

THE STRUCTURE OF
MRC OX-2 MEMBRANE GLYCOPROTEIN



Melanie Jane Clark

St Hugh's College

Oxford

A thesis submitted for the degree of D.Phil

Sir William Dunn School of Pathology

Oxford

Trinity Term 1985

CONTENTS

Abstract

Acknowledgements

Publications

Abbreviations

Page number

<u>CHAPTER 1</u>	<u>INTRODUCTION TO THE THESIS</u>	1
1.1	General introduction	1
1.2	The use of monoclonal antibodies in the identification of cell surface molecules	2
1.3	Cell surface interactions in the lymphoid system	5
1.4	Cell surface interactions in the brain	12
1.5	The immunoglobulin superfamily	15
1.6	OX-2 antigen	22
1.7	Strategies for determining the sequence of cell surface molecules	24
1.8	Discussion to Chapter 1	29
<u>CHAPTER 2</u>	<u>MATERIALS AND METHODS</u>	31
2.1	Materials	31
2.2	Methods : PROTEIN CHEMISTRY	35
2.2.1	Purification of OX-2 antigen from rat brain	35
2.2.2	Chemical modification of OX-2 antigen	37
2.2.3	Trypsin digestion of reduced, alkylated and succinylated OX-2	38
2.2.4	Purification of OX-2 tryptic peptides	38
2.2.5	Lowry assay	39
2.2.6	Amino acid analysis	40
2.2.7	Automated amino acid sequence determination	40

2.2.8	Indirect binding assay to quantify OX-2 antigen	41
2.2.9	Polyacrylamide gel electrophoresis in SDS	42
2.3	Methods : RECOMBINANT DNA TECHNIQUES	43
2.3.1	Microbiological techniques	43
2.3.2	Preparation of phage and plasmid DNA	45
2.3.3	Preparation of genomic DNA	46
2.4.4	Digestion of DNA with restriction enzymes.	47
2.3.5	³² P radiolabelling of DNA	48
2.3.6	Analysis of DNA by gel electrophoresis	50
2.3.7	Electroelution of DNA fragments from agarose gels	51
2.3.8	Identification of recombinant clones	52
2.3.9.	Hybridisation of OX-2 oligonucleotide with dried down agarose gels	53
2.3.10	Southern blot analysis of DNA	54
2.3.11	Preparation of RNA	55
2.3.12	Northern blot analysis of RNA	56
2.3.13	S1 nuclease mapping of RNA	58
2.3.14	Sequencing DNA by the method of Maxam and Gilbert	58
2.3.15	Oligonucleotide synthesis	59
2.3.16	Construction of the rat thymocyte cDNA library	60
2.3.17	Construction of the human genomic library λ DA Mbo 1	60
<u>CHAPTER 3 PARTIAL AMINO ACID SEQUENCE OF OX-2 ANTIGEN</u>		62
3.1	Introduction to Chapter 3	62
3.2	Purification of OX-2 antigen from rat brain	63
3.3	Preparation and purification of tryptic peptides	64
3.4	The sequences of OX-2 tryptic peptides	67
3.5	Analysis of the sequence of OX-2 tryptic peptides	69
3.6	Discussion to Chapter 3	69

<u>CHAPTER 4</u>	<u>ISOLATION AND SEQUENCING OF cDNA CLONES FOR OX-2</u>	71
4.1	Introduction to Chapter 4	71
4.2	Design and synthesis of the oligonucleotide probe for OX-2 antigen	71
4.3	The rat thymocyte cDNA library	73
4.4	Screening the thymocyte cDNA library with the OX-2 oligonucleotide	74
4.5	Identification of the oligonucleotide site in postive clones	75
4.6	Strategy for determining the complete OX-2 cDNA sequence	76
4.7	Northern blot analysis of OX-2 thymocyte mRNA	78
4.8	Translation and interpretation of the OX-2 cDNA sequence	79
4.9	Discussion to Chapter 4	82
<u>CHAPTER 5</u>	<u>SEQUENCE HOMOLOGY OF OX-2 ANTIGEN WITH IMMUNOGLOBULIN-RELATED MOLECULES</u>	85
5.1	Introduction to Chapter 5	85
5.2	What to look for in Ig homologies	85
5.3	Computer search for sequence homologies	88
5.4	A model for the structure of OX-2 antigen	88
5.6	Discussion to Chapter 5	89
<u>CHAPTER 6</u>	<u>CLONING THE GENE FOR THE HUMAN HOMOLOGUE OF OX-2</u>	91
6.1	Introduction to Chapter 6	91
6.2	Southern blot analysis of the gene for OX-2	92
6.3	Isolation of a genomic clone for the human homologue of OX-2	94
6.4	Preliminary sequencing of the gene for the human homologue of OX-2	95
6.6	Discussion to Chapter 6	97

<u>CHAPTER 7 CHARACTERISATION OF ABNORMAL OX-2 TRANSCRIPTS IN DIFFERENT</u>		
<u>RAT TISSUES</u>		100
7.1	Introduction to Chapter 7	100
7.2	The abnormal features of thymocyte cDNA clones coincide with intron/exon boundaries	100
7.3	S1 nuclease mapping OX-2 mRNA from different tissues	102
7.4	Northern Blot analysis of mRNA from different tissues.	104
7.5	Discussion to Chapter 7	105
<u>CHAPTER 8 FINAL DISCUSSION</u>		109
8.1	Introduction to the final discussion	109
8.2	The structure of OX-2 at the cell surface	109
8.3	The evolutionary relationship between OX-2 and members of the Ig superfamily	111
8.4	Possible functions of OX-2 and Thy-1	114
8.5	Possible OX-2 counterparts on human thymocytes	116
8.6	Future applications for the human genomic clone	118
8.7	Families of structurally related cell surface molecules	119
8.8	Conclusions	120
<u>REFERENCES</u>		121

ABSTRACT

The structure of MRC OX-2 glycoprotein

Melanie Jane Clark

St Hugh's College, Oxford

Submitted for the degree of D.Phil

Trinity term, 1985

The mouse monoclonal antibody MRC OX-2 recognises a rat cell surface glycoprotein of M_r about 45,000 which is expressed on brain and lymphoid tissues. The OX-2 antigen is of interest because its tissue distribution and biochemical properties are similar to Thy-1, a cell surface antigen which is known to have a structure resembling an immunoglobulin domain. In this thesis, the structure of OX-2 antigen is determined to evaluate whether it shares sequence homology with Thy-1 and immunoglobulins.

OX-2 antigen was purified from rat brain by monoclonal antibody affinity chromatography, and several tryptic peptides were isolated and sequenced. An oligonucleotide probe was designed from the peptide sequence and used to isolate OX-2 cDNA clones from a thymocyte cDNA library. Several cDNA clones were analysed to derive the complete sequence of OX-2 antigen.

The mature protein is 248 amino acids long, comprising of two extracellular domains that show homology with immunoglobulin, followed by a transmembrane segment and a short cytoplasmic region. The N-terminal domain fits best with V-regions while the second domain is like an Ig C-domain. Thus the structure is similar to an Ig light chain or one chain of the T cell receptor.

Using Southern blot analysis, no rearrangement of the OX-2 gene could be detected in rat tissues which express the protein. Homologues of the OX-2 gene were identified in Southern blots of mouse and human DNA, and the human OX-2 gene was isolated from a genomic library. Partial characterisation of the human gene showed that domain I is encoded by a single exon, the sequence of which is highly conserved between rat and human.

