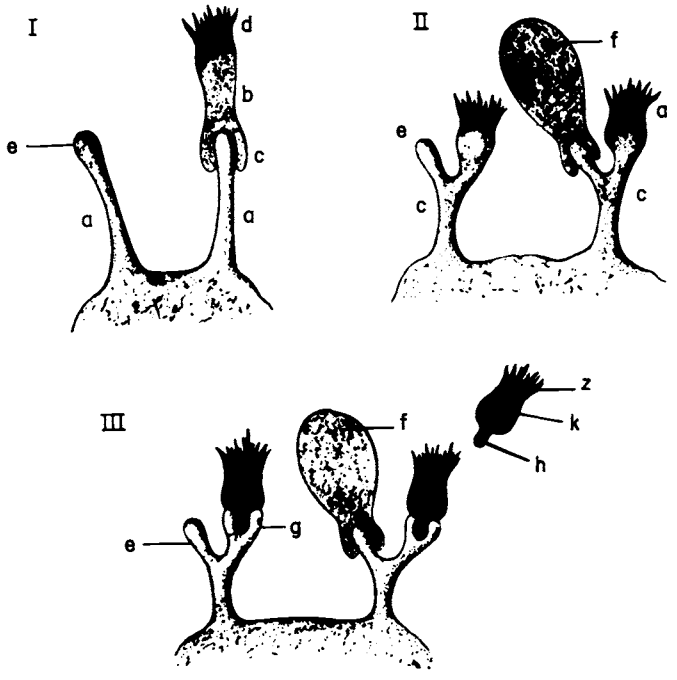


FIG. 1. Ehrlich's concept of "side-chains" or receptors. (I) Receptor of first order: (a) haptophore complex; (b) absorbed toxin molecule with (c) haptophore group; (d) toxophore group. (II) Receptor of second order with (e) haptophore group, (d) zymophore group, and (f) absorbed nutritive molecule. (III) Receptor of the third order: (e) haptophore; (g) complementophile group; (k) complement with (h) haptophore; (z) zymotoxic group; (f) nutritive molecule. (From Ehrlich, 1906).

to toxin. Yet Ehrlich did not believe that the body normally reacts against itself (his dictum of horror autotoxicus) because he had been unable to generate autoantibodies by immunizing goats with their own red blood cells. However, within several years Donath and Landsteiner (1904) discovered the first autoantibody, which lysed red blood cells after binding with complement to a surface blood group antigen.

C. Autoimmunity: The Universal Set of Receptor Antibodies

Autoimmune processes are now recognized as a major cause of human disease, and autoantibodies are strongly implicated as prime effectors of pathological change in autoimmune diseases. Autoantibodies have traditionally been identified by indirect immunofluorescence staining or re-

TABLE I

Autoimmune Diseases With or Without Known Receptor Antibodies

Disorder	Tissue	Receptor
Graves' disease (goiter and hyperthyroidism)	Thyroid	Thyrotropin (TSH)
Myasthenia gravis	Neuromuscular endplate	Acetylcholine (ACh)
Insulin-resistant diabetes	Liver, muscle, fat, etc.	Insulin
Hypoglycemia	Liver, muscle, fat, etc.	Insulin
Allergic asthma, "hayfever"	Respiratory tract, vascular smooth muscle	β-adrenergic
Pernicious anemia (atrophic gastritis)	Stomach	Gastrin
Hypopituitarism	Pituitary	?
Hypoparathyroidism	Parathyroid	?
Addison's disease (hypoadrenalism)	Adrenal	?
Insulin-dependent diabetes	Pancreatic islet β cell	?
Premature ovarian failure	Ovary	?
Infertility	Sperm	?
Alopecia (hair loss)	Hair follicle	?
Vitiligo (depigmentation)	Skin	?
Dermatitis herpetiformis	Skin	?
Pemphigus vulgaris	Skin	?
Bullous pemphigoid	Skin	?
Idiopathic thrombocytopenic purpura (ITP)	Platelets	?
Coeliac disease	Small intestine mucosa	?
Rheumatoid arthritis	Connective and vascular tissues	?
Systemic lupus erythematous	Connective and vascular tissues	?
Chronic active hepatitis	Liver	?
Primary biliary cirrhosis	Liver	?
Glomerulonephritis	Kidney	?

lated techniques, following exposure of tissues to serum. This approach is relatively insensitive and unspecific and does not reveal the functional properties of antibodies or the submicroscopic nature of their targets. However, with the development of receptor-binding techniques, it became apparent that cell-surface receptors were prime targets for autoimmune reactions and that the study of receptor autoantibodies would define specific molecular mechanisms of autoimmunity (Lennon and Carnegie, 1971). Antireceptor antibodies are the hallmark of a subset of autoimmune diseases in which immune reactions are relatively restricted to defined cell surface receptors (Table I). Some organ-specific autoimmune dis-

eases (e.g., thyroid/gastric) have now been redefined in terms of receptor rather than tissue or cell targets, and it seems reasonable to propose that receptor antibodies represent the tip of an autoimmune iceberg. Most, if not all, autoimmune diseases might eventually be reclassified as receptor-specified immune reactions. The discovery of receptor antibodies is probably limited only by our inability to define relevant receptors or receptor-related functions associated with autoimmune phenomena.

II. IMMUNE MECHANISMS AND RECEPTOR ANTIBODIES

A. PREAMBLE

The details of immune regulation are dealt with more appropriately elsewhere. In this section an endocrinologist's viewpoint on several possible mechanisms for the genesis of receptor antibodies is presented. At the outset, several points are worth making. First, there may be a number of mechanisms involved in the genesis of receptor antibodies that are not mutually exclusive. Second, autoimmune diseases are selective and are not associated with a generalized breakdown in self-tolerance. This is seen most clearly with the receptor antibody diseases. Therefore, theories based on general defects in suppressor T-cell function or on polyclonal B-lymphocyte activation are untenable. Third, autoantibodies react with perfectly normal antigens from all sorts of humans and animals, and there is no evidence that the defect is at the level of the (auto)antigen. Finally, low levels of autoantibodies are present in normal people and, like autoimmune diseases, increase with age. We have recently shown, for example, that normal IgG can precipitate the thyrotropin (TSH) receptor, albeit at lower titers than IgG from patients with Graves' disease (Heyma and Harrison, 1984), a finding in accord with previous evidence that normal IgG inhibits TSH binding (Beall *et al.*, 1978; McKenzie *et al.*, 1978; Kleinmann *et al.*, 1980; Brown *et al.*, 1983). At first glance, the assertion that normal individuals have low levels of autoantibodies appears to conflict with a prevailing view in immunology that autoimmunity is due to the emergence of "forbidden clones," due perhaps to somatic mutations in the *V* region genes of immunocytes. However, there is no reason why somatic mutations, which are always occurring, especially at times of immune stress (e.g., infection); could not be regulated or clonally deleted (e.g., by an antiidiotypic response from a preexisting set of germ-line-coded antibodies).

B. HLA Linkage

There is a strong association between the major histocompatibility antigens and autoimmune disease, (HLA-DR3 and/or -4 and insulin-dependent diabetes mellitus (IDDM), Graves' disease hyperthyroidism, and Addison's hypoadrenalism, postulated to be due to linkage dysequilibrium between histocompatibility antigen genes and putative neighboring pathogenic immune response genes (McDevitt and Bodmer, 1974). The mechanisms underlying this association are still the subject of speculation. Immune responses are thought to be initiated by cells bearing the class II major histocompatibility antigens, termed Ia in animals or HLA-DR in man, which present antigen to T cells. The epithelial cells involved in the receptor antibody diseases described below do not normally express HLA antigens. However, Botazzo and his colleagues (Pujol-Borrell *et al.*, 1983) reported that normal thyroid epithelial cells in culture can be induced by mitogenic lectins to express DR antigens and that thyroid cells from patients with Graves' disease actually exhibit a patchy distribution of DR antigen (Hanafusa *et al.*, 1983). These findings suggest a role for aberrant expression of DR antigen together with TSH receptors in the pathogenesis of Graves' disease. On the other hand, Campbell *et al.* (1984), in our laboratory, has shown that pure interferon γ induces the appearance of the Class I H-2K histocompatibility antigen (equivalent to HLA-B in human cells) in cultured mouse pancreatic islets, a finding now extended to the β cells of human islets. Aberrant expression of Class I antigens, in response to lymphokines such as interferon γ, would focus cytotoxic T cells and could be implicated in the β-cell destruction of IDDM.

C. Infection

Notkins and colleagues (Haspel *et al.*, 1983) have shown that autoantibodies are formed in the wake of virus infections. The random but age-related occurrence of autoimmune disease would be consistent with virus-induced autoreactivity in genetically susceptible individuals. The increased frequency of HLA-DR3 and/or -4 antigens in autoimmune disease could then reflect an increased susceptibility or responsivity to virus infection, increased sensitivity to interferon, or a greater capacity to synthesize/express DR antigen. If we assume that repeated transient induction of autoimmune reactivity is not uncommon, then the persistence and/or pathogenicity of autoantibodies must depend on other factors such as the number of insults, the duration and level of exposure, the rate of somatic mutation, the inherent or induced function of T-cell regulators, and the regenerative capacity of affected target cells.

Another mechanism to initiate autoimmune reactions and one that bypasses the need for antigen presentation on DR+ cells involves cross-reaction between self-antigens and exogenous agents such as bacteria or viruses. Such a mechanism has been invoked to explain the association between infection with group A streptococci and rheumatic heart disease (Kaplan and Meyeserian, 1962), *Klebsiella* and ankylosing spondylitis (Edmonds *et al.*, 1981), and *Shigella, Salmonella, Yersinia,* and *Campylobacter* and reactive arthritis (Edmonds, 1984), as well as the very strong association in Scandanavia between agglutinins to *Y. enterocolitica* and Graves' disease (Lidman *et al.*, 1976). With respect to the latter, we have recently demonstrated that *Y. enterocolitica* possesses low-affinity binding sites for TSH and that globulins from patients with Graves' disease that inhibit TSH binding to human thyroid membranes also inhibit the binding of TSH to *Yersinia* (Heyma *et al.*, 1984). Such findings do not provide a causal link between bacterial infection and thyroid autoimmunity, but they do demonstrate at least a common molecular antigen.

D. The Idiotypic Network

Antibodies themselves possess antigenic determinants called idiotopes, whose specificity is described by the term idiotype, on the variable region adjacent to or within their combining site (paratope). Specific immune responses can be inhibited by antiidiotypic antibodies. It was first shown that the mouse myeloma protein T15 bound the hapten phosphorylcholine (Cosenza and Kohler, 1972) and that mice immunized with phosphorylcholine not only produced antibodies to the hapten, as expected, but also produced antibodies against T15 (Kluskens and Kohler, 1974). This finding indicated that induced antibodies to phosphorylcholine had acted as immunogens, inducing secondary antibodies against their idiotopes. Subsequently, Binz and Wigzell (1975) found that T and B cells that are reactive to the same antigen have shared idiotypes. A network of idiotype–antiidiotype interactions was hypothesized by Jerne (1974) to be a fundamental property of the immune system and to be responsible for clonal regulation, either suppression or amplification. Thus, stimulation with antigen would lead to antibody Ab1 and then to anti-Ab1 antiidiotypes, Ab2. Immunization of an animal with Ab2 would induce anti-antiidiotypes, Ab3. Three subsets can be defined within Ab3: the first, which recognizes only Ab2; the second, which shares idiotypic specificities with Ab1 but does not bind antigen; and a third, which binds antigen (Urbain, 1983). In structural terms these antibodies can be considered as complementary images. The idiotypes that mimic the antigen binding site (epi-

tope) have been called internal images (Jerne *et al.*, 1982), and since they can substitute for antigen they are potentially important tools for manipulating immune responses and could be used for vaccination.

The natural occurrence of antibodies that bind to nonimmune receptors and mimic or block the effects of hormones and neurotransmitters has a number of important implications. Thus, the "shapes" for receptor "fit" are not the prerogative of any one communication system and may be shared between systems that evolved discretely. The activation of receptors by pseudoligands supports the concept that the receptor is the critical effector molecule with all the "information" for ligand action. One can also infer that the immune system is incredibly plastic in learning how to make new shapes to fit the receptors of another communication system. On the other hand, one might raise the question whether receptor antibodies could have a physiological function in regulating cellular activity or in "cleaning up" shed receptors. Finally, the existence of receptor autoantibodies begs the question whether they develop as antiidiotypes via a network response to ligand antigen (Fig. 2). There is in fact considerable evidence for such a mechanism (Strosberg, 1983).

Sege and Peterson (1978) raised antibodies to bovine insulin in rats and purified them on an insulin affinity column. Rat antiinsulin IgG was then injected into rabbits. Rabbit IgG was purified and antibodies against common determinants on rat IgG were removed, as were antiinsulin antibodies and insulin. Antiidiotypic antibodies in the rabbit IgG were then demonstrated by their ability to compete with insulin for binding to antiinsulin

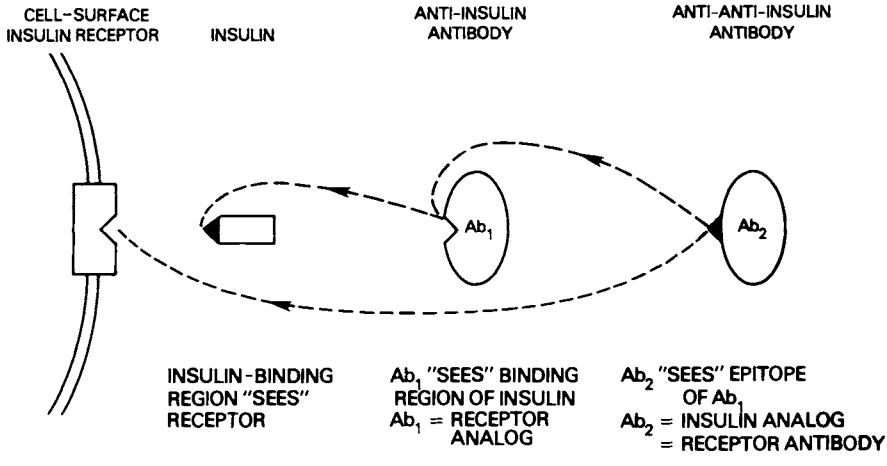


FIG. 2. Hypothetical scheme for generating complementary antiidiotypic antibodies to the insulin receptor.

Fab fragments and to insulin receptors. Shechter *et al.* (1982) showed that immunization of mice with insulin produced not only antiinsulin antibodies but antibodies that blocked insulin binding and mimicked the biological effects of insulin on fat cells. Antiidiotypic antibodies have been similarly produced again receptors for β-adrenergic catecholamines (Hoebeke *et al.*, 1977; Schreiber *et al.*, 1980), acetylcholine (Wasserman *et al.*, 1982), TSH (Farid *et al.*, 1982), and the neutrophil formyl (Met-Leu-Phe) chemoattractant (Marasco and Becker, 1982). Such studies indicate that antireceptor antibodies do not have to be induced by receptors whose concentrations and locations, compared to those of their ligands, are probably less than optimal for inducing immune responses.

Whether receptor antibodies in autoimmune diseases develop as antiidiotypes and by what means is unknown. Perhaps another way in which infectious agents could induce receptor antibodies is via antiidiotypic responses to antibodies to those structures (e.g., on a virus) that might need to interact with surface receptors in order to gain entry to cells. It has been reported that antiidiotypic antibodies raised against a monoclonal antireovirus antibody directed at the region of the virus that binds to host cells both mimic and inhibit the virus binding to host cells (Fields and Greene, 1982). Viruses would have to recognize receptors (e.g., for TSH, ACh, insulin) or structures with shared determinants. Certainly in lupus erythematosis the antigens to which autoantibodies are directed (e.g., RNA–protein complexes) are structures on which the viruses might be expected to depend on for their reproduction or genome insertion.

The occurrence of polyglandular autoimmunity supports the idea that endocrine cells share a common antigen (? virus receptor) and indeed Notkins and colleagues (Onodera *et al.*, 1981) found that newborn mice infected with reovirus type 1 develop polyglandular autoantibodies. As pointed out by Plotz (1983), the virus-induced antiidiotypic autoantibody hypothesis is perfectly testable. Would this virus mechanism incorporate all theories and observations on autoimmunity? T-cell control of idiotypes is said not to be Ia restricted (Bottomly and Mosier, 1981). Therefore idiotype dysregulation per se would not explain the HLA-DR associations of polyglandular and receptor autoimmunity. However, the HLA-DR linkage could confer susceptibility to virus infections or to the ability of viruses to interact with genetically determined cell surface structures. Again, Notkins and colleagues (Yoon *et al.*, 1980) have shown that the ability of mice to develop virus-induced autoimmune disease such as pancreatic diabetes is genetically determined. Finally, one could propose that T-cell defects reported to accompany autoimmune disease might result from antiidiotypic feedback on lymphocyte function.

III. RECEPTOR ANTIBODY DISEASES

A. GRAVES' DISEASE: TSH RECEPTOR ANTIBODIES

1. Preamble

Graves' disease (thyrotoxicosis) is hyperthyroidism and goiter (enlarged thyroid), sometimes associated with eye changes (exopthalmos) due to retroorbital infiltration by lymphocytes and fat and, rarely, skin changes (pretibial myxedema). Antibodies that mimic TSH, directed to sites on or close to the TSH receptor, are considered to be responsible for Graves' hyperthyroidism and goiter (Major and Munro, 1962; Kriss *et al.*, 1964; Manley *et al.*, 1974; Smith and Hall, 1974; Kendall-Taylor *et al.*, 1972; McKenzie and Zakarija, 1976; reviewed by Manley *et al.*, 1982; and by Davies and Bernardo, 1983).

Thyroid-stimulating hormone (TSH) is secreted by basophilic cells in the anterior pituitary gland. It is a glycoprotein of M_r 28,000, composed of two peptide subunits designated α and β. The α subunit is common to the other glycoprotein hormones FSH, LH, and hCG, whereas the β subunit is specific for TSH. Thyroid-stimulating hormone initiates thyroid hormone synthesis, leading to the release of thyroxine (T_4) and smaller amounts of triiodothyronine (T_3) (Fig. 3). At least one-third of the T_4 produced by the thyroid is converted to T_3 by deiodination in peripheral tissues such as liver and kidneys. It is still debatable whether T_4 acts only after conversion to T_3; T_3 initiates its actions by binding to specific recep-

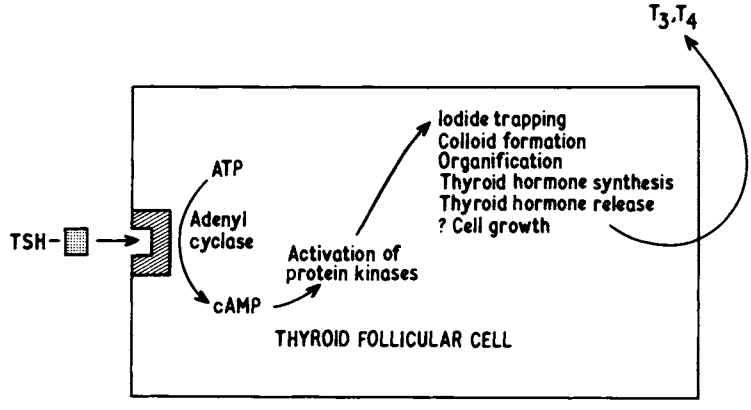


FIG. 3. TSH-Receptor-mediated thyroid hormone biosynthesis.

tors in the nucleus of target cells. In the pituitary, TSH secretion is regulated by negative feedback by T_3, derived mostly from intrapituitary deiodination of T_4.

2. The TSH Receptor

Thyroid-stimulating hormone binds to a receptor on the surface of the thyroid cell (Pastan *et al.*, 1966) and activates adenyl cyclase to increase intracellular 3′, 5′-cyclic adenosine monophosphate (cAMP) (Yamashita and Field, 1970). The coupling mechanism that activates binding probably involves an intermediate GTP binding protein similar to that involved in coupling the β-adrenergic receptor to cyclase. Cyclic AMP acts as a second messenger, activating protein kinases that phosphorylate and thereby activate the enzymes necessary for thyroid hormone synthesis (Fig. 3).

Studies of the binding of TSH to its receptor have been performed under a variety of conditions (Moore and Wolff, 1974; Manley *et al.*, 1974; Smith and Hall, 1974; Tate *et al.*, 1975; Amir *et al.*, 1976; Silverberg *et al.*, 1978; Pekonen and Weintraub, 1979). Earlier workers measured TSH binding to thyroid tissues under nonphysiological conditions and found a high degree of cross-reactivity with the related glycoprotein hormones, as well as with cholera toxin, gamma globulin, and thyroglobulin. However, careful studies by Pekonen and Weintraub (1979), performed under physiological conditions, have now defined high-affinity binding sites on thyroid plasma membranes specific for TSH.

Despite the fact that thyroid autoimmunity has been studied for over 2 decades with the emphasis on the TSH receptor as the "autoantigen," there is still little information on the structure of the TSH receptor. Using TSH affinity chromatography to "purify" the TSH receptor, investigators have reported proteins ranging in size from 15,000 to 500,000 daltons (Rickards *et al.*, 1981; Czarnocka *et al.*, 1979; Iida *et al.*, 1981; Koizumi *et al.*, 1982). In our hands (Heyma and Harrison, 1984) purification of ^{125}I-labeled thyroid membranes by elution from a TSH affinity column yielded reduced proteins of M_r 100,000–110,000, 80,000–90,000, and 60,000–70,000. It is noteworthy that two of these bands are similar in size to those of M_r 88,000 and 66,000 isolated by Koizumi *et al.* (1982) using TSH affinity chromotography.

3. Thyroid-Stimulating Immunoglobulins

Adams and Purves (1956) first detected abnormal thyroid stimulators in the sera of patients with Graves' disease. They found that the release of radioiodine from ^{131}I-prelabeled guinea pig thyroid glands could be stimulated by sera from patients with Graves' disease. McKenzie (1958) then

developed a similar bioassay in the mouse and suggested that the stimulator, designated long-acting thyroid stimulator or LATS, may have a etiological role in Graves' disease. After several years, LATS was shown to be an immunoglobulin (Kriss *et al.*, 1964). Further studies revealed, however, that no more than 50% of patients with untreated Graves' disease were LATS positive, and that LATS titers did not correlate with the severity of thyrotoxicosis in these subjects (Major and Munro, 1962).

Adams and Kennedy (1967) discovered another thyroid stimulator, which they designated long-acting thyroid stimulator/protector or LATS protector (LATS-P). They found that LATS activity could be removed from serum by incubation with human thyroid, but if the thyroid was first incubated with LATS-negative serum from a patient with active Graves' disease, LATS activity was retained. Thus LATS-P was assayed by its ability to protect LATS (measured in the mouse bioassay) from depletion by human thyroid tissue. Activity of LATS-P was shown to be in the IgG fraction of serum and was present in the sera nearly all patients with Graves' disease; moreover, it correlated well with other indices of hyperthyroidism (Adams and Kennedy, 1971; Adams *et al.*, 1974). These findings suggested that the immunoglobulins (Igs) probably responsible for stimulating the thyroid gland were heterogeneous. Also, LATS-P was species specific in that it stimulated human but not mouse thyroid yet blocked the binding of LATS to human thyroid. It is comforting to know that assays for LATS and LATS-P have now been superseded by a variety of direct techniques discussed below. Nevertheless it was the discovery of these thyroid stimulating Igs by *in vivo* techniques that provided the first insights into the immunological basis of Graves' disease.

In 1974 several investigators (Manley *et al.*, 1974; Smith and Hall, 1974; Mehdi and Nussey, 1975) demonstrated that binding of TSH to thyroid membranes could be inhibited by Graves' Igs. This capacity to displace TSH binding is now widely used to detect Graves' Igs, but it does not correlate well with the capacity of Graves' Igs to stimulate thyroid tissue *in vitro* and hence mimic the actions of TSH (Endo *et al.*, 1978; Kuzuya *et al.*, 1979).

A variety of methods for detecting thyroid "stimulators" have been developed, employing isolated human or bovine thyroid cells or membranes or slices of human and other mammalian thyroids. Endpoints used to measure thyroid stimulating activity include activation of adenyl cyclase or production of cAMP, release of T_3, and incorporation of ^{131}I or ^{32}P. The assays and their nomenclature are summarized in Table II. As shown, the percentage of patients positive for Graves' Igs ranges from 50 to 100%, depending on the technique. The discrepancies obviously reflect variations in assay conditions or tissue preparations. However, even if

TABLE II

TSH Receptor Autoantibodies in Graves' Disease: Assays, Nomenclature, and Prevalence

Description	Assay	Abbreviation	Positive (%)	Authors
Radioiodide release from labeled guinea pig thyroid	Long-acting thyroid stimulator	LATS	50	Adams and Purves (1956)
Radioiodide release from labeled mouse thyroid	Long-acting thyroid stimulator	LATS	—	McKenzie (1958)
Inhibition of binding of LATS to human thyroid	Long-acting thyroid stimulator-protector	LATS-P	90	Adams and Kennedy (1967)
[^{32}P] incorporation into dog thyroid slices	Long-acting thyroid stimulator	LATS	—	Field *et al.* (1968)
Inhibition of binding of labeled bovine TSH to human thyroid membranes	TSH-displacing activity	TDA	—	Manley *et al.* (1974)
Inhibition of binding of labeled bovine TSH to human thyroid membranes	TSH-binding inhibitory immunoglobulins	TBII	—	Smith and Hall (1974)
Release of T3 from thyroid slices	Human thyroid stimulator	HTS	—	Laurberg and Weeke (1975)
Inhibition of TSH binding to human thyroid membranes	TSH-binding inhibitory immunoglobulins	TBII	100	Mukhtar *et al.* (1975)
Inhibition of TSH binding to human thyroid membranes	TSH-binding inhibitory immunoglobulins	TBII	59	Endo *et al.* (1978)

Adenylcyclase stimulation in human thyroid membranes	Human thyroid adenylcyclase stimulator	HTACS	60	Orgiazzi *et al.* (1976)
Adenylcyclase stimulation in human thyroid slices	Thyroid-stimulating antibodies	TsAb	—	McKenzie and Zakarija (1976)
Adenylcyclase stimulation in human thyroid slices	Thyroid-stimulating immunoglobulins	TSI	—	McKenzie and Zakarija (1976)
Inhibition of TSH binding to human thyroid membranes	TSH-displacing activity	TDA	76	O'Donnell *et al.* (1978)
Inhibition of TSH binding to human thyroid membranes	Thyroid-stimulating antibodies	TsAb	78	Strackosch *et al.* (1978)
Inhibition of TSH binding to human thyroid membranes	Thyroid-stimulating antibodies	TsAb	50	Bryson *et al.* (1976)
Stimulation of DNA synthesis in guinea pig thyroid slices	Thyroid growth-stimulating immunoglobulins	TGI	—	Drexhage *et al.* (1980)
Several of above	TSH receptor antibodies	TRA	—	Hall (1980)
Adenylcyclase stimulation in human thyroid membranes	Thyroid-stimulating antibodies	TsAb	82	Bech and Nistrup-Madsen (1979)
Adenylcyclase stimulation in human thyroid cells	Thyroid-stimulating immunoglobulins	TSI	80	Etienne-Decerf and Winand (1981)

the conditions were physiological, it would be unreasonable to expect perfect agreement, because Graves' Igs are polyclonal and functionally heterospecific and the sites for binding and stimulation may not be identical (Manley *et al.*, 1982; Valente *et al.*, 1983).

