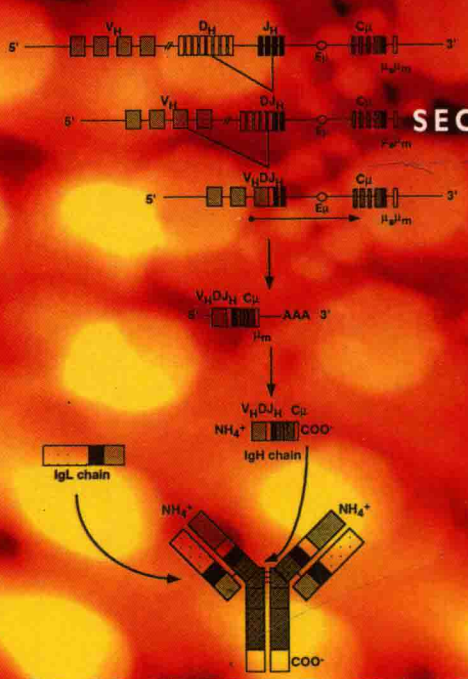

SECOND EDITION

IMMUNOGLOBULIN GENES

EDITED BY

Tasuku Honjo

Frederick W Alt

Immunoglobulin Genes
Second edition

Edited by

T. Honjo
Kyoto University, Japan

F. W. Alt
Howard Hughes Medical Institute Research Laboratories,
The Children's Hospital, Boston

ACADEMIC PRESS

Harcourt Brace & Company, Publishers
London, San Diego, New York, Boston, Sydney, Tokyo, Toronto

ACADEMIC PRESS LIMITED
24–28 Oval Road
LONDON NW1 7DX, UK

U. S. Edition Published by
ACADEMIC PRESS INC.
San Diego, CA 92101, USA

This book is printed on acid-free paper

A catalogue record for this book is available from the British Library

ISBN 0 12 053640 4

Typeset by Phoenix Photosetting, Chatham
Printed and bound by Hartnolls Ltd, Bodmin, Cornwall

Immunoglobulin Genes
Second edition

Dedicated to the memory of Georges Kohler

Contributors

FW Alt
Howard Hughes Medical Institute Research Laboratories, The Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA

MK Anderson
Department of Molecular Genetics, All Children's Hospital, 801 Sixth Street South, St Petersburg, FL 33701-4899, USA

PD Burrows
Division of Developmental and Clinical Immunology, Departments of Microbiology, Medicine and Pediatrics, University of Alabama at Birmingham, 263 Wallace Tumor Institute, UAB Station, Birmingham, AL, USA

K Calame
Columbia University College of Physicians and Surgeons, Department of Microbiology, 701 West 168th Street, New York, NY 10032-2704, USA

JD Capra
Southwestern Medical School, Department of Microbiology, 5323 Harry Hines Boulevard, Dallas, TX 75235, USA

MD Cooper
Howard Hughes Medical Institute, Department of Pediatrics and Microbiology, University of Alabama at Birmingham, 263 Wallace Tumor Institute, UAB Station, Birmingham, AL 35294, USA

LE Daitch
Tufts University School of Medicine, Department of Pathology, 136 Harrison Avenue, Boston, MA 02111, USA

EA Faust
Howard Hughes Medical Institute Research Laboratories, University of California, Los Angeles, 5-748 MacDonald Building, 10833 Le Conte Avenue, Los Angeles, CA 90024-1662, USA

S Ghosh
Howard Hughes Medical Institute, Yale University, 295 Congress Avenue, 154 BCMM, New Haven, CT 06510, USA

LA Herzenberg
Stanford University School of Medicine, Department of Genetics, Immunogenetics and Cell Sorting Laboratory, Stanford, CA 94305-1662, USA

T Honjo
Faculty of Medicine, Department of Medical Chemistry, Kyoto University, Yoshida, Sakyo-ku, Kyoto 606, JAPAN

AB Kantor
Department of Microbiology, College of Physicians & Surgeons, Columbia University, New York, NY, USA

KL Knight
Loyola University Chicago, Stritch School of Medicine, Department of Microbiology and Immunology, 2160 South First Avenue, Maywood, IL 60153, USA

GW Litman
Department of Molecular Genetics, All Children's Hospital, 801 Sixth Street South, St Petersburg, FL 33701-4899, USA

F Matsuda
Kyoto University, Center for Molecular Biology and Genetics, Shogo-in Kawaharacho 53, Sakyo-ku, Kyoto 606, JAPAN

F Melchers
Basel Institute for Immunology, Grenzacherstrasse 487, Postfach, CH-4005 Basel, SWITZERLAND

C Milstein
Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, UK

D Nemazee
National Jewish Center, Department of Pediatrics, 1400 Jackson Street, Denver, CO 80206, USA