Preliminary S1 nuclease mapping of rat OX-2 mRNA indicates that not all transcripts are capable of encoding the normal OX-2 protein. Abnormal transcripts which differ from the normal message at exon splice sites are abundant in RNA from lymphoid tissues, but absent from brain.

To my parents

ACKNOWLEDGEMENTS

I would especially like to thank my supervisor Dr Neil Barclay for his continual help, encouragement and advice throughout this project. I am also grateful to Dr Alan Williams for numerous constructive and interesting discussions. Many thanks also to other members of the MRC CIU for ideas and advice, to Dr Geoff McCaughan for collaboration on studies of the OX-2 gene, to Dr Jean Gagnon and Tony Willis for sequencing the peptides, to Tony Gascoyne for performing the amino acid analyses and to Keith Gould for assistance with the oligonucleotide synthesis.

I am indebted to many members of Professor George Brownlee's group and the MRC Immunochemistry Unit for valuable advice and practical assistance on recombinant DNA techniques; in particular I would like to thank Drs David and Anne Bentley, Dr Emma Whitelaw, and Colin Sharpe. In addition, my thanks are extended to Stan Buckingham and Cathy Lee for photographic services, and to Caroline Griffin whose help with using the word processor and assistance in completing the typing of this thesis is greatly appreciated.

It has been a pleasure to work in the MRC Unit of Cellular Immunology. I thank all members of the Unit for making my three years here so enjoyable, and I am grateful also to Colin Sharpe for his understanding and companionship throughout my time in Oxford.

I thank the MRC for financial support during this study.

PUBLICATIONS

Williams, A.F., Barclay, A.N., Clark, M. and Gagnon, J.
(1984) Cell surface glycoproteins and the origins of
immunity. In Proceedings of the Sigrid Juselius Symposium:
Gene Expression during normal and Malignant Differentiation.
Eds. Andersson, L.C., Gahmberg, C.G. and Ekblom, P.
Academic Press pp. 125-138.

Williams, A.F., Barclay, A.N., Clark, M. and Gagnon, J.
(1985) Annales de l'Institut Pasteur 136C, 283-290. "Cell
surface glycoproteins and the origins of immunity."

- * Clark, M.J., Gagnon, J., Williams, A.F. and Barclay, A.N.
(1985) MRC OX-2 antigen: a lymphoid/neuronal membrane
glycoprotein with a structure like a single immunoglobulin
light chain. EMBO J. 4, 113-118.
- * Barclay, A.N., Clark, M.J. and Thomas, M.L. (1985) Amino
acid sequence of lymphocyte surface antigens. Biochem. Soc.
Trans. 13, 418-419.

Clark, M.J. (1985) MRC OX-2 lymphoid/brain glycoprotein: S1
mapping suggests higher levels of abnormal RNA in the thymus
than in the brain. Biochem. Soc. Trans. 614th Meeting,
Oxford, in press.

- * A copy of this publication is included at the end of this thesis

ABBREVIATIONS

A_{260} , A_{280}	Absorbance at 260 or 280 nm
	Amino acid
APC	Antigen presenting cell
B cell	Lymphocyte derived from the mammalian bursa equivalent
bp	Nucleotide base pair
BSA	Bovine serum albumin
cDNA	Complementary DNA
cpm, cps	Counts per minute, counts per second
DAB	Dulbecco's phosphate buffered saline
dH_2O	Distilled water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DOC	Sodium deoxycholate
DTT	Dithiothreitol
FACS	Fluorescence activated cell sorter
EDTA	Ethylenediamine tetra-acetic acid
EMBL3	Cloning vector used for human genomic library
HPLC	High performance liquid chromatography
IAA	Iodoacetic acid
Ig	Immunoglobulin
IL-2	Interleukin-2
kb	Kilobase
LA	L-broth plus ampicillin
LB	L-broth

ABBREVIATIONS, CONTINUED

L-CA	leucocyte-common antigen
LSGP	Leucocyte sialoglycoprotein
M Ab	Monoclonal antibody
MC 1061	Bacterial host strain for pAT153
β_2 M	β_2 microglobulin
β -ME	β -mercaptoethanol
MHC	Major histocompatibility complex
MOPS	Morpholinopropane sulphonic acid
M_R	Relative molecular mass
mRNA	Messenger RNA
MSNT	1-mesitylenesulphonyl-3-nitro 1,2,4, triazole
NT	Nucleotide
OD _{280 nm}	Absorbance of a sample of 1cm pathlength at wavelength 280nm
O/N	Overnight
OX-2	Antigen recognised by MRC OX-2 antibody
pAT153	Plasmid vector pAT153/Pvu11/8
PBS	Phosphate buffered saline
PIPES	Piperazine-N,N'-bis[2-ethanesulfonic acid]; 1,4-Piperazinediethanesulfonic acid
pfu	Plaque forming unit
PMSF	Phenyl methyl sulphonyl fluoride
Q 359	Bacterial host strain for \EMBL3
RAM	Rabbit F(ab') ₂ anti-mouse F(ab') ₂
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
tRNA	Transfer RNA

ABBREVIATIONS, CONTINUED

SCMC	S-carboxymethyl cysteine
SDS	Sodium dodecyl sulphate
SDS-PAGE	Polyacrylamide gel electrophoresis in SDS
SRBC	Sheep red blood cell
T cell	Thymus derived lymphocyte
TcR	T cell receptor
TEMED	N,N,N',N'-tetramethylethylenediamine
TPCK	Tos-Phe-CH ₂ Cl
Tris	Tris (hydroxyethyl) aminoethane
w/v	Weight/volume
w/w	Weight/weight

CHAPTER 1

INTRODUCTION TO THE THESIS

1.1 GENERAL INTRODUCTION

Cell surface interactions are both diverse and fundamental in living organisms. The interactions may involve contact between cells, or between cells and soluble factors. The molecules at the cell surface are predominantly glycoproteins and glycolipids, together with other lipids. Given the potential complexity of glycoproteins, they are attractive candidates for mediating direct cell-cell interactions. In the past decade, the development of monoclonal antibodies has provided invaluable reagents for the identification of membrane molecules and for the subsequent isolation of these antigens for biochemical studies (Williams et al., 1977; Williams, 1980).

This thesis is concerned with Ig-related molecules. Initially, Ig-like structures were found only on receptors involved with the immune system, but this relationship no longer held when Thy-1 antigen was found to have a structure resembling an Ig domain (Cohen et al., 1981). Thy-1 is unlikely to have an immune function as it is not conserved on lymphoid cells across species, but is expressed on neuronal tissue. The Thy-1 data led Cohen et al. to suggest that the Ig domain has a more basic function than in immunity, and is a recognition unit for cell interactions. Other cell surface molecules, including some not involved in immunology, might exist which share sequence homology with Ig. A possible candidate is the rat cell surface antigen recognised by monoclonal antibody MRC OX-2, which resembles Thy-1 in that it has an unusual brain/lymphoid distribution and similar biochemical properties

(Barclay and Ward, 1982). In this thesis, the sequence of OX-2 antigen is determined to evaluate whether it has an Ig-related structure.

The Introduction describes the types of functions of the cell surface of brain and lymphoid tissues, and the molecules characterised on these tissues in the rat, including OX-2 and Thy-1. Approaches for determining the sequence of cell surface proteins, in particular the application of recombinant DNA methods, will be discussed.

1.2 THE USE OF MONOCLONAL ANTIBODIES IN THE IDENTIFICATION OF CELL SURFACE MOLECULES

Many developmental events involve direct cell-cell contact, yet comparatively little is known about the underlying molecular mechanisms. As there are few simple assays applicable to these intricate and subtle interactions, an alternative approach is to identify and characterise the cell surface molecules which are likely mediators of these events.