As might be expected there is only a poor correlation between these *in vitro* assays, measurements of LATS and LATS-P, and the clinical indices of thyroid dysfunction. While LATS-positive sera usually contain TSH-binding inhibitory immunoglobulins (TBII) (Ozawa *et al.*, 1979), they may also be negative for TBII (Schleusener *et al.*, 1978); TBII-positive sera are frequently LATS negative. Therefore, TBII-positive serum may or may not stimulate the mouse thyroid. Further evidence that Igs can bind to the TSH receptor but not stimulate the human thyroid is the finding that TBII are found in the sera of some patients with ophthalmic Graves' disease, who have no obvious thyroid gland dysfunction, and in hyperthyroid patients made euthyroid by antithyroid drugs (Teng *et al.*, 1977). The TBII have also been detected in the cord blood of mothers with Graves' disease whose infants did not develop thyrotoxicosis (Hales and Luttrell, 1980), in some patients with Hashimoto's thyroiditis and hypothyroidism (Orgiazzi *et al.*, 1976), and in patients with other autoimmune diseases (Strakosch *et al.*, 1978). Furthermore, thyroid-stimulating immunoglobulins measured by cAMP stimulation in human thyroid slices (Onaya *et al.*, 1973) or primary cultures of human thyroid (Etienne-Decerf and Winand, 1981) do not correlate with LATS.

Sera that are LATS-P positive have been shown to contain thyroid-stimulating immunoglobulin (TSI) activity *in vitro* (Shishiba *et al.*, 1973; Holmes *et al.*, 1979); in the original studies by Adams and co-workers (1974), LATS-P showed a positive correlation with thyroid uptake of ^{131}I in untreated Graves' patients, and infusion of LATS-P plasma into normal human volunteers resulted in stimulation of thyroid function. These data indicate that LATS-P serum must contain stimulating Igs. However, in view of recent evidence for the nonspecificity of TBII (Davies, 1981), the nature of LATS-P as originally conceived by Adams and Kennedy (1971)—an Ig specific for human but not mouse thyroid—is open to question. An alternative explanation could be that LATS-P is a non-species-specific TBII, which in itself has no stimulating activity but which is associated with TSI.

Although assays for LATS and LATS-P have been replaced by more convenient *in vitro* assays (Table II), the latter are neither standardized nor subject to interlaboratory control. A recent modification to the standard TBII assay, giving improved precision, involves the use of Triton-solubilized TSH receptors (Shewring and Rees-Smith, 1982). However, until the molecular specificities of the antibodies are defined, bioassays

such as those using primary cultures of thyroid cells would appear to be the most appropriate for measuring TSI activity *in vitro*. By analogy with the studies of autoantibodies to acetylcholine (ACh) and insulin receptors discussed in Section III,B and C, it should be possible to address the question of the specificity of Graves' Igs and to define the structure of the TSH receptor, if indeed it is the autoantigen. To this end we have recently been able to demonstrate (Heyma and Harrison, 1984) that Graves' Igs precipitate solubilized [^{125}I]TSH binding sites and electrophoretically defined proteins that are similar in size to those purified by TSH affinity chromatography. These findings demonstrate the feasibility of immunoprecipitation as an assay for TSH receptor antibodies and the utility of Graves' Igs as probes of thyroid autoantigen structure. They support the view that the TSH receptor is the target for Graves' Igs.

Drexhage and co-workers (1981) have described a further class of autoantibodies that appear to block the trophic effect of TSH on the thyroid in primary hypothyroidism. In contrast, other antibodies with the ability to mimic the trophic effects of TSH (stimulation of [^{3}H]thymidine uptake into thyroid DNA) have been detected in the sera of patients with goiter and with or without Graves' disease (Drexhage *et al.*, 1980; Valente *et al.*, 1983). Thus there exists a range of functionally defined thyroid autoantibodies and perhaps immune complexes (Van Der Heide *et al.*, 1980) within which TSI is but one example.

4. TSI and Antithyroid Drugs

The thiocarbamide drugs used for the treatment of Graves' thyrotoxicosis, namely carbimazole, methimazole, and propylthiouracil, were first introduced in the 1940s. These drugs block thyroid hormone synthesis by inhibiting incorporation of iodine into an organic form, thyroid peroxidase, and coupling of iodotyrosines to form iodothyronines. In addition there is now evidence that the thiocarbamide drugs have immunosuppressive effects. They impair the mitogenic activity of lymphocytes and the production of TSI by lymphocytes from thyrotoxic patients (Hallengren *et al.*, 1980; McGregor *et al.*, 1980a).

Most thyrotoxic patients treated with these drugs show improvement, but permanent remission after therapy is unpredictable and varies from 14 to 70% (Marchant *et al.*, 1978; Wartofsky, 1973). The question is whether remission and relapse can be predicted in subjects on antithyroid therapy, because alternative treatment (^{131}I or surgery) can be offered. Most investigators feel that the disappearance of TSI on drug treatment is associated with an increased likelihood of remission. For example, only 1 of 37 patients negative for TSI (as measured by cAMP stimulation) after treat-

ment subsequently relapsed, whereas 34 of 36 patients still positive for TSI relapsed (McKenzie *et al.*, 1978). Similar trends were noted using TBII as a marker (Schernthaner *et al.*, 1980). Using LATS-P, Hardisty and Munro (1980) reported relapse in 21 of 24 patients positive at the cessation of carbimazole treatment, compared with 8 relapses in 23 patients negative for LATS-P.

A significant association has been found between the presence of the human leukocyte antigen HLA-DR3 and Graves' disease (Grumet *et al.*, 1974; Farid *et al.*, 1976), especially between HLA-DR3 and the eye and skin complications, but no correlation has been found between HLA-DR3 and the titer of TSI (Bech and Nistrup-Madsen, 1981). However, persistence of hyperthyroidism after a course of antithyroid drug is associated with both HLA-DR3 and the closely linked locus *HLA-B8* (Irvine *et al.*, 1977). Furthermore, HLA-DR3-positive patients are more resistant to the effect of radioiodine therapy (Farid *et al.*, 1980). The HLA antigens appear to be markers for an immune-response gene that influences the susceptibility to and the persistence of Graves' disease. A combination of HLA typing and a suitable assay for TSI appears to be useful in predicting relapse after antithyroid drug treatment (McGregor *et al.*, 1980b).

B. Myasthenia Gravis: Acetylcholine Receptor Antibodies

1. Preamble

Myasthenia gravis is a disorder of neuromuscular transmission resulting in weakness and easy fatigability, especially of facial, ocular, pharyngeal, laryngeal, and respiratory muscles. A characteristic feature in myasthenia is a decremental electromyographic response to successive low frequency stimuli (2–3 Hz), the electrophysiological counterpart of muscle fatigability. This phenomenon, while not completely understood, can be produced by several mechanisms including the simple functional blockade of ACh receptors with specific antagonists. It is reversed by the use of anticholinesterase drugs, which prolong the action of ACh (reviewed by Drachman *et al.*, 1978; and Engel, 1979).

Striated muscle fibers are innervated by myelinated nerve axons extending from motor neuron cell bodies in the spinal cord. The synapse between the motor neuron and the muscle fiber is known as an end plate. In the nerve ending, ACh is stored in 500-Å vesicles, each containing up to 10,000 ACh molecules. Depolarization of the normally negative membrane potential in the nerve terminal by a descending action potential wave causes exocytosis of 50–200 ACh vesicles into a 900-Å synaptic

gap, opposite the postsynaptic membrane where ACh receptors are concentrated on folds at 20,000 binding sites/μm^2. Binding of ACh to these receptors triggers the opening of a cation channel through which sodium, calcium, and potassium ions flow down concentration gradients. The summation of potential changes from these channel openings generates the end-plate potential. When this exceeds a certain threshold level, an action potential is generated and propagated to activate muscle contractile processes. Normally, the amount of ACh released and the occupancy of receptors greatly exceeds the minimum necessary for activation (the "safety factor"), but the potential decays rapidly as ACh levels fall due to destruction by acetylcholinesterase and to diffusion out of the synapse. Even in the resting state, miniature endplate potentials (MEPPS) of ~1 mV are recordable due to the spontaneous release of the contents of single vesicles. The early observation by Elmqvist and co-workers (1964) that MEPPS were deficient in myasthenic patients led many investigators to conclude that the defect was an inadequate release of ACh. However, the structure of the presynaptic membrane and the amount of Ach released in myasthenia are normal. The abnormality is a defect in the action of ACh, due to reduced numbers of ACh receptors and to an altered structure of the postsynaptic membrane that is induced by antibodies to the ACh receptor (Fambrough *et al.*, 1973; Engel *et al.*, 1977; Lindstrom and Lambert, 1978).

2. The ACh Receptor

The ACh receptor was a prototype for studies on receptor structure and function, due to the intense effort directed toward understanding the pathogenesis of myasthenia gravis and because ACh receptors were readily available in a concentrated form in the electric organs of *Electrophorus electricus,* the electric eel, and various species of *Torpedo,* the electric ray. The receptor is an integral transmembrane glycoprotein, and in *Torpedo californica* it consists of four peptide chains of M_r 40,000 (α), 50,000 (β), 60,000 (γ), and 65,000 (δ) present in a mole ratio of 2:1:1:1 that form a macromolecular complex of M_r 250,000 (Raftery *et al.*, 1980). The ACh receptor from *Torpedo* spp. (but not from eel or mammalian muscle) exists in the postsynaptic membrane predominantly as a dimer, linked via disulfide bonds in the M_r 65,000 chains (Hamilton *et al.*, 1977). The monomer contains the ion channel and two ACh binding sites, the latter formed at least in part by the 40,000 chains, since these chains are affinity labeled by ACh analogs (Karlin *et al.*, 1975). As shown by Raftery and co-workers (1980), the four subunits have distinct but homologous sequences suggesting that they descended from a single ancestral gene. Electronmi-

croscopy of freeze-fractured and negatively stained synaptic membranes containing ACh receptors reveals structures ~85 Å in diameter with a central "pit" that is ~25 Å in diameter. The relationship of the channel to these structures and to the central pit is still unclear.

3. Autoimmunity in Myasthenia Gravis

A number of lines of evidence, including the presence of thymic hyperplasia and antibodies to striated muscle, changes in serum complement, and the occurrence of transient weakness in the newborn of some myasthenic mothers, led to the suggestion by Simpson (1960) and Nastuk and co-workers (1960) that myasthenia gravis was an autoimmune disease, possibly mediated by autoantibodies to the ACh receptor. The mechanisms of impaired neurotransmission in myasthenia remained controversial for another decade until ultrastructural studies by Engel and Santa (1971) convincingly demonstrated that the lesion in myasthenia was in the postsynaptic membrane. The missing link in the pathogenesis was then provided with the fortuitous observation by Patrick and Lindstrom (1973) that rabbits immunized with ACh receptors purified from the electric organs of *Electrophorus* developed muscular weakness similar to that seen in humans with myasthenia gravis. This experimental model, now called experimental autoimmune myasthenia gravis (EAMG), was associated with autoantibodies to the ACh receptor and provided the clue to the identification of such antibodies in the human syndrome.

Most studies have demonstrated that autoantibodies to the ACh receptor do not act mainly as competitive antagonists of ACh binding, but lead to receptor loss by accelerating the internalization and degradation of the receptor and by inducing complement-mediated lysis of the postsynaptic membrane (Appel *et al.*, 1977; Kao and Drachman, 1977; Heinemann *et al.*, 1977). Antibody-induced receptor degradation, or antigenic modulation, is equivalent to down-regulation of receptors induced by exposure to homologous ligands. Engel *et al.* (1981) also showed that receptor degradation and synthesis are both increased, but that the insertion of newly synthesized receptors into the end plate is impaired. The ability of ACh receptor antibodies to induce receptor loss requires antibody bivalency and therefore may depend on cross-linking of receptor subunits (Drachman *et al.*, 1978). Studies of EAMG in animals also reveal that neurotransmission is affected less by direct antibody binding than by loss of ACh receptor (Engel, 1979). Comprehensive descriptions of the immunopathology of ACh receptors in myasthenia gravis and EAMG may be found in reviews by Lindstrom (1979) and Vincent (1980).

Animals immunized with purified ACh receptor in Freund's and *Borde-*

tella pertussis adjuvants develop an acute myasthenic syndrome after 8–12 days. This is associated with measurable levels of ACh receptor antibody, and most strikingly, with intense phagocytic infiltration of the end plate, complement fixation, and membrane lysis, leading to rapid denervation. This acute phase may be an artifact or an enhanced reaction due to the *B. pertussis* since it is not usually seen in animals immunized without this adjuvant. The chronic phase of EAMG occurs after 30–40 days and is associated with high titers of ACh receptor antibody, loss of receptors, "simplification" of the postsynaptic membrane folds, and complement-dependent lysis, but no phagocytic infiltration. This pathology resembles that seen in human myasthenia. In both situations the amount of ACh receptor, bound antibody, and bound C3 complement is greatest in the least severely affected subjects, consistent with the evidence that loss of ACh receptor causes impaired transmission.

Experimental autoimmune myasthenic gravis can be passively transfered by globulins from an affected to a normal rat (Lindstrom *et al.*, 1976a). Transfer is rapid and within a day or two there is an intense phagocytic infiltration resembling EAMG. This indicates that antibodies alone (with complement) are capable of inducing a cellular response at the endplate. Passive transfer from human to mouse has also been accomplished (Toyka *et al.*, 1975). Passive transfer and acute EAMG do not occur if C3 complement is depleted (by treatment with cobra venom factor), even though antibodies are bound (Lennon *et al.*, 1978). This demonstrates the importance of complement fixation in opsonizing the membrane for destruction by phagocytes. Antigenic modulation does not appear to depend on complement but simply on the presence of bivalent antibodies. The human counterpart of the passive transfer experiment is the transplacental passage of antibodies from some myasthenic mothers to their offspring; weakness in the neonate declines during 4–6 weeks as antibody levels decrease.

4. ACh Receptor Antibodies

Acetylcholine receptor antibodies in animals with chronic EAMG or in humans with MG are 7 S IgG, the major subclass being IgG-III (Lindstrom, 1979). Their production is T-lymphocyte dependent and is not seen in rats subjected to neonatal thymectomy and X irradiation (Lennon *et al.*, 1976). The responses to immunization in these animals can be reconstituted by B plus T lymphocytes, but not by B lymphocytes alone. Thymetomy of adult immunized rats does not, however, prevent EAMG.

Antisera against each of the four subunits of the *Torpedo californica* receptor cross-react with receptors in human muscle and with the same

subunits from other species (Lindstrom *et al.*, 1979), suggesting conservation of common determinants exposed on the extracellular surface. However, species specificity is also present, as indicated by the limited cross-reactivity of MG patients sera with ACh receptors from rat muscle or electric organs. Thus, antibodies in MG sera do not react with eel ACh receptor, although antibodies to eel receptor do cross-react with human muscle ACh receptor. In both EAMG and MG, most antibodies are against a main immunogenic region on the 40,000-dalton (α) subunit, but antibodies to many other antigenic determinants are also produced (Tzartos *et al.*, 1981).

Lindstrom and co-workers (1979) found that EAMG in rats will develop after immunization with any single receptor subunit. This suggests that there is no single myasthenogenic determinant and is in keeping with the evidence that the myasthenic state is due to the accelerated degradation of receptors triggered by antibody crosslinking, rather than to antibody interference with the function of a specific determinant. Most of the antibodies are directed at determinants other than the ACh binding site. The binding of α bungarotoxin to human ACh receptor is inhibited by only about one-third of MG sera. Furthermore, Lindstrom and colleagues have shown that antibody–ACh receptor complexes extracted from affected muscle can still bind toxin, and that it makes little difference whether antibodies are measured by using ^{125}I-labeled ACh receptor as antigen or [^{3}H]acetyl ACh receptor, in which the toxin binding sites are free.

A further interesting finding concerns the differences between junctional receptors concentrated at the end plate and extrajunctional receptors found at lower densities over the entire surface of embryonic or denervated adult muscle fibers. Although both forms of the receptor have the same gross subunit structure, they differ in several respects including immunoreactivity. Weinberg and Hall (1979) found that rat extrajunctional receptors have determinants not present on junctional receptors that are recognized by some human MG antibodies but not by several antisera to purified ACh receptors. One interpretation therefore might be that the autoimmune reaction in MG is against an embryonic receptor, perhaps resembling the ACh receptor, on myoid cells of the thymus (Fuchs *et al.*, 1980).

Tzartos and Lindstrom (1980) and Lennon and Lambert (1980) raised monoclonal antibodies against purified ACh receptors and found that infusion of these Abs into normal rats produced myasthenia. This indicates that myasthenia can result from a defect in a single clone of immunocytes producing an antibody presumably against a single determinant on the receptor. Lindstrom and his colleagues have also used monoclonal antibodies to map the receptor structure (Tzartos *et al.*, 1981). Approximately

50% of monoclonals are species nonspecific, and many react with a main immunogenic region on the 40,000-dalton (α) subunit outside the ACh binding site. Tzartos *et al.* suggest that these antibodies may be directed against a highly conserved immunogenic region important in receptor function. Delineation of the antibody specificities at the monoclonal level could lead to specific forms of immunotherapy for MG, as discussed in Section IV.

5. Humoral versus Cellular Immunity

Both humoral and cellular immune responses can be demonstrated in myasthenia, but the chief effector mechanism appears to be humoral. The humoral response is T-cell dependent, and animals with EAMG demonstrate delayed-type hypersensitivity to the ACh receptor. Lymph node cells can transfer myasthenia, but the result is not as dramatic as after passive transfer with antibodies (Lennon *et al.*, 1976). Weakness is also much more delayed and probably results from antibody production in the recipient. Purified ACh receptors stimulate [^{3}H]thymidine uptake in peripheral blood lymphocytes from MG patients, but this is not a universal finding (Lennon and Lambert, 1980). Cellular responses to the ACh receptor have also been observed in some cases of polymyositis with the clinical features of MG (Conti-Tronconi *et al.*, 1978), but ACh receptor antibodies have not been found in polymyositis and are present in $<1\%$ of patients with other autoimmune and neurological disorders (Lindstrom *et al.*, 1976b). The evidence for a primary role for ACh receptor antibodies in myasthenia can be summarized thus: (1) antibodies, not cells, are found at the end plates in MG and chronic EAMG; (2) EAMG can be induced by passive transfer of antibodies from human to mouse or from rat to rat or by infusion into normal rats of monoclonal antibodies against ACh receptor; and (3) removal of antibodies by plasmapheresis results in clinical and electrophysiological improvement.

6. Measurement of ACh Receptor Antibodies

The frequency of antibody detection in MG patients depends on the assay employed and the source of ACh receptor. The most sensitive assay is the so-called radioimmunoassay, the indirect immunoprecipitation of solubilized receptor from human muscle labeled with [^{125}I]bungarotoxin (Appel *et al.*, 1975; Lindstrom *et al.*, 1976b; Mittag *et al.*, 1976; Monnier and Fulpius, 1977). The immunoprecipitation assay is positive in $>90\%$ of patients. It does not detect antibodies against the binding site, but this is not a major problem since most antibodies are against other determinants. In fact, receptor–antibody complexes can be extracted from muscle, la-

beled with [^{125}I]toxin, and then precipitated with antihuman IgG (Lindstrom *et al.*, 1976b).

Although the titer of antibodies is generally low in patients with localized (e.g., ocular) MG, the correlation between titer and the severity of generalized MG is poor (Lindstrom *et al.*, 1976b), and patients in remission are reported to have titers within the range seen in active disease (Vincent, 1980). Thus the ability of antibodies to precipitate solubilized receptors does not always match their ability to impair function. Sera could contain additional antibodies to solubilized receptors not recognized by receptors *in situ*, especially if the autoimmune response was maintained by the shedding of postsynaptic membrane. Additional factors possibly responsible for the discrepancy between antibody titers and clinical status might include the degradation on storage of specific antibody subclasses (e.g., IgG3) responsible for functional effects, the activity of the complement system in individual patients, and the ability to repair end-plate damage.

7. Factors Bearing on the Treatment of Myasthenia

Rational therapy for myasthenia gravis and for other autoimmune diseases would aim to eradicate the clones producing autoantibodies or to restore immunoregulation to normal. Unfortunately, because of our ignorance of the precipitating and sustaining factors in autoimmunity and our lack of detailed knowledge of the molecular bases of autoimmune phenomena, therapy remains largely empirical. There is no evidence for an intrinsic receptor defect in myasthenia gravis or the other receptor antibody diseases. However, there is evidence for impaired immunoregulation.

Caucasians with MG fall into two broad classes. The majority are young females with a peak age of onset in the third decade; ~65% have thymic hyperplasia, and tissue typing reveals a significantly higher frequency of the HLA antigens A1, B8, and DR3. In a minority are older males in whom thymoma is common; they have a high frequency of antibodies to striated muscle and an increased frequency of HLA antigens A2 and A3. In Japanese, thymic hyperplasia is associated with HLA-B12 and thymoma with HLA-B5 (Yoshida *et al.*, 1977). Thymectomy, particularly in young females, is generally associated with clinical improvement, but the mechanism is unknown and this procedure is not usually associated with a dramatic fall in receptor antibody titers. The role of the thymus in MG is still poorly understood. The thymus contains ACh receptors on myoid cells (Fuchs *et al.*, 1980), and it has been postulated that MG may be initiated there. However, thymectomy is not always beneficial

and so the thymus is probably not the sole source of autoantigenic stimulation. It is possible that in established MG the shedding of ACh receptors from damaged end plates could sustain the autoimmune process.

Penicillamine treatment of rheumatoid arthritis has been associated with the reversible development of MG (Bucknall *et al.*, 1975) and ACh receptor antibodies (Masters *et al.*, 1977). This drug has also been associated with other autoimmune responses indicating that it induces a defect in immunoregulation (Dawkins *et al.*, 1981).

Corticosteroids and other immunosuppressive agents given to patients with MG result in clinical improvement and a decrease of antibody titer (Engel, 1979; Lindstrom, 1979; Vincent, 1980). Antibody concentration can be decreased and a temporary clinical improvement observed after plasmapheresis (exchange of the patient's plasma with normal plasma) (Vincent, 1980). Drainage of the thoracic duct, which removes lymphocytes as well as antiboides, is also effective (Matell *et al.*, 1976), but these procedures, which physically remove globulins, must be combined with immunosuppressive drug treatment to prevent rebound synthesis of receptor antibodies.

C. Insulin-Resistant Diabetes: Insulin Receptor Antibodies

1. Preamble

Diabetes mellitus is a condition characterized by chronic hyperglycemia due to defects in insulin secretion and/or action and, in most cases, by the development of obstructive lesions in small and large blood vessels leading to complications especially in the eyes, kidneys, and nerves. Clinically, diabetes is classified into two types: type 1 or insulin-dependent diabetes (IDDM, 20% of diabetics) occurs predominantly in young persons and is due to a deficiency of pancreatic insulin-producing β cells; type 2 or non-insulin-dependent diabetes (NIDDM, 80% of diabetics) occurs predominantly in middle-aged persons who tend to be obese and is associated with a relative lack of insulin secretion in the face of tissue resistance to the action of insulin. A large body of evidence indicates type 1 diabetes is due to the autoimmune destruction of pancreatic β cells triggered by virus infection or possibly in some cases by exposure to chemical toxins (Nerup and Lernmark, 1981). This form of diabetes is likely to emerge as a receptor antibody disease when the structure and function of the beta cell surface antigen is defined (see Section IV.)

Under the mantle of non-insulin-dependent diabetes are an increasing number of genetic and acquired disorders associated with "secondary

diabetes" in which tissue resistance to insulin is the common characteristic feature (Harrison and Flier, 1980). Notwithstanding, the great majority of type 2 diabetics do not have an obvious predisposing disease. Their disorder, impaired glucose utilization through insulin-sensitive pathways, could be inherent to cells (e.g., an inborn error) or due to the effect of a circulating inhibitor. There is no evidence for an inherent defect in cultured diabetic cells (Howard *et al.*, 1981). On the other hand, we recently have provided evidence for a circulating (nonantibody) inhibitor of postreceptor insulin action in many type 2 diabetics (Dean *et al.*, 1984).

In this section I will describe an uncommon form of non-insulin-dependent diabetes, an autoimmune syndrome of insulin resistance due to autoantibodies to the insulin receptor. This syndrome has been studied in considerable detail and the insulin receptor autoantibodies have served as unique probes of receptor structure and function.

2. The Insulin Receptor

The structure of the receptor has been elucidated by labeling the binding sites with photoreactive ^{125}I-labeled insulins (Yip *et al.*, 1980; Jacobs *et al.*, 1979; Wisher *et al.*, 1980) or chemically cross-linked ^{125}I-labeled insulin (Pilch and Czech, 1980; Kasuga *et al.*, 1981) or by autoantibody precipitation of receptors surface-labeled with ^{125}I (Lang *et al.*, 1980; Harrison and Itin, 1982) or biosynthetically labeled with ^{35}S (Van Obberghen *et al.*, 1981) (Fig. 4). In addition, the receptor protein has been analyzed after sequential chromatographic purification on columns of immobilized lectins and either insulin (Jacobs *et al.*, 1977; Fujita-Yamaguchi *et al.*, 1983) or receptor antibodies (Harrison and Itin, 1980). The consensus from these studies is that the insulin receptor in several human and rat tissues is a multivalent oligomer of M_r ~350,000 composed of disulfide-linked M_r 130,000 (α) and 90,000 (β) subunits. The M_r 130,000 subunit is the insulin-binding site; the M_r 90,000 subunit possesses tyrosine kinase activity stimulated by insulin and is autophosphorylated (Kasuga *et al.*, 1982; Van Obberghen *et al.*, 1983). Both subunits are surface glycoproteins whose disappearance parallels the down-regulation of insulin binding after exposure of cells to supraphysiologic concentrations of insulin (Harrison *et al.*, 1982a). The receptor is synthesized as a single polypeptide chain, which is glycosylated and proteolytically cleaved to yield the two disulfide-linked subunits (Ronnett *et al.*, 1984).

The stoichiometry of the disulfide-linked oligomer has not been determined directly, but analysis of graded reduction products after chemically cross-linking ^{125}I-labeled insulin suggests a combination of two pairs of M_r 130,000 and 90,000 subunits (Massague *et al.*, 1980). Several high molecu-

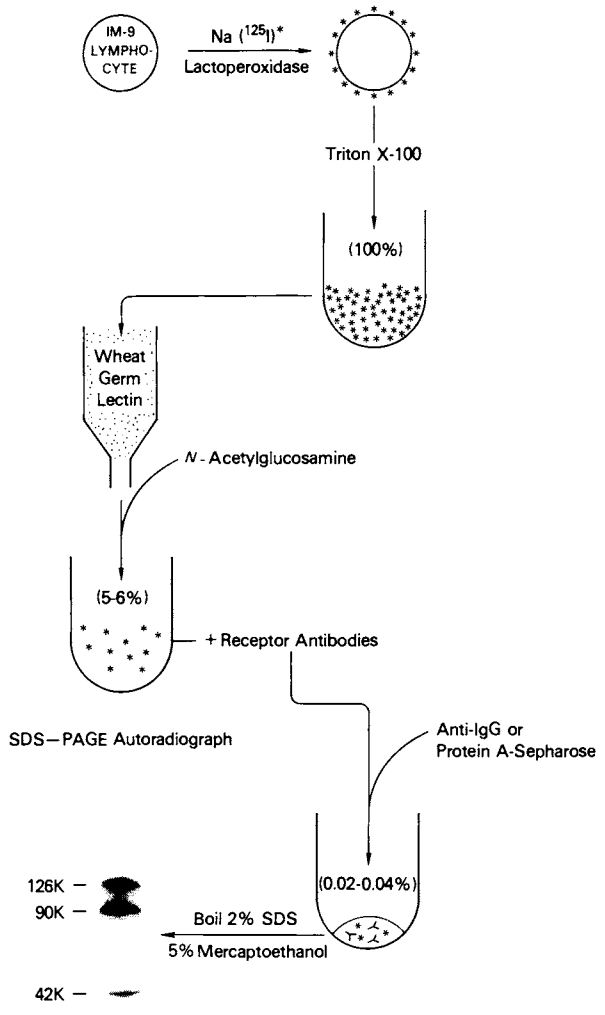


FIG. 4. Isolation of insulin receptor subunits from ^{125}I-surface-labeled lymphocytes using receptor autoantibodies.

lar weight, partially reduced oligomers are also observed, indicating that the receptor exists in a redox equilibrium (Yip and Moule, 1983; Heidenreich *et al.*, 1983). We have shown that a small fraction of insulin specifically bound to receptor in adipocytes becomes covalently attached via a sulfhydryl–disulfide interchange (Clark and Harrison, 1982). Disulfide binding occurs to the M_r 130,000 subunit within partially reduced oligomeric receptors (Clark and Harrison, 1985), and a number of lines of

evidence suggest that this covalent modification is a necessary step in insulin action (Clark and Harrison, 1983).

Insulin binding is coupled to a variety of postreceptor events in different tissues, and while the gross structure of the receptor may be the same in most tissues, it seems likely that fine differences must exist to subserve different postbinding functions. This is suggested by the fact that solubilized receptors from different species and from different tissues within the same species exhibit different reactivities with polyclonal antisera in a receptor radioimmunoassay (Harrison *et al.*, 1979a).

3. Autoimmunity and Insulin Resistance

Autoantibodies to the insulin receptor were discovered by Flier and coworkers in 1975 in the sera of three nonobese women with severe insulin resistance and the skin disorder acanthosis nigricans (Flier *et al.*, 1975). These women had evidence of generalized autoimmune disease and were classified as having the type B syndrome of insulin resistance and acanthosis nigricans, to distinguish them from type A patients with insulin resistance and acanthosis nigricans who also had hirsutism and polycystic ovarian disease but who lacked receptor autoantibodies (Kahn *et al.*, 1976). A total of no more than 25 patients with insulin receptor autoantibodies have been documented thus far, making this the least common of the known receptor antibody diseases.

Insulin receptor antibodies are polyclonal and predominantly of the IgG class (Flier *et al.*, 1976). They inhibit insulin binding (Flier *et al.*, 1975,1976,1977), acutely mimic the actions of insulin (Kahn *et al.*, 1977,1978), and immunoprecipitate solubilized insulin receptors (Harrison *et al.*, 1979b). The clinical features of the patients have been reviewed (Kahn and Harrison, 1981) and are summarized in Table III. Patients have presented with symptomatic diabetes resistant to extremely high doses of insulin (up to 24,000 U/day). As with other autoimmune diseases females predominate, but there are notably few Caucasians. Nearly all patients had acanthosis nigricans, a pigmented thickening of the skin over the extensor surfaces of the neck and joints, frequently extending over the trunk and face and associated with multiple skin tags. The cause of acanthosis is uncertain, but the lesion seems to mirror the severity of insulin resistance and the degree of hyperinsulinemia and thus could possibly represent a growth effect of supraphysiological concentrations of insulin. Most patients had evidence of generalized autoimmune disease usually manifested by nonspecific serological markers, and five patients had the classic features of systemic lupus erythematous. In contrast, generalized features of autoimmunity are uncommon in other receptor antibody dis-

TABLE III
Features of Patients with Insulin Receptor Antibodies[a]

Total number of patients documented	18
Females/males	14/4
Age range: 15–62 years	
Race: 12 Black, 3 Caucasian, 2 Japanese, 1 Mexican-American	
Presenting features:	
Symptomatic diabetes	11
"Reactive" hypoglycemia	1
Lupus erythematosis	1
Sjögren's syndrome	2
Polyarthralgia	1
Asymptomatic associated with ataxia–telangiectasia	2
Acanthosis nigricans	13
Severe glucose intolerance	15
Tendancy to ketosis	8
Severe insulin resistance	17
Generalized autoimmune features at some time (e.g., vitiligo, alopecia, submandibular gland enlargement, thyroiditis, glomerulonephritis, hemolytic anemia, arthralgias, splenomegaly, antinuclear antibodies)	16
Development of hypoglycemia	3
Improvement or remission (with or without therapy)	9

[a] From Kahn and Harrison (1981).

eases. None of the patients had evidence of organ or other receptor-specific autoimmunity, and in screening sera from many patients with other autoimmune diseases insulin receptor antibodies have not been detected (L. C. Harrison and J. S. Flier, unpublished).

The clinical course has been variable, although the tendency has been for insulin resistance to improve and antibody titers to fall, with or without therapy (Kibata *et al.*, 1975; Kawanishi *et al.*, 1977; Blackard *et al.*, 1977; Flier *et al.*, 1978). At least half of the patients have been given immunosuppressive drugs, usually glucocorticoids with or without cyclophosphamide, but in only two cases could a remission be reasonably attributed to treatment (Kibata *et al.*, 1975; Flier *et al.*, 1978). Two patients have been treated by plasma exchange, and in one case this procedure resulted in a progressive decrease in antibody titer with a concomitant increase in insulin binding to the patient's cells *in vitro* (Muggeo *et al.*, 1979a). These changes were accompanied by disappearance of ketosis and by a slight increase in insulin sensitivity *in vivo*, but the results were short lived and the antibody levels rebounded within several days of the last exchange. To be of practical benefit, plasma exchange should be

combined with immunosuppressive treatment to inhibit antibody synthesis.

Three patients have developed hypoglycemia, in one case severe and intractable and leading to death (Flier *et al.*, 1978). Although the insulin receptor antibodies have acute insulin-like actions *in vitro*, they are usually associated with clinical insulin resistance, consistent with their ability to desensitize cells chronically *in vitro* (Karlsson *et al.*, 1979). The development of hypoglycemia in the fatal case was not accompanied by any apparent change in the properties of the receptor autoantibodies on normal cells and could not be attributed to excessive levels of insulin or insulin-like growth factors. There was, however, a large increase in the insulin binding capacity of the patient's cells due to an increase in low-affinity binding sites that were not blocked by autoantibodies *in vitro*. It seems likely that this receptor proliferation was responsible for hypoglycemia and that the receptor antibodies failed to down-regulate or desensitize receptors, continuing to act as agonists in an analogous fashion to thyroid-stimulating antibodies in Graves' disease. Another patient with receptor autoantibodies presenting with hypoglycemia has also been reported (Taylor *et al.*, 1982) but is not included here.

Antibodies to the insulin receptor have also been found in the New Zealand obese (NZO) mouse, an animal model of obesity and insulin-resistant diabetes (Harrison and Itin, 1979). These NZO mice were inbred from a stock colony that had originally given rise to the NZB and NZB/W models of autoimmune hemolytic anemia and systemic lupus erythematous, both of which have been the subject of extensive immunological investigations. When Marion Bielschowsky first characterized the metabolic syndrome in NZO mice, she suggested that the insulin resistance might be due to an antagonist of insulin action (Bielschowsky and Bielschowsky, 1956), but the logical possibility that an antagonist might be an insulin receptor autoantibody was not previously considered. In contrast to the human antibodies, those in the sera from NZO mice are of low titer and are IgM class. They only partially inhibit insulin binding to human receptors but can be detected by their ability to immunoprecipitate solubilized receptors (Harrison and Itin, 1979). Further evidence for an immune basis for the NZO syndrome is the finding of autoantibodies to single- and double strand DNA in NZO sera and the demonstration that NZO kidneys contain dense glomerular deposits of IgM (Melez *et al.*, 1980).

4. Receptor Antibodies and Insulin Binding

Binding of insulin to receptors on fresh circulating monocytes from type B patients is markedly impaired. Monocytes are a convenient source

of receptors and probably reflect the status of insulin receptors generally (Bar *et al.*, 1979). Analysis of the binding defect may be difficult because of the low level of binding and this itself suggests a loss of receptors. However, in the few patients studied, a decrease in receptor affinity has been the predominant feature; that is, the antibodies act like competitive antagonists. Greater concentrations of unlabeled insulin are required to decrease tracer ^{125}I-labeled insulin binding, and the Scatchard plot of binding shows a loss of curvilinearity due to absence of high-affinity binding at low insulin concentrations (Fig. 5) (Muggeo *et al.*, 1979b; Bar *et al.*, 1980). An "average affinity profile" shows that the receptors are "locked" in a low-affinity state (Fig. 5). Total binding capacity (receptor concentration) is usually not reduced in these analyses. Similar binding defects can be reproduced *in vitro* by incubating normal cells with patient's serum or globulins and are partially reversed by an acid wash of cells designed to elute surface immunoglobulins (Flier *et al.*, 1977; Muggeo *et al.*, 1979b; Bar *et al.*, 1980).

The lack of any major change in apparent receptor concentration in these studies is surprising, since the antibodies are themselves rapidly internalized (Carpentier *et al.*, 1979) and might be expected to down-regulate receptors. It is assumed that insulin receptors on monocytes are sensitive to regulation (e.g., by insulin) because their concentration usually mirrors that on classical target tissues such as liver and fat under

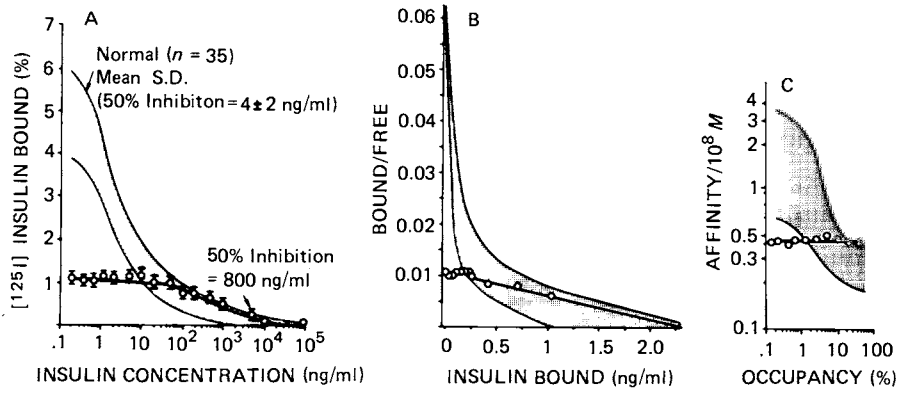


FIG. 5. [^{125}I]insulin binding to circulating monocytes in a patient (B-6) with the type B syndrome of insulin resistance and acanthosis nigricans (mean ± SEM of four basal studies). The impairment of insulin binding to the patient's cells (o——o) is due to a decrease in receptor affinity, indicated by the high concentration of unlabeled insulin required to inhibit [^{125}I]insulin binding by 50% (A), by the decrease in curvilinearity of the Scatchard plot (B), and by the flat affinity profile (C). (Data supplied by Dr. M. Muggeo: studies of 2/13/78–3/17/78.)

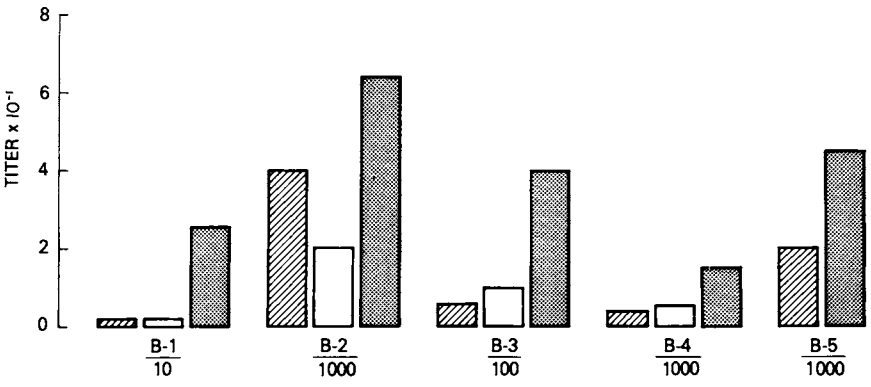


FIG. 6. Comparison of three methods used to measure titers of autoantibodies to the insulin receptor (antireceptor sera from patients B1–B5). ▨, Inhibition of [^{125}I]insulin binding; □, stimulation of glucose oxidation; ▩, immunoprecipitation of soluble receptor.

various conditions *in vivo*, but direct studies *in vitro* to confirm this point are lacking. Incubation of human lymphocytes (IM-9 line) for 60 min at 22°C with antireceptor sera causes a decrease in receptor affinity (Flier *et al.*, 1977). However, when these experiments are performed for 3 h at 37°C, antireceptor sera cause a marked decrease in receptor concentration (L. C. Harrison and A. Itin, unpublished). It has also been shown that an affinity defect induced by brief exposure of lymphocytes to an antireceptor serum progresses to an apparent defect in receptor concentration with further exposure at 37°C. These findings suggest that insulin receptor autoantibodies do indeed down-regulate the receptor under physiological conditions, via an effect on receptor affinity. Down-regulation has been demonstrated in hepatocytes exposed to rabbit antiserum against partially purified insulin receptors (Caro and Amatruda, 1980). The observation that binding to patients' monocytes shows predominantly an affinity defect is therefore difficult to reconcile. Monocyte receptors may not be modulated by antibodies in similar fashion to those on other cells, but this seems unlikely. Alternatively, the methods of data analysis may be misleading, since several assumptions underlying the Scatchard plot are not met in receptor binding studies, and negative cooperativity is but one model of insulin binding. An alternative explanation might be that patients' monocytes do in fact exhibit a loss of a separate class high-affinity receptors and a residual population of low-affinity receptors whose binding function is not altered by antibodies.

Any generalization about the mechanisms underlying the effects of receptor antibodies must account for the fact that they are polyclonal and

functionally heterospecific. Flier *et al.* (1977) found that inhibition of insulin binding to lymphocytes after short-term exposure to antireceptor sera was due to several mechanisms that included competitive occupancy of binding sites, decreased rates of association, increased rates of dissociation, and negatively cooperative effects, depending on the particular antiserum and its concentration. The molecular basis of each of these effects is unknown, but two observations suggested that they may largely be due to allosteric mechanisms with changes in binding site conformation, rather than to direct "competitive" effects on the binding site. First, one antireceptor serum was shown to reduce receptor affinity in particulate placental membranes but to decrease apparent receptor number in solubilized placental membranes (Harrison *et al.*, 1979b). Second, acute exposure of lymphocytes *in vitro* to certain antireceptor sera produced a right shift in the binding competition curve without a decrease in binding of tracer ^{125}I-labeled insulin (Harrison *et al.*, 1979c), consistent with an increase in low-affinity binding sites presumably due to conformational change and exposure of cryptic binding sites. Irrespective of these acute effects, as outlined above chronic occupancy of receptors by insulin receptor autoantibodies probably induces receptor loss.

Antireceptor sera are approximately equipotent in inhibiting binding to a variety of human and other mammalian tissues but are less potent against nonmammalian receptors (Flier *et al.*, 1975). There is no apparent preferential effect of the patients' sera on binding to their own receptors, and patients' fibroblasts in culture have normal binding properties (Muggeo *et al.*, 1979b). This suggests that the underlying binding site in affected patients is normal.

Insulin receptor autoantibodies are specific for the insulin receptor in that they do not inhibit the binding of other hormones or precipitate receptors for other hormones (Harrison *et al.*, 1979b). However, some antibodies do cross-react with and precipitate receptors for the related insulinlike growth factor-1 (IGF-I) receptor (Jonas *et al.*, 1982). When purified antireceptor IgG was labeled with ^{125}I and then receptor purified by binding and elution from cultured IM-9 lymphocytes, the eluted [^{125}I]IgG, enriched for ^{125}I-labeled antireceptor IgG, bound to a variety of cells in indirect proportion to their insulin receptor concentration, and a major part of the binding was specifically competed for by insulin and insulin analogs (Jarrett *et al.*, 1976). This indicates that a subpopulation of the antibodies, at least, binds at or very close to the binding site for insulin. Further studies have clearly demonstrated that a major proportion of the IgG antibodies in several antireceptor sera recognize the insulin binding subunits in the native, oligomeric receptor (Kasuga *et al.*, 1981; Harrison *et al.*, 1982a). On the other hand, it is likely that some

TABLE IV
Insulinlike Effects of Insulin Receptor Autoantibodies

Adipocytes
Stimulation of glucose transport, incorporation into lipid and glycogen, and oxidation to CO_2
Stimulation of amino acid transport and incorporation into protein
Inhibition of lipolysis
Activation of glycogen synthase
Inhibition of phosphorylase
Activation of pyruvate dehydrogenase and acetyl-CoA carboxylase
3T3-L1 fatty fibroblasts
Stimulation of glucose transport and oxidation to CO_2
Activation of lipoprotein lipase
Muscle
Stimulation of glucose transport and incorporation into glycogen
Activation of glycogen synthase
Liver
Stimulation of amino acid transport
Activation of glycogen synthase
Down-regulation of receptors
IM9-Lymphocytes
Down-regulation of receptors

antibodies bind to determinants away from the insulin binding site (but on the same subunits), because antireceptor sera will still recognize the solubilized receptor (Harrison *et al.*, 1979b) and mimic the actions of insulin (Kahn *et al.*, 1977) after destruction of insulin binding activity by trypsin.

5. Receptor Antibodies and Insulin Action

The clinical state associated with insulin receptor autoantibodies has generally been one of insulin resistance, in keeping with the ability of the antibodies to impair insulin binding. However, the acute effect of the antibodies in most systems *in vitro* is to mimic the actions of insulin (Table IV). The spectrum of antibody bioactivity covers all of insulin's actions except the direct growth effect of high concentration of insulin. It includes effects dependent and nondependent on glucose transport and on membrane and cytoplasmic systems, and chronic effects that require protein synthesis. The titer of most antisera with respect to stimulation of glucose uptake in fat cells is higher than that determined by binding inhibition and approaches the titer for immunoprecipitation (Fig. 6), consistent with the idea that some populations of antibodies are directed against determinants outside the insulin binding site.

The studies on antibody bioactivity have led to several important in-

sights into the nature of insulin action. Antibody bioactivity, unlike antibody inhibition of insulin binding, requires bivalency (Kahn *et al.*, 1978). Purified IgG or $F(ab)_2$ fragments are equally effective in their ability to inhibit insulin binding and mimic the actions of insulin; monovalent Fab fragments inhibit binding but have no intrinsic bioactivity. The bioactivity of Fab fragments on cells is, however, restored by the addition of anti-$F(ab)_2$ IgG (Fig. 7). These findings indicate that bioactivity requires not only receptor occupancy but also bivalent binding or cross-linking (? of receptor subunits). The importance of this requirement for the modulation of acetylcholine receptors by antibodies has already been discussed. A further example is the IgE-mediated degranulation of mast cells leading to histamine release. This event requires cross-linking of the IgE receptor with either IgE and second antibody, antibodies to the receptor itself, or chemically cross-linked IgE dimers (Metzger, 1978).

The suggestion was made by Kahn *et al.* (1978) that cross-linking might be involved in activation by insulin. They showed that antibodies to insulin sometimes enhanced the effect of suboptimal concentrations of insulin prebound to its receptors. In addition, fluorescence microscopy revealed that insulin, like receptor autoantibodies, can induce patching and capping on the cell surface (Schlessinger *et al.*, 1980). Anticytoskeletal agents, however, have no effect on the bioactivity of either receptor antibodies or insulin, and cell-surface events such as macroscopic aggregation and internalization (of fluorescent-labeled ligands) do not appear to be necessary for cellular activation. Also, it is important to note that the rates of internalization of bivalent and monovalent receptor antibodies are identical, although only the former possess bioactivity (Kasuga *et al.*,

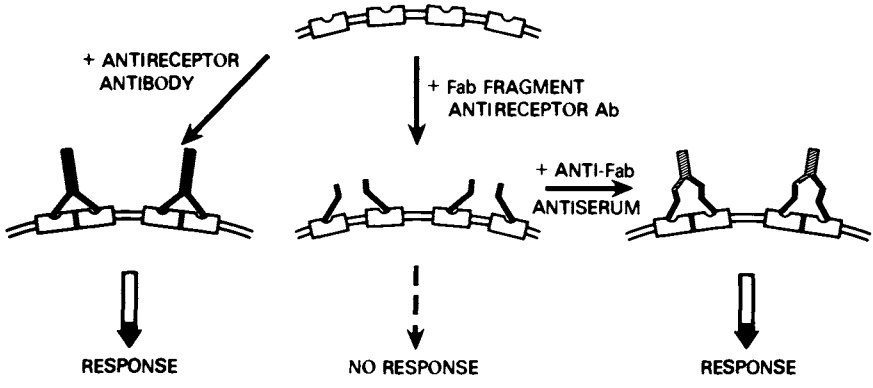


FIG. 7. Schematic representation of the role of cross-linking in insulin receptor activation by antibody. (Courtesy Dr. C. R. Kahn.)

1983). None of the evidence cited may be directly relevant to cross-linking at a molecular level. Insulin circulates as a monomer, but the possibility that it could dimerize on a multivalent receptor is not unreasonable. Alternatively, the monovalent binding of insulin may be enough to induce an equivalent conformational change. One presumes that receptor antibodies and insulin exert their bioeffects through common postreceptor pathways. Evidence that their binding interaction is not identical comes from studies of the disulfide binding of insulin. The ability of insulin to covalently attach to its receptor via a disulfide exchange process and to trigger glucose transport and several other bioeffects is inhibited by an externally acting thiol blocking agent that is without effect on the insulinomimetic actions of receptor antibody (Clark and Harrison, 1983).

The ability of cross-linking ligands other than insulin (receptor antibodies and certain lectins) to mimic the actions of insulin indicates that the program for insulin action is contained within the receptor. It also suggests that insulin degradation or a product thereof is not required (e.g., as a second messenger) for insulin action. The only action of insulin not mimicked by receptor autoantibodies is growth stimuation ([^{3}H]thymidine incorporation into fibroblast DNA). King and co-workers (1980) showed that this effect of insulin, usually seen only at supraphysiological concentrations in nontransformed cells, is mediated through receptors for insulinlike growth factors (IGFs) rather than through the insulin receptor.

The paradox of receptor autoantibodies that mimic insulin *in vitro* and donor patients who are insulin resistant and hyperglycemic is not explained by species differences, since the antibodies are bioactive in both human and rodent adipocytes. The explanation appears to be that the bioactivity is only short lived, as shown by Karlsson and co-workers (1979) in studies using cultured 3T3-L1 fatty fibroblasts. In these cells receptor antibodies inhibited insulin binding and stimulated glucose uptake and oxidation, as in freshly isolated adipocytes, but the insulinlike effects were transient, reaching a maximum by 1 h and then decreasing. After several hours basal activity returned to normal and was resistant to stimulation by insulin, concanavalin A, and further addition of receptor antibodies. Insulin binding remained inhibited. Although co-occupancy of receptors by antibodies shifts the dose–response for insulin action to the right, there was also a decrease in maximum responsiveness. This desensitization required antibody bivalence, and although the mechanism has not been clearly defined, the locus of desensitization is probably close to the binding site since the insulinlike actions of spermine and vitamin K_5 were unaltered (Grunfeld *et al.*, 1980). Densensitization was not affected by anticytoskeletal and antilysosomal agents, but this required that the

TABLE V
Assays for Insulin Receptor Antibodies

Direct
Inhibition of insulin binding
Immunoprecipitation of the solubilized receptor
Stimulation of insulinlike effects
Indirect
Reversal of binding defect after acid wash of cells
Binding of [^{125}I]protein A to cells

cells have a source of energy such as glucose, pyruvate, or a glucose derivative capable of being phosphorylated (Grunfeld *et al.*, 1980).