A Okada
Howard Hughes Medical Institute Research Laboratories, The Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA

KN Potter
Southwestern Medical School, Department of Microbiology, 5323 Harry Hines Boulevard, Dallas, TX 75235, USA

C Rada
Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, UK

JP Rast
Department of Molecular Genetics, All Children's Hospital, 801 Sixth Street South, St Petersburg, FL 33701-4899, USA

M Reth
MPI fur Immunbiologie, Postfach 79011, D-7800 Freiburg, GERMANY

C-A Reynaud
Institut Necker, Faculté de Medicine, Necker-Enfants Malades, 156 rue de Vaugirard, 75730 Paris cedex 15, FRANCE

DC Saffran
Howard Hughes Medical Institute Research Laboratories, University of California, Los Angeles, 5-748 MacDonald Building, 10833 Le Conte Avenue, Los Angeles, CA 90024-1662, USA

HW Schroeder, Jr
Division of Developmental and Clinical Immunology, Departments of Microbiology, Medicine and Pediatrics, University of Alabama at Birmingham, 263 Wallace Tumor Institute, UAB Station, Birmingham, AL, USA

E Selsing
Tufts University School of Medicine, Department of Pathology, 136 Harrison Avenue, Boston, MA 02111, USA

AM Stall
Department of Microbiology, College of Physicians & Surgeons, Columbia University, New York, NY, USA

U Storb
University of Chicago, Department of Molecular Genetics and Cell Biology, 920 East 58th Street, Chicago, IL 60637, USA

C Tunyaplin
Loyola University Chicago, Stritch School of Medicine, Department of Microbiology and Immunology, 2160 South First Avenue, Maywood, IL 60153, USA

J-C Weill
Institut Necker, Faculté de Medicine, Necker-Enfants Malades, 156 rue de Vaugirard, 75730 Paris cedex 15, FRANCE

SM Wells
Department of Microbiology, College of Physicians & Surgeons, Columbia University, New York, NY, USA

ON Witte
Howard Hughes Medical Institute Research Laboratories, University of California, Los Angeles, 5-748 MacDonald Building, 10833 Le Conte Avenue, Los Angeles, CA 90024-1662, USA

HG Zachau
Institüt für Physiologische Chemie der Universität München, München, Germany.

J Zhang
The Rockefeller University, 1230 York Avenue, Box 167, New York, NY 10021, USA.

Foreword

When Behring and Kitasato discovered antibodies back in 1890, they could hardly have imagined that these factors were prime examples of both the chemical and biological bases of molecular recognition. The specificity of antibodies was soon recognized as a biological puzzle by Ehrlich, who in 1905 proposed the first (selective) theory of antibody specificity. Even so, Ehrlich considered it inconceivable that there could be specific substances ready to recognize and neutralize toxins that the animal species had never encountered before. So he produced the most ingenious idea that the toxins were the ones which by pure coincidence were capable of recognizing 'side chains', located on the surface of cells, and required for the utilization 'of foodstuffs'. Ehrlich therefore made the remarkable prediction of the existence of receptors on the cell surface, but could not conceive the capacity of the organisms to recognize the unknown and to learn to improve such recognition. While he laid the foundations of what was going to be a major preoccupation of immunologists for a long time, namely quantitative immunochemistry, it was the role of Landsteiner to demonstrate that antibodies were indeed capable of recognizing substances, naturally occurring or otherwise, which the animal had never seen before, through his classical studies with haptens. This remarkable property of the immune system became a dominating intellectual challenge to basic immunologists. The additional conviction that such specific recognition was capable of further improvement through a process of maturation of the immune response, demonstrated by Heidelberger and Kendall in the late 1930s, was an added complication for which no rational explanation could be proposed at the time.

At first it was proposed that antigens act as a form of template, around which the antibody is synthesized or folded. Prominent proponents of such an 'instructive' hypothesis were Haurowitz and Pauling. But this soon ran into difficulties, for a number of reasons. A major one was the progress in the understanding of the molecular basis of protein structure and folding and, more generally speaking, the molecular basis of biological specificity.

The new ideas and the developments of new methods which characterize the birth of molecular biology, had a direct impact on immunology. The emerging techniques of protein chemistry were immediately applied to the studies of antibodies by Rodney Porter, soon after his PhD supervisor, Sanger, developed the methods which culminated with the demonstration that proteins had defined amino acid sequences. This early inroad into the protein chemistry and early amino acid sequences of the N-terminus of antibodies, in the early 1950s, gave support to the instruction theories, in that no heterogeneity could be discovered. On the contrary, a single N-terminal

sequence could be discerned in rabbit immunoglobulins, leading to calculations into the highly improbable possibility of different molecules having identical N-terminal sequences. This may sound strange today, when we have become so accustomed to the idea of protein families, and closely related tandem arrays of genes. However, in those days, the idea that genes could arise by gene duplication only gained acceptance in the 1960s, following Braunitzer's comparison of the α and β chains of haemoglobin. That antibodies were indeed heterogeneous was very difficult to demonstrate, and was the result of a variety of studies coming from different directions that slowly built up into an inescapable conclusion at a later date. Earliest among them was the discovery of idiotypes made by Oudin in the early 1960s. The clonal selection theory proposed by Burnett and inspired by an alternative selective theory made by Niels Jerne, provided a very sound theoretical basis to the generation of specificity through protein microheterogeneity.

The common structural architecture of antibody molecules made up of two heavy and two light chains could be established in the early 1960s by Edelman and Porter, with the heterogeneous population of antibodies, because of the very fact that they represented an invariant character of all molecules. The critical element which revealed the essential character of the antibody diversity did not come from studies of antibodies themselves, but from myeloma proteins. Myeloma proteins have been known for a very long time, so much so that what turned out to be the light chains of myeloma proteins were discovered by Bence-Jones in 1847. The relationship between human myeloma proteins and antibodies arose largely from the careful antigenic analysis performed by Henry Kunkel, and extended to the mouse counterparts from the mouse plasmacytomas discovered by Michael Potter. It was the structural analysis of such molecules which brought about a further understanding of the underlying diversity within the frame of a general common architecture. The early peptide maps of human and mouse Bence-Jones proteins performed, respectively, by Putnam and colleagues and by Dreyer and his colleagues, were quickly superseded by the demonstration in 1965 by Hilschmann and Craig that such light chains consisted of a common segment and a variable segment. The existence of allotypic markers which appeared to be localized in the V region of rabbit heavy chains, and shared by IgG, IgM and IgA (the Todd phenomenon), and my own demonstration that variable segments of human kappa light chains consisted of at least three non-allelic sets of V_κ regions in association with a single C_κ region, provided experimental evidence that the variable and constant domains must be encoded by separate genes, as proposed by Dreyer and Bennett.

The comparison between the rapidly expanding sequences of V segments of myeloma proteins disclosed the existence of the hypervariable regions, which were to be called complementarity determining regions (CDRs) by Kabat, to imply that those were the residues directly involved in the antigenic recognition. The generalized common architecture, predicted by our analysis of disulphide bonds, and conceptualized by the domain structure proposal of Edelman, with the hypervariable segments predictably located at the tips of the Y-shaped molecule (as seen in electron micrographs) received confirmation from the structural studies which followed the

crystallization of Fab fragments of myeloma proteins by Poljak and Nisonoff. This exciting period characterized by studies of myeloma proteins using protein chemistry techniques, was to be enriched by the first glimpse of the somatic hypermutation, which resulted from the comparison of different lambda chains made by Cohn, Weigert and co-workers.

All this was soon to be superseded by the application of the DNA recombinant technology, in myelomas and in the newly derived hybridomas. The spectacular confirmation of the two genes/one polypeptide made at the DNA level by Tonegawa in 1976 led, within a period of less than 10 years, to our present understanding of the genetic arrangement and rearrangement of the antibody genes. Further success was provided by the attack on the problem of the T-cell recognition system. The long-drawn out controversy concerning the T-cell receptor was finally solved, to close the chapter of basic understanding of the genetic nature of the origin of diversity and of the structures involved in antigen recognition. The connection between the major histocompatibility complex and immunology became established and the observations of Zinkernagel and Doherty (which later culminated with the crystallographic explanation of the phenomenon) opened a new area with the first glimpse of the molecular bases of T–B cell collaboration.

The complexity of cell–cell and receptor–ligand interactions were soon to become a major focus of research. The International Workshops of Human Leucocyte Differentiation Antigens (now organizing its sixth conference) gave impetus and introduced order in efforts towards the dissection and characterization of the relevant cell surface molecules and receptors. Other techniques of molecular and cell biology, like transgenic and knockout mice at the one extreme, and gene amplification and fast DNA sequencing techniques at the other, provided new approaches to reassess old problems and to uncover and investigate previous intractable questions. For example, the bases of antigen recognition and of the maturation of the immune response have been generally clarified to uncover unique molecular mechanisms and interactions which underpin the processes.

As in years gone by, the impact of technological advances is not only revolutionizing but also reshaping the way we think about immunology. The pace of progress and change is faster and faster all the time. I wonder how many young colleagues realize that there are still a good number of active scientists who in their own youth, were asking the question: does a single amino acid sequence code for the world of antibodies? and were yet to learn that lymphocytes were not only large or small but also T or B.

C. Milstein

Preface to first edition

Since Kitasato and Behring discovered antibodies in animal serum in the late 19th century, the structure, function and expression of antibodies or immunoglobulins have posed exciting and important questions in immunology. There is no doubt that immunoglobulins are essential molecules in the immune system since most infectious diseases can be prevented or cured by appropriate specific antibodies.

Protein chemical studies on the immunoglobulin structure showed, firstly, that the light chain of an immunoglobulin molecule is composed of variable and constant regions. This discovery, however, served only as the vanguard to further questions: How can single polypeptides with variable and constant regions be synthesized? How are the immunoglobulin genes organized? How can so many variable regions be produced by a limited number of genes?

We had to wait until the next major technical development, namely DNA cloning, to elucidate the basic framework of dynamic rearrangement of the immunoglobulin genes. During the period up to this discovery, a variety of models were proposed, most of which did, however, turn out to be partially correct. A new technology had to become available to solve questions which emerged from immunology, but are fundamental to molecular genetics in the eukaryote. The development of the recombinant DNA techniques allowed us to explore the above questions in a straightforward manner.

During the past decade, we have accumulated an enormous amount of information on the organization, structure, rearrangement and expression of immunoglobulin genes in a variety of organisms. Studies on immunoglobulin genes have had a great impact not only on immunology but also on molecular biology in general. Such studies have provided many precedents for new concepts in eukaryotic molecular biology: exon–intron organization, differential splicing, site-specific as well as region-specific recombination, gene deletion and somatic mutation are examples.
This book provides up-to-date overviews of various aspects of immunoglobulin genes by authors who have actively participated in the accumulation of our knowledge on this subject. The editors hope that this book will serve as a prelude to further advancement and look forward to new developments in the field.

Tasuku Honjo
Frederick W. Alt
Terry H. Rabbitts

Preface to second edition

Since the first edition of *Immunoglobulin Genes* appeared five years ago, we have accumulated a large amount of knowledge about the structure and function of immunoglobulin (Ig) genes and the mechanisms that regulate their assembly and expression. We now have substantial information concerning the organization and structure of the human Ig heavy chain and κ light chain loci. In particular, the complete human Ig heavy-chain variable region (V_H) locus has been isolated in the form of linked cosmids and yeast artificial chromosomes; all of the V_H gene segments have been precisely arrayed within the locus, the 5′ end of which is linked to the telomeric sequence of human chromosome 14. A fundamental question in immunology has been the nature of the enzymatic machinery involved in the assembly of Ig variable region genes, and the related TCR variable region genes, from their germ-line V_H, D and J_H segments and the mechanisms by which this process (V(D)J recombination) is regulated. This area has witnessed remarkable progress following the seminal discovery of the Recombination Activating Genes (RAG) 1 and 2, the simultaneous expression of which generates V(D)J recombination activity in any mammalian cell type. The RAG gene discovery has facilitated numerous ongoing studies which have identified novel genes which encode products involved in the V(D)J recombination and DNA double strand break repair processes, at least one of which has been implicated as the target of the Murine Scid mutation. An additional major advance has been the rapid development of the gene targeted mutation technology which has been particularly useful to address questions concerning the regulation of Ig gene assembly and expression. 'Knock-out' experiments which have focused on various sequences within the Ig gene loci have elucidated both expected and unexpected functions of various exons and sequences.

In spite of the rapid progress in the aforementioned areas, there still remain many fundamental areas of investigation which are wide open for future in-depth studies. The molecular and cellular mechanisms that regulate assembly of Ig heavy and light chain genes remain a subject of intense investigation. In particular, the long debated phenomena of allelic and isotype exclusion and the mechanisms that regulate this process are still not well understood. Despite considerable progress in the V(D)J recombination field, a similar level of mechanistic understanding of the class switch recombination and somatic hypermutation processes awaits the further development of model systems and the identification of involved molecules. The function of Ig as a surface receptor and mechanisms and processes involved in its signaling has been another rapidly emerging area of new discoveries and progress, but much remains to be done. Finally, we expect that the complete nucleotide sequences of the Ig heavy

and light chain loci will be determined by the time we complete the next edition of this book. Such information, coupled with the increasing ability to perform targeted genetic manipulations in animals, should allow very rapid advances with respect to elucidation of mechanisms that regulate Ig gene assembly and expression.

The second edition of *Immunoglobulin Genes* provides updated reviews of various aspects of the structure, function, and regulation of Ig Genes, and the related aspects of B-lymphocytes development and function, by authors who have actively participated in the accumulation of this knowledge. We hope that this book will continue to serve as a reference to past progress and as a stimulus for future studies. We look forward to many exciting developments in Ig gene research during the next five years.

Tasuku Honjo
Fred Alt
June, 1995

Contents

PART I: B CELLS