The best way to identify surface molecules is to produce specific antibodies against them. In the past, this involved the raising of allo- and xeno-antisera. Immunisations between individuals within a species produces alloantisera, but the disadvantage of this method is that the antisera tend to be of low affinity and titre, and may not recognise the antigen of interest if it is not polymorphic within the species. Immunisations between species produce xenoantisera of higher affinity but a more complex mixture of antibodies, so adsorptions with other cells/tissues were often required to get a highly specific reagent (reviewed by Williams, 1977).

The introduction of monoclonal antibodies has revolutionised the recognition and characterisation of membrane molecules, as monoclonals

are highly specific reagents capable of resolving individual components from the wide range of molecules expressed at the cell surface.

1.2.1 Production and identification of monoclonal antibodies against cell surface molecules

Most membrane proteins will be antigenic in a cross-species immunisation due to genetic divergence in evolution. Xenogenic immunisations generally give rise to anti-protein antibodies (more rarely anti-carbohydrate), and each antibody is extremely specific for a small determinant of 3 to 5 amino acids or saccharide units (Nisonoff et al., 1975).

Thus immunisations with whole cells or crude mixtures of membrane glycoproteins results in a mixture of antibodies against different glycoproteins and different determinants. The complexity is resolved by fusing cells with a myeloma cell line and cloning out the hybrids to give hybridomas, each secreting a single, specific monoclonal antibody (Kohler and Milstein, 1976). The next step is to identify those monoclonals directed against proteins which might be involved in surface interactions. When monoclonals were first introduced, it was thought that surface molecules would be specific for a certain cell lineage and thus the search concentrated on tissue-specific molecules. As it became apparent that few antigens were truly tissue-specific (section 1.3.3) the criteria were extended to molecules with interesting but restricted distributions. The following selection procedures can be used to identify likely candidates for cell interaction functions.

- 1) Screening of cell suspensions by sorting on the FACS to search for monoclonals which label only one type of cell or a subset of previously defined cells. This is particularly suitable for the study of lymphocyte surface antigens as the cells can be isolated easily. An

example of such an antibody is W3/25 which labels only thymocytes and 60% of mature T cells (Williams et al., 1977).

2) Investigation of the pattern of binding of the monoclonal to tissue sections using immunoperoxidase or fluorescence staining. This selection procedure does not require prior knowledge of the potential function of the antigen, and permits the identification of antigens with interesting patterns of expression in different tissues or at different stages of development. For example, MRC OX-2 antibody has an unusual distribution on neuronal and lymphoid tissue in the rat (McMaster and Williams, 1979; Barclay 1981). This technique is suited to the screening of large numbers of hybridomas which result from each fusion.

3) Perturbation of a specific cell function with the monoclonal, as measured in an in vivo or in vitro assay. The lack of functional assays is a severe limitation to this type of screen, although assays for cellular immune functions have enabled important surface molecules in the immune response to be identified, e.g. monoclonals which inhibit cytotoxic T cell mediated killing recognise the antigens LFA-1 and Lyt-2,3 on mouse $T_{\text{CYTOTOXIC}}$ cells (Springer et al., 1982).

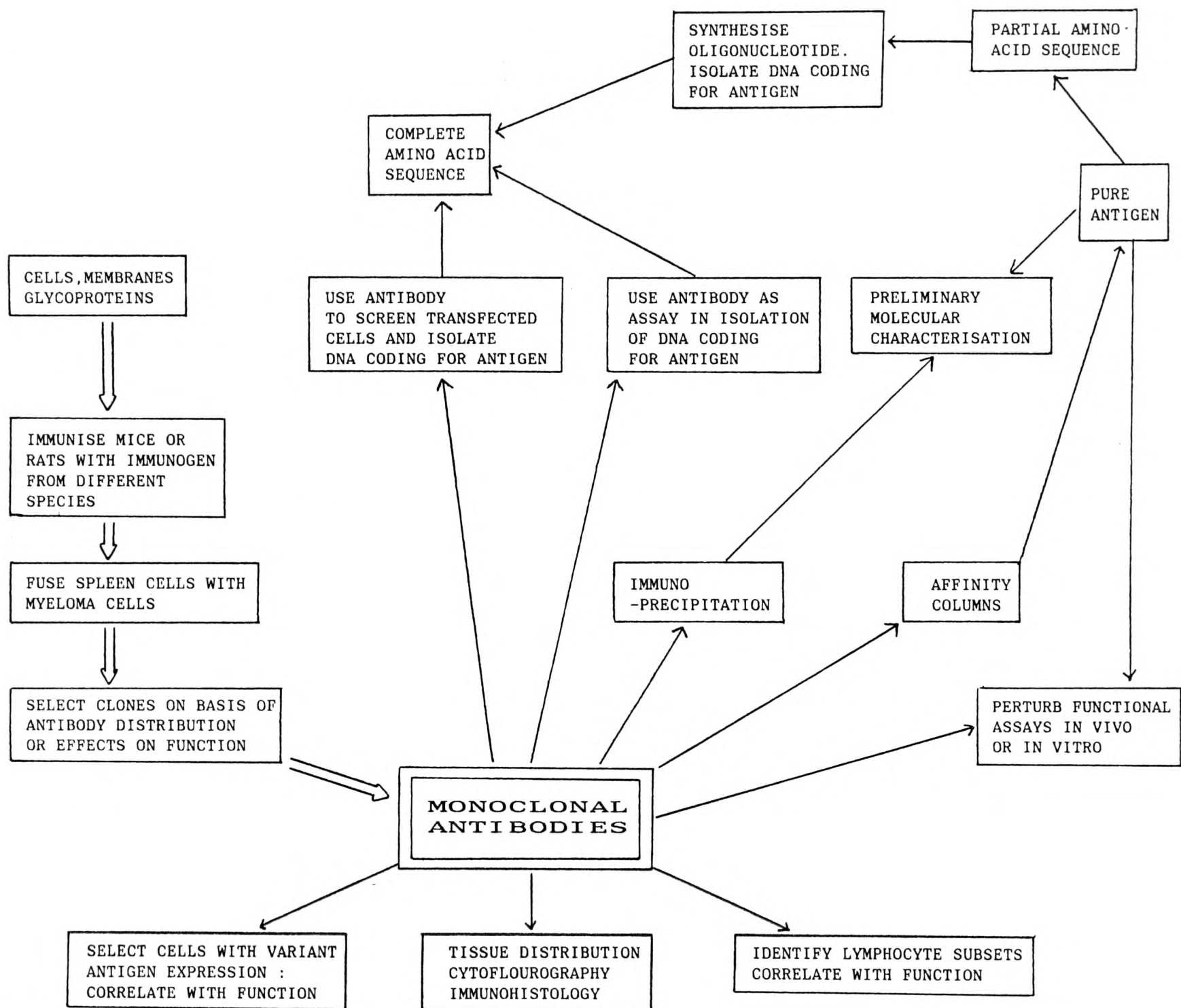
1.2.2 Biochemical characterisation of membrane proteins

Monoclonal antibodies can be used in a variety of ways to define and characterise the antigen of interest, as illustrated for lymphocyte surface molecules in figure 1.1.

For preliminary studies, the glycoprotein can be radiolabelled and its molecular weight and antigenic heterogeneity investigated by immunoprecipitation. On a preparative scale, antibody coupled to Sepharose beads can be used as an affinity column, and combined with a gel filtration step often yields pure antigen in sufficient quantities for biochemical analysis (reviewed by Williams and Barclay, 1985). The

Figure 1.1

MONOCLONAL ANTIBODY APPROACH TO THE STRUCTURE AND FUNCTION OF LYMPHOCYTE SURFACE MOLECULES



amino acid sequence of the antigen can be determined by direct protein sequencing, or by one of the strategies discussed in section 1.7 which involve the isolation and identification of cDNA clones encoding the antigen.