6. Measurement of Insulin Receptor Antibodies

The assays for insulin receptor antibodies are listed in Table V. The ability of the antibodies to inhibit insulin binding specifically and to elicit insulinlike effects confers unique specificity and is the basis for two direct, functional assays. The other type of direct assay is based on the ability of the receptor antibodies to specifically immunoprecipitate Triton-solubilized receptors labeled with [^{125}I]insulin following the addition of second antibody. Binding inhibition and immunoprecipitation are the two most practical and specific assays currently employed. Their salient features are shown in Fig. 8.

Immunoprecipitation measures all receptor antibodies, not only those that happen to impair binding or evoke insulinlike effects. For this reason, immunoprecipitation is the most sensitive technique (Fig. 6); however, the immunoprecipitation titer has not been found to exceed the binding inhibition or bioactivity titers by more than fourfold, and we have only seen one patient whose antireceptor serum immunoprecipitated the receptor but did not inhibit binding. It is of interest that antibodies raised in rabbits by immunization with purified receptor immunoprecipitate the receptor but do not impair insulin binding.

Many of the naturally occurring autoantibodies are directed at determinants close to the insulin binding site and have probably been selected because of this. Their ability to compete with insulin and insulin analogs for binding has been discussed above. In addition, significant occupancy of the receptor by insulin impairs the ability of the autoantibodies to immunoprecipitate the receptor. In the immunoprecipitation assay the solubilized receptor labeled with only a tracer amount of ^{125}I-labeled insu-

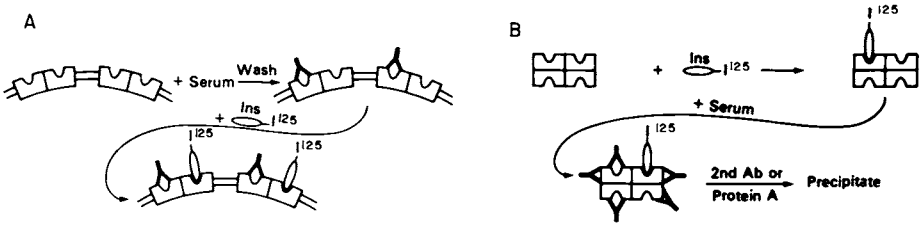


FIG. 8. Direct assays for (insulin) receptor antibodies. (A) Binding inhibition assay. Advantages: fast and easy. Disadvantages: separate assay needed to verify Ab nature; less sensitive; and detects only Ab that impairs insulin binding. (B) Immunoprecipitation assay. Advantages: more sensitive; detects Ab directed at determinants other than binding site; and proves Ab nature in a single assay. Disadvantages: more difficult; requires solubilized receptor; and will not detect non-Ab inhibitors.

lin can be precipitated quantitatively. Immunoprecipitation of the [^{125}I]insulin–receptor complex implies therefore that the receptor has more than one binding site or that some antibodies bind to determinants outside of the insulin binding region. The insulin receptor molecule has a complex subunit structure and will probably turn out to have functionally distinct components subserving cooperativity, affinity regulation, and biological signal transduction, in addition to an insulin-binding site. It seems reasonable to propose that the antibodies recognize a variety of determinants on the receptor, some of which are demonstrably involved in insulin binding.

Table V also includes two indirect methods for detecting receptor antibodies, employed for special purposes. Thus the presence of receptor antibodies may be inferred by the reversal of the binding defect by acid-washing the patients cells, provided that underlying receptor binding is normal and that chronic occupancy by antibody does not accelerate receptor loss. Neither seems to be the case in circulating monocytes obtained from patients with the type B syndrome, where acid washing restores binding toward normal. The presence of antibodies may also be inferred by uptake of [^{125}I]protein A onto the patients' cells or onto normal cells preincubated with the patients' sera. Although uptake of [^{125}I]protein A is a sensitive means of detecting cell surface immunoglobulins, it gives no information about their target or functional specificity.

D. Allergy; β-Adrenergic Receptor Antibodies

1. Preamble

Asthma, allergic rhinitis (hay fever), eczema, and urticaria (skin wheals), together with increased levels of serum IgE, represent the various manifestations of familial allergy or atopy (Coca and Cooke, 1923). Patients with atopy also have abnormal autonomic reactivity with decreased sensitivity to β-adrenergic agonists and increased sensitivity to α-adrenergic and cholinergic agonists (Szentivanyi, 1980).

Beginning two decades ago, studies performed *in vivo* have documented that asthmatic subjects have an impaired response to β-adrenergic agonists. The effect of epinephrine or isoproterenol on plasma levels of glucose, free fatty acids, and cAMP, or on pulse pressure, is less in asthmatics and, moreover, is inversely correlated with an enhanced effect of acethylcholine or histamine on bronchial constriction (Parker, 1973; Reed, 1974; Trembath and Shaw, 1976; Apold and Aksnes, 1977). The β-adrenergic theory of the pathogenesis of asthma (Szentivanyi, 1968) proposes an imbalance in autonomic control due to decreased β-adrenergic sensitivity of bronchial smooth muscle, mucous glands, and mucosal vessels and is supported by studies showing impaired cyclic AMP responses specifically to β-adrenergic agonists in leukocytes from asthmatics (Logsdon *et al.*, 1972; Parker and Smith, 1973; Conolly and Greenacre, 1976; Makino *et al.*, 1977).

Attempts to study β-adrenergic receptors by binding the natural agonist [^{3}H]epinephrine were fraught by technical problems, but with the advent of the high-affinity antagonists [^{3}H]dihydroalprenolol and [^{125}I]hydroxybenzylpindolol (IHYP), it became possible in the mid-1970s to identify directly sites that fulfilled the criteria for specific β-adrenergic receptors (for review, see Hoffman and Lefkowitz, 1980). In keeping with the previous bioresponse studies, it was shown that cells from asthmatics exhibited a reduction in the apparent number of β receptors (Galant *et al.*, 1978; Brooks *et al.*, 1979), although many attributed this solely to the down-regulation of receptors induced by exposure to pharmacological concentrations of β-adrenergic drugs administered for the treatment of asthma. Drug-induced down-regulation (and desensitization) clearly occurs, but there is also evidence from studies of untreated and asymptomatic subjects to indicate that a defect in β-receptor binding is a basic feature of asthma (see Szentivanyi, 1980). Similar conclusions also apply to atopic eczema, confirming earlier reports of impaired catecholamine responses in lymphocytes from such subjects (Busse and Lee, 1976). The decrease in β-receptor binding in atopic subjects has recently been linked to an increase in α-adrenergic receptor binding, and the autonomic hypothesis

has been extended to encompass a dual receptor hypothesis (Szentivanyi, 1980).

2. Atopy, Autonomic Function, and Autoimmunity

Coca and Cooke (1923) reported the clinical and familial picture of atopy and suspected a common immunopathological basis. Consistent with this notion was the fact that the immediate hypersensitivity of (atopic) individuals to extrinsic allergens could be passively transferred by blood or serum (Prausnitz and Kustner, 1921). However, although we know that certain antigens provoke IgE-mediated release of smooth muscle constrictors such as histamine from mast cells, the effector role of the immune system in atopy has not been defined. A possible connection between allergy and immunity is the discovery of autoantibodies to the β_2-adrenergic receptor in atopic subjects (Venter *et al.*, 1980) and the subsequent demonstration that these antibodies are related to autonomic dysfunction (Fraser *et al.*, 1981).

Autoantibodies of the IgG class that precipitated Triton-solubilized canine lung ^{125}IHYP-labeled β receptors were first found in the sera of a patient with allergic rhinitis (Venter *et al.*, 1980). The IgG from the patient also inhibited the binding of IHYP to lung and placenta but not cardiac membranes, indicating that at least some of the antibodies recognized determinants on β_2 receptors *in situ* at or near the ligand binding site. The inability to totally immunoprecipitate all available IHYP-labeled receptors, and the lower titer with immunoprecipitation versus binding inhibition, suggested that the antibodies in this one serum were predominantly directed at the IHYP binding site in the solubilized receptor.

In a follow-up study, IgG antibodies were detected in 9 of 60 coded sera: 3 of 19 control, 1 of 9 preallergic, 4 of 17 asthma, 0 of 8 allergic rhinitis, and 1 of 7 cystic fibrosis (Fraser *et al.*, 1980). Most importantly, the presence of antibodies was associated with autonomic abnormalities (Fraser *et al.*, 1980, 1981). Subjects with antibodies required higher doses of infused isoproterenol to elevate their pulse pressure or plasma cAMP. Moreover, this evidence for β-adrenergic resistance was accompanied by increased sensitivity to the effects of the α-adrenergic agonist phenylephrine on pupillary dilation and with increased sensitivity to the effects of the cholinergic agent carbachol on pupillary constriction (Table VI). It is interesting to note that the 3 control subjects with antibodies required a greater dose of isoproterenol to raise pulse pressure (12 ± 1.73 ng/kg/min) than the 16 control subjects without antibodies (7.5 ± 0.55).

Further studies on the molecular actions of β-receptor antibodies are

TABLE VI

β-Adrenergic Receptor Antibody Status and Autonomic Sensitivity

Infusion rate of isoproterenol in increase (ng/kg/min)	Antibody		
	Positive	Negative	
Pulse pressure >22 mm Hg	15.0 ± 1.90 (n = 9)	7.7 ± 0.40 (n = 20)	$p < .001$
Plasma cAMP >50%	12.4 ± 1.80 (n = 9)	8.1 ± 0.62 (n = 13)	$p < .02$
Phenylephrine to dilate pupils >0.5 mm (%)	2.06 ± 0.30 (n = 9)	2.55 ± 0.08 (n = 57)	$p < .05$
Carbachol to constrict pupils >0.5 mm (%)	0.61 ± 0.08 (n = 9)	0.78 ± 0.03 (n = 57)	$p < .05$

consistent with their role as β-receptor antagonists (Harrison *et al.*, 1982b). Thus, after preincubation of cultured human lung cells (VA-13 line) with antireceptor IgG, there was a marked decrease in sensitivity and responsiveness of cAMP to stimulation by isoproterenol. It remains to be shown whether this effect is only due to blockade at the receptor level, or whether other mechanisms such as desensitization at a postreceptor step or accelerated receptor loss are also involved.

3. Significance of β-Receptor Autoantibodies

These preliminary findings would be consistent with a role for β-receptor autoantibodies in mediating autonomic dysfunction in at least some atopic subjects. The simplest interpretation is that they impair β-receptor-mediated relaxation of airways smooth muscle, unmask the opposing influence of α-receptor agonists and acetylcholine, and lower the threshold for release of mediators such as histamine, prostaglandins, and leukotrienes. It would be important to understand the influence of β-adrenergic dysfunction on these biochemical pathways and on the classical effector mechanism of mast cell histamine release triggered by IgE, as well as to determine whether other receptor autoantibodies are present. Furthermore, an explanation is required for the anatomical localization of atopic manifestations to the lower airways in asthma, to the upper airways in allergic rhinitis, and to the skin in urticaria and atopic eczema. Most importantly, the prevalance of β-receptor antibodies needs to be established and their clinical correlates determined in larger populations. We have found (unpublished) that of chronic asthmatics only 15–20% have

circulating globulins that impair binding to particulate human lung membranes, and the prevalence of "β-blocking autoantibodies" in children with bronchial asthma has been reported as being of the same order (Blecher *et al.*, 1984). It must also be considered whether β-receptor antibodies might develop as a secondary manifestation of the atopic state, particularly as antiidiotypes in response to treatment with β-receptor ligands (Homcy *et al.*, 1982; see Section II,D).

To consider atopy within the spectrum of autoimmune disorders may be somewhat premature. On the other hand, if autoimmunity is a naturally regulated state, it may be seen as a continuum: from a normal state of regulated autoimmunity, through the emergence of (?) nonpathogenic autoantibodies with advancing age or after exposure to certain drugs, and to the emergence of pathogenic autoantibodies in classic autoimmune states. In the latter there is usually tissue destruction by cells or complement, but there may be less pathological situations exemplified by atopy where the manifestations are due solely to the specific functional effects of autoantibodies. The implications for the treatment of atopic diseases are the same as discussed for the other receptor antibody diseases. Corticosteroids and immunosuppressive drugs that are used for intractable asthma or atopic eczema might act by impairing autoantibody production. Plasmapheresis, has been reported to be effective in severe asthma (Gartmann *et al.*, 1978).

IV. CONCLUDING REMARKS: FUTURE PROSPECTS

The studies of autoantibodies to TSH, ACh, and insulin receptors illustrate the extent to which processes subserving autoimmune disease can be dissected and demonstrate the utility of autoantibodies as probes of receptor structure and function. The recognition of these receptor antibody diseases and the premise that autoimmunity is initiated at the cell surface suggests that autoantibodies against cell surface receptors will be found at some stage in all autoimmune diseases. The question is how to identify potential receptor targets. Where the primary target tissue is not even known, this is clearly difficult. However, given that the target tissue has been specified (e.g., by immunofluorescence studies) and that some knowledge of its function is available, it might be possible to make an educated guess about the receptor involved.

Loveridge and co-workers (1980) found that sera from patients with pernicious anemia contained antibodies to the surface of gastric parietal cells and that these sera inhibited acid production by parietal cells. It is

not too difficult then to deduce that the loss of parietal cells and their intrinsic factor (with subsequent vitamin B_{12} malabsorption and anemia) could be initiated by autoantibodies to surface gastrin or histamine receptors known to regulate acid secretion. Lernmark and coworkers (1978) showed that sera from subjects with insulin-dependent diabetes contained antibodies to the surface of isolated pancreatic islet cells and that these sera specifically impaired insulin biosynthesis (Kanatsuna *et al.*, 1981). Are islet cell-surface antibodies directed against specific receptors involved in regulating insulin synthesis/secretion? Candidates might include receptors for glucagon, somatostatin, or even insulin itself, glucose transport protein, or a surface molecule yet to be defined. Irrespective of whether a receptor label can be assigned, it is important to define the molecular structure of autoantigens and to generate libraries of such structures. Elsewhere (Harrison, 1984) I have referred to autoantibodies as "nature's gift to the receptor biochemist" because of their potential to pinpoint the molecular structures and events underlying autoimmunity. There is no reason why receptor antibodies should not be sought in systemic autoimmune diseases in which the notion of receptor autoimmunity has received little credence. One could propose, for example, that sera from subjects with early rheumatoid arthritis or reactive arthritis might contain antibodies against a surface component of synovial cells.

Despite our ignorance of the pathogenesis of autoimmune disease, the receptor antibody diseases with their restricted repertoire of autoreactive immune cells would appear to be the most suitable candidates for evaluating specific forms of immunotherapy. It is now possible to obtain amino acid sequence information from picomolar amounts of purified membrane receptors and use this information to identify genomic or cloned receptor DNA. Alternatively, polyclonal autoantibodies can be used to screen DNA libraries in expression vectors (for reviews, see Fernley *et al.*, 1984; Kaufman and Tobin, 1984). These approaches should permit the isolation of relatively large amounts of cloned antigen and are currently being used to obtain parasite surface antigens for conventional vaccines (Kemp *et al.*, 1983). However, specific therapy for autoimmune diseases presents a challenge seemingly much greater than that which faced Jenner last century in first vaccinating against the smallpox disease. Conventional vaccination deliberately sets out to stimulate immunity to a defined (exo) antigen. Autoimmunologists would like deliberately to suppress immunity against self-antigens. How can one immunize against clones of receptor-reactive lymphocytes? Current strategies are based mainly on the notion of an antiidiotypic network (Jerne, 1974), in which autoreactivity is seen as a normal, regulated function of the immune system and, by implication,

antireceptor antibody disease as a failure of regulation by antiidiotypic antibodies.

There is both experimental and natural evidence for the presence of antiidiotypic, antireceptor antibodies capable of modulating cell function, as outlined under Section II,D. In situations in which the idiotype is on the antireceptor antibody, the antiidiotype is theoretically capable of abrogating the antireceptor effect by blocking or complexing antireceptor antibodies in the circulation, by down-regulating or suppressing the reactive B-cell clones, or by inducing the formation of antigen-specific suppressor T cells. Naturally occurring antiidiotypic antibodies have been found in the sera of patients with myasthenia gravis (Dwyer *et al.*, 1983), and their level is inversely related to the level of acetylcholine receptor antibodies, which suggests a suppressive effect of antiidiotype on antireceptor antibodies or B-cell progenitors. It has also been reported that patients with lupus erythematosis who are in remission have antiidiotypes against anti-DNA antibodies (Abdou *et al.*, 1981).

The question is whether antiidiotypes can be engineered for specific immunotherapy. This might depend on several factors such as the range of specificities involved, the degree of cross-reactivity, and the degree to which the disease is solely dependent on antiidiotype network control. Experimental myasthenia gravis can, for example, result from the action of a monoclonal antibody (Lennon and Lambert, 1980), but the human disease is associated with a polyclonal reaction to a receptor known to have hundreds of immunogenic determinants (Tzartos, 1984). It will be interesting to discover whether antiidiotypes must encompass the whole range of determinants. One approach in humans might be to obtain a library of monoclonal antireceptor antibodies by hybridization of peripheral blood lymphocytes with myeloma cells. These antireceptor antibodies would then be used to obtain libraries of monoclonal antiidiotypes for injection into patients with receptor antibody disease. Alternatively, one could start with purified or cloned receptors.

It is possible that, compared with B cells, the number of idiotypes recognized by T-cell receptors is limited. When a single clone of T lymphocytes with specificity for myelin basic protein was irradiated and used to immunize rats, the animals resisted the induction of experimental autoallergic encephalomyelitis (Ben-Nun *et al.*, 1981). This implies that the suppressor T-cell response to the idiotype of the immunizing clone effectively combated the entire set of effector T cells stimulated by immunization with myelin basic protein. Cohen (1984) has pointed out that T cells, as the likely effectors, could be attenuated and used as immunogens to induce resistance in a number of autoimmune diseases.

There are many facets to the complex control of the immune system that bear on the question of specific therapy for autoimmune disease but that are beyond the scope of this chapter (and its author). The reader is referred particularly to a recent Springer treatise on the immunopathology of idiotypic interactions edited by Lambert (1983).

The rapid development of knowledge of the receptor antibody diseases and the insights into cellular biochemistry and immunology afforded by the study and application of receptor antibodies represent a powerful argument for the marriage of clinical medical and basic sciences, which I hope has been exemplified in this chapter.

ACKNOWLEDGMENTS

Expert secretarial assistance was provided by Mrs. Linda Stafford.

Studies reported from the author's laboratory have been supported by the National Health and Medical Research Council of Australia, the Kroc Foundation (United States), the Victor Hurley Medical Research Fund of The Royal Melbourne Hospital, and the Asthma Foundation of Victoria.

REFERENCES

Abdou, N. I., Wall, H., Lindsley, H. B., Halsey, J. F., and Suzuki, T. (1981). *J. Clin. Invest.* **67,** 1297–1304.

Adams, D. D., and Kennedy, T. H. (1967). *J. Clin. Endocrinol. Metab.* **27,** 173–177.

Adams, D. D., and Kennedy, T. H. (1971). *J. Clin. Endocrinol. Metab.* **33,** 47–51.

Adams, D. D., and Purves, H. D. (1956). *Proc. Univ. Otago Med. Sch.* **34,** 11–12.

Adams, D. D., Kennedy, T. H., and Stewart, R. D. H. (1974). *Br. Med. J.* **2,** 199–201.

Amir, S. M., Goldfine, I. D., and Ingbar, S. H. (1976). *J. Biol. Chem.* **251,** 4693–4699.

Apold, J., and Aksnes, L. (1977). *J. Allergy Clin. Immunol.* **59,** 343–347.

Appel, S. H., Almon, R. R., and Levy, N. (1975). *N. Engl. J. Med.* **293,** 760–761.

Appel, S. H., Anwyl, R., McAdams, M. W., and Elias, (1977). *Proc. Natl. Acad. Sci. U.S.A.* **74,** 2130–2134.

Bar, R. S., Harrison, L. C., Muggeo, M., Gorden, P., Kahn, C. R., and Roth, J. (1979). *Adv. Int. Med.* **24,** 23–52.

Bar, R. S., Muggeo, M., Kahn, C. R., Gorden, P., and Roth, J. (1980). *Diabetologia* **18,** 209–216.

Beall, G. N., Chopra, I. J., Solomon, D. H., and Kruger, S. R. (1978). *J. Clin. Endocrinol. Metab.* **47,** 967–972.

Bech, K., and Nistrup-Madsen, S. (1979). *Clin. Endocrinol.* **11,** 47–58.

Bech, K., and Nistrup-Madsen, S. (1981). *Clin. Endocrinol.* **13,** 417–424.

Ben-Nun, A., Werkerle, H., and Cohen, I. R. (1981). *Nature (London)* **292,** 60–61.

Bielschowsky, M., and Bielschowsky, F. (1956). *J. Exp. Biol.* **34,** 181–198.

Binz, H., and Wigzell, H. (1975). *J. Exp. Med.* **142,** 1231–1240.

Blackard, W. G., Anderson, J. H., and Mullinax, F. (1977). *Ann. Int. Med.* **86,** 584–585.

Blecher, M., Lewis, Hicks, J. M., and Josephs, S. (1984). *J. Allergy Clin. Immunol.* **74,** 246–251.

Bottomly, K., and Mosier, D. E. (1981). *J. Exp. Med.* **154,** 411–421.

Brooks, S. M., McGowan, K., Bernstein, I. L., Altenau, P., and Peagler, J. (1979). *J. Allergy Clin. Immunol.* **63,** 401–406.

Brown, R. S., Kertiles, L. P., and Reichlin, S. (1983). *J. Clin. Endocrinol. Metab.* **56,** 156–163.

Bryson, J. M., Joasoo, A., and Turtle, J. R. (1976). *Acta Endocrinol. (Copenhagen)* **83,** 528–538.

Bucknall, R. C., Dixon, A. St. J., Glick, E. N., Woodland, J., and Zutshi, D. W. (1975). *Br. Med. J.* **1,** 600–602.

Busse, W. W., and Lee, T. P. (1976). *J. Allergy Clin. Immunol.* **58,** 586–596.

Campbell, I. L., Wong, G. H. W., Schrader, J. W., and Harrison, L. C. (1984). *Diabetes* (in press).

Caro, J. F., and Amatruda, J. M. (1980). *Science* **210,** 1029–1031.

Carpentier, J.-L., Van Obberghen, E., Gorden, P., and Orci, L. (1979). *Diabetes* **28,** 345. (Abstr. 2.)

Clark, S., and Harrison, L. C. (1982). *J. Biol. Chem.* **257,** 12239–12244.

Clark, S., and Harrison, L. C. (1983). *J. Biol. Chem.* **258,** 11434–11437.

Clark, S., and Harrison, L. C. (1985). *Biochem. J.* **229,** 513–519.

Coca, A. F., and Cooke, R. A. (1923). *J. Immunol.* **8,** 163–168.

Cohen, I. R. (1984). *Clinics Immunol. Allergy* **4,** 593–605.

Conolly, M. E., and Greenacre, J. K. (1976). *J. Clin. Invest.* **58,** 1307–1316.

Conti-Tronconi, B., Morgutti, M., Albizzati, M. G., and Clementi, F. (1978). *Z. Neurol.* **217,** 281–286.

Cosenza, H., and Kohler, H. (1972). *Science* **176,** 1027–1029.

Czarnocka, B., Nauman, J., Adler, G. and Kielczynski, W. (1979). *Acta Endocrinol. (Copenhagen)* **92,** 512–521.

Davies, T. F. (1981). *J. Clin. Endocrinol. Metab.* **52,** 426–430.

Davies, T. F., and De Bernardo, E. (1983). *In* "Autoimmune Endocrine Disease" (T. F. Davies, ed.), pp. 127–137. Wiley, New York.

Dawkins, R. L., Zilko, P. L., Carrano, J., Garlepp, M. I., and McDonald, B. L. (1981). *J. Rheumatol.* **8,** 56–61.

Dean, B., Peluso, I., and Harrison, L. C. (1984). *Diabetes* **33,** 450–454.

Donath, J., and Landsteiner, K. (1904). *Munch. Med. Wschr.* **51,** 1590–1593.

Drachman, D. B., Angus, C. W., Adams, R. N., Michelson, J. D., and Hoffman, G. J. (1978). *N. Engl. J. Med.* **298,** 1116–1122.

Drexhage, H. A., Bottazzo, G. F., Doniach, D., Bitensky, L., and Chayen, J. (1980). *Lancet* **2,** 287–292.

Drexhage, H. A., Bottazzo, G. F., Bitensky, L., Chayen, J., and Doniach, D. (1981). *Nature (London)* **289,** 594–596.

Dwyer, D. S., Bradley, R. J., Urquhart, C. K., and Kearney, J. F. (1983). *Nature (London)* **301,** 611–614.

Edmonds, J. (1984). *Aust. N. Z. J. Med.* **14,** 81–87.

Edmonds, J., Macaulay, D., Tyndall, A., Liew, M., Alexander, K., Geczy, A., and Bashir, A. (1981). *Arthritis Rheum.* **24,** 1–7.

Ehrlich, P. (1906). "Studies on Immunity," 1st ed. Wiley, London.

Elmqvist, D., Hofmann, W. W., Kugelberg, J., and Quastel, D. M. J. (1964). *J. Physiol. (London)* **174,** 417–434.

Endo, K., Kasagi, K., Konishi, J., Ikekubo, K., Okuno, T., Takeda, Y., Mori, T., and Torizuka, K. (1978). *J. Clin. Endocrinol. Metab.* **46,** 734–739.

Engel, A. G. (1979). *In* "Handbook of Clinical Neurology" (P. J. Winken and G. W. Bruyn, eds.), Vol. 41, Part II, pp. 95–145. Elsevier, Amsterdam.
Engel, A. G., and Santa, T. (1971). *Ann. N. Y. Acad. Sci.* **183,** 46–63.
Engel, A. G., Lindstrom, J., Lambert, E. H., and Lennon, V. A. (1977). *Neurology* **27,** 307–315.
Engel, A. G., Sahashi, K., and Fumagalli, G. (1981). *Ann. N. Y. Acad. Sci.* **377,** 158–174.
Etienne-Decerf, J., and Winand, R. J. (1981). *Clin. Endocrinol.* **14,** 83–91.
Fambrough, D. M., Drachman, D. B., and Satyamurti, S. (1973). *Science* **182,** 293–295.
Farid, N. R., Barnard, J. M., and Marshall, W. H. (1976). *Tissue Antigens* **8,** 181–189.
Farid, N. R., Stone, E., and Johnson, G. (1980). *Clin. Endocrinol.* **13,** 535–544.
Farid, N. R., Pepper, B., Urbina-Briones, R., and Islam, N. R. (1982). *J. Cell Biochem.* **19,** 305–313.
Fernley, R. T., Gorman, J. J., Niall, H. D., and Coghlan, J. P. (1984). *In* "Molecular and Chemical Characterization of Membrane Receptors" (J. C. Venter and L. C. Harrison, eds.), pp. 261–284. Alan R. Liss, New York.
Field, J. B., Remer, A., Bloom, G., and Kriss, J. P. (1968). *J. Clin. Invest.* **47,** 1553–1560.
Fields, B. N., and Greene, M. I. (1982). *Nature (London)* **300,** 19–23.
Flier, J. S., Kahn, C. R., Roth, J., and Bar, R. S. (1975). *Science* **190,** 63–65.
Flier, J. S., Kahn, C. R., Jarrett, D. B., and Roth, J. (1976). *J. Clin. Invest.* **58,** 1442–1449.
Flier, J. S., Kahn, C. R., Jarrett, D. B., and Roth, J. (1977). *J. Clin. Invest.* **60,** 784–794.
Flier, J. S., Bar, R. S., Muggeo, M., Kahn, C. R., Roth, J., and Gorden, P. (1978). *J. Clin. Endocrinol. Metab.* **47,** 985–995.
Fraser, C. M., Harrison, L. C., Kaliner, M. C., and Venter, J. C. (1980). *Clin. Res.* **28,** 236A.
Fraser, C. M., Venter, J. C., and Kaliner, M. (1981). *N. Engl. J. Med.* **305,** 1165–1170.
Fuchs, S., Schmidt-Hopfeld, I., Tridente, G., and Tarrab-Hazdai, R. (1980). *Nature (London)* **287,** 162–164.
Fujita-Yamaguchi, Y., Choi, S., Sakamoto, Y., and Itakura, K. (1983). *J. Biol. Chem.* **258,** 5045–5049.
Galant, S. P., Duriseti, L., Underwood, S., and Insel, P. A. (1978). *N. Engl. J. Med.* **299,** 933–936.
Gartmann, J., Grob, P., and Frey, M. (1978). *Lancet* **2,** 40. (Letter.)
Grumet, F. C., Payne, R. O., Konishi, J., and Kriss, J. P. (1974). *J. Clin. Endocrinol. Metab.* **39,** 1115–1119.
Grunfeld, C., Van Obberghen, E., Karlsson, F. A., and Kahn, C. R. (1980). *J. Clin. Invest.* **66,** 1124–1134.
Hales, I. B., and Luttrell, B. M. (1980). *Proc. 8th Int. Thyroid Congr.*, p. 755.
Hall, R. (1980). *Proc. 8th Int. Thyroid Congr.*, pp. 9–12.
Hallengren, B., Forsgren, A., and Melander, A. (1980). *J. Clin. Endocrinol. Metab.* **51,** 298–301.
Hamilton, S. L., McLaughlin, M., and Karlin, A. (1977). *Biochem. Biophys. Res. Commun.* **79,** 692–699.
Hanafusa, T., Pujol-Borrell, R., Chiorato, R., Russell, R. C. G., Domiach, D., and Bottazzo, G. F. (1983). *Lancet* **2,** 1111–1115.
Hardisty, C. A., and Munro, D. S. (1980). *Proc. 8th Int. Thyroid Congr.*, p. 613.
Harrison, L. C. (1984). *In* "Receptor Purification Procedures" (J. C. Venter and L. C. Harrison, ed.), Vol. 2, pp. 125–137. Alan R. Liss, New York.
Harrison, L. C., and Flier, J. S. (1980). *In* "Secondary Diabetes: The Spectrum of the Diabetic Syndrome" (S. Pokosky and M. Viswanahan, eds.), pp. 269–286. Raven, New York.

Harrison, L. C., and Itin, A. (1979). *Nature (London)* **279,** 334–336.
Harrison, L. C., and Itin, A. (1980). *J. Biol. Chem.* **255,** 12066–12072.
Harrison, L. C., and Itin, A. (1982). *In* "Current Views on Insulin Receptors" (D. Andreani, R. De Pirro, R. Lauro, J. Olefsky, and J. Roth, eds.), pp. 29–36. Academic Press, London.
Harrison, L. C., Flier, J. S., Itin, A., Kahn, C. R., and Roth, J. (1979a). *Science* **203,** 544–547.
Harrison, L. C., Flier, J. S., Roth, J., Karlsson, F. A., and Kahn, C. R. (1979b). *J. Clin. Endocrinol. Metab.* **48,** 59–65.
Harrison, L. C., Muggeo, M., Bar, R. S., Flier, J. S., Waldmann, T., and Roth, J. (1979c). *Clin. Res.* **27,** 252A.
Harrison, L. C., Itin, A., Kasuga, M., and Van Obberghen, E. (1982a). *Diabetologia* **22,** 233–238.
Harrison, L. C., Callaghan, J., Venter, J. C., Fraser, C. M., and Kaliner, M. (1982b). *In* "Receptors, Antibodies and Disease" pp. 248–262. Pitman, London.
Haspel, M. V., Onodera, T., Prabhakar, B. S., Korita, M., Suzuki, H., and Notkins, A. L. (1983). *Science* **220,** 304–306.
Heindenreich, K. A., Berhanu, P., Brandenburg, D., and Olefsky, J. M. (1983). *Diabetes* **32,** 1001–1009.
Heinemann, S., Bevan, S., Kullberg, R., Lindstrom, J., and Rice, J. (1977). *Proc. Natl. Acad. Sci. U.S.A.* **74,** 3090–3094.
Heyma, P., and Harrison, L. C. (1984). *J. Clin. Invest* **74,** 1090–1097.
Heyma, P., Robbins-Browne, R., and Harrison, L. C. (1984). (Submitted.)
Hoebeke, J., Vanquelin, G., and Strosberg, A. D. (1977). *Biochem. Pharmacol.* **27,** 1527–1533.
Hoffman, B. B., and Lefkowitz, R. J. (1980). *Annu. Rev. Pharmacol. Toxicol.* **20,** 581–608.
Holmes, S. D., Dirmikis, S. M., Martin, T. J., and Munro, D. S. (1979). *J. Endocrinol.* **80,** 215–221.
Homcy, C. J., Rockson, S. G., and Haber, E. (1982). *J. Clin. Invest.* **69,** 1147–1154.
Howard, B., Hidaka, H., Ishibashi, F., Fields, R. M., and Bennett, P. H. (1981). *Diabetes* **30,** 562–567.
Iida, Y., Konishi, J., Kasagi, K., Ikekubo, K., Kuma, K., and Tomzuka, K. (1981). *Acta Endocrinol. (Copenhagen)* **98,** 50–56.
Irvine, W. J., Gray, R. S., Morris, P. J., and Ting, A. (1977). *Lancet* **2,** 898–900.
Jacobs, S., Shechter, Y., Bissell, K., and Cuatrecasas, P. (1977). *Biochem. Biophys. Res. Commun.* **77,** 981–988.
Jacobs, S., Hazum, E., Shechter, Y., and Cuatrecasas, P. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76,** 4918–4921.
Jarrett, D. B., Roth, J., Kahn, C. R., and Flier, J. S. (1976). *Proc. Natl. Acad. Sci. U.S.A.* **73,** 4115–4119.
Jerne, N. K. (1974). *Ann. Immunol. (Paris)* **125C,** 373–389.
Jerne, N. K., Roland, J., and Cazenave, P. A. (1982). *EMBO J.* **1,** 243–247.
Jonas, H., Baxter, R. C., and Harrison, L. C. (1982). *Biochem. Biophys. Res. Commun.* **109,** 463–470.
Kahn, C. R., and Harrison, L. C. (1981). *In* "Carbohydrate Metabolism and Its Disorders" (P. J. Randle, D. F. Steiner, and W. J. Whelan, eds.), Vol. 3, pp. 278–330. Academic Press, London.
Kahn, C. R., Flier, J. S., Bar, R. S., Archer, J. A., Gorden, P., Martin, M. M., and Roth, J. (1976). *N. Engl. J. Med.* **294,** 739–745.
Kahn, C. R., Baird, K. L., Flier, J. S., and Jarrett, D. B. (1977). *J. Clin. Invest.* **60,** 1094–1106.

Kahn, C. R., Baird, K. L., Jarrett, D. B., and Flier, J. (1978). *Proc. Natl. Acad. Sci. U.S.A.* **75,** 4209–4213.

Kanatsuna, T., Lernmark, A., Rubenstein, A. H., and Steiner, D. F. (1981). *Diabetes* **30,** 231–234.

Kao, I., and Drachman, D. B. (1977). *Science* **196,** 527–529.

Kaplan, M., and Meyeserian, M. (1962). *Lancet* **2,** 706–710.

Karlin, A., Weill, C. L., McNamee, M. G., and Valderrama, R. (1975). *Cold Spring Harbor Symp. Quant. Biol.* **60,** 203–210.

Karlsson, F. A., Van Obberghen, E., Grunfeld, C., and Kahn, C. R. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76,** 809–813.

Kasuga, M., Van Obberghen, E., Yamada, K., and Harrison, L. C. (1981). *Diabetes* **30,** 354–357.

Kasuga, M., Karlsson, F. A., and Kahn, C. R. (1982). *Science* **215,** 185–186.

Kasuga, M., Carpentier, J.-L., Van Obberghen, E., Orci, L., and Gorden, P. (1983). *Biochem. Biophys. Res. Commun.* **114,** 230–233.

Kaufman, D. L., and Tobin, A. J. (1984). *In* "Molecular and Chemical Characterization of Membrane Receptors" (J. C. Venter and L. C. Harrison, eds.), pp. 241–259. Alan R. Liss, New York.

Kawanishi, K., Kawamura, K., Nishina, Y., Goto, A., Okada, S., Ishida, T., Ofuji, T., Kahn, C. R., and Flier, J. S. (1977). *J. Clin. Endocrinol. Metab.* **44,** 15–21.

Kemp, D. J., Coppel, R. L., Cowman, A. F., Saint, R. B., Brown, G. V., and Anders, R. F. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80,** 3787–3791.

Kendall–Taylor, P., Dirmikis, S. M., and Munro, D. C. (1972). *Proc. R. Soc. Med.* **68,** 252–254.

Kibata, M., Hiramatsu, K., Shimizu, Y., Fuchimoto, T., Sasaki, M., Shimono, M., Miyake. K., Flier, J. S., and Kahn, C. R. (1975). *Proc. Symp. Chem. Physiol. Pathol.* **15,** 58.

King, G. L., Kahn, C. R., Rechler, M. M., and Nissley, S. P. (1980). *J. Clin. Invest.* **66,** 130–140.

Kleinmann, R. E., Braverman, L. E., Vagenakis, A. G., Butcher, R. W., and Clark, R. B. (1980). *J. Lab. Clin. Med.* **95,** 581–583.

Kluskens, L., and Kohler, H. (1974). *Proc. Natl. Acad. Sci. U.S.A.* **71,** 5083–5087.

Kriss, J. P., Pleshakou, V., and Chien, J. R. (1964). *J. Clin. Endocrinol. Metab.* **24,** 1005–1028.

Koizumi, Y., Zakarija, M., and McKenzie, J. M. (1982). *Endocrinology (Baltimore)* **110,** 1381–1391.

Kuzuya, N., Chiu, S. C., Ikeda, H., Uchimura, H., Ito, K., and Nagataki, S. (1979). *J. Clin. Endocrinol. Metab.* **48,** 706–711.

Lambert, E. H. (1983). *Springer Semin. Immunopathol.* **6,** 1–115.

Lang, U., Kahn, C. R., and Harrison, L. C. (1980). *Biochemistry* **19,** 64–70.

Langley, J. N. (1878). *J. Physiol. (London)* **1,** 339–369.

Langley, J. N. (1905). *J. Physiol. (London)* **33,** 375–380.

Laurberg, P., and Weeke, J. (1975). *Scand. J. Clin. Lab. Invest.* **35,** 723–727.

Lennon, V. A., and Carnegie, P. R. (1971). *Lancet* **1,** 630–633.

Lennon, V. A., and Lambert, E. H. (1980). *Nature (London)* **285,** 238–240.

Lennon, V. A., Lindstrom, J., and Seybold, M. E. (1976). *Ann. N. Y. Acad. Sci.* **274,** 283–299.

Lennon, V. A., Seybold, M. E., Lindstrom, J. M., Cochrane, C., and Yulevitch, R. (1978). *J. Exp. Med.* **147,** 973–983.

Lernmark, A., Freedman, Z. R., Hofmann, C., Rubenstein, A. H., Steiner, D. F., Jackson, R. L., Winter, R. J., and Traisman, H. S. (1978). *N. Engl. J. Med.* **299,** 375–380.

Lidman, K., Eriksson, V., Norberg, R., and Fabraeus, A. (1976). *Clin. Exp. Immunol.* **23,** 429–433.
Lindstrom, J. (1979). *Adv. Immunol.* **27,** 1–50.
Lindstrom, J. M., and Lambert, E. H. (1978). *Neurology* **28,** 130–138.
Lindstrom, J. M., Engel, A. G., Seybold, M. E., Lennon, V. A., and Lambert, E. H. (1976a). *J. Exp. Med.* **144,** 739–753.
Lindstrom, J. M., Seybold, M. E., Lennon, V. A., Whittingham, S., and Duane, D. (1976b). *Neurology* **26,** 1054–1059.
Lindstrom, J., Walter, B., and Einarson, B. (1979). *Biochemistry* **18,** 4470–4480.
Logsdon, P. J., Middleton, E., and Coffey, R. G. (1972). *J. Allergy Clin. Immunol.* **50,** 45–56.
Loveridge, N., Bitensky, L., Chayen, J., Hausamen, J. U., Fischer, J. M., Taylor, K. B., Gardener, J. D., Bottazzo, G. S., and Domiach, D. (1980). *Clin. Exp. Immunol.* **41,** 264–270.
McDevitt, H. O., and Bodmer, W. F. (1974). *Lancet* **1,** 1269–1271.
McGregor, A. M., Petersen, M. M., McLachlan, S. M., Rooke, P., Rees-Smith, B., and Hall, R. (1980a). *N. Engl. J. Med.* **303,** 302–307.
McGregor, A. M., Smith, B. R., Hall, R., Petersen, M. M., Miller, M., and Dewar, P. J. (1980b). *Lancet* **1,** 1101–1103.
McKenzie, J. M. (1958). *Endocrinology (Baltimore)* **63,** 372–382.
McKenzie, J. M., and Zakarija, M. (1976). *J. Clin. Endocrinol. Metab.* **42,** 778–781.
McKenzie, J. M., Zakarija, M., and Sato, A. (1978). *Clin. Endocrinol. Metab.* **7,** 31–45.
Major, P. W., and Munro, D. S. (1962). *Clin. Sci.* **23,** 463–475.
Makino, S., Ikemori, K., Kashima, T., and Fukuda, T. (1977). *J. Allergy Clin. Immunol.* **59,** 348–352.
Manley, S. W., Bourke, J. R., and Hawker, R. W. (1974). *J. Endocrinol.* **61,** 437–445.
Manley, S. W. W., Knight, A., and Adams, D. D. (1982). *Springer Semin. Immunopathol.* **5,** 413–431.
Marasco, W. A., and Becker, E. L. (1982). *J. Immunol.* **128,** 963–968.
Marchant, B., Lees, J. F. H., and Alexander W. D. (1978). *Pharmacol. Ther., Part B* **3,** 305–348.
Massague, J., Pilch, P., and Czech, M. P. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 7137–7141.
Masters, C. L., Dawkins, R. L., Zilko, P. J., Simpson, J. A., Leedman, R. J., and Lindstrom, J. M. (1977). *Am. J. Med.* **63,** 689–694.
Matell, G., Bergstrom, K., Frankssen, C., Hammarstrom, L., Lefvent, A. K., Moller, E., Von Reis, G., and Smith, E. (1976). *Ann. N. Y. Acad. Sci.* **274,** 659–676.
Mehdi, S. Q., and Nussey, S. S. (1975). *Biochem. J.* **145,** 105–111.
Melez, K., Harrison, L. C., Gilliam, J. N., and Steinberg, A. D. (1980). *Diabetes* **29,** 835–840.
Metzger, H. (1978). *Immunol. Rev.* **41,** 186–199.
Mittag, T. W., Kornfeld, P., Tormay, A., and Woo, C. (1976). *N. Engl. J. Med.* **294,** 691–694.
Monnier, V. M., and Fulpius, B. W. (1977). *Clin. Exp. Immunol.* **29,** 16–22.
Moore, W. V., and Wolff, J. (1974). *J. Biol. Chem.* **249,** 6255–6263.
Muggeo, M., Flier, J. S., Abrams, R. S., Harrison, L. C., Deisserroth, A. B., and Kahn, C. R. (1979a). *N. Engl. J. Med.* **300,** 477–480.
Muggeo, M., Kahn, C. R., Bar, R. S., Rechler, M., Flier, J. S., and Roth, J. (1979b). *J. Clin. Endocrinol. Metab.* **49,** 110–119.
Mukhtar, E. D., Smith, B. R., Pyle, G. A., Hall, R., and Vice, P. (1975). *Lancet* **1,** 713–715.

Nastuk, W. L., Plescia, O. J., and Osserman, K. E. (1960). *Proc. Soc. Exp. Biol. Med.* **105,** 177–184.

Nerup, J., and Lernmark, A. (1981). *Am. J. Med.* **70,** 135–141.

O'Donnell, J., Trokoudes, K., Silverberg, J., Row, V., and Volpe, R. (1978). *J. Clin. Endocrinol. Metab.* **46,** 770–777.

Onaya, T., Kotani, M., Tamada, T., and Ochi, Y. (1973). *J. Clin. Endocrinol. Metab.* **36,** 859–866.

Onodera, T., Toniolo, A., Ray, V. R., Jensen, A. B., Knazek, R. A., and Notkins, A. L. (1981). *J. Exp. Med.* **153,** 1457–1473.

Orgiazzi, J., Williams, D. E., Chopra, I. J., and Solomon, D. H. (1976). *J. Clin. Endocrinol. Metab.* **42,** 341–354.

Ozawa, Y., Maciel, R. M. B., Chopra, I. J., Solomon, D. H., and Beall, G. N. (1979). *J. Clin. Endocrinol. Metab.* **48,** 381–387.

Parker, C. W. (1973). *In* "Asthma: Physiology, Immunopharmacology, and Treatment" (K. F. Austen and L. M. Lichtenstein, eds.), pp. 185–210. Academic Press, New York.

Parker, C. W., and Smith, J. W. (1973). *J. Clin. Invest* **52,** 48–59.

Pastan, I., Roth, J., and Macchia, V. (1966). *Proc. Natl. Acad. Sci. U.S.A.* **56,** 1802–1809.

Patrick, J., and Lindstrom, J. (1973). *Science* **180,** 871–872.

Pekonen, F., and Weintraub, B. D. (1979). *Endocrinology (Baltimore)* **105,** 352–359.

Pilch, P. F., and Czech, M. P. (1980). *J. Biol. Chem.* **255,** 1722–1731.

Plotz, P. H. (1983). *Lancet* **2,** 824–826.

Prausnitz, C., and Kustner, H. (1921). *Bakt. Orig.* **86,** 160–165.

Pujol-Borrell, R., Hanafusa, T., Chiovato, L., and Bottazzo, G. F. (1983). *Nature (London)* **304,** 71–76.

Raftery, M. A., Hunkapiller, M. W., Strader, C. D., and Hood, L. E. (1980). *Science* **208,** 1454–1456.

Reed, C. E. (1974). *J. Allergy Clin. Immunol.* **53,** 34–41.

Rickards, C., Buckland, P., Smith, B. R., and Hall, R. (1981). *FEBS Lett.* **127,** 17–21.

Ronnett, G., Krutson, V. P., Kohanski, R. A., Simpson, T. L., and Lane, M. D. (1984). *J. Biol. Chem.* **259,** 4566–4575.

Schernthaner, G., Schleusner, H., Kotulla, P., Finke, R., Wenzel, B., and Mayr, W. R. (1980). *Lancet* **2,** 373–374.

Schlessinger, J., Van Obberghen, E., and Kahn, C. R. (1980). *Nature (London)* **286,** 729–731.

Schleusener, H., Kotulla, R., Finke, R., Sorje, H., Meinhold, H., Adlkoffer, F., and Wenzel, K. W. (1978). *J. Clin. Endocrinol. Metab.* **47,** 379–384.

Schreiber, A. B., Couraud, P. O., Andre, C., Vray, B., and Strosberg, A. D. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 7385–7389.

Sege, K., and Peterson, P. A. (1978). *Proc. Natl. Acad. Sci. U.S.A.* **75,** 2443–2447.

Schechter, Y., Maron, R., Elias, D., and Cohen, I. R. (1982). *Science* **216,** 542–545.

Shewring, G., and Rees-Smith, B. (1982). *Clin. Endocrinol.* **17,** 409–417.

Shishiba, Y., Shimuzo, T., Yoshimura, S., and Shizume, K. (1973). *J. Clin. Endocrinol. Metab.* **36,** 517–521.

Silverberg, J., O'Donnell, J., Sugenoya, A., Row, V. V., and Volpe, R. (1978). *J. Clin. Endocrinol. Metab.* **46,** 420–426.

Simpson, J. A. (1960). *Scott. Med. J.* **5,** 419–436.

Smith, B. R., and Hall, R. (1974). *Lancet* **2,** 427–431.

Strakosch, C. R., Toyner, D., and Wall, D. R. (1978). *J. Clin. Endocrinol. Metab.* **47,** 361–365.

Strosberg, D. (1983). *Springer Semin. Immunopathol.* **6,** 67–78.

Szentivanyi, A. (1968). *J. Allergy* **42,** 203–231.
Szentivanyi, A. (1980). *J. Allergy Clin. Immunol.* **65,** 5–11.
Tate, R. L., Schwartz, H. I., Holmes, J. M., and Kohn, L. D. (1975). *J. Biol. Chem.* **250,** 6509–6515.
Taylor, S., Grunberger, G., Marcus-Samuels, B., Underhill, L. H., Dons, R. F., Ryan, J., Roddam, R. F., Rupe, C. E., and Gorden, P. (1982). *N. Engl. J. Med.* **307,** 1422–1426.
Teng, C. S., Smith, B. R., Clayton, B., Evered, D. C., Clark, F., and Hall, R. (1977). *Clin. Endocrinol.* **6,** 207–211.
Toyka, K. V., Drachman, D. B., Pestronk, A., and Kao, I. (1975). *Science* **190,** 397–399.
Trembath, P. W., and Shaw, J. (1976). *Br. J. Clin. Pharmacol.* **3,** 1001–1005.
Tzartos, S. J. (1984). *TIBS* **9,** 63–67.
Tzartos, S. J., and Lindstrom, J. (1980). *Proc. Natl. Acad. Sci. U.S.A.* **77,** 755–759.
Tzartos, S. J., Rand, D. E., Einarson, B. L., and Lindstrom. J. M. (1981). *J. Biol. Chem.* **256,** 8635–8654.
Urbain, L. (1983). *Springer Semin. Immunopathol.* **6,** 1–5.
Valente, W. A., Vitti, P., Rotella, C. M., Vaughan, M. M., Aloj, S. M., Grollman, E. F., Ambesi-Impiobato, F. S., and Kohn, L. D. (1983). *N. Engl. J. Med.* **309,** 1028–1034.
Van Der Heide, D., Daha, M. R., Bolk, J. H., Bussemaker, J. K., De Bruin, T. W. A., Goslings, B. M., Van Es, L. A., and Querido, A. (1980). *Lancet* **1,** 1376–1380.
Van Obberghen, E., Kasuga, M., LeCam, A., Hedo, J., Itin, A., and Harrison, L. C. (1981). *Proc. Natl. Acad. Sci. U.S.A.* **78,** 1052–1056.
Van Obberghen, E., Rossi, B., Kowalski, A., Gazzano, H., and Ponzio, G. (1983). *Proc. Natl. Acad. Sci. U.S.A.* **80,** 945–949.
Venter, J. C., Fraser, C. M., and Harrison, L. C. (1980). *Science* **207,** 1361–1363.
Vincent, A. (1980). *Physiol. Rev.* **60,** 756–824.
Wartofsky, L. (1973). *JAMA, J. Am. Med. Assoc.* **226,** 1083–1088.
Wasserman, N. H., Penn, A. S., Freimuth, P. I., Treptow, N., Wentzel, S., Cleveland, W. L., and Erlanger, B. F. (1982). *Proc. Natl. Acad. Sci. U.S.A.* **79,** 4810–4814.
Weinberg, C. B., and Hall, Z. W. (1979). *Proc. Natl. Acad. Sci. U.S.A.* **76,** 504–508.
Wisher, M. H., Baron, M. D., Jones, R. H., Sonksen, P. H., Saunders, D. J., Thamm, P., and Brandenburg, D. (1980). *Biochem. Biophys. Res. Commun.* **92,** 492–498.
Yamashita, K., and Field, J. B. (1970). *Biochem. Biophys. Res. Commun.* **40,** 171–178.
Yip, C. C., and Moule, M. L. (1983). *Diabetes* **32,** 760–767.
Yip, C. C., Yeung, C. W. T., and Moule, M. L. (1980). *Biochemistry* **19,** 70–76.
Yoon, J.-W., McClintock, P. R., Onodera, T., and Notkins, A. L. (1980). *J. Exp. Med.* **152,** 878–892.
Yoshida, T., Tsuchiya, M., Ono, A., Yoshimatsu, H., Satoyoshi, E., and Tsuji, K. (1977). *J. Neurol. Sci.* **32,** 195–201.

CHAPTER 24

Sensitivity and Specificity of Autoantibody Testing*

ROGER L. DAWKINS
Departments of Clinical Immunology
Royal Perth Hospital
and The Queen Elizabeth II Medical Centre
Perth, Western Australia, Australia

I. INTRODUCTION

Although tests for autoantibodies are performed by countless laboratories and appear to be useful in diagnosis, there is relatively little information that establishes their practical value. In this chapter, data on several autoantibodies will be reviewed from the point of view of the clinician who wishes to use the results for diagnostic purposes. It will be shown that autoantibody testing can be useful but that application of some simple principles could lead to major improvements. At the same time, some of the basic characteristics of autoantibody testing pose some interesting conceptual problems for the immunologist (Dawkins and Peter, 1980).

* This is publication 8237 of the Departments of Clinical Immunology, Royal Perth Hospital, and the Queen Elizabeth II Medical Centre, Perth, Western Australia.

II. INFORMATION NECESSARY FOR THE INTERPRETATION OF AUTOANTIBODY TESTING

As in other areas of laboratory medicine, it is critical for the clinician to have some information on the sensitivity and specificity of autoantibody testing (Galen and Gambino, 1975). Sensitivity can be defined as "positivity in disease," or the probability of a positive test (T) given the presence of the disease (D), that is, $p(T/D)$. Specificity, on the other hand, is often referred to as "negativity in health" or the probability of a negative test ($\bar{T}$) given the absence of the disease in question ($\bar{D}$), or $p(\bar{T}/\bar{D})$. These terms are best understood by reference to the 2 × 2 format in Table I. It should be clear that this table raises some critical issues:

1. Positive and negative results must be defined so that they can be contrasted with each other.
2. The disease in question must be defined very clearly and must have some relevance to decisions made in practice. For example, very little is gained if disease is defined in such terms that no testing is necessary to establish the diagnosis. The clinician would like to know the sensitivity and specificity of antinuclear antibody (ANA) in relation to systemic lupus erythematosus (SLE) but would be misled if the 2 × 2 table were constructed by using data based only on patients who satisfied essentially all of the ARA criteria for SLE (Cohen *et al.*, 1971) or who possessed clinical features that were essentially pathognomonic. If the operational definition of SLE, that is, that used by the clinician who will make decisions with respect to diagnosis and management, relates to the presence or absence of particular clinical features, the 2 × 2 table should be constructed accordingly and not as dictated by the application of epidemiological criteria.
3. The calculation of specificity should not be attempted until it is clear whether the reference group has been selected so as to exclude SLE or, alternatively, so as to obtain a "healthy" group defined by the absence of any *major* illness.
4. It is quite clear that the most critical information obtained in the 2 × 2 table relates to the predictive value of positive or negative results rather than to sensitivity and specificity themselves. Predictive value of a positive result or $p(D/T)$ will of course be very much influenced by the relationship between the two row totals. For example, if the table simply compares 100 patients with SLE and 100 healthy subjects who have been shown to lack SLE, the $p(D/T)$ will approach 100%. If, on the other hand, the table has been constructed so as to allow for the prevalence of SLE (say, 1 : 1000 of the population as a whole), it will be necessary to give data

TABLE I

Formulae for the Evaluation of Tests[a]

	Test result		
	Positive (T)	Negative ($\bar{T}$)	Total
Disease group (D)			
Healthy Reference	a	b	a + b
group ($\bar{D}$)	c	d	c + d
Total	a + c	b + d	a + b + c + d

[a] $p(T/D)$ or probability of positive test given disease

$$= \frac{a}{a + b} = \text{sensitivity},$$

$p(\bar{T}/\bar{D})$ or probability of negative test given health

$$= \frac{d}{c + d} = \text{specificity},$$

$p(D/T)$ or probability of disease given positive test

$$= \frac{a}{a + c} = \text{predictive value of a positive},$$

$p(\bar{D}/\bar{T})$ or probability of health given negative test

$$= \frac{d}{b + d} = \text{predictive value of a negative}.$$

on 100,000 healthy subjects if 100 patients are to be included. Even if false-positive ANA occurs in only 1% of the healthy population, the p(D/T) will only be ~10%. Quite clearly the clinician needs to be able to construct a 2 × 2 table that is relevant to the use of the test. For example, if the test is to be used to screen a population for SLE, account must be taken of the prevalence of the disease. Alternatively, if ANA testing is to be used to distinguish between a limited number of options on the differential diagnosis (e.g., SLE versus rheumatoid arthritis), then the *relative* prevalence of these two conditions in that clinician's practice is much more important than the prevalence of either condition in the population as a whole.

It follows from these simple considerations that autoantibody testing may have been used incorrectly because the relevant information has not been collected, or because of substantial difficulties in defining positive and negative tests and the presence or absence of the relevant diseases. Examples of some of these difficulties will be given below, but the immunologist who can solve these problems will still have to provide the requestor with yet further information if the results are to have any substantial value. It will be shown below that most if not all autoantibodies should be regarded as continuous variables that can be measured over a quantitative range. If it is possible to define a useful cut-off between positive and

negative, it will also be necessary to provide some information on the distribution of results on either side of this cut-off. The requestor will need to know the confidence with which he can accept a particular quantitative result (e.g., a titer) as positive or negative. In this context he will need information relating to reproducibility. Within-laboratory precision must be known but so too must the relevance of the result to that obtained by other laboratories. It should be possible to make an estimate of accuracy (or the degree to which the given result relates to the "true" result).

The laboratory should also be able to provide information on the degree to which a particular result should influence decision making processes. It seems likely that many tests for autoantibodies are done in situations in which the result will not have much impact on the decision made. If the test is of value in excluding or confirming a diagnosis, it should be used accordingly and differently from a test that is of value in monitoring disease progress or in subclassifying a disease in relation to prognosis. In some cases the same test may give results which can be used profitably in different ways.

Finally, the laboratory will also have to be prepared to provide information on the cost-effectiveness of the test.

III. AN APPROACH TO THE EVALUATION OF AUTOANTIBODY TESTING

For simplicity it will be assumed that

1. A given autoantibody can be treated as a continuous variable and that it will be present in any given serum to a greater or lesser degree;
2. The test used to measure this antibody is capable of distinguishing between at least several amounts of the antibody;
3. Health is associated with less antibody than disease;
4. There will be some overlap since the amount of autoantibody will not be the sole determinant of whether or not disease is present at that particular time.

These assumptions have been largely or completely validated in some of the particular instances discussed below.

Given these assumptions it is convenient to approach the evaluation of a particular test by use of the very simple model given in Fig. 1. Essentially this figure represents frequency distributions of the antibody in health and disease. Depending upon the cut-off for positive and negative, $p(D/T)$ and $p(\bar{D}/\bar{T})$, etc., can be derived. The model can be adjusted so that it is more appropriate for a screening test as shown in Fig. 2. Here the

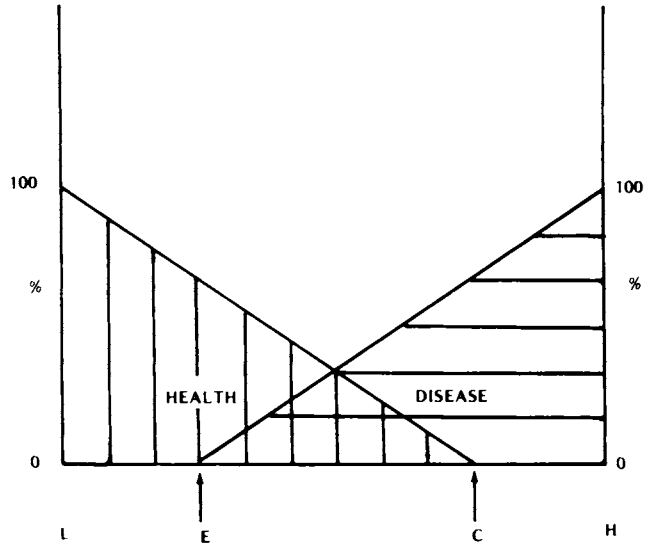


FIG. 1. A simple model illustrating a method for the assessment of a diagnostic assay that measures a continuous variable from low (L) to high (H). Cut-off points E and C refer to the excluding and confirming points. The overlap zone is between E and C. (Reproduced with permission from Garlepp *et al.*, 1982, *Journal of Neuroimmunology*.)

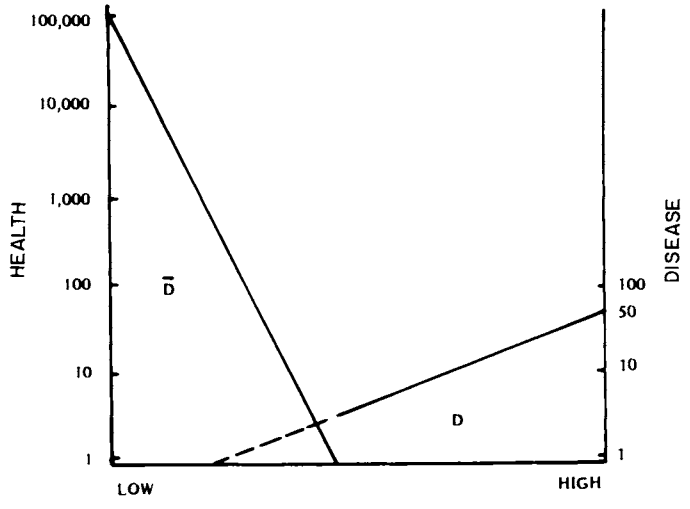


FIG. 2. Simple model illustrating a method for the assessment of a diagnostic assay after correcting for prevalence of disease (D) in a population. In this case the prevalence of the disease is taken as 50 per 100,000. The distribution in health ($\bar{D}$) is extrapolated to 100,000.

healthy population is extrapolated to 100,000, and the number of diseased subjects that might be expected to exist within that population is given on the right of the figure.

While these models necessarily represent oversimplifications, they do provide a convenient method for the display of the data and allow the clinician to deduce probabilities. A further adaptation is provided by comparing the frequency distributions in the diseases that might be included in the differential diagnosis.

All these versions have their defects, but they do allow the laboratory to concentrate its activity on the critical area of the measuring range. Obviously the laboratory will wish to achieve maximal reproducibility around its cut-off for positive and negative or at least around the range of results that are most discriminating. In this context it is important to remember that a result may be used for several purposes. Thus, although the test may be set up in such a way as to maximize its value in distinguishing between health and disease, the same result may subsequently be used for monitoring or some other purpose. If used for monitoring, a greater or lesser degree of precision may be required. Clearly the laboratory will have to consider the appropriate uses of the results.

The application of these models and the practical problems encountered will be discussed further by reference to selected examples.

IV. ANTINUCLEAR ANTIBODY

The term "antinuclear antibody" is used in two quite distinct ways. It may refer to the test for antibodies reactive with nuclei and detected by immunofluorescence, that is, the antinuclear factor (ANF) test. The same term is also used to describe all antibodies reactive with nuclei, irrespective of whether they can be detected by immunofluorescence. Those antibodies that react with soluble or extractable nuclear antigens may not be detected in the routine immunofluorescence test. In this section, ANF is used in the original sense and refers to ANA that are demonstrated by immunofluorescence.

A. ANF: SENSITIVITY FOR SLE

A review of the literature leads to confusion rather than clarity. In the early 1970s most workers found that ANF was present in essentially all cases with SLE, and presumably many of these workers required the presence of ANF as a sine qua non for the diagnosis (Davis, 1981). It was recognized that there were several circumstances in which the ANF test

TABLE II
Explanations for ANF-Negative Systemic Lupus Erythematosus

Laboratory
Wrong cut-off for positive and negative
Failure to detect "other ANA" (e.g., anti-DNA and anti-RNP
Reporting error, etc.
Clinical
Previous corticosteroid therapy
Hypogammaglobulinemia including nephrotic syndrome
Lupuslike syndromes (e.g., complement deficiency)

could be negative in the presence of definite SLE; these are listed in Table II. Therapy in the form of corticosteroids can certainly lead to a fall in titer, and it was well known that treated SLE can be ANF negative. Thus, as a diagnostic criterion, positive ANF referred to the untreated patient. In practical terms, however, it is not always clear to the patient and the physician whether or not there has been previous corticosteroid therapy. Antibody may be lost from the serum, as for example after plasmapheresis, or via the kidney, in the case of renal involvement in SLE. Such cases can be ANF negative and may remain negative for some time after the major episode of immunoglobulin loss. Substantial information is not available but there is the impression that a positive ANF should be qualified by a statement such as "unless the serum IgG is low." It has been reported that some lupuslike syndromes are not necessarily ANF positive. For example, a disease very similar to SLE can occur with inherited C2 deficiency but with a negative ANF. Further, on at least theoretical grounds, it is conceivable that ANF could be negative when antigen and antibody are at equivalence, but this phenomenon does not appear to be of practical importance.

These and other situations might account for a negative ANF in the presence of SLE but undoubtedly the most important consideration relates to the definition of positive and negative. It is quite extraordinary that many papers describing ANF-negative SLE have not considered the above situations and have not sought to justify the cut-off used by that laboratory. Even if the appropriate cut-off had been employed, it is likely (see Section IV,C) that imprecision at this level of detection would lead to some false negatives.

For all of these reasons it might be justifiable to consider that the sensitivity *should* be essentially 100% in untreated idiopathic SLE. High sensitivity at least provides a valuable starting point for the assessment of the specificity and predictive value of ANF testing.

TABLE III
Predictive Value Antinuclear Factor in Systemic Lupus Erythematosus[a]

	ANF		
	Positive (T)	negative ($\bar{T}$)	Total
Untreated SLE (D)	100	0	100
Health ($\bar{D}$)	2900	97000	99900
Total	3000	97000	100000

[a] Specificity $[p(\bar{T}/\bar{D})] = \frac{97000}{99900} = 97\%$;

Sensitivity $[p(T/D)] = \frac{100}{100} = 100\%$;

Predictive value $[p(D/T)] = \frac{100}{3000} = 3\%$.

B. ANF: Specificity for SLE

Several years ago it was assumed in our laboratory that untreated SLE was associated with ANF. Forty-four cases were studied, and it was found that almost all were obviously positive. However, it was necessary to adjust the assay to ensure positivity of the entire 100% (Rigby *et al.*, 1978). Such adjustment was made by varying the optical conditions. Using these adjustments a study of the Busselton population was undertaken (Hawkins *et al.*, 1979). It was found that ~3% of the total population was ANF positive. Thus, using a 2 × 2 table corrected for a prevalence of 1 per 1000, the specificity or $p(\bar{T}/\bar{D})$ was ~97% but the predictive value of a positive, or $p(D/T)$, was only 100 per 3000 or ~3% (Table III).

Several years later the same study was repeated with essentially the same results in the patients (i.e., a sensitivity of 100%), but the specificity had fallen to 89%. This change was shown to reflect two aspects. First, there was imprecision around the cut-off point, as will be discussed further below. Second, there had been an increase in the frequency of ANF positivity in the population as a whole. An analysis of the population revealed that ANF positivity was strongly correlated with prior antihypertensive therapy, and it was relevant that the community had been subjected to an intensive campaign aimed at the treatment of early hypertension. Thus, any assessment of the specificity and predictive value of ANF will be influenced by exposure to at least those drugs that are known to induce ANF positivity. Other factors associated with ANF were also examined. Interestingly there was no evidence of mild SLE, but there was

TABLE IV
Relationship between Antiviral Titers and ANF in the Population

	Percentage of group[a]		
	ANF-Negative (n = 90)	ANF-Positive (n = 50)	p^b
Cytomegalovirus	4	20	0.05
Herpes simplex (Type I)	41	72	0.001
Measles	26	36	NS[c]

[a] Percentage with antiviral titers above arbitrary cutoffs in subjects with and without ANF using a cutoff of ~10 WHO units.
[b] Probability of χ^2.
[c] NS, not significant.

a weak association between ANF positivity and a questionnaire response that suggested the presence of Raynaud's phenomenon. Further studies of this type are required.

The associations of ANF in the population as a whole may prove to be of value to the primary care physician, but it is not yet clear whether the demonstration of ANF should influence management decisions. Although antihypertensive and other therapy may induce ANF positivity, only a small fraction of these cases will develop SLE requiring treatment and it is not known whether ANF positivity should lead to any particular action on the part of the physician. Studies such as those conducted by Batchelor's group should be pursued (Batchelor *et al.*, 1980). If it could be shown that there was an identifiable genetic factor that predisposed subjects to the development of drug-induced SLE, ANF monitoring of such subjects may be warranted. Further studies are also necessary to determine whether the presence of ANF is of practical consequence to the patient with viral infection; certainly there is an association between ANF positivity and elevated titers to some viruses (Table IV).

It must be clear that any consideration of the specificity of ANF will be complicated by changes with time. Factors such as drug therapy and viral infections will come and go, and many subjects will be positive only transiently. Indeed in an earlier study we showed that ANF positivity varied substantially when subjects were followed sequentially (Hawkins *et al.*, 1979). The predictive value of persistent ANF could well be quite different from the predictive value of an ANF that is positive at only one particular time (see also Section VIII).

C. Quality Control

This term is used to refer to the problems related to precision and accuracy. Precision relates to scatter within or between runs, as depicted in Fig. 3. Accuracy should be used to describe the departure from the true result. In the case of autoantibodies, the truth is somewhat difficult to define and must generally be conceived in terms of internationally agreed standards. As shown in Fig. 3, accuracy may vary with time.

In an attempt to measure precision and accuracy for ANF testing, we have accepted the WHO ANA standard and attempted to develop internal standards that will be particularly relevant to the critical measurement zone. An international exchange program had shown very clearly that there was no substantial difficulty in achieving a consensus if the titer were high (e.g., 100 WHO units). There was much greater difficulty, however, with results around the cut-off for positive and negative as used by most laboratories. We therefore developed four standards by diluting a serum that was judged to have 10 WHO units with serum that appeared to have very little if any ANF. By this means we derived 1.25-, 2.5-, 5-, and 7.5-unit standards in addition to the 10-unit standard. When these stan-

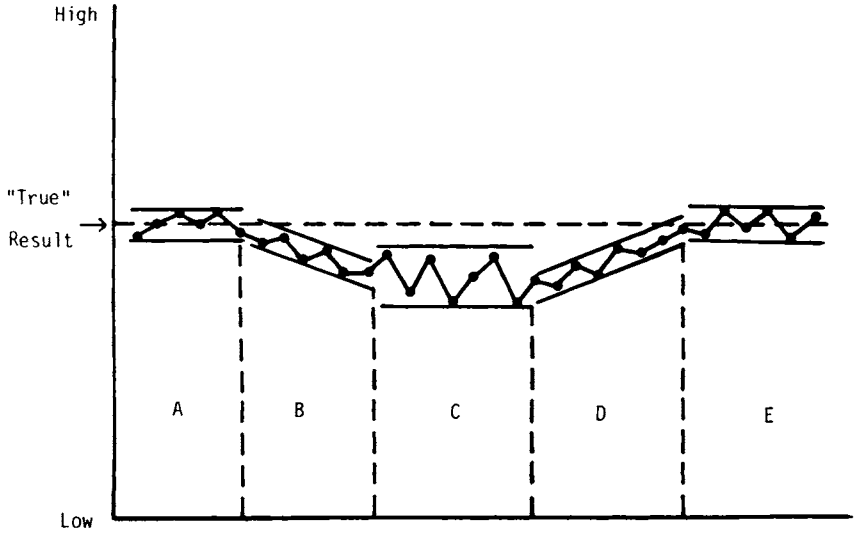


FIG. 3. Model illustrating precision and accuracy as seen in a plot of consecutive results for a specimen. Region A shows both precision and accuracy. Region B shows precision but a drift in accuracy demonstrated by the downward trend of the results away from the "true" result. Region C shows a lack of both precision and accuracy. Region D shows a return to precision and a drift back towards the "true" result. Region E shows a return to both precise and accurate results.

TABLE V
Differential Staining of Tissue Nuclei by Local Standards

Standard units/ml	Nuclei positive			Intensity grade[a]
	Heart	Kidney	Liver	
1.25	±[b]	0[b]	0	0.5
2.5	+[b]	0	0	1
5	+	+	0	2
7.5	+	+	±	3
10	+	+	+	4

[a] See Fig. 4.
[b] Composite staining patterns distinguish the standards: 0, negative; ±, weak positive; +, positive.

dards were tested at the usual screening dilution of 1:10, it was found that there were perceptible differences among them (see Table V). Using these differences it was then possible to score test sera and to relate these scores to those obtained with the four standards on the particular day. A curve was drawn as shown in Fig. 4 and results read off in units.

Using this approach, 200 of the Busselton group were retested; the frequency distribution is shown in Fig. 5. It can be seen that the population frequency of ANF will depend very much on whether positive is regarded as 10, 7.5, or 5 units. The fact that there was a predictable gradation provided some confirmation for the belief that it is possible to determine the amount of ANF present, that ANF could be treated as a continuous variable, and that distinctions could be made without expensive and time-consuming titrations.

As expected there was some difficulty with precision in the 5- to 10-unit range. Most of the difficulty could be attributed to day-to-day variation, which could be largely eliminated by defining the standard curve each day and reading off the results. While this approach could only be an approximation of ANF, it did appear justified by its success with other autoantibodies (see Sections V and VII). Further refinement appears possible.

When these 2.5 to 10-unit standards were circulated in an interlaboratory autoantibody exchange, some between-laboratory variation became apparent. The results are shown in Fig. 6. On the most recent exchange a higher degree of consensus was achieved, and it appears quite practical to attempt standardization within this critical range.

Returning to the problem of ANF in SLE, all patients' sera were well above 10 units. An approximation of the data as it would relate to the simple models is shown in Fig. 7. Based on our total experience to date

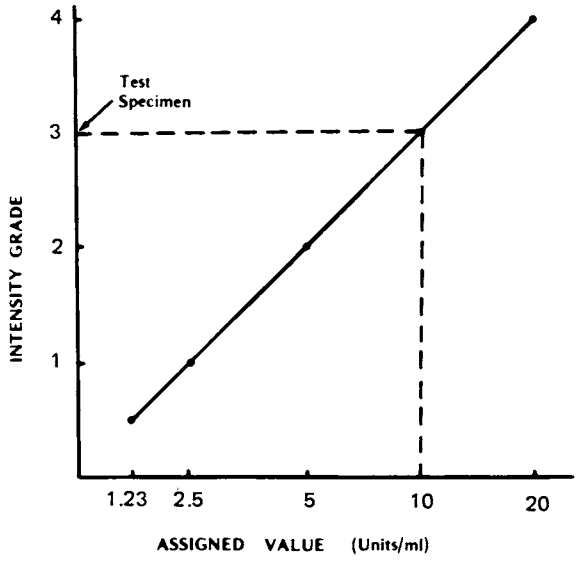


FIG. 4. Standard curve obtained by plotting assigned values against the intensity grades obtained. A test specimen with a particular grade (e.g., 3) may be reported in units (e.g., 7.5).

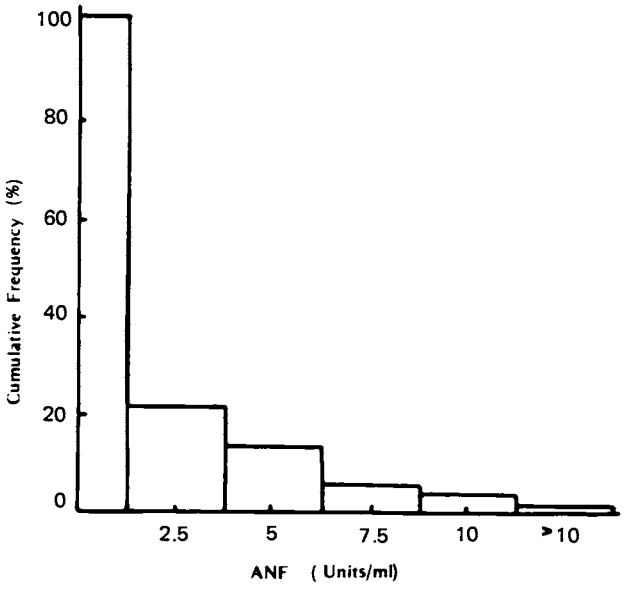


FIG. 5. Cumulative frequency plot of ANF after conversion to units. Approximately 11% of the population have ANF of 5 units/ml or more.

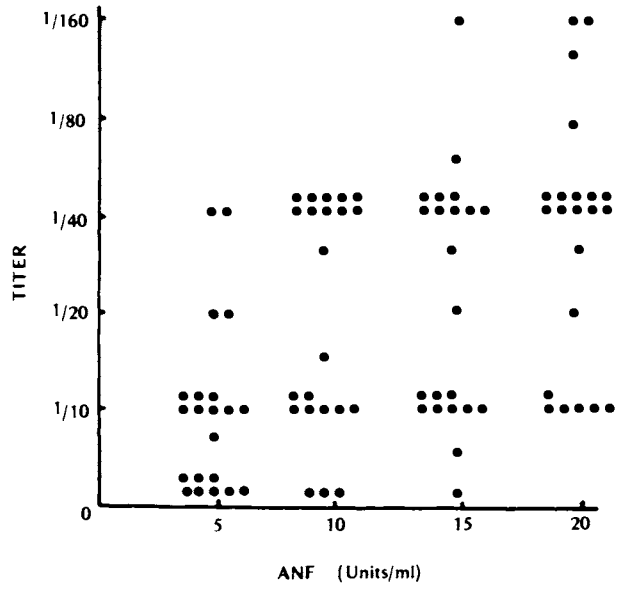


FIG. 6. Titers obtained when ANF standards were tested by laboratories involved in a quality control exchange program.

we regard 7.5 WHO units as a conservative excluding cutoff for SLE. Most if not all false negatives can be accounted for by the explanations given in Table II or by the application of ill-defined diaprostic criteria.

V. ANTI-DNA

Textbooks frequently state that ~70% of patients with SLE will have an elevated anti-DNA, although the proportion may be higher in those with active glomerulonephritis. False-positives (insofar as SLE is concerned) are found in chronic active hepatitis, rheumatoid arthritis, myasthenia gravis, and other autoimmune diseases.

Information of this type is of very limited value to the clinician. It would seem that a low result does not exclude SLE and that an elevated result does not confirm the diagnosis. Some workers claim that the result does not relate to activity and is of no value in monitoring. Some of the different opinions might reflect technical and statistical differences. There are numerous assays available, and some use quite different antigens. If indeed the critical antibody is reactive with only double strand DNA, improvement might be relatively simple and consensus might be achieved quite quickly.

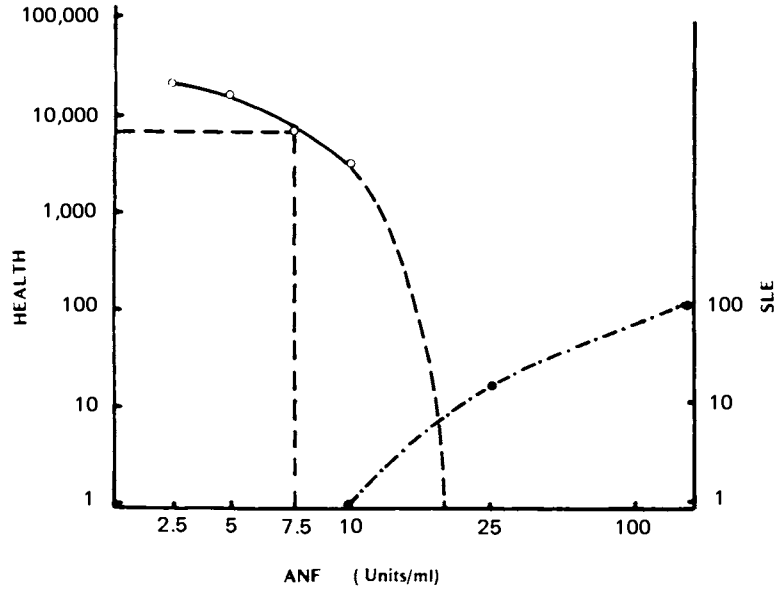


FIG. 7. Approximate cumulative frequency histograms for ANF in health (open circles) and in SLE (solid circles). The prevalence of SLE in the population is estimated at 100 per 100,000. The distribution in health is extrapolated to 100,000. The line drawn at 7.5 units/ml shows that at this ANF concentration there will be ~5000 individuals with a positive ANF of whom only 100 will have SLE.

Undoubtedly one problem relates to the definition of normal and abnormal, that is, the cut-off for positive and negative. Some authors have treated their data as if normally distributed, but this is rarely if ever appropriate in the case of autoantibodies, if only because lower results in healthy subjects are skewed since the assays are not discriminating at the lower end of the range of measurement. There would appear to be no justification to define a "normal range" in terms of mean ± 2 SD. Quite apart from the fact that the distribution is not Gaussian, this approach fails to take account of the relative infrequency of SLE and the large portion (e.g., 2.5%) of positive results in otherwise healthy subjects.

We have adopted an alternative approach based on the models discussed above. As shown in Fig. 8, an approximation of the frequency distribution of anti-DNA in the healthy population can be calculated so that this can be compared with results obtained from patients with untreated SLE. Accordingly, it might be reasonable to ascribe significance to results of 20% or higher, as determined by the membrane assay (Ginsberg and Keiser, 1973). Such results are not compatible with health but

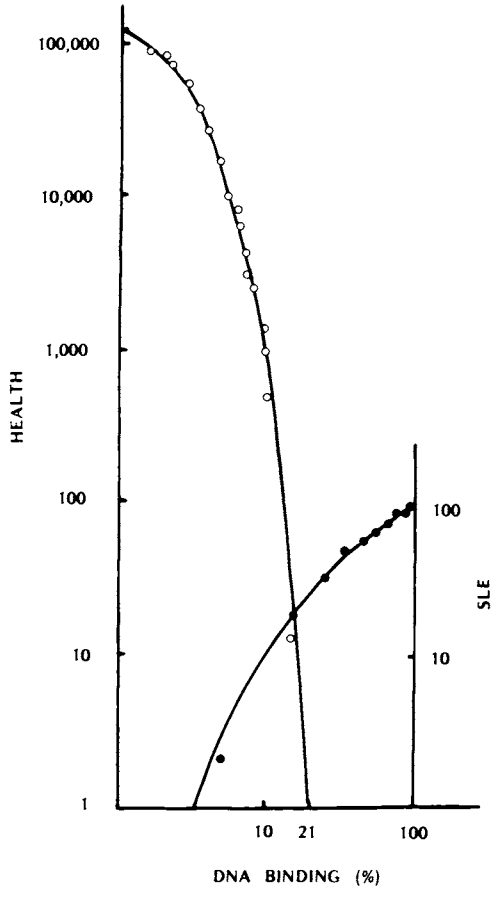


FIG. 8. Approximate cumulative frequency histograms for anti-DNA in health (open circles) and (solid circles) SLE. The prevalence of SLE in the population is estimated at 100 per 100,000. The distribution in health is extrapolated to 100,000.

suggest or confirm some diagnosis such as SLE. It is not possible to define the lower limit of the gray zone and therefore not possible to use the anti-DNA test to exclude SLE. Further studies of the distribution of results in patients are warranted.

The results given above relate only to those obtained by the membrane assay and perhaps only to those obtained within this particular laboratory. The results of a recent exchange are of some interest. It was found (Fig. 9) that most laboratories were able to distinguish different amounts of anti-DNA in that they obtained linear plots when their results were plotted against the mean or assigned results. However, the slopes of the

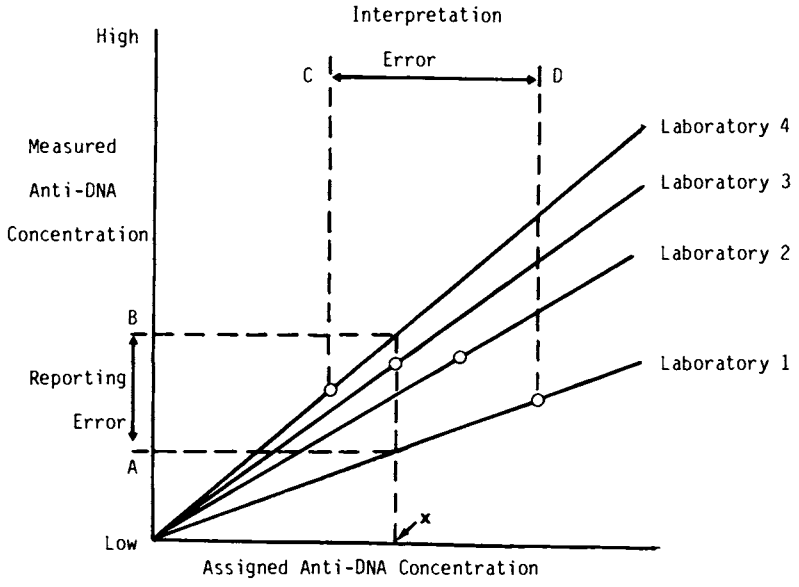


FIG. 9. Interlaboratory results for anti-DNA. Although linearity was obtained when each laboratory's results were plotted against the assigned results, the slope of each line varied substantially, leading to a reporting error (AB) as shown. The interpretation error (CD) is due to the vastly different reference ranges used (0 = cut-off point for each laboratory). If a serum (x) (arrow) were measured by each of the four laboratories, laboratories 1 and 2 would regard the result as negative whereas laboratory 4 would regard the result as positive.

lines varied substantially, Thus serum containing a given amount of anti-DNA antibody would be assigned quite different results by different laboratories. This problem could be avoided quite simply by standardization and agreement on the numbers to be reported. More importantly, however, the laboratories attributed vastly different significance to the results for a given serum. This was especially because they used such different reference ranges. As can be seen in Fig. 9, there was no relationship between the slope of the line and the actual reference range used. Presumably most laboratories used a reference range from the literature and had not made any local assessment. This problem can be resolved quite simply. Indeed sera that describe the distribution in health are available; such sera could be tested by other laboratories without the need for an exhaustive population study.

The predictive value of particular anti-DNA results should be determined. To our knowledge, this has not been done as yet. "False-positive" results do occur in myasthenia gravis and other conditions, but it is quite possible that at least most of these patients truly do have SLE in addition

to any other disease. It is interesting that the association between myasthenia gravis and SLE is well recognized.

VI. OTHER ANA

Although numerous other antinuclear antibodies have been described and some of these do appear to be associated with particular diseases or variants of SLE, most have not been submitted to careful scrutiny. In general terms, the assays are rather inadequate in that they only detect very high titers. Although results can appear to be of some interest (e.g., by showing a significant Chi square when the presence or absence of antibody is compared with some particular feature or diagnosis), the sensitivity and specificity may be low. It would be useful if the assay could be adjusted so that either *all* patients were positive or *only* patients were positive. Papers that describe antinuclear antibodies and their associations should give the distribution of results and should derive the predictive values.

VII. ANTI-ACETYLCHOLINE-RECEPTOR ANTIBODY

Several different methods have been used to detect and quantitate autoantibody to the acetylcholine receptor (anti-AChR). Although there might appear to be some rather different estimates of sensitivity and specificity, most of these probably reflect differences in methodology. For example, the immunofluorescence test of Sondag-Tschroots *et al.* (1979) quite clearly gives different results from those of the radioimmunoassay.

The results shown in Figs. 10 and 11 have been obtained using crude human AChR labeled with α bungarotoxin (Lindstrom, 1977; Garlepp and Dawkins, 1982). Similar assays have been used by most but not all (for review, see Vincent, 1980).

The sensitivity of anti-AChR for the adult-onset form of myasthenia gravis approaches 100%, but this is only true if the disease is defined so as to include patients with generalized involvement and clinical evidence of activity at the time of testing. Thus to achieve high p(T/D) it is necessary to use a specific subset of D. It will also be necessary to use an appropriate definition of T, that is, an appropriate cut-off for positive.

As one method of defining this cut-off, we have examined the distribution of results in patients with active adult-onset generalized myasthenia gravis and used a representative subset of the Busselton population as a reference group. Results (Fig. 12) have been expressed in the manner described above. It can be seen that the C cut-off is of the order of 1–2

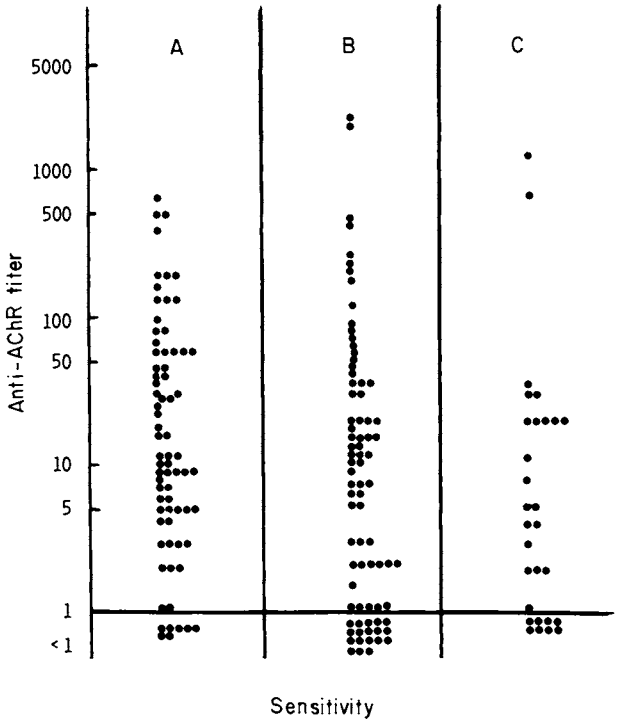


FIG. 10. Anti-AChR titers in patients with adult-onset generalized myasthenia gravis. Patients have been grouped according to disease activity at the time of assay. (A) Active disease; $n = 72$, sensitivity = 90%. (B) Normal daily activity; $n = 79$, sensitivity = 78%. (C) Asymptomatic; $n = 29$, sensitivity = 72%.

units but that some patients with disease may have slightly lower results. For practical purposes, a cut-off of 1 unit can be used except in situations where the test is applied to the screening of populations (Garlepp *et al.*, 1982).

It is implicit in this approach that anti-AChR can be detected in healthy subjects without myasthenia gravis. Very low results (e.g., <0.5 units) might be largely or wholly attributable to nonspecific binding to the AChR. On the other hand, results >1 unit are largely if not entirely due to specific binding of anti-AChR, and it seems likely that at least those subjects with results between 0.5 and 1.0 units do have specific antibody. It follows that the clinical features of myasthenia gravis reflect a certain amount of antibody rather than the mere presence of any antibody at all. Presumably some subjects are more or less susceptible to a given amount, possible reflecting genetic susceptibility, as has been shown by passive transfer to mice (Marzo, Garlepp, and Dawkins, in preparation).

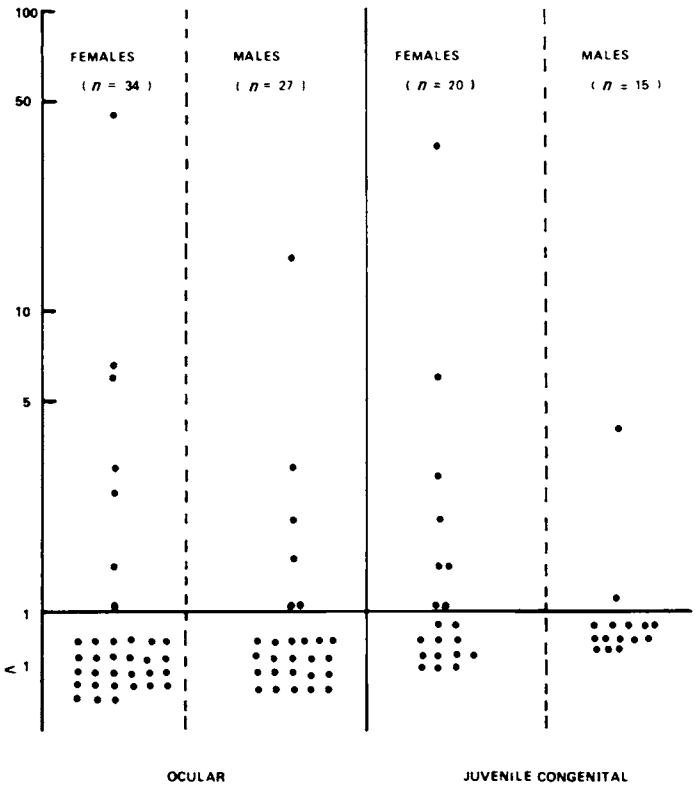


FIG. 11. Anti-AChR titers in patients with ocular myasthenia gravis or congenital/juvenile myasthenia gravis (onset <12 years). Sensitivity for ocular MG was 21%, and for juvenile/congenital, 29%. (Reprinted with permission from *Toxicon* Suppl. 3. Garlepp and Dawkins, "Elapid alpha-neurotoxins and myasthenia gravis," 1982, Pergamon Press.)

There are several other ways that a practically useful cut-off might be determined. For example, it is valuable to examine the distribution in diseases that are considered in the differential diagnosis of myasthenia gravis. Illustrative results are shown in Fig. 13. It can be seen that in some disorders there is a tendency for higher titers that are not necessarily associated with clinically apparent myasthenia gravis.

We have had the opportunity to study some patients who have developed myasthenia gravis some years after the collection and storage of serum for other reasons. An illustrative case is shown in Fig. 14. Retrospective examination of the sera of two other patients also showed elevated titers prior to the development of clinically obvious myasthenia gravis. It would appear that the threshold in the three cases studied se-

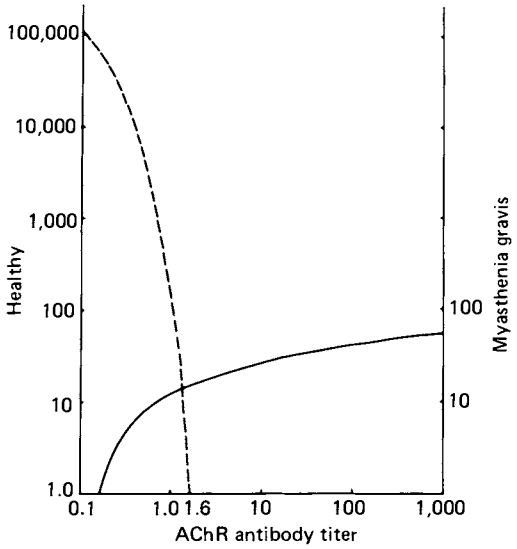


FIG. 12. Cumulative frequency histogram of anti-AChR titers in healthy individuals (dashed line), after extrapolation to 100,000 individuals, superimposed on that for active adult-onset generalized MG (solid line). (From Garlepp *et al.*, 1982, with permission from *Journal of Neuroimmunology*.)

quentially is of the order of 1–5 units. Thus the cut-off for positive, if defined in terms of association with clinically apparent disease, might have to be somewhat higher than the cut-off deduced from population studies. While it may be difficult to recognize myasthenia gravis in the presence of a second disease such as thyrotoxicosis or systemic lupus erythematosus, it does seem likely that patients with these disorders may have somewhat higher titers than the population as a whole and that they may be asymptomatic—with respect to myasthenia gravis—for at least some years after the acquisition of such titers.

There is some controversy as to whether anti-AChR is associated with restricted persistent ocular myasthenia gravis. Our own results are shown in Fig. 11. We have argued that those patients with relatively high titers may have associated mild generalised myasthenia gravis. At least most patients with restricted persistent ocular MG are negative with respect to anti-AChR, that is, they are no different from the population as a whole (Garlepp *et al.*, 1981). Others have suggested that there may be some detectable anti-AChR, especially if extraocular muscle is used as a source of antigen. Further studies are required.

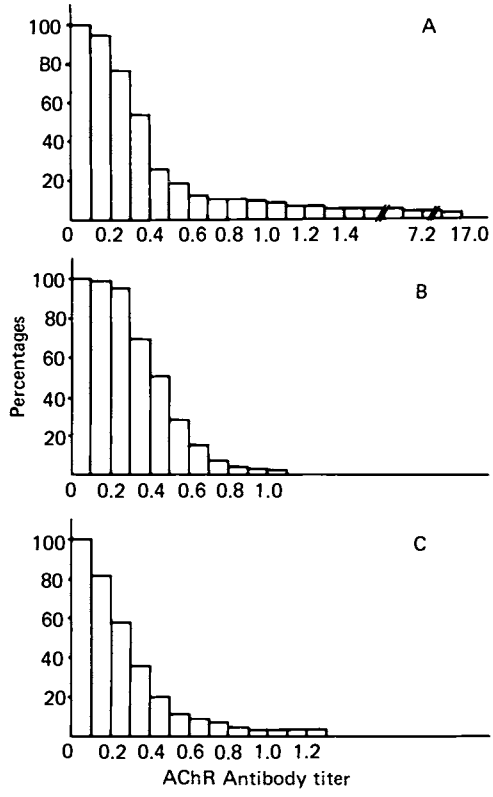


FIG. 13. Percent cumulative frequency histograms of anti-AChR titers in SLE (A), rheumatoid arthritis (B), and patients with high-titer thyroid autoantibodies (C). (From Garlepp *et al.*, 1982, with permission from *Journal of Neuroimmunology*)

As shown in Fig. 11, the congenital and juvenile forms of myasthenia gravis are generally, if not always, negative. These subcategories of the disease have not been well defined. It is quite clear that some patients with onset soon after puberty will be positive for anti-AChR. The disease with an onset in the first 2 years of life is negative in our experience, although some of these patients have been inactive at the time they were tested. Congenital myasthenia gravis contrasts with the neonatal type. The disease transferred from mother to baby is associated with anti-AChR, as might be expected. However only a proportion of those babies who receive maternal anti-AChR develop clinical evidence of the disease, again suggesting that there is some variation in susceptibility to a given amount of antibody.

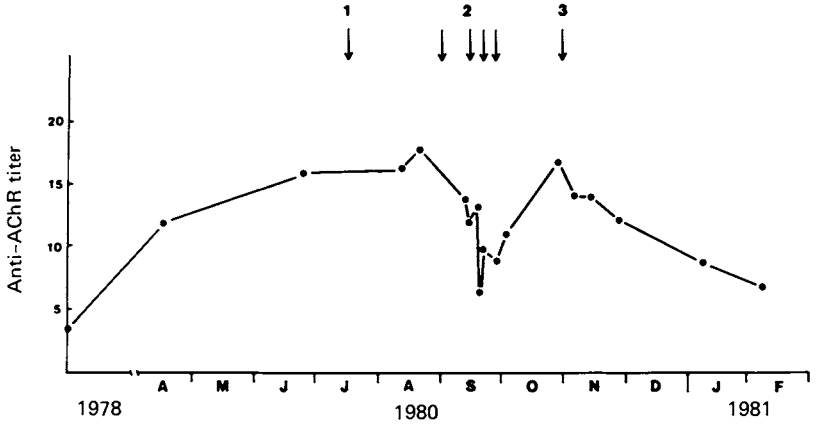


FIG. 14. Serial anti-AChR titers in a patient who presented with thyrotoxicosis in 1978. Retrospective analysis of the 1978 serum revealed an anti-AChR titer of 4 units. Myasthenia gravis was diagnosed in February 1980, and treatment included (see arrows) thymectomy (1), plasmapheresis (2), and alternate-day prednisone (3). (From Garlepp *et al.*, 1982, with permission from *Journal of Neuroimmunology*.)

Although there is general agreement that anti-AChR has relatively high sensitivity, specificity, and predictive value for myasthenia gravis, there is some controversy as to the value of the test for monitoring. In our experience changes in titer relate very well to changes in clinical status in the individual case, but there is no correlation between titer and severity if these two parameters are compared in a group of patients. This lack of correlation is at least partly due to the fact that young females with thymic hyperplasia and HLA A1, B8, DR3 have relatively high titers but often mild or at least remitting disease. Presumably these patients are high producers of antibody but have a relatively high biological threshold. Even in these cases, however, sequential testing has shown that changes in titer run parallel to changes in clinical severity.

In conclusion, a titer of anti-AChR >5 units essentially confirms a diagnosis of myasthenia gravis. A result <1 unit largely excludes this diagnosis, although not inactive, treated, congenital or ocular disease. A result between 1 and 5 will generally be found to be associated with active myasthenia gravis, but some patients, and especially those with thyrotoxicosis and systemic lupus erythematosus, will not have clinical evidence of myasthenia gravis. At this stage we would recommend a periodic review of such patients in the expectation that clinically apparent myasthenia gravis will develop.

The relatively high sensitivity and specificity of anti-AChR for active

generalized myasthenia gravis contrasts with the situation for most autoantibodies. Presumably this difference reflects the fact that anti-AChR is directly involved in the pathogenesis of the disease. By contrast ANA may simply be a marker for a whole family of antinuclear antibodies, most of which are not associated with recognized disease. Anti-DNA may or may not be important in the pathogenesis of SLE, but presumably it is dependent on other factors such as the availability of the appropriate antigen and the degree of complement activation.

VIII. ANTITHYROID ANTIBODY

A recent review (Strakosch *et al.*, 1982) illustrates some of the problems. Although most of the article is devoted to the presumed defect of immunoregulation, there is some discussion of thyrotropin receptor antibodies. It is stated that the measurement of these antibodies is important but difficult and of little assistance to the clinician in diagnosis and management. It is then stated that a high titer of thyroglobulin antibody is found in 55% of patients with Hashimoto's thyroiditis and that a low titer is found in 10–20% of normal subjects. No quantitative data are given. Thyroid microsomal antibody is said to occur in 95% of patients with Hashimoto's thyroiditis, 90% with idiopathic myxedema, and 80% with Graves' disease. This antibody is apparently detected less often in normal subjects. Again, no quantitative data are given.

In our own studies of patients with Hashimoto's thyroiditis proven by needle aspiration, it did appear that antimicrosomal antibody was positive in at least 90% (Gutteridge and Orrell, 1978.) Of some practical interest, most of these patients with negative antithyroid microsomal antibodies do have antibodies to antithyroglobulin. Again, quantitative data are not available.

Given that the sensitivity of thyroid microsomal antibodies for Hashimoto's thyroiditis approaches 100%, it is of some interest to determine the specificity and predictive value. In examining the frequency of these antibodies in the population as a whole, it is clear that age and sex have a major impact (Hawkins *et al.*, 1980). So also does persistence of the antibody (Hawkins *et al.*, 1979). It has generally been assumed that antibody-positive subjects in the population should be regarded as "false positive." However, our own data suggest that a large proportion of those with persistent antimicrosomal antibodies have or will develop hypothyroidism. Thus the major difficulty would appear to be in defining hypothyroidism, especially in relation to its evolution with time.

IX. OTHER AUTOANTIBODIES

Until careful quantitative studies are undertaken it will be difficult to determine the predictive value of most other autoantibodies. Meanwhile, only oversimplifications are possible. For example, antibodies reactive with the intercellular substance of the epidermis are found in essentially all patients with pemphigus vulgaris but also in other situations such as extensive burns. Deposition of immunoglobulin in biopsies is found in <100% of patients with pemphigus vulgaris, but this finding is of high specificity and high predictive value. By contrast the absence of serum autoantibody helps to exclude pemphigus vulgaris.

The situation in bullous pemphigoid is apparently different. Positive biopsy findings are more or less universal, whereas the serum antibody occurs with a sensitivity of somewhat less than 100%. A negative biopsy helps to exclude bullous pemphigoid whereas the presence of serum antibody to the epidermal–dermal junction is of relatively high specificity and predictive value.

These two diseases of the skin illustrate the diagnostic promise of autoantibody testing, but undoubtedly better data are required.

ACKNOWLEDGMENTS

I am grateful to my colleagues M. Garlepp, P. Kay, E. Bonifacio, P. Hollingsworth, and A. Miller.

REFERENCES

Batchelor, J. R., Welsh, K. I., Tinoco, R. M., Dollery, C. T., Hughes, G. R. V., Bernstein, R., Ryan, P., Naish, P. F., Aber, G. M., Bing, R. F., and Russell, G. I. (1980). *Lancet* **1,** 1107–1109.

Cohen, A. S., Reynolds, W. E., Franklin, E. C., Kulka, J. P., Ropes, M. L., Shulman, L. E., and Wallace, S. M. (1971). *Bull. Rheum. Dis.* **21** 643.

Davis, J. A. (1981). *In* "Textbook of Rheumatology" (W. N. Kelley, E. D. Harris, Jr., S. Ruddy, and C. B. Sledge, eds.), pp. 691–709. Saunders, Philadelphia, Pennsylvania.

Dawkins, R. L., and Peter, J. B. (1980). *Am. J. Med.* **68,** 3–5.

Galen, R. S., and Gambino, S. R. (1975). *In* "Beyond Normality: The Predictive Value and Efficiency of Medical Diagnosis" Wiley, New York.

Garlepp, M. J., and Dawkins, R. L. (1982). *Toxicon* Suppl. 3, 169–172.

Garlepp, M. J., Dawkins, R. L., Christiansen, F. T., Lawton, J., Luciani, G., McLeod, J., Bradley, J., Genkins, G., and Teng, C. S. (1981). *J. Neuroimmunol.* **1** 325–332.

Garlepp, M. J., Kay, P., and Dawkins, R. L. (1982). *J. Neuroimmunol.* **3,** 337–350.

Ginsberg, B., and Keiser, H. (1973). *Arthritis Rheum.* **16,** 199–207.

Gutteridge, D. H., and Orrell, S. R. (1978). *Clin. Endocrinol. (Oxford)* **9,** 505–514.

Hawkins, B. R., O'Connor, K. J., Dawkins, R. L., Dawkins, B., and Rodger, B. (1979). *J. Clin. Lab. Immunol.* **2,** 211–215.

Hawkins, B. R., Dawkins, R. L., Burger, H. G., Mackay, I. R., Cheah, P. S., Whittingham, S., Patel, Y., and Welborn, T. A. (1980). *Lancet* **2,** 1057–1060.

Lindstrom, J. (1977). *J. Clin. Immunol. Immunopathol.* **7,** 36–43.

Rigby, R. J., Dawkins, R. L., Kay, P. H., Matz, L. R., Papadimitriou, J. M., Quinter, J., and Haywood, E. F. (1978). *Aust. N. Z. J. Med.* **8,** 29–33.

Sondag-Tschroots, I. R. J. M., Schulz-Raateland, R. C. M., Van Walbeek, H. K., and Feltkamp, T. E. W. (1979). *Clin. Exp. Immunol.* **37,** 323–327.

Strakosch, C. R., Wenzel, P. E., Row, V. V., and Volpe, R. (1982). *N. Engl. J. Med.* **307,** 1499–1507.

Vincent, A. (1980). *Physiol. Rev.* **60,** 756–823.

CHAPTER 25

Autoimmunity of the Future

GUSTAV J. V. NOSSAL
The Walter and Eliza Hall Institute of Medical Research
Melbourne, Australia

I. INTRODUCTION

A look at the future makes sense only if one studies the past. I had the good fortune to "be around," if not at the very beginning of the saga of autoimmunity, then at least at the time that the field began to crystallize into a substantial body of scholarship. This was in the mid- to late 1950s when the logical interconnections between three areas of research, namely cellular immunology, embracing the field of immunological tolerance, experimental autoimmunization, and human autoimmune diseases came into full focus. It is interesting to recall the widespread animosity to the idea of autoimmunity at that time, particularly among clinicians, who, in the main, were content to see clinical immunology as a recondite field confined to uncommon diseases such as systemic lupus erythematosus, hemolytic anemia, or Hashimoto's thyroiditis. In those days, I found

myself frequently as an advocate for autoimmunity as being truly important, but even the most optimistic of us at The Walter and Eliza Hall Institute did not predict the range of diseases that would, to a greater or lesser degree, become entrapped within the web of autoimmunity inside a quarter of a century.

Looking at this resistance with the benefit of hindsight, I can see three reasons for it beyond the innate conservatism of the medical profession. First, I believe we made too absolute a distinction between intrinsic or "self" antigens and extrinsic or "not self" antigens, and thus took Ehrlich's "horror autotoxicus" too literally. Of course, the carrier effect had not yet been discovered, nor had its corollary—the breakage of tolerance, including self-tolerance, by cross-reacting antigens. In a conceptual sense, we had failed to pick up the cues from diseases like rheumatic fever and Chagas' disease, which clearly told us that significant cross-reactions occurred between intrinsic and extrinsic antigens, antibodies to the latter being capable of provoking serious damage to the former.

Second, we failed to see clearly enough that immunological tolerance was a quantitative rather than an absolute concept. We were all deeply under the influence of P. B. Medawar (Billingham *et al.*, 1953) and what he later termed the doctrinal tyranny of the skin graft. If a skin graft was rejected, even 30 or 40 days after transplantation, that was not "real" tolerance. Failing to see that experimentally induced or self-tolerance could be partial and yet operationally effective in many circumstances, we also largely failed to see the contingent truth, namely that some autoantibodies could be made by healthy people without causing disease or signifying a total breakdown of immunoregulation. Early thinkers like Najjar and Grabar (Grabar, 1957) got onto the point but failed to come to grips with the issue of what differentiated "normal" autoantibodies (such as those all mice make to bromelain-treated erythrocytes) from disease-causing autoimmunity.

Third, immunologists of the late 1950s strongly resisted the then exotic clonal selection theory of antibody formation (Burnet, 1957), and this being so, were even less impressed with his "forbidden clone" idea of autoimmunity (Burnet, 1959), which saw potentially antiself clones of immunocytes escaping from the homeostatic mechanisms that normally kept them silent. In autoimmunity, these clones were seen as proliferating wildly in response to the sea of self-antigen surrounding them. The forbidden clone notion has not survived as such, but, given that B-lymphocyte clones to tolerated antigens are clearly not physically deleted from the B-cell repertoire (Nossal, 1983), it behooves us to ponder the regulatory loops that might permit activation of such tolerant cells. We now realize that a certain degree of polyclonal activation of B cells can be achieved,

whether by mitogens like bacterial endotoxins, by lymphokine growth factors, or by molecules released from parasites, and so we have come to realize that antigen is by far not the only force involved in immunoregulation.

Resistance gradually gave way to compliance, so much so, in fact, that we as early promoters of the concept of autoimmunity sometimes had to blow the whistle as virtually every obscure chronic disease was promoted by one author or another as a candidate addition to the growing list of autoimmune conditions! This increased interest in the area has brought a wealth of knowledge over the past quarter century but also its share of unanswered questions.

II. UNANSWERED QUESTIONS IN THE WAKE OF THE GREAT DISCOVERIES

A. The Growing List of Autoimmune Diseases

The growing list of autoimmune diseases has posed serious questions about where the line should be drawn. In diseases like autoimmune hemolytic anemia or myasthenia gravis where the association between autoantibodies and mechanism of disease is so clear cut, there are no problems. At the other end of the spectrum are diseases like rheumatoid arthritis and multiple sclerosis, where the number of patients exhibiting the relevant autoantibodies but clearly not having the disease poses a real dilemma as to the meaning of the antibody. Insulin-dependent diabetes sits somewhere in between, and the lengthy period during which relatives of patients can have anti-islet-cell antibodies introduces the new concept of a prodromal or preclinical stage of the disease at which immunologically mediated damage perhaps remains clinically undetectable because of the reserve capacity of the target organ and/or compensatory division among target cells.

To this must be added the issue of diseases of target organs in which autoantibodies have only recently been discovered. Idiopathic hypoparathyroidism, partial hypopituitarism, and idiopathic diabetes insipidus (with antivasopressin antibodies) fall in this camp. The situation in these conditions is that (arguing by analogy) they will probably fall into the autoimmune camp, but before this can be generally accepted, a sufficient weight of confirmatory evidence is required.

In the future as in the past, each disease will have to be studied as an entity in its own right, with the autoimmune component seen as only one in a web of causative factors. The grouping of the diseases into the

autoimmune category should not be allowed to impede the search for particularities that might clarify etiology in each specific case.

B. The Genetic Risk Factors

Undoubtedly, the unraveling of genetically transmitted risk factors, particularly those associated with major histocompatibility (MHC) and immune regulatory genes, has been one of the major discoveries in the field. However, I believe no one guessed just how complicated the genetics of autoimmunity would turn out to be. This is referred to in many chapters in this volume. It represents an area of research where clinical and animal studies have greatly complemented each other. Linkages to MHC or to immunoglobulin allotype are relatively easy to understand in that immunoregulatory or antibody structural genes are involved. But, as Drs. Mackay and Rose point out in their chapter (Chapter 1), a close dissection of animal models of autoimmunity may reveal four separate MHC-linked and one non-MHC-linked immunoregulatory genes involved without nearing the end of the story! Other genes affecting target organ structure, B-lymphocyte-activation levels, or thymic physiology are in the process of being identified, and the track record to date indicates that still other genes remain to be defined. So, for the majority of autoimmune diseases, we shall have to face up to highly multifactorial, multigenic effects.

C. Mechanisms of Immunological Damage

In the early years of autoimmunity research, life was rather simple because you could take your pick between an antibody's damaging a target organ directly or immune complexes' depositing and initiating an inflammatory cascade involving leukocytes, platelets and the complement system. One of our problems today is that we have so many competing models for damage to target cells! The simple view of complement-fixing, cytolytic antibody has to be supplemented with those of antibodies' causing aggregation and pinocytosis of vital cell-surface macromolecules; antibodies' preparing cells for antibody-dependent cellular cytoxicity; antibodies that stimulate receptors; and antibodies that block either a physiological signal or the damaging effects of other antibodies. On the cellular side, we have not only cytotoxic T lymphocytes capable of inflicting target cell damage, but also K cells, NK cells, ALK cells, and their relatives. Furthermore, T cells of the helper/inducer phenotype are clearly important in chronic inflammatory states, delayed hypersensitivity reactions, and graft rejection, and therefore need investigation within the

context of autoimmunity. The whole question of how to relate the many *in vitro* tests to the *in vivo* immunopathology requires a new conceptual approach.

D. The Importance of Immunological Cross-Reactions

The monoclonal revolution—both monoclonal antibodies and antigen-specific, functional T-lymphocyte clones—has illustrated more dramatically than any theoretical argumentation by a clonal selectionist the degree to which the immune system is degenerate and redundant. One antibody can recognize more than one antigen; one antigen can react with more than one antibody. The same is true for T-lymphocyte recognition. Hybridomas have been made using the B lymphocytes of both humans and mice with autoimmune diseases (Schwartz, 1983). These have illustrated both the extraordinary polyclonality of autoimmune responses and the importance of cross-reactivities. For example, anti-DNA autoantibodies from lupus patients cross-reacted with cardiolipin and also with phosphorylcholine-containing bacterial antigens, suggesting that some lupus autoantibodies might be primarily antibacterial antibodies that just *happen* to register with the phosphodiester-containing molecule DNA. Furthermore, several of the anti-DNA monoclonal antibodies share a cross-reactive idiotype, an interesting point for proponents of network theories of immune regulation. Other autoantibodies are reactive to cellular cytoskeletal molecules like actin and vimentin. These could be due to cross-reactivity with viruses, as extremely similar molecules are present in the virions of measles and parainfluenza.

As already mentioned in the introduction, the marked cross-reactivities among antigens and antibodies have profound implications for the field of immunological tolerance and hence for the background concepts shaping our views on autoimmunity. From the standpoint of clonal selection, the simplest way of preventing autoimmunity would be to silence all potential antiself lymphocytes. This could be accomplished via an actual killing of the cell, either through a direct encounter between the antigen and the cellular receptor for it or through the action of an antiidiotypic cytotoxic T lymphocyte, which might easily read out as a suppressor T cell. Alternatively, the antiself cell might be rendered anergic through early contact with antigen, but not be physically deleted from the repertoire. Whatever variant of these ideas is proposed, the end result would be a functional purging of some elements of the B- or T-lymphocyte repertoires. If that purging were absolute, even for cells with quite a low affinity for the tolerated antigen, cross-reactivity patterns would ensure that each self-antigen would create a sizable "hole" in the repertoire. Given the very

large number of self-antigens, as an ensemble they could well delete the whole repertoire!

I might add that this whole argument is unchanged if one places the onus for self-tolerance on suppressor T cells: At what level of B-cell cross-reactivity should they stop suppressing? This line of reasoning suggests that tolerance will *never* be absolutely complete, but rather that it is a concept requiring definition in terms of the numbers of immunocompetent precursor cells available for activation and their affinities for the tolerated epitope. Conversely, autoimmunity must be similarly defined, the strong probability being that small amounts of low-affinity antiself antibodies constitute a normal reaction to pulsatile release of self-antigens, aberrant antigen presentation, or polyclonal lymphocyte activation. Autoimmunity, on this view, must be seen as a perturbation of the regulatory processes that constrain these events.

III. FUTURE DIRECTIONS OF RESEARCH IN AUTOIMMUNITY

A. Genetic Aspects

I anticipate that research on the genetic aspects of autoimmunity will take a new turn as a result of the recombinant DNA revolution. Particularly in the human, the serological approach to the definition of the products of the *D*-locus-related immunoregulatory genes has had severe limitations. Now, the study of *HLA* gene polymorphism by restriction endonuclease mapping has become a reality (Cohen and Dausset, 1983). Using the Southern blotting technique, genomic DNA can be digested with various restriction endonucleases and labeled with a Class I or Class II *MHC* gene probe. Each person gives a characteristic, complex banding pattern, and family studies have shown that the gene fragments segregate only with *HLA* haplotype. Moreover, systematic correlations with serologically defined alleles reveal both the expected concordance and interesting exceptions, which should provide more definitive proof of *HLA* gene organization and polymorphism.

The approach is already giving clues to gene segments of importance as risk factors for particular diseases. For example, when the highly polymorphic β-chain genes of the class II *HLA* cluster are analyzed with *Eco*RII and a βDC probe, 15 of 16 normal healthy individuals possessing the serologically defined allele *DR2* display a characteristic 2.2-kb band. Individuals lacking *DR2* lack the band. *DR2* is a risk factor for multiple sclerosis but a protective factor for insulin-dependent diabetes. In fact, all

DR2-positive multiple sclerosis patients were found to possess the band, but when seven *DR2*-positive patients with diabetes were identified—a rare group—none of them possessed this band! So *DR2*-positive individuals with the band appear to be absolutely protected against insulin-dependent diabetes.

It is a safe prediction that many more polymorphic differences will be found through this form of mapping that will correlate more closely with disease susceptibility than the serological associations presently available. Painstaking work in this direction will pinpoint not only which *HLA* genes, but which sequences within them, are responsible for the disease association. Through this, a technology of enormous diagnostic value but also one capable of shedding new light on disease mechanisms will be born.

B. The Role of Antireceptor Antibodies

Having taken some joy in the role that The Walter and Eliza Hall Institute played in the early days of antireceptor antibodies in autoimmunity (Lennon and Carnegie, 1971; Lindstrom *et al.*, 1976), I can now see that this field is entering a new era. Once again, the two new biotechnologies, gene cloning and monoclonal antibodies, are the reasons for rapid progress. Three future trends can readily be identified. First, monoclonal antibodies, including those prepared from animals or humans with autoimmune diseases, will identify more and more cell-surface macromolecules as the clinically important targets for autoimmune attack. The TSH receptor, the acetylcholine receptor, and the insulin receptor will surely be followed by a long string of like examples. Moreover, the case of differing bioeffects of autoantibodies to receptors will become much more numerous. In thyroid disease, particular autoantibodies to the TSH receptor can block hormone action or mimic it, and in the latter case there is the suggestion that there may be differential effects on thyroid cell growth and thyroid cell differentiation. This raises the interesting possibility of different sorts of TSH receptors, capable of transmitting different signals back to the cell.

I anticipate that both the structure and the function of the most important target molecules for pancreatic islet-cell autoantibodies will soon be defined. Antiinsulin receptor antibodies are operative only in rare, insulin-resistant cases of the disease, as discussed by Dr. Harrison elsewhere in this volume (Chapter 23), but logical considerations suggest that antibodies to cell-surface macromolecules would be of greater pathogenetic significance than those detected by staining cytoplasmic components in fixed cells or sections. The latter type of antibody is still of great impor-

tance in the serological diagnosis of autoimmune diseases, and frequently, in organ-specific autoimmune disease, a family of tissue-specific antibodies is present. In such cases, a continued search for anti-cell-surface antibodies should prove worthwhile. As more examples are found, antireceptor autoantibodies will add to our knowledge of normal cellular physiology, which in turn will allow us to understand the disease mechanisms more fully.

Second, gene cloning is revealing the structure of many important receptors with breathtaking speed. This will sharpen up concepts of how receptors function and therefore of how autoantibodies can perturb that function. For the more important receptors, a longer term goal will be to produce enough material to allow crystallization and hence X-ray crystallographic analysis.

Third, as structural information about receptors increases so will attempts to use this new knowledge to interfere in the disease process. One could envisage computer-aided drug design creating agonists and antagonists specific for receptors, which could counterbalance whatever was the damaging effect of the antibody. It might even not be too fanciful to imagine active attempts to induce tolerance or immune suppression using synthetic epitopes that possess or mimic the receptor's structure. In this longer range work, models dependent on spontaneous autoimmune diseases in mice or rats will be particularly valuable.

While on the topic of possible therapeutic interventions, it is worth mentioning two areas that will certainly continue to be explored, namely monoclonal antibodies to particular MHC gene products that function as restriction elements for autoimmune responses and antiidiotypic monoclonal antibodies directed against cross-reactive idiotypes of autoantibodies. Such "brave new world" approaches may lack credibility today, but they deserve real research attention when one considers how devastating some of the more serious autoimmune disorders are.

C. Autoimmune T-Lymphocyte Lines

The capacity to transmit autoimmunity passively and transiently by antibody (e.g., in the case of the neonate, from an autoimmune mother) helped to cement the notion of autoantibodies as genuine mediators of disease. A recent area of research (Cohen, 1983) promises as much for the T lymphocyte. Long-term lines of T lymphocytes can be generated that possess reactivities enabling them to cause autoimmune disease on adoptive transfer. Thus, new insights into experimental allergic encephalomyelitis, experimental neuritis, autoimmune thyroiditis, and adjuvant-in-

duced arthritis are being obtained. The T-cell lines are generated from autoimmunized mice through alternative cycles of treatment with the autoantigen and with T-cell growth factor in the absence of antigen. Surprisingly small numbers of these cells, 10,000 or so, can cause significant symptomatology. The T cells, which are predominantly of the helper/inducer phenotype, are antigen specific and MHC restricted. Their infusion into normal mice raises fascinating problems. In the case of EAE, how do they reach the brain? There may be a concentration gradient of autoantigen at work. How is autoantigen presented to them? Endothelial cells or certain astroglia may possess the necessary class II MHC restriction molecules. How do they breach the blood–brain barrier? There are even some lines with suppressive potential, that is, the capacity to "switch off" autoimmunity in an antigen-specific, MHC-restricted way. In the arthritis model, these can successfully prevent arthritis from developing and can lessen its severity once established. Artificial manipulation of the disease state by *in vitro* cellular engineering represents a valid long-term goal.

D. Induction of HLA-DR Expression as a Trigger of Autoimmunity

An early event in the induction of the immune cascade is the presentation of antigen to a helper/inducer T lymphocyte by a cell bearing class II MHC molecules on it. While macrophages, dendritic cells, and B lymphocytes are the cell types most readily seen as expressing such molecules (e.g., HLA-DR in the human), it has recently been noted that a variety of other cells can be induced to become DR postive. The chief mediator of this aberrant DR expression is γ interferon. This has led to a novel hypothesis for the mechanism leading to organ-specific autoimmunity (Bottazzo *et al.*, 1983). A local virus infection (e.g., of an endocrine organ) causes interferon production. This leads to aberrant DR expression on the relevant epithelial cell. Cellular autoantigens, for which no major degree of self-tolerance exists because of their low concentration in the circulation, are now correctly presented to helper/inducer T cells within the damaged organ. These T cells lead to the activation of effector precursors, both T cells and B cells, and thus to autoimmunity.

This hypothesis is of considerable interest, if only because it focuses the mind on an interesting possible explanation for self-tolerance. Organ-specific autoantigens, such as cell-surface receptors, may simply be tolerated by the body because normally they do not engage the interest of the immune system. Aberrant DR expression places the restriction element

and the self-antigen into a proximity not normally possible. It may be that some failure of suppressor mechanisms is also required before transient autoantibody production turns into frank autoimmune disease.

The examples of abnormal class II antigen expression in epithelia in various disease states are increasing and warrant careful examination in the light of these new ideas.

E. Hybridization Histochemistry in the Search for Viral and Other Genes

A consideration of viruses as possible causative agents in autoimmunity leads on to one final technical point. Immunofluorescence has been a tool of immense value in autoimmunity research and is a linchpin of our present diagnostic technology. Essentially, immunohistochemistry detects *gene products* in and on cells. A potentially even more powerful technique is now in its infancy. It is the use of genetic probes labeled in various ways as tools for the search for specific DNA or RNA sequences in cells or tissues. Specifically, radioactively labeled probes are being used extensively to search for given mRNA molecules in the cytoplasm and thus to identify the tissue of origin of the relevant gene product. The technique can equally validly be used to search for DNA viruses, or cDNA copies of viral RNA, that have been integrated into the host genome. The general name for this technology is hybridization histochemistry.

It is not yet clear how widespread a use this technology will have in pathology and what practical advantages it will possess over immunohistochemistry. It is certainly already proving of value in the search for a possible viral etiology of chronic diseases, including some possibly due to autoimmunity. It has occasioned surprises in terms of production of hormones, brain peptides, and other molecules in previously unsuspected locations. The technique may find imaginative uses in immunopathology. If the structure of an antigen of interest is known, it is now possible to make a labeled probe for its mRNA far more rapidly than it is to make an antibody! Quite apart from its relevance to a search for viral sequences, it may come to be regarded as an adjunct to immunohistochemistry for practical reasons.

IV. SUMMARY AND CONCLUSIONS

The former paradigm of the immune response centered on a first injection of antigen creating a small amount of antibody and a pool of memory

cells, and a booster injection giving a secondary response with an explosive division and differentiation of the primed immunocytes. While nothing has invalidated this simple notion, it has had to accommodate to a large number of immunoregulatory facts, including the modulating influence of helper and suppressor T cells, the markedly different effects of sudden pulses of an antigen versus its constant presence, the positive and negative feedback potential of antibody and of antiidiotype, and the realities of immune regulatory genes. These considerations have shaded the differences between self and nonself antigens, which now in many respects seems more an operational than an absolute one. If to this degree immunological tolerance has lost its magic, so has its converse, autoimmunity.

If autoimmunity then is to be viewed not as due to some single, striking causative mechanism, but rather as a serious failure of several elements in a highly complex control system, we have somehow to gear our future research away from an overly reductionistic pathway. The student of autoimmunity will not be able to afford to neglect any strand of cellular immunology or immunogenetics and furthermore should be constantly alert to clues coming from other fields where persisting antigen and an effective immune system coexist; parasitic diseases and transplantation biology are the two obvious examples. More than at any previous time, these fields will be technology sensitive. Their moving edges will shift and adapt as new techniques redefine the questions that can be asked. Therefore, the clinical researcher faced with the real-life problems of disease will have to maintain a high level of alertness about what the new techniques can do and a high degree of connectivity to basic research scientists able to master them fully.

The biotechnological revolution will affect the laboratory diagnosis of the autoimmune diseases just as surely. Indeed, there is a high probability that any peek at the autoimmunity of the future will miss the most exciting developments, which will depend on profound discoveries not yet made. I have no doubt that the next quarter century will be just as stunning, both for the field of autoimmunity as a whole and for each of the major diseases that constitute the loosely knit group.

ACKNOWLEDGMENTS

The author's own work was supported by The National Health and Medical Research Council, Canberra, Australia, by grant number AI-03958 from the National Institute for Allergy and Infectious Diseases, United States Public Health Service, and by generous private donors to the Hall Institute.

REFERENCES

Billingham, R. E., Brent, L., and Medawar, P. B. (1953). *Nature (London)* **172,** 603.
Bottazzo, G. F., Pujol-Borrell, R., Hanafusa, T., and Feldman, M. (1983). *Lancet* **2,** 1115.
Burnet, F. M. (1957). *Aust. J. Sci.* **20,** 67.
Burnet, F. M. (1959). *Br. Med. J.* **2,** 720.
Cohen, I. R. (1983). *In* "Progress in Immunology V" (Y. Yamamura and T. Tada, eds.), pp. 1129–1137. Academic Press, New York.
Cohen, D., and Dausset, J. (1983). *In* "Progress in Immunology V" (Y. Yamamura and T. Tada, eds.), pp. 1–12. Academic Press, New York.
Grabar, P. (1957). *Tex. Rep. Biol. Med.* **15,** 1.
Lennon, V. A., and Carnegie, P. R. (1971). *Lancet* **1,** 630.
Lindstrom, J. M., Seybold, M. E., Lennon, V. A., Whittingham, S., and Duane, D. D. (1976). *Neurology* **26,** 1054.
Nossal, G. J. V. (1983). *Annu. Rev. Immunol.* **1,** 33.
Schwartz, R. S. (1983). *In* "Progress in Immunology V," (Y. Yamamura and T. Tada, eds.), pp. 1111–1114. Academic Press, New York.

Index

A

B

C

D

E

F

G

H

J

K

L

M

N

O

P

Q

R

S

T

U

V

W