

general editors:  
A. NEUBERGER & L.L.M. VAN DEENEN

# Molecular Genetics of Immunoglobulin

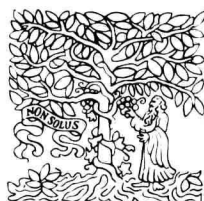
edited by  
F. Calabi & M.S. Neuberger

# Molecular Genetics of Immunoglobulin

*Editors*

F. CALABI and M.S. NEUBERGER

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



1987

ELSEVIER

Amsterdam · New York · Oxford

© 1987, Elsevier Science Publishers B.V. (Biomedical Division)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the Publisher, Elsevier Science Publishers B.V. (Biomedical Division), P.O. Box 1527, 1000 BM Amsterdam, The Netherlands.

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of the rapid advances in the medical sciences, the Publisher recommends that independent verification of diagnoses and drug dosages should be made.

*Special regulations for readers in the USA.* This publication has been registered with the Copyright Clearance Center, Inc. (CCC), Salem, Massachusetts. Information can be obtained from the CCC about conditions under which the photocopying of parts of this publication may be made in the USA. All other copyright questions, including photocopying outside of the USA, should be referred to the Publisher.

ISBN 0-444-80915-5 (volume)

ISBN 0-444-80303-3 (series)

Published by:

Elsevier Science Publishers B.V. (Biomedical Division)  
P.O. Box 211  
1000 AE Amsterdam  
The Netherlands

Sole distributors for the USA and Canada:

Elsevier Science Publishing Company, Inc.  
52 Vanderbilt Avenue  
New York, NY 10017  
USA

#### **Library of Congress Cataloging in Publication Data**

Molecular genetics of immunoglobulin.

(New comprehensive biochemistry ; v. 17)

Includes bibliographies and index.

1. Immunoglobulins--Genetics. 2. Gene expression.

I. Calabi, F. (Franco) II. Neuberger, M.S. (Michael S.)

III. Series. [DNLM: 1. Gene Expression Regulation.

2. Immunoglobulins--genetics. W1 NE372F v.17 / QW 601 M718]

QD415.N48 vol.17 574.19'2 s [616.07'9] 87-24302  
[QR186.7]

ISBN 0-444-80915-5 (U.S.)

Printed in The Netherlands

## MOLECULAR GENETICS OF IMMUNOGLOBULIN

# New Comprehensive Biochemistry

Volume 17

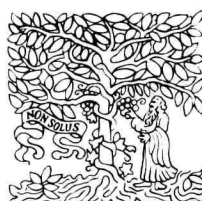
*General Editors*

A. NEUBERGER

*London*

L.L.M. van DEENEN

*Utrecht*



ELSEVIER

Amsterdam · New York · Oxford

## Preface

Immunoglobulin genes are not just of interest to immunologists. An understanding of the way in which DNA rearrangement and somatic mutation contribute to antibody diversity is of importance to a wide range of biologists. The cell-type specificity of immunoglobulin gene expression is of concern to many who are interested in gene expression in mammals. Furthermore, the immunoglobulin superfamily itself presents important questions to those interested in evolution.

The analysis of immunoglobulins and of their genetics has advanced rapidly since the mid-1970s, mainly as a result of the application of recombinant DNA and monoclonal antibody technologies. The essential features of the molecular anatomy of both antibodies and their genes have been largely identified; this has resulted in significant insights into the way antibody diversity is generated. Clearly, much still remains to be elucidated in these areas, whilst studies both of regulation and of phylogeny are still in their infancy. We felt nevertheless that it was a good time to draw together what we do know about the molecular genetics of immunoglobulin.

We wish to thank the authors for contributing to this volume and the publisher for prompt publication.

Cambridge

Franco Calabi  
Michael S. Neuberger

## List of abbreviations

Ars	<i>p</i> -azophenylarsonate
BiP	immunoglobulin heavy chain binding protein
C region	constant region
CDR	complementarity determining region
D segment	diversity segment
Fab	antigen binding fragment from papain digestion of immunoglobulin
F(ab') <sub>2</sub>	antigen binding fragment from pepsin digestion of immunoglobulin
Fc	crystallizable fragment from papain digestion of immunoglobulin
FcR	immunoglobulin Fc receptor
FR	framework region
H	immunoglobulin heavy chain
H <sub>m</sub>	membrane form of immunoglobulin heavy chain
H <sub>s</sub>	membrane form of immunoglobulin heavy chain
HAT	hypoxanthine, aminopterin, thymidine medium
HIV	human immunodeficiency virus
Ig	immunoglobulin
IL-4	interleukin 4
J segment	joining segment
J chain	joining chain
L	immunoglobulin light chain
LPS	bacterial lipopolysaccharide
LTR	long terminal repeat
MAG	myelin associated glycoprotein
MHC	major histocompatibility complex
N region	nucleotide region
NCAM	neuronal cell adhesion molecule
NK cell	natural killer cell

NP	4-hydroxy-3-nitrophenylacetyl
Ox	2-phenyl-5-oxazolone
PC	phosphorylcholine
Poly-IgR	poly-immunoglobulin receptor
RFLP	restriction fragment length polymorphism
RS	rearranging sequence ( $\kappa$ locus)
S region	switch region
TcR	T cell receptor
tk	thymidine kinase
V region/segment	variable region/segment



# Contents

Preface .....	V
List of abbreviations .....	XII

## Chapter 1

### *Structure and function of antibodies*

D.R. Burton (Sheffield, UK) .....	1
1. Introduction .....	1
2. Structure of IgG .....	3
2.1. General considerations .....	3
2.2. Domain structure .....	6
2.3. Structure of Fab .....	6
2.4. Antigen recognition .....	8
2.5. Structure of Fc .....	11
2.6. The hinge: IgG subclasses .....	18
2.7. Isotypes, allotypes and idiotypes .....	21
3. Functions of IgG .....	21
3.1. Introduction .....	21
3.2. Interaction with protein A .....	21
3.3. Complement activation .....	22
3.4. Interaction with cellular Fc receptors .....	24
3.5. Other functions of IgG .....	26
3.6. Rheumatoid factors .....	27
3.7. Membrane or surface IgG .....	27
3.8. Structure-function relationships in IgG: domain hypothesis .....	28
4. Structure of other immunoglobulins in relation to IgG .....	28
5. Structure and function of IgM .....	36
5.1. Structure of IgM .....	36
5.2. Functions of IgM .....	37
5.3. Membrane IgM .....	39
6. Structure and function of IgA .....	39
6.1. Structure of serum IgA .....	39
6.2. Structure of secretory IgA .....	41
6.3. Functions of IgA .....	41

## VIII

7. Structure and function of IgD . . . . .	41
8. Structure and function of IgE . . . . .	43
8.1. Structure of IgE . . . . .	43
8.2. Functions of IgE . . . . .	44
9. Summary . . . . .	46
Acknowledgements . . . . .	46
References . . . . .	47

## Chapter 2

### *Genes encoding the immunoglobulin constant regions*

M. Brüggenmann (Cambridge, UK) . . . . .	51
--	----

1. Introduction . . . . .	51
2. Chromosomal localization . . . . .	51
3. Organization of constant region genes . . . . .	52
3.1. Mouse heavy chain genes . . . . .	52
3.2. Mouse light chain genes . . . . .	53
3.3. Human heavy chain genes . . . . .	54
3.4. Human light chain genes . . . . .	55
3.5. Other species . . . . .	55
3.6. Switch regions . . . . .	58
3.7. Membrane exons . . . . .	61
3.8. Immunoglobulin J chain . . . . .	62
3.9. DNA motifs . . . . .	62
4. Structure of individual constant region genes . . . . .	64
4.1. Heavy chain genes . . . . .	64
4.2. Light chain genes . . . . .	67
4.3. Polymorphism . . . . .	68
4.4. Pseudogenes . . . . .	69
4.5. Evolution . . . . .	70
4.6. Aberrations and malignancies . . . . .	72
Acknowledgements . . . . .	75
References . . . . .	75

## Chapter 3

### *Genes encoding the immunoglobulin variable regions*

P.H. Brodeur (Boston, MA, USA) . . . . .	81
--	----

1. Introduction . . . . .	81
2. V gene structure . . . . .	81
3. Gene families . . . . .	84
3.1. Mouse $V_H$ families . . . . .	85
3.2. Human $V_H$ families . . . . .	86
3.3. Mouse $V_\kappa$ families . . . . .	87
3.4. Human $V_\kappa$ families . . . . .	88
3.5. Mouse $V_\lambda$ families . . . . .	89
3.6. Human $V_\lambda$ families . . . . .	89

4. Gene number . . . . .	90
4.1. Number of mouse V genes . . . . .	90
4.2. Number of human V genes . . . . .	92
5. Chromosome assignment . . . . .	94
6. Gene organization . . . . .	95
6.1. Igh locus organization . . . . .	95
6.2. Igk locus organization . . . . .	99
6.3. Igλ locus organization . . . . .	101
7. Polymorphism . . . . .	101
8. Conclusions . . . . .	105
Acknowledgements . . . . .	105
References . . . . .	106

## Chapter 4

### *Assembly of immunoglobulin variable region gene segments*

M. Reth and L. Leclercq (Cologne, FRG and Paris, France) . . . . .	111
--	-----

1. Introduction . . . . .	111
2. The rearrangement mechanism . . . . .	112
2.1. Joining signals . . . . .	112
2.2. Joining models . . . . .	114
2.3. Control of joining . . . . .	117
3. Order of rearrangement events during B cell development . . . . .	119
3.1. Rearrangements at the IgH locus . . . . .	119
3.2. Rearrangements at the light chain loci . . . . .	126
4. Allelic exclusion of immunoglobulin gene expression . . . . .	129
Acknowledgements . . . . .	131
References . . . . .	131

## Chapter 5

### *Immunoglobulin heavy chain class switching*

U. Krawinkel and A. Radbruch (Cologne, FRG). . . . .	135
--	-----

1. Introduction . . . . .	135
2. Frequency of class switching . . . . .	136
2.1. Avian B cells . . . . .	136
2.2. Mammalian B cells . . . . .	136
3. Switch competence and regulation . . . . .	138
3.1. Switch commitment . . . . .	139
3.2. Isotype commitment . . . . .	139
4. Molecular analysis . . . . .	140
4.1. Early steps . . . . .	140
4.2. Long transcripts . . . . .	140
4.3. Class switch recombination . . . . .	142
4.4. Switch sequences . . . . .	145
4.5. Switch recombination sites . . . . .	146

5. Conclusion .....	147
Reviews .....	149
References .....	149

## Chapter 6

### *Immunoglobulin gene expression*

G.P. Cook, J.O. Mason and M.S. Neuberger (Cambridge, UK) .....	153
1. Introduction .....	153
2. Tumours as models .....	153
3. Patterns of immunoglobulin gene expression during B cell ontogeny .....	154
3.1. Changes in chromatin structure .....	154
3.2. Expression of productively rearranged loci .....	155
3.3. Expression of aberrantly rearranged loci .....	156
3.4. Expression of unrearranged loci .....	156
4. Processes regulating immunoglobulin gene expression .....	157
4.1. Promoter upstream elements .....	157
4.2. Enhancer elements .....	159
4.3. Other promoter elements .....	164
4.4. Transcription termination .....	164
4.5. RNA cleavage/polyadenylation .....	165
4.6. RNA splicing .....	166
4.7. Messenger RNA turnover .....	167
4.8. Translational and posttranslational regulation .....	167
5. Major aspects of cell-type specificity .....	168
5.1. Restricted cell-type specificity of immunoglobulin gene transcription .....	168
5.2. Control of the difference in mRNA abundance between B and plasma cells .....	169
5.3. The relative abundance of membrane and secreted immunoglobulin .....	170
5.4. Co-expression of two immunoglobulin classes .....	173
References .....	173

## Chapter 7

### *The generation and utilization of antibody variable region diversity*

T. Manser (Princeton, NJ, USA) .....	177
1. Introduction .....	177
2. Antigen independent diversity .....	178
2.1. Combinatorial diversity .....	178
2.2. Junctional diversity .....	182
2.3. The multiplicative potential of combinatorial and junctional diversity .....	185
3. Antigen dependent diversity .....	185
3.1. The evidence for somatic mutation: a historical perspective .....	186
3.2. Somatic mutation and the immune response .....	188
3.3. Mechanistic considerations regarding somatic mutation .....	190
3.4. Somatic mutation and clonal selection theory .....	194
3.5. T cells and somatic mutation .....	197
3.6. Antibody diversity and B cell subsets .....	197

4. Summary . . . . .	198
Acknowledgements . . . . .	198
References . . . . .	198

## Chapter 8

### *The immunoglobulin superfamily*

F. Calabi (Cambridge, UK) . . . . .	203
1. Introduction . . . . .	203
2. The immunoglobulin homology unit. . . . .	203
3. The T cell receptor . . . . .	206
3.1. The $\alpha/\beta$ T cell receptor . . . . .	207
3.2. The $\gamma/\delta$ T cell receptor . . . . .	215
3.3. Relationship between the expression of the $\alpha/\beta$ and of the $\gamma/\delta$ T cell receptors . . . . .	218
4. The major histocompatibility complex . . . . .	219
4.1. Overall structure. . . . .	221
4.2. Structure of the variable region . . . . .	221
4.3. Structure of the constant region . . . . .	224
4.4. Genetic basis of diversity . . . . .	224
5. Accessory molecules of the $\alpha/\beta$ T cell receptor (CD4 and CD8) . . . . .	226
6. Other members . . . . .	228
6.1. Thy-1 . . . . .	229
6.2. Poly-immunoglobulin receptor . . . . .	230
6.3. Neuronal cell adhesion molecule (NCAM) . . . . .	230
6.4. Myelin associated glycoprotein/neuronal cytoplasmic protein 3 (MAG/Ncp 3) . . . . .	231
6.5. Viral members . . . . .	231
7. Evolutionary considerations . . . . .	231
Acknowledgements . . . . .	233
References . . . . .	233
Subject index. . . . .	241

# Structure and function of antibodies

DENNIS R. BURTON

*Department of Biochemistry, University of Sheffield, Sheffield S10 2TN, UK*

## 1. Introduction

Antibody molecules are essentially required to carry out two principal roles in immune defence:

(i) to recognise and bind to foreign material (antigen). In molecular terms this generally means binding to structures on the surface of the foreign material (antigenic determinants) which differ from those of the host. Such antigenic determinants are usually expressed in multiple copies on the foreign material, e.g. proteins on a bacterial cell surface. The host needs to be able to recognise a wide variety of different structures – it has been estimated that a human being is capable of producing antibodies against more than  $10^6$  different molecular structures. This is described as antibody diversity.

(ii) to trigger the elimination of foreign material. In molecular terms this involves the binding of certain molecules (effector molecules) to antibody-coated foreign material to trigger complex elimination mechanisms, e.g. the complement system of proteins, phagocytosis by cells such as neutrophils and macrophages. The effector systems are generally triggered only by antibody molecules clustered together as on a foreign cell surface and not by free unliganded antibody. This is crucial considering the high serum concentration of some antibodies.

The requirements imposed on the antibody molecule by the functions (i) and (ii) are in a sense quite opposite. Function (i) requires great antibody diversity. Function (ii) requires commonality, i.e. it is not practical for Nature to devise a different molecular solution for the problem of elimination for each different antibody molecule. In fact the conflicting requirements are elegantly met by the antibody structure represented in Fig. 1. The structure consists of three units. Two of the units are identical and involved in binding to antigen – the Fab (fragment antigen binding) arms of the molecule. These units contain regions of sequence which vary greatly from one antibody to another and confer on a given antibody its unique binding specificity. The existence of two Fab arms greatly enhances the affinity of antibody for antigen in the normal situation where multiple copies of antigenic determinants are presented to the host. The third unit – Fc (fragment crystalline) – is involved in binding to effector molecules. As shown in Fig. 1, the antibody molecule has a four-chain structure consisting of two identical heavy chains spanning Fab and Fc and two identical light chains associated only with Fab.

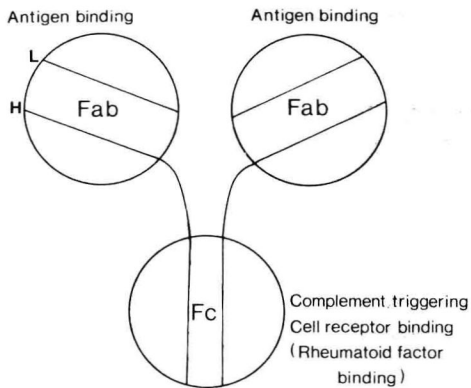


Fig. 1. A schematic representation of antibody structure emphasising the relationship between structure and function. The antibody molecule can be thought of in terms of three structural units. Two Fab arms bind antigen and are therefore crucial for antigen recognition. The third unit (Fc) binds effector molecules triggering antigen elimination. The antibody molecule thus links antigen recognition and antigen elimination. The structure is composed of four chains. Two identical heavy (H) chains span Fab and Fc regions and two identical light (L) chains are associated with Fab alone.

The five classes of antibodies or immunoglobulins termed immunoglobulin G (IgG), IgM, IgA, IgD and IgE differ in their heavy chains termed  $\gamma$ ,  $\mu$ ,  $\alpha$ ,  $\delta$  and  $\epsilon$ , respectively. The differences are most pronounced in the Fc regions of the antibody classes and this leads to the triggering of different effector functions on binding to antigen, e.g. IgM recognition of antigen might lead to complement activation whereas IgE recognition (possibly of the same antigen) might lead to mast cell degranulation and anaphylaxis (increased vascular permeability and smooth muscle contraction). Structural differences also lead to differences in the polymerisation state of the monomer unit shown in Fig. 1. Thus, IgG and IgE are generally monomeric whereas IgM occurs as a pentamer. IgA occurs predominantly as a monomer in serum and as a dimer in seromucous secretions.

The major antibody in the serum is IgG and as this is the best-understood antibody in terms of structure and function we shall consider it shortly. The other antibody classes will then be considered in relation to IgG. First, however, a very brief overview of the structure and function of the different immunoglobulins will be presented [1].

*IgG* is the major antibody class in normal human serum forming about 70% of the total immunoglobulin. It is evenly distributed between intra- and extravascular pools. IgG is a monomeric protein and can be divided into four subclasses in humans. It is the major antibody of secondary immune responses.

*IgM* represents about 10% of total serum immunoglobulin and is largely confined to the intravascular pool. It forms a pentameric structure and is the predominant antibody produced early in an immune response, serving as the first line of defence against bacteraemia. As a membrane-bound molecule on the surface of B lymphocytes it is important as an antigen receptor in mediating the response of these cells to antigenic stimulation.

*IgA* forms about 15–20% of total serum immunoglobulin where it occurs largely as a monomer. In a dimeric complex known as secretory IgA (sIgA) it is the major antibody in seromucous secretions such as saliva, tracheobronchial secretions, colostrum, milk and genitourinary secretions.

*IgD* represents less than 1% of serum immunoglobulin but is widely found on the cell surfaces of B lymphocytes where it probably acts as an antigen receptor analogously to IgM.

*IgE* though a trace immunoglobulin in serum, is found bound through specific receptors on the cell surface of mast cells and basophils in all individuals. It is involved in protection against helminthic parasites but is most commonly associated with atopic allergies.

## 2. Structure of IgG

### 2.1. General considerations

In IgG the Fab arms are linked to the Fc via a region of polypeptide chain known as the hinge. This region tends to be sensitive to proteolytic attack generating the basic units of the molecule as distinct fragments. The discovery of this action by Porter in 1959 [2] provided the first great insight into antibody structure. In 1962 Porter proposed a four-chain structure for the IgG molecule [3]. Since then, chemical and sequence analyses, notably by Edelman [4], have confirmed a four-chain structure consisting of two identical heavy (H) chains of molecular weight approximately 50 000 and two identical light (L) chains of molecular weight approximately 25 000. The molecular weight of IgG is thus typically approximately 150 000. The light chains are solely associated with the Fab arms of the molecule whereas the heavy chains span Fab and Fc parts as shown in Fig. 2. A single disulphide bond connects light and heavy chains and a variable number, depending on IgG subclass (see below), connects the two heavy chains. The latter connection is made in the hinge region of the molecule. Papain cleaves heavy chains to the amino-terminal side of these hinge disulphides producing two Fab and one Fc fragment. Pepsin cleaves to the carboxy-terminal side producing a single  $F(ab')_2$  fragment and smaller fragments of Fc including a carboxy-terminal pFc' fragment.

The light chains exist in two forms known as kappa ( $\kappa$ ) and lambda ( $\lambda$ ); the forms are distinguished by their reaction with specific antisera. In humans,  $\kappa$  chains are somewhat more prevalent than  $\lambda$ , in mice,  $\lambda$  chains are rare [5]. The heavy chains can also be grouped into different subclasses, the number depending upon the species under consideration. In humans there are four subclasses having heavy chains labelled  $\gamma 1$ ,  $\gamma 2$ ,  $\gamma 3$  and  $\gamma 4$  which give rise to the IgG1, IgG2, IgG3 and IgG4 subclasses. In mouse there are again four subclasses denoted IgG1, IgG2a, IgG2b and IgG3. The subclasses – particularly in humans – have very similar primary sequences, the greatest differences being observed in the hinge region. The existence of subclasses is an important feature as they show marked differences in their ability to trigger effector functions. In a single molecule, the two heavy chains are identical as are the two light chains; hybrid molecules are not found.