1.2.3 Identifying the protein products of tissue-specific genes with antibodies to chemically synthesised peptides

An alternative strategy for characterising rare tissue-specific molecules for which there are no known functional assays has been used by Sutcliffe et al.(1983). cDNA clones chosen at random from a brain cDNA library were screened in turn on northern blots of RNA from various tissues. The brain-specific clones were sequenced and antisera raised against short synthetic peptides corresponding to the sequence of open reading frames. The sera were used to localise the proteins in fixed brain sections, and identified two previously unknown proteins. The limitation of this technique for isolating proteins involved in cell-cell interactions is that transcripts for cytoplasmic and secreted brain-specific proteins will also be selected, while RNAs which are present in a restricted number of other tissues, e.g. thymus, may be discarded.

1.3 CELL SURFACE INTERACTIONS IN THE LYMPHOID SYSTEM

Underlying the complex events in embryogenesis, organogenesis, and neural development are the primary processes of cell division, migration, differentiation and death. Such processes are also crucial to the functioning of the immune system in mammals, where they continue throughout adult life. The lymphocyte surface membrane is of considerable functional interest in the above interactions, in addition to antigen recognition and effector functions. The lymphocyte is an

especially favourable cell type to study because it can be readily isolated intact and viable, in sufficiently large numbers ($>10^{11}$ thymocytes per 50 rats) to permit purification of its membrane molecules.

1.3.1 Cell interactions of lymphocytes

The immune system of adult mammals comprises of large numbers of small lymphocytes which recirculate between blood and lymph and reside, together with non-lymphoid accessory cells, in discrete lymphoid organs such as spleen and lymph node. The lymphocyte is continuously making contacts with other cells throughout the following stages (General text-Paul, 1984) :

a) Differentiation The two types of lymphocytes, B and T, are derived from pluripotential stem cells whose progeny have followed distinct differentiation pathways in different anatomical sites. The environment in which the cells differentiate has a profound effect on their subsequent function.

B lymphocytes give rise to cells which secrete antibody. In mammals, B cells are produced in hemopoietic tissues along with the other types of blood cells except for T cells. The bone marrow is the site of B cell differentiation in the adult, and the first identifiable member of the B cell series is the pre-B cell which expresses cytoplasmic μ chains but lacks membrane Ig. These develop into mature cells which express membrane IgM or IgD and can react in response to foreign antigens.

T lymphocytes develop in the thymus and are subsequently responsible for cell mediated immunity. The thymus consists of two morphologically distinct areas, the medulla and the cortex, which contain a dense 3-dimensional network of epithelial cells joined by

desmosomes. There are thymocyte-epithelial cell contacts throughout the thymus, and thymocyte-dendritic cell contacts within the medulla. Events occurring within the thymus determine the specificity of the peripheral T cells of an individual. During T cell differentiation, antigen-specific receptors are expressed, and a selection process occurs which selects for thymocytes which co-recognise self MHC products together with foreign antigen (the phenomenon of 'histocompatibility restriction'). This probably results from the recognition of self MHC molecules expressed on the surface of the epithelial cells in the thymus. Thymocytes capable of reacting with autologous antigen are eliminated, leading to T cell tolerance to self antigens. Also during maturation in the thymus, subpopulations of T cells are formed with different effector functions, namely T_{HELPER} and T_{CYTOTOXIC} cells, and surface receptors responsible for the subsequent 'homing' of lymphocytes are induced. It is still not fully understood how the thymus mediates these T cell differentiation events. A great deal of cell death is associated with this proliferation and differentiation.

b) Recirculation Mature resting B and T cells recirculate between blood and lymph (Gowans and Knight, 1964), a process which maximises the opportunity for lymphocytes to encounter foreign antigen. Recirculation involves specific recognition of the specialised high walled endothelium in lymph nodes, movement between endothelial cells, and migration to positions in the lymph node specific for B and T cells (Ford, 1975). All these processes must be triggered by cell surface contacts.

c) Interaction with antigen Both B and T cells carry specific receptors on their surface to recognise foreign antigen and the response to antigen requires interactions between B and T cells and antigen presenting cells (APC). B lymphocytes interact directly with antigen

through their surface receptors which are membrane bound immunoglobulin (Warner, 1974). B cell responses to most protein antigens are dependent on help from the T_{HELPER} subset of T lymphocytes and this may be delivered in two ways. The T_{HELPER} cell can either interact directly with the B cell, recognising determinants on antigenic molecules already bound on the B cell and in association with a MHC Class II molecule (Klein et al., 1981). The initial activation of the T_{HELPER} cell depends on its co-recognition of antigen and MHC Class II on the surface of an APC. The T_{HELPER} subset expresses the CD4 antigen on its surface, and it is thought that CD4 interacts with Class II on target cells (Watts et al., 1984). These interactions result in cell division and differentiation into large activated blast cells which undertake further migration and interaction with antigen.

The T cell receptor is a dimer of two transmembrane proteins which resemble membrane bound Ig light chains (Hedrick et al 1984b; Chien et al., 1984b) - see section 1.5.2.3.

T cells also act as effectors of the immune response by lysing target cells bearing antigens for which they are specific. The $T_{\text{CYTOTOXIC}}$ subset of T cells recognise antigen (e.g.virus) in association with MHC Class I, and lyse their target cells by a complex process (Klein et al., 1981). The $T_{\text{CYTOTOXIC}}$ subset expresses the CD8 antigen (reviewed by Williams, 1985) which is believed to bias cytotoxic cells towards interaction with foreign antigen in association with MHC Class I (MacDonald et al., 1982).

1.3.2 Surface molecules of lymphocytes

It seems reasonable to assume that the majority of the major surface molecules of lymphocytes are involved in the recognition phenomena discussed above. Resting lymphocytes have low levels of

metabolism, so metabolite transport molecules and ion pumps are likely to be minor components of the membrane. Cells stimulated to divide and differentiate after contact with antigen, however, presumably need to express additional surface molecules involved with growth regulation which reflect their heightened metabolic activity rather than their role in the immune response.

Table 1.1 lists the major surface antigens which have been identified to date on rat lymphoid cells, together with their molecular characteristics and distribution pattern on lymphoid and non-lymphoid cells. All the antigens in table 1.1 are glycoproteins or multimers of protein and glycoprotein chains. The extent of glycosylation varies enormously between 6% for MHC class I and 60% for LSGP. As the three major surface glycoproteins of thymocytes (Thy-1, L-CA and LSGP) are all extensively glycosylated, then carbohydrate structures must be an important feature of the thymocyte surface.

Functions are known for only some of the molecules listed in table 1.1, and these are summarised later in table 1.6. Immunoglobulins, the T cell receptor, and MHC class I and II play a direct role in antigen recognition, while OX-8 (CD8) and W3/25 (CD4) perform as yet poorly understood accessory functions, guiding the interaction between T cells and targets expressing MHC class I and II (see previous section, 1.3.1). The receptors for transferrin and IL-2, which are expressed on activation, are in the category of molecules required for the increased metabolic state of the cell. The functions of the other molecules in table 1.1 are not known.

Table 1.1

PROPERTIES OF MEMBRANE ANTIGENS ON LYMPHOID CELLS IN THE RAT

	Name	Apparent M _r on SDS-PAGE x 10 ⁻³	Site number	% carbo- hydrate	Cell distribution	Reference
MAJOR GLYCOPROTEINS	Thy-1	gp25	10 ⁶	33%	Neuronal, fibroblasts, connective tissue, thymocytes but not mature T cells	(1) (2)
	LSGP (W3/13)	gp95	10 ⁵	60%	T, thymocytes, polymorphs, stem cells, plasma cells, brain	(3)
	LC-A	gp 170- 220	7 x 10 ⁴ (thymocytes)	25%	Leucocytes only	(4)
	IgM	2 x gp 78 (H) 2 x p 23 (L)	7 x 10 ⁴	13%	B cells only	(5) (6)
MHC ANTIGENS	Class I (RT1A)	gp 45 p12 (β ₂ M)	2 x 10 ⁵	6%	all cells except those of the central nervous system	(7) (8) (9)
	Class II (RT1B)	gp 33 (α) gp 28 (β)	2 x 10 ⁵	8% (α) 19% (β)	B cells, dendritic, macrophage, epithelial.	(10) (11) (12)
MINOR GLYCOPROTEINS	W3/25 (CD4)	gp 55-62	3 x 10 ⁴ (T _H)	n.k.	Thymocytes, T _{HELPER} , macrophages	(13)
	OX-8 (CD8)	gp 30-35 gp 32-43	1 x 10 ⁵ (T _{CYT})	n.k.	Thymocytes, T _{CYTOTOXIC} /SUPPRESSOR natural killer cells.	(13)
	OX-2	gp 47	2 x 10 ⁴	33%	Thymocytes but not mature T, follicular dendritic cells, endothelia, brain.	(14)
	T cell receptor	n.k. in rat (gp 43) (gp 40) in other species	n.k.	n.k.	Thymocytes, T, in other species No Abs against rat TcR known, but identified at the DNA level with a cDNA probe for the mouse β chain.	(15)
	OX-19 (CD2)	gp 69	2 x 10 ⁴ (thymocyte)	n.k.	Thymocytes, T	(13)
EXPRESSED ON ACTIVATED LYMPHOCYTES	OX-39 = IL-2 receptor	gp 47-53	7 x 10 ⁴	n.k.	Activated T.	(16) (17)
	OX-26 = transferrin receptor	2 x gp 95	n.k.	n.k.	Dividing cells, and some non- dividing normal cells, e.g. brain endothelium.	(18) (19)

Key: gp = glycoprotein
n.k. = not known

- References: (1) Williams et al., 1976
(2) Barclay and Hyden, 1978
(3) Brown et al., 1981
(4) Standring et al., 1978
(5) Jensenius and Williams, 1974b
(6) Kehry et al., 1980
(7) Hart and Fabre, 1981
(8) Fukumoto et al., 1982
(9) Parham et al., 1977
(10) McMaster and Williams, 1979
(11) Bayer and Reske, 1983
(12) Williams and Barclay, 1985
(13) Mason et al., 1983
(14) Barclay and Ward, 1982
(15) K. Morgan, unpublished; M. Morris, unpublished
(16) Osawa and Diamanstein, 1983
(17) Leonard et al., 1982
(18) Jefferies et al., 1985a
(19) Jefferies et al., 1984

1.3.3 Distribution of lymphocyte surface molecules on other tissues

It might be expected that certain molecules would be specific to a particular cell lineage, mediating specific functions for this cell type, whilst others, such as ion pumps, would be widely distributed as they are required by all cell types. However, it is surprising that the majority of lymphocyte surface antigens in table 1.1 are expressed in an unpredictable pattern on a variety of non-lymphoid tissues. Only immunoglobulin of B cells, and the T cell receptor, are truly cell-type specific. Even surface antigens commonly used to define T lymphocytes are expressed on other cells, for example the W3/25 antigen is present on macrophages (Jefferies et al., 1985b). Yet stranger patterns of expression are seen for other molecules in table 1.1. Amongst these are the MHC Class I products and the OX-2 and Thy-1 antigens. Although cross-reactions of monoclonal antibodies between unrelated proteins have been noted, e.g. one anti-Thy-1 antibody recognises vimentin (Dulbecco et al., 1981), it is unlikely that the patterns in table 1.1 result from serological cross-reactions. For example, Class II (Klareskog et al., 1980) and OX-2 (Barclay and Ward, 1982) have been isolated from different tissues and found to have very similar biochemical properties. The sequence of the rat Thy-1 antigen from brain (Campbell et al., 1981) and thymocyte (Moriuchi et al., 1983) is identical (despite controversy over the mode of attachment to the membrane - see Chapter 4.9) and human fibroblast and brain Thy-1 show biochemical similarities (Cotmore et al., 1980).

A recurrent pattern appears to be shared expression on lymphoid cells and brain tissue, seen for OX-2, Thy-1 and W3/13 antigens in the rat, and also noted for antigens in other species as listed in table

Table 1.2

SURFACE ANTIGENS EXPRESSED ON LYMPHOID AND BRAIN TISSUE

Antigen	Comment	Reference
<u>A ANTIGENS WHICH HAVE BEEN BIOCHEMICALLY CHARACTERISED</u>		
Thy-1	Abundant brain/lymphoid antigen in rodents. Variable tissue expression in other species but always conserved in neuronal tissue.	(1)
OX-2	Rat brain/lymphoid/epithelia. Homologues not yet characterised in other species.	(2)
F10-44-2	Human brain and leucocytes. Biochemically distinct from W3/13 on rats.	(3) (4)
Thy-2	Mouse thymocyte and brain, bone marrow, spleen	(5)
44 H4 } 44 D7 } 44 D10 }	Recognise human non-T non-B acute lymphocytic leukemia line. In addition : 44 D10 recognises white matter of brain 44 H4 recognises white and grey matter 44 D7 " " " " "	(6)
<u>B ANTIGENS IDENTIFIED SEROLOGICALLY BUT NOT CHARACTERISED BIOCHEMICALLY</u>		
Leu-7	recognises a carbohydrate epitope on myelin- associated glycoprotein and cross-reacts with carbohydrate on human large granular lymphocytes and chick and rat embryonic nervous system.	(7)
LSGP	M Ab W3/13 recognises rat leucocytes, stem cells and brain but antigen in brain not yet shown to be biochemically related to lymphoid form.	(8)

References

- | | |
|--------------------------------|-----------------------------|
| (1) reviewed by Williams, 1982 | (5) Siodak & Nowinski, 1981 |
| (2) Barclay and Ward, 1982 | (6) Quakenbush et al., 1985 |
| (3) Dalchau et al., 1980 | (7) McGarry et al., 1982 |
| (4) McKenzie et al., 1982 | (8) Brown et al., 1981 |

1.2. Another fairly common relationship is the leucocyte/epithelial or endothelial expression observed for MHC Class II, Thy-1, and OX-2. This is also seen for the antigens recognised by monoclonal antibodies OX-43 (Robinson et al., 1985) and OX-45 (Arvieux et al., 1985) in the rat, and OKM5 in humans (Knowles et al., 1984). It is not known why these shared expression patterns should be common while other combinations, such as lymphoid/liver, are not encountered.

1.3.4 Expression of lymphocyte surface antigens across species

The majority of the rat lymphoid surface antigens in table 1.1 have also been identified on mouse and human cells (Williams and Barclay, 1985). A few molecules, for example the OX-2 antigen, have not been characterised in other species. Conversely, counterparts of a few structures on lymphoid cells of other species have not yet been found in the rat, for example the T3 complex which is associated with the T cell receptor in human thymocytes and T cells (Meuer et al., 1983).

Some of the surface molecules listed in table 1.1 show surprising variations in tissue distribution between species. A striking example is the Thy-1 antigen, and the pattern of expression of this molecule in rat, mouse, dog and human tissues is summarised in table 1.3. Thy-1 is a major component of all neural tissues studied so far, and has been identified in chickens (Rostas et al., 1983) and frog (Mansour and Cooper, 1984) in addition to the four species listed in table 1.3. Its expression on lymphoid tissues, however, is not conserved across species. It is the most abundant cell surface glycoprotein of rat and mouse thymocytes, but is expressed in lower amounts on dog thymocytes, and not at all on human thymocytes. Thy-1 is present on rat stem cells and absent from rat mature T cells, whereas the converse expression is seen in mice (reviewed by Campbell et al., 1981).

Table 1.3

PATTERNS OF EXPRESSION OF THY-1 ANTIGEN IN SOME MAMMALIAN SPECIES

Tissue or cell type	SPECIES			
	Mouse	Rat	Dog	Human
Brain (neuronal)	++++	++++	++++	++++
Fibroblasts	++	++	n.k.	++
Thymocytes	++++	++++	+	-
T lymphocytes	+	-	+	-
Bone marrow lymphoid cells	-	++	+	-
Haemopoetic stem cells	-	++	n.k.	n.k.

n.k. = not known

The number of crosses gives an approximate indication of the amount of antigen on different cell types.

The table is taken from Williams (1982).

Other rat surface antigens show an inexplicable but less dramatic variation in distribution pattern. Class II antigens, for example, are present on kidney endothelium of human but not of rat (Hart et al., 1981). Also the major human Class II homologues are of the rodent E type, yet both A and E types are also found in the mouse and some mouse strains do not express the E type at all (Jones et al., 1981).

1.4 CELL SURFACE INTERACTIONS IN THE BRAIN

The mammalian nervous system consists of an intricate network of multiple intercellular connections between heterogeneous cell types. The brain is less easy to study at the molecular level than the immune system because of its great complexity, the inability to separate viable mature cells, and the lack of suitable functional assays.

1.4.1 The complexity of cell surface interactions in the brain

The vertebrate nervous system is comprised of nerve cells (neurons) together with specialised neuroglial cells. Neurons are the active elements of the nervous system. They have extended cytoplasmic processes (the axons or dendrites) and directly initiate and conduct nerve impulses. The axons of one neuron touch the cell membrane of other neurons at synapses, and these contacts establish a multitude of complex pathways in the nervous system. Neuroglia interact closely with neurons but do not conduct impulses. They are subdivided into three classes (astrocytes, oligodendroglia and microglia) of which oligodendrocytes are the counterparts of the peripheral nervous system Schwann cells, as they surround the nerve cell and synthesise the myelin sheath. The role of astrocytes is less well defined, and it is thought that they might transport substances to neurons. Microglia are mobile and phagocytic. (reviewed by Rhodin, 1974; Romanes, 1981).

Direct cell-cell contact is thought to be essential for various stages of development in the brain. During embryogenesis, there is an initial separation of neuronal from non-neuronal precursors in the ectodermal layer, and subsequent migration of neurons to their ultimate destinations. Cell surface interactions between immature neurons and neighbouring cells must be crucial in the selection of the migratory pathway and in the identification of the correct target cell. Contact with the cell initiates a new series of interactions which involve competition of outgrowing fibres and finally differentiation and formation of the synapse.

In the adult brain, there is considerable diversity in the types of neurons, as recognised by the following criteria :

- a) Type of neurotransmitter(s) released.
- b) Type of neurotransmitter(s) received.
- c) Type of connections made with target cell.
- d) Variation in signalling capacity.
- e) Structure and size.
- f) Variation in amount and type of cell surface markers.

Direct analysis of the mRNA population in the rodent brain using methods of renaturation kinetics has indicated that there are about 10^5 different transcripts specific for the nervous system (Chickaraishi, 1979). The abundance of some of these messages is so low that they can only be expressed on subsets of neurons, but it has not been established whether they all encode proteins.

1.4.2 The use of monoclonals to study cell surface interactions in the brain

Monoclonal antibodies have proved to be valuable reagents for many aspects of neurobiological research (reviewed by Brockes, 1984;

Reichardt, 1984). The strategy for investigating the molecular basis for neuron heterogeneity has been to raise monoclonals against whole brain or regions, then screen on sections of neural tissue to identify subsets of neurons which are otherwise morphologically indistinguishable. Some examples of these are listed in table 1.4 A, which also includes data from invertebrate nervous systems.

In several cases, antibody blocking experiments have suggested functions for antigens (table 1.4B). For example, the N-CAM molecule has been identified as a key component in cell adhesion (Rutishauser, 1983), and antibodies against the L1 antigen block neuronal migration during development as well as in cell adhesion (Schachner et al., 1983).

The purification of neuronal surface antigens, and the determination of their structure will undoubtedly contribute to understanding of the neuron at the molecular level. Thus far, only a few neuronal proteins have been biochemically characterised and examples are listed in table 1.5. At the outset of this project, only two of these membrane proteins had been sequenced. The first is the acetylcholine receptor, which had been well characterised prior to the introduction of monoclonals because it is present in large quantities in the electric organs of Electrophorus species, and snake α neurotoxins bind specifically to the acetylcholine binding site enabling the molecule to be identified and isolated. The second molecule for which sequence data is known is Thy-1, which has already been discussed with respect to its tissue distribution in the rat and other species (sections 1.3.3 and 1.3.4). Thy-1 is an abundant surface molecule of neuronal tissue of all species studied so far, and on a weight-for-weight basis, there is approximately the same amount on brain as on thymus. (Williams et al.,

Table 1.4

APPLICATIONS OF MONOCLONAL ANTIBODIES TO THE STUDY OF NEURONAL CELL SURFACE INTERACTIONS

A ANTIBODIES SPECIFIC FOR NEURAL CELL TYPES AND SUBCLASSES

Type of information	M Ab	Raised against	Notes on distribution	Reference
DISTINGUISH SUBCLASSES OF NEURONE WHICH MAY BE FUNCTIONALLY RELATED	various	leech CNS	Distinguish subsets of neurons in each ganglion. Some subsets functionally related e.g. sensory neurons	(1)
	CAT 301	cat spinal cord	Reacts with limited classes of neurons in different areas of cat and monkey CNS. In visual cortex, labels rows of cells which might be associated with optical functions.	(2) (3)
DISTINGUISH CELL POSITION	TOP	chick retina	Recognises surface protein distributed in 35-fold concentration gradient across cells of avian retina. Gradient increases from ventro-anterior to dorso-posterior margin and is established early in retinal development.	(4)
AXON PATHFINDING	Lan3-2	leech	Recognises a subset of neuronal cell bodies that are distantly located in the ganglia, but whose axons selectively fasciculate.	(5)

B ANTIBODIES USED IN FUNCTIONAL STUDIES

Function	M Ab against	Functional studies	Reference
CELL ADHESION	N-CAM (table 1.5)	1) Isolate mg quantities for structural and functional studies 2) Fab fragment of anti N-CAM antibody in <u>in vitro</u> assay : a) Inhibits cell aggregation (polyclonal sera blocks completely, monoclonal blocks partially. b) Inhibits development of chick retina. c) Decreases size of fascicles of outgrowing nerves. d) Disrupts synaptogenesis.	(6)
CELL ADHESION AND MIGRATION	L1 (table 1.5)	Fab fragment of anti L1 antibody : a) Inhibits cell aggregation (polyclonal sera block but monoclonal do not) b) Inhibit migration of granule cells in cultured cerebellar explants.	(7)
NEUROTRANSMITTER RECEPTOR ACTION	Acetyl choline receptor (AChR) (table 1.5)	1) Competition between different M Abs distinguishes a predominant immunogenic region on the α subunit.	(8)
		2) Fab fragments used in high resolution electron microscopy to map positions of the subunits.	(9)
		3) M Ab used in inhibition binding studies with serum from patients with myasthenia gravis shows that the autoimmune disease is stimulated by the human AChR and not by a cross-reacting antigen.	(10)
NEURAL GROWTH	Thy-1	M Ab coated on glass enhance the growth of rat retinal ganglion cells.	(11)

REFERENCES :

- | | |
|-----------------------------|-----------------------------------|
| (1) Zipser et al., 1983 | (6) reviewed by Rutishauser, 1983 |
| (2) Hockfield & McKay, 1983 | (7) Schachner et al., 1983 |
| (3) Hendry et al., 1984 | (8) Tzartos et al., 1981 |
| (4) Trisler et al., 1981 | (9) Fairclough et al., 1983 |
| (5) McKay et al., 1983 | (10) Tzartos et al., 1982 |
| | (11) Leifer et al., 1984 |

Table 1.5

EXAMPLES OF BIOCHEMICALLY CHARACTERISED NEURONAL CELL SURFACE PROTEINS

Protein	Molecular properties	Comments	Reference
<u>A AMINO ACID SEQUENCE KNOWN</u>			
Acetyl Choline Receptor	pentameric complex $\alpha_2\beta\gamma\delta$	4 different but homologous subunits form transmembrane channel which is opened by binding ACh. Subunits can be reconstituted to form functional receptor in artificial vesicles. cDNA clones isolated for each subunit.	(1)
Thy-1	gp 25K	Abundant neuronal protein of all species studied so far. (Discussed in section 1.5.2)	(2) (3)
<u>B AMINO ACID SEQUENCE UNKNOWN</u>			
N-CAM	gp 120K-200K	Different mw forms with differing amounts of sialic acid selectively expressed in different regions of brain during development. Involved in cell adhesion (Table 1.4). N-CAM gene cloned but sequence not yet reported.	(4) (5)
L1	gp 200K gp 140K	Expressed on most but not all neurons in mouse embryo. Involved with cell adhesion (Table 1.4) but biochemically distinct from N-CAM.	(6)
F10-44-2	gp 90K	Human cerebral white matter and lymphoid tissues (Table 1.2).	(7)
OX-2	gp 41K	Rat neurons but not glia. Also expressed on lymphoid and other tissues (Discussed in section 1.6).	(8) (9)
F3-87-8	gp 130K gp 100K	Human and rat CNS. Absent from peripheral nerve and other tissues.	(10)

gp = glycoprotein

References:

- | | |
|--|----------------------------|
| (1) Reviewed by Conti-Tronconi and Raferty, 1982 | |
| (2) Campbell et al., 1981 | |
| (3) Williams and Gagnon, 1982 | (7) McKenzie et al., 1982 |
| (4) Edelman et al., 1983 | (8) Barclay and Ward, 1982 |
| (5) Goridis et al., 1985 | (9) Webb and Barclay, 1984 |
| (6) Schachner et al., 1983 | (10) Lakin et al., 1983 |

1976). The molecule has sequence homology with immunoglobulin (Campbell et al., 1981) and this is discussed in section 1.5 below.

1.5 THE IMMUNOGLOBULIN SUPERFAMILY

Studies of serine proteases, globins and immunoglobulins have shown that molecules with similar functions often have homologous protein sequences (Dayhoff et al., 1972) and gene organisation (Blake, 1983). These protein families are believed to have evolved from primitive genes by duplication and divergence. Hence one aspect of structural studies of membrane molecules is the discovery of homologies which indicate an evolutionary relationship. An example is the similarities between many surface antigens and immunoglobulins which have led to the concept of the 'Immunoglobulin Superfamily'.

1.5.1 Characteristics of the Ig fold

Immunoglobulin molecules comprise of domains of about 100 amino acids which share sequence homology. Ig light (L) chains consist of two such domains, and heavy (H) chains of four or five domains depending on the class (figure 1.2A). X-ray crystallography confirmed that both the variable (V) and constant (C) domains of heavy and light chains have similar secondary and tertiary structures, and established a characteristic Ig folding pattern, reviewed by Amzel and Poljak, (1979), and illustrated in figures 1.2B and 1.2C. The key features of the Ig domain are as follows:

- 1) The size of the homology unit is about 100 amino acids.
- 2) A disulphide bond always occurs between β strands B and F as indicated in figure 1.2C .

Figure 1.2

THE FOLDING PATTERN OF THE Ig DOMAIN

A : Schematic diagram of IgM at the cell surface. Intrachain disulphide bonds are shown by the S-S symbols and interchain disulphide bonds by dotted lines. N-linked carbohydrate structures are shown by .

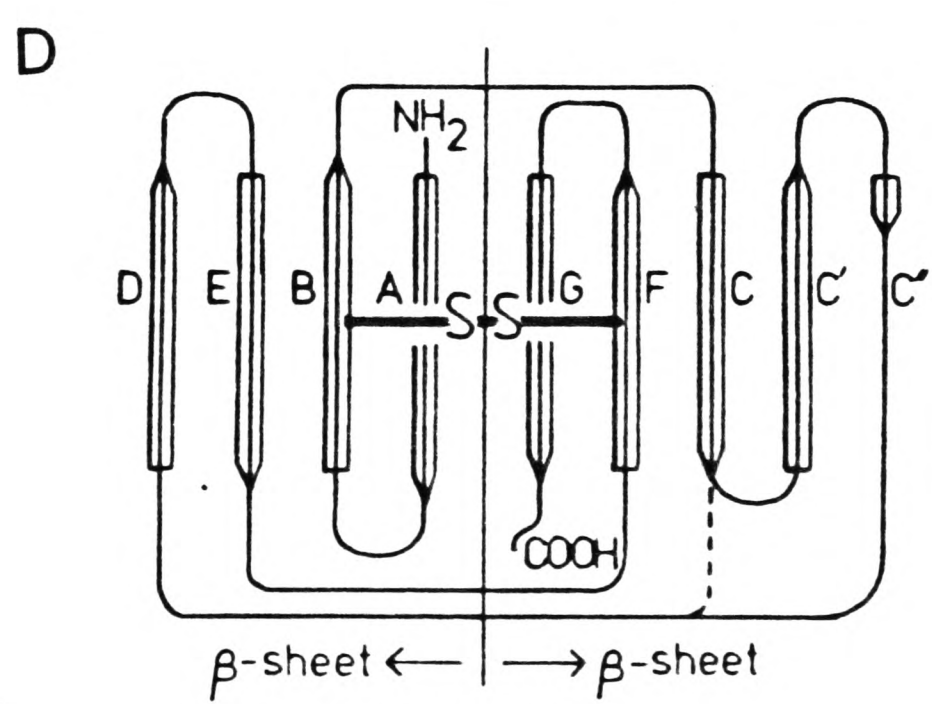
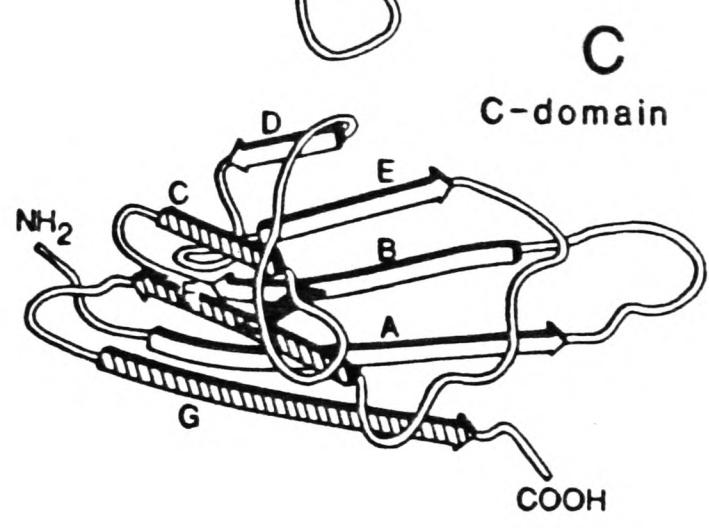
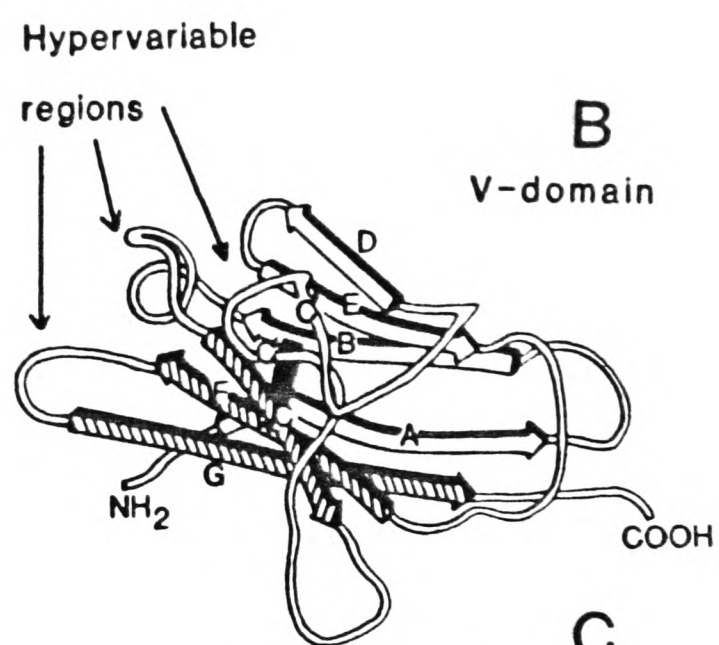
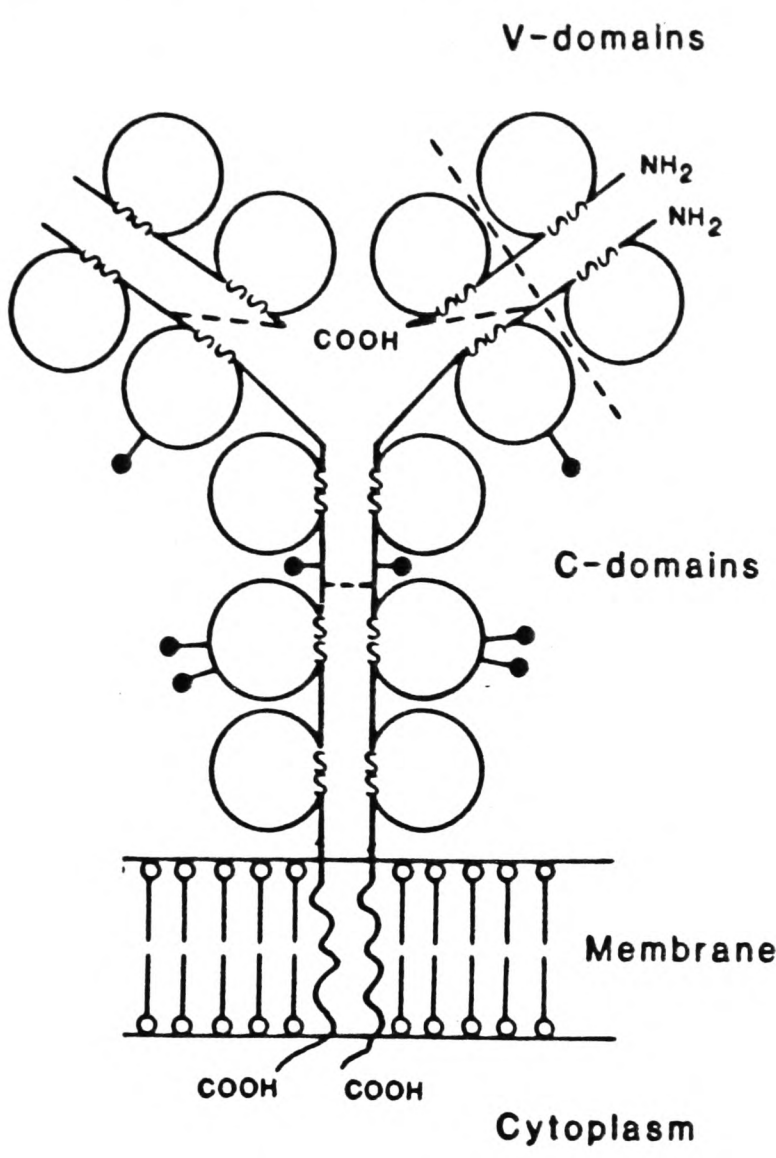
B,C : 3-dimensional folding pattern of V domain (B) and C domain (C).

D : Schematic diagram of the folding pattern in Ig domains. The broad arrows represent β strand segments and the narrow lines sequences connecting β strands. The pattern for a C domain is indicated by the dotted line directly connecting β strands C and D.

Figures A and B are adapted from Williams and Gagnon, 1982.

Figures B and C are adapted from Amzel and Poljak, 1979.

A IgM AT A CELL SURFACE



Folding pattern for V and C domains

3) Antiparallel β strands labelled A to G in figure 1.2 B fold into two β sheets that are held together by inpointing hydrophobic residues and the disulphide bond (figure 1.2C).

4) Conserved patterns of sequence are seen between Ig domains, particularly in β strand segments. These are discussed in more detail in Chapter 5.

There are two main variants of folding pattern characteristic for V and C domains. V domains have an extra loop of sequence between β strands C and D, designated C' and C'' in figure 1.2B. In addition, some sequence patterns are characteristic for V or C domains (Williams and Gagnon, 1982).

Hill et al. (1966) postulated on the basis of the sequence similarities between V and C domains that all immunoglobulins have descended from a primordial molecule homologous to a single domain, by gene duplication and divergence. The organisation of Ig genes support this, as all Ig C domains are encoded by separate exons (Sakano et al., 1979b). The majority of each V domain sequence is also a separate exon, which becomes joined to one small J exon (in light chains) or one small D exon and one small J exon (in heavy chains) by DNA rearrangements during the differentiation of B lymphocytes (Tonegawa, 1983).

1.5.2 Molecules with the Ig fold

Several molecules other than antibody have been found to contain regions of sequence which indicate that their structure is Ig-like. Ideally, the 3-dimensional structure is required to confirm the homology, but this has not yet been determined for molecules other than antibody fragments. The sequences of the following molecules display all the key features of the Ig domain as discussed in the previous section.

1.5.2.2 β_2 microglobulin and MHC Class I

The first of these homologies discovered was β_2 microglobulin, which resembles a C domain (Peterson et al., 1972). Subsequently, the membrane-proximal domain of the MHC Class I molecule was also found to be homologous with Ig C domains (Ploegh et al., 1981).

1.5.2.2 Thy-1

When Thy-1 was found to share sequence homