VOLUME 22

CANCER SURVEYS

Molecular Mechanisms of the Immune Response

Guest Editors
W F Bodmer
M J Owen

Series Editor E Sidebottom

Published for the
Imperial Cancer Research Fund

CSHL PRESS

COLD SPRING HARBOR LABORATORY PRESS

Molecular Mechanisms of the Immune Response

Guest Editors
W F Bodmer
M J Owen

COLD SPRING HARBOR LABORATORY PRESS 1995

CANCER SURVEYS
Molecular Mechanisms of the Immune Response
Volume 22

Published by Cold Spring Harbor Laboratory Press
Printed in the United States of America
ISBN 0-87969-442-4
ISSN 0261-2429

Cover and book design by Leon Bolognese & Associates, Inc.

All Cold Spring Harbor Laboratory Press publications may be ordered directly from Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803-9729. Phone: Continental US & Canada 1-800-843-4388; all other locations (516) 349-1930. FAX: (516) 349-1946.

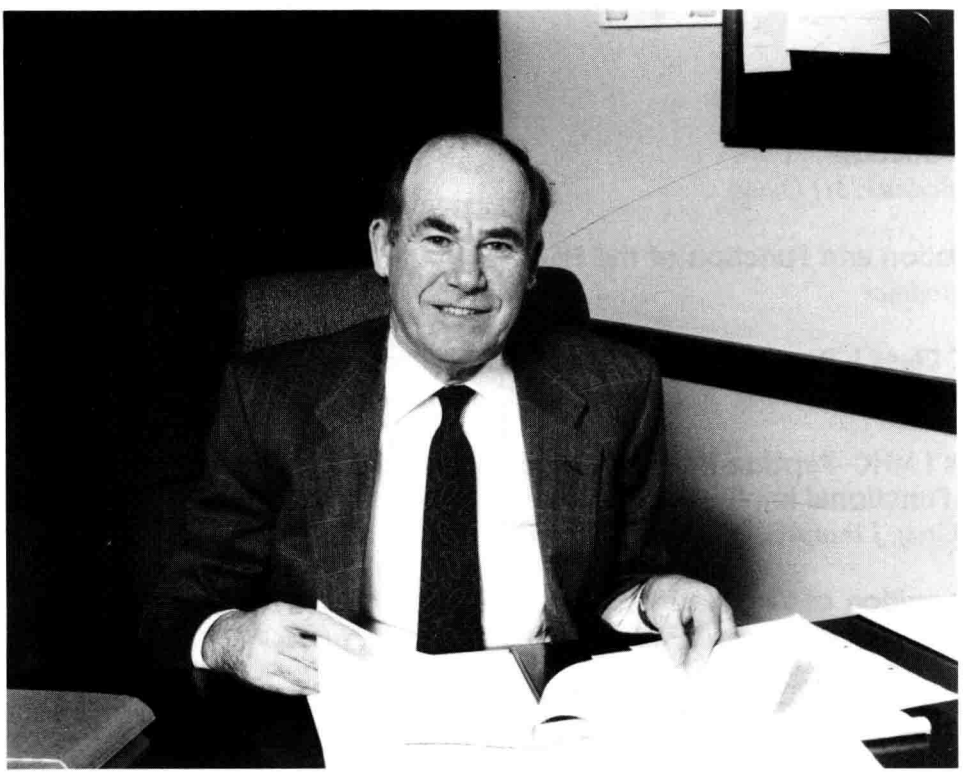
Molecular Mechanisms of the Immune Response

CANCER SURVEYS

Advances and Prospects in Clinical, Epidemiological and Laboratory Oncology

Published for the

Imperial Cancer Research Fund



This volume of Cancer Surveys is dedicated to Michael Crumpton, Ph.D., CBE., FRS, on the occasion of his retirement as Director of Research (Laboratories) at the Imperial Cancer Research Fund, after many years of outstanding and dedicated service.

Contents

Introduction

W F BODMER • M J OWEN
Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX

This volume of *Cancer Surveys* honours Dr Mike Crumpton on the occasion of his retirement as Director of Research (Laboratories) at the Imperial Cancer Research Fund. It is based on a scientific symposium held at the Royal Society on 11 October 1993. The title of the symposium, "The Lymphocyte Cell Surface", reflected Mike's major area of interest for much of his scientific life, and one to which he has made so many contributions.

As an undergraduate at the end of the war, Mike read chemistry at Southampton University (then a college of London University) on a scholarship. At that time, most school-leavers were required to undertake a period of National Service, most university places being reserved for forces' educational training, although an exception was made for students who attained university scholarships. During this time, Mike developed an interest in biochemistry, despite the disapproval of the head of school, Professor NK Adam, who was an eminent surface chemist. Fortunately, Mike was introduced to the Lister Institute by Margary Ish-Campbell and undertook a PhD with Walter Morgan in carbohydrate chemistry, working on the chemical basis of blood group ABO specificity.

After completing his PhD studies, Mike could no longer evade National Service and was seconded first to the David Bruce laboratories on Salisbury Plain. There, Mike received a training in microbiology, working with James Howard. He was then posted to DAL Davies' laboratory at the Microbiological Research Establishment at Porton, working on immune responses to plague.

At the end of his period of National Service, Mike stayed on briefly as a member of the scientific civil service before undertaking a postdoctoral period in the USA. Even at that formative stage of his career, Mike realized the potential power of a rigorous quantitative biochemical approach to immunology and consequently undertook a postdoctoral period in the laboratory of one of the world's leading protein chemists, Christian Anfinsen. There, Mike worked with Dreyer and Streisinger, who had initiated studies on phage lysozyme mutants in experiments designed to establish the co-linearity of the genetic code. Although these experiments met with mixed success, they gave him a grounding in protein chemistry that was to become such a feature of his scientific work.

After his sojourn at the National Institutes of Health, Mike returned to the UK, joining Professor Rodney Porter's department at St Mary's Hospital

Cancer Surveys Volume 22: *Molecular Mechanisms of the Immune Response*

Medical School. Rod Porter exerted a major influence over Mike's career, and indeed over British immunology, until his tragic and untimely death in 1985. There, Mike undertook a study designed to elucidate the molecular basis of antigenicity. He chose to dissect the rabbit antibody response to sperm whale myoglobin, a small globular protein that was in plentiful supply and the X ray structure of which had been determined to 2.2 Å resolution by Kendrew's group in Cambridge. In a meticulous series of experiments, Mike localized some of the major antigenic regions of myoglobin and, in a short paper to the *Journal of Molecular Biology*, proposed the concept that antigenic peptides could adopt an equilibrium of conformations and that an antibody binding site could select a particular conformation, thus displacing the equilibrium. This concept of protein "breathing" was proposed independently by Anfinsen in his Nobel prize winning studies on the denaturation and refolding of RNase S.

By now firmly established as a biochemical immunologist, Mike Crumpton's next move was to the division of biochemistry at the National Institute for Medical Research at Mill Hill. At that time, NIMR was a Mecca for immunological research, with such luminaries as Av Mitchison, John Humphrey and Ite Askonas, and also for protein chemistry. It was thus a natural choice. His time at NIMR was associated with two major areas of research. He initiated an enduring interest in T cell activation, and, together with Michael Green, was first to show using the A23187 calcium ionophore that a rise in intracellular [Ca^{2+}] was sufficient to initiate the complex process of T cell activation. He also began his studies on the purification of lymphocyte plasma membranes which would form the basis of his work on the characterization of the human class I and class II histocompatibility antigens. These latter studies were undertaken as a long collaboration with one of us (WFB) that was a major influence on the remainder of Mike's scientific career. The availability of highly purified plasma membrane from large quantities of B lymphoblastoid cells and the demonstration that the lectin purified from lentils could be used to purify detergent solubilized membrane glycoproteins were key developments in the successful purification of the HLA-A, -B, -C and -DR antigens.

During Mike Crumpton's tenure at NIMR, progress in our understanding of the major histocompatibility complex (MHC) was rapid. The overall structures of class I and class II antigens had been elucidated and the functions of these molecules as guidance systems for T cells had been firmly established. Indeed, the pioneering experiment by Zinkernagel and Doherty that defined MHC restriction had been carried out at the John Curtin School of Medicine in Australia when Mike was on sabbatical leave there in 1973. The elegant experiments of the Crumpton/Bodmer laboratories on the purification of HLA-A, -B, -C and -DR antigens in amounts sufficient for biochemical analysis were carried out with the firm conviction that structural knowledge of a protein would illuminate its function. How right this conviction was, as so dramatically illustrated by the subsequent work, over a decade later, by Wiley's group on the three dimensional structures of class I and class II antigens.

Perhaps Mike's major regret in his scientific life was that he chose not to

follow up the protein chemistry with detailed structural analysis of histocompatibility antigens. This decision was in part due to the philosophy that a scientist should eschew "obvious" experiments and constantly blaze new conceptual trails, leaving others to follow up. The impact of our knowledge of the three dimensional structures of histocompatibility antigens on immunology surely shows how simplistic this philosophy can sometimes be.

Mike's next move was to the Imperial Cancer Research Fund as Deputy Director of Research in 1979. With WFB as Director of Research, the Oxford/NIMR collaboration was united. However, active, scientific collaboration between the two laboratories was less than when they were situated in different cities. A number of reasons contributed to this: recombinant DNA technology had begun to make an impact on MHC genetics, and Mike had developed an interest in the structure and function of surface molecules, such as the CD2 and CD3 antigens involved in T cell activation. Mike also continued studies on the intracellular events associated with T cell activation, concentrating on phosphorylation events and on a family of calcium binding proteins known as the annexins. He continued to make major contributions in these areas despite shouldering an increasingly large burden of administrative responsibility both at the ICRF and in a variety of outside contexts.

There are a number of hallmarks of excellence in a scientific career. A scientist is often judged by the subsequent careers of people who have trained in his or her laboratory. There are a variety of successful scientists around the world who have passed through Mike's lab or who have been influenced by him in collaborative studies. Perhaps the most satisfying hallmark is to be able to have made fundamental contributions to one's field and to see these contributions form a platform for much of the subsequent work in the field. The articles that comprise this volume of *Cancer Surveys* provide ample evidence of the impact of Mike Crumpton's research and underscore the power of molecular approaches to the study of complex biological systems such as the immune response.

Evolution and Function of the HLA Region

WALTER F BODMER
Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX

INTRODUCTION

The HLA, H2 and other major histocompatibility systems were discovered, by analogy with the human red cell bloodgroups, in the search for polymorphic antigens to match for transplantation. Subsequently, they have been shown to have a key functional role in immune response, and their membership of the immunoglobulin supergene family suggested the name "histoglobulins" (Bodmer, 1989).

The first experimental clue to a relationship between the HLA gene products and immunoglobulins came from the fact that the HLA class I associated molecule β_2 microglobulin had a sequence that predicted an immunoglobulin related structure. Subsequently, partial aminoacid sequence analysis of HLA class I gene products, and complete sequences derived from the cloned genes, showed clearly that the membrane proximal domains of HLA class I and class II products were members of the immunoglobulin supergene family. The N-terminal domains, however, did not fit this structure, and it was not until the determination of the HLA-A2 structure by Bjorkman *et al* (1987) that the significance of this domain as a peptide clamp became clear. Thus, the function of the HLA molecules was shown to be to hold peptides derived from endogenous or exogenous proteins by specific proteolytic degradation in a position such that they could be specifically recognized by the T cell receptor. The structural data furthermore confirmed the immunoglobulin homology of the membrane proximal domain, and similar data have now been obtained from Wiley and Strominger and their colleagues for the HLA class II structure (Brown *et al*, 1993).

Cancer Surveys Volume 22: *Molecular Mechanisms of the Immune Response*
 0-87969-442-4/95. $5.00 + .00

The immunoglobulin homology supported speculative suggestions that the evolutionary precursor of the HLA region products may have a role in cellular recognition and cell-cell interaction during differentiation and development (Bodmer, 1972; Gally and Edelman, 1972). This idea is now further supported by the discovery by Edelman and others of a variety of cellular adhesion molecules, which are indeed also members of the immunoglobulin supergene family (Edelman, 1992).

EVOLUTION OF THE HLA REGION

The key event during the very early evolution of the HLA system genes must have been the fusion of the gene for an immunoglobulin domain, possibly derived from a cellular adhesion molecular function, with a gene for a peptide binding domain, perhaps derived from a gene for a heat shock protein or one of its ancestors (Fig. 1) (Bodmer, 1992). This event, which must have taken place at least 5–7 hundred million years ago in early Cambrian or even pre-Cambrian organisms, was followed by a series of duplications and divergence, leading eventually to the present set of genes. The early stages of this process were at a time when there appears to have been a very rapid evolution of these and related functional genes, perhaps shortly after the major initial evolutionary development of multicellular eukaryotic organisms. The first HLA molecule was probably a homodimer of a fused immunoglobulin and peptide binding domain analogous to the HLA class II molecules. Class I molecules could then have arisen in two different ways, either by duplication of the N-terminal peptide binding domain of such an initial primitive histoglobulin molecule or by an independent fusion event between the genes for an immunoglobulin domain and an already duplicated peptide binding domain. In either case, the complete class I molecule needs an additional free immunoglobulin domain, now represented by β_2 microglobulin.

Duplication and divergence of the original histoglobulin gene formed by an immunoglobulin and peptide binding domain fusion will have given rise to the original class II α/β heterodimer. Further divergence of the separate class II A and B genes, followed by duplication, then led to the *DPB, DQB, DRB, DPA, DQA* and *DRA* sets of genes. An analogous series of steps following the formation of the initial gene for the class I precursor will have given rise to the class I genes now seen in the HLA region.

Periods of expansion and contraction of closely related sets of genes, such as *DRB* and *DPA* and *DPB,* during mammalian evolution and no doubt also earlier have resulted in differences in the representation of these various genes in different species including, in particular, humans, and mice (for earlier discussion of these evolutionary ideas, see Travers *et al,* 1984; Bodmer *et al,* 1986; Bodmer, 1992).

The recently described *DMA* and *DMB* genes (Kelly and Trowsdale, 1994) are intriguing in that they have an overall class II like structure but have se-

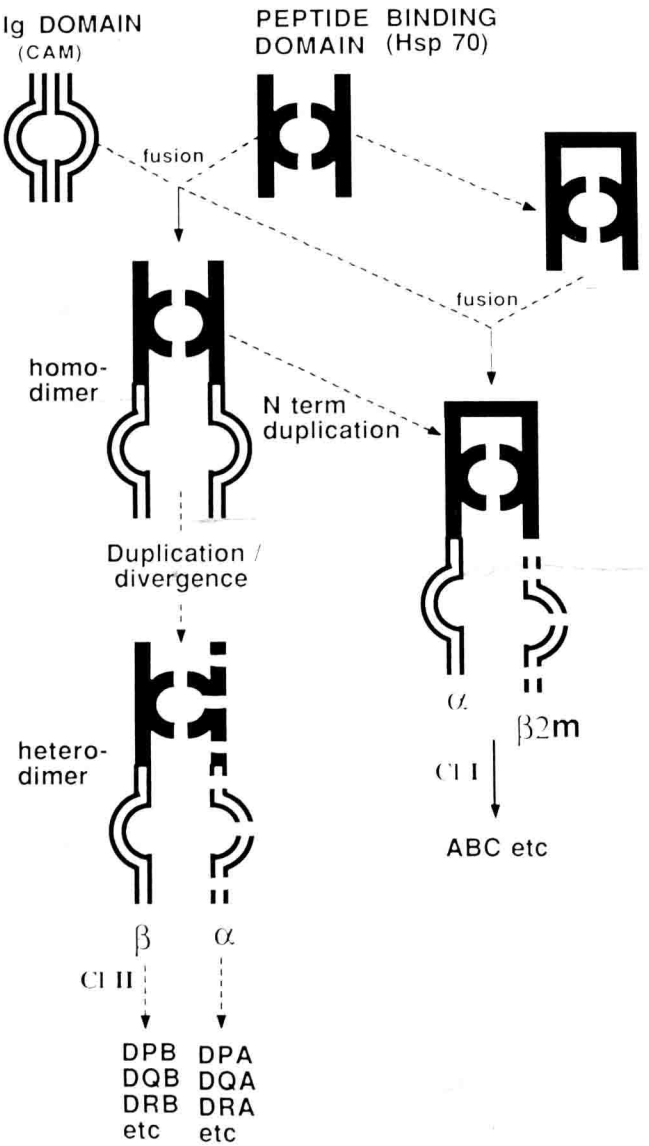


Fig. 1. Scheme for early evolution of the HLA system

quences that show no greater homology with class II than with class I. This suggests a very early separation between this *DM* pair and the others, after the initial evolution of class II A and B genes. These genes are much less polymorphic than the classical HLA class II genes, other than *DRA*, and may have some rather specific function connected with class II based recognition that perhaps does not require interaction with a very wide range of peptides. Kasahara *et al* (1992) have isolated a typical mammalian like class II α chain from a nurse shark (*Ginglymostoma cirratum*), suggesting that the basic divergence of histoglobulins into class I and class II as we now know them most probably predated the appearance of cartilaginous fishes. They find that

HLA *DMA* has the strongest sequence homology with their shark class II *A* gene.

The HLA Region Map: An Evolutionary Interpretation

More and more genes are being found in the HLA region, especially between the class I and class II clusters (see Campbell and Trowsdale, 1993). The region as a whole consists of some 3.5 million base pairs, extending from somewhere between the collagen gene for type IIAII and *DPB2* to the end of the class I region, so far defined by HLA-F. This size estimate, now made precise by genomic analysis using yeast artificial chromosomes, has remained remarkably stable over a period of nearly 25 years. The overall genomic organization in humans and mice is extraordinarily similar, but there is polymorphism in the HLA region for the number of genes and their arrangement around *DRB2* and the C4 complement gene cluster.

There are three major categories of genes in the HLA region:

(a) The HLA class I genes, including several in addition to those for A, B and C, the class II genes, including *DMA, DMB* and *DNA,* and various related pseudogenes.
(b) Other genes with immune related functions. The first to be found were the complement genes for C4, C2 and Bf and then *TNFA* and *TNFB* in the region between the class I and class II clusters. More recently, and strikingly, there has been the discovery through genomic analysis of the class II region of the transporter and proteasome component genes *TAP1* and *TAP2* and *LMP2* and *LMP7*. The discovery of a small cluster of *Hsp70* genes (*HSPA*) in the middle of the HLA region is intriguing, although their apparent specificity for the testis makes a role for them in immune response unlikely. The class II equivalents of the transporters also remain to be clearly identified.
(c) Many other genes with functions unrelated to the immune system. These include, for example, the 21-hydroxylase genes, a gene for a tenascin related molecule, the gene for valyl-t-RNA synthetase and a number of other genes whose function has yet to be determined.

The strong human-mouse homology of the map implies that the overall arrangement is common, at least to mammals, apart from minor perturbations due to local expansion and contraction of duplicated genes. It seems possible that this overall organization reflects the collection of genes that surrounded the primitive precursor of the histoglobulin genes formed from the immunoglobulin and peptide binding domain fusion. The various duplications followed by divergence are then presumed to have occurred within and around this collection of genes, leading to the present organization. Duplicates could easily have been spread locally by, for example, association with an intrachromosomal inversion giving rise to the gap between the class I and the

class II genes. It is notable that there are relatively few, if any, non-immune function genes within the class II region. There is also a suggestion of a further major rearrangement which separated HLA-A and the associated cluster of class I genes from HLA-B and C. It remains to be seen whether the gap between A and B,C will contain as divergent a set of genes with respect to function as does the gap between B,C and the class II region.

There is no reason why the duplicate genes that led to the evolution of the HLA class I and class II genes should have been separated during evolution. The separation would have required specific selection, either for a translocation or for a transposed gene. The duplicate genes have presumably "co-adapted" following their divergence, and favourable combinations have no doubt been held together from time to time by interactive selection, resulting in linkage disequilibrium.

It seems likely, as suggested originally for the complement genes, that some of the immune function genes that are not HLA class I or class II genes could have been "brought in" to the HLA region by selection. Thus, selection favouring a particular combination of one of the "brought in" genes, say a transporter, and a particular pre-existing HLA region gene would favour, through linkage disequilibrium, a transposition that brought the two together. Once this combination had swept through a population, the reorganization would be established and there would need to be no residual selective interaction between the two genes to maintain the new arrangement.

There remain at least two major puzzles. Where are the human equivalents of the mouse *TL* and *QA* genes, which appear to be peptide receptors, although perhaps of more limited specificity (Rotzschke *et al,* 1993) and perhaps specialized to recognizing δ/γ T cells, as proposed by Tonegawa and colleagues (Wu *et al,* 1991)? If these molecules have a special role in recognizing a more specific restricted class of determinants, or do so in selected tissues, then where is the equivalent human function? It is not clear whether the atypical class I genes such as those for HLA E, F and G necessarily fulfil this role.

The second puzzle relates to the origin of the *CD1* genes and their functions. These are clearly class I like genes associated with β_2 microglobulin but which map outside the HLA region (Calabi and Milstein, 1986). Do these also have a function that is tissue specific or more restricted than the other HLA class I genes? And what is their origin? Were they an early transposition of class I region genes to another part of the genome, or were the original class I genes transposed into what we now call the HLA region during a very early stage in its evolution?

HLA CLASS II FUNCTION AND POLYMORPHISM

The basic function of class II molecules is to present peptides, derived from exogenous proteins, to T cell receptors on CD4 positive cells. The class II molecules are expressed constitutively at a high level on specialized antigen presenting cells, especially dendritic cells, as well as on B lymphocytes, where

their function is presumably to present peptides from proteins recognized by the specific antibody on the surface of the B lymphocyte to promote T-B interaction for the stimulation of antibody production. It is clear now that class II molecules are more promiscuous than class I molecules with respect to the range of peptides that can be bound by any particular allelic combination (Brown *et al,* 1993). From a functional point of view, it would appear likely that class II/T cell receptor interactions were associated with the evolution of the antibody production system, and that is likely to have followed the more primitive direct T cell cytotoxic recognition system. This would not be consistent with an evolution of class I from class II molecules but leaves open the question of whether the homodimer postulated in Fig. 1 was in fact the common precursor of both class I and class II molecules.

The fact that most of the DNA sequence differences between polymorphic alleles, whether of class II or class I products, give rise to aminoacid substitutions and occur in restricted positions that are now seen to correspond to the peptide binding regions is strong evidence for the role of selection in generating these polymorphisms (Bodmer *et al,* 1986). This follows from a principle first clearly enunciated by Luca Cavalli-Sforza that, when there is differential variation, this cannot be explained by population structure effects, namely the combination of mutation and random drift (Cavalli-Sforza and Bodmer, 1972). This is because the effects of population structure will apply equally to all regions of a gene, whatever their functional significance, and so the sort of differential variation associated with the peptide binding region cannot be explained by the combination of mutation and random drift.

It is striking that for the DR molecules, essentially only the DR β chain varies, whereas for DQ and DP, both α and β chains vary. There is no doubt that there must be some functional differentiation between these three sets of molecules, although whether it is simply that DR is involved in help and DQ in suppression is, I believe, unlikely. It seems more probable that the different sets of molecules provide different spectra of efficiency of antigen recognition. The diversification of both class II and class I molecules may thus represent a fine tuning of the immune system, providing an increase in the efficiency with which a broad spectrum of peptides can be handled. There may also be an element of tissue specificity, since it seems likely that the induction of class II expression by, for example, γ interferon in a wide variety of tissues provides a mechanism whereby a cell from any tissue may at first act as an antigen presenting cell. The class II *DP* genes may be switched on earlier by γ interferon in at least some cell types, such as epithelial cells, than are the other class II genes (Balkwill *et al,* 1987). Perhaps, therefore, *DP* genes have a specific role in antigen presentation to CD4 T helper cells when they are expressed on "non-professional" antigen presenting cells, such as epithelial cells.

It is important to distinguish the origin of a polymorphic variant, by mutation, recombination or gene conversion, from its propagation into a population, namely its increase in frequency. As already pointed out, mutation and random drift are in general most unlikely to be the forces that have driven HLA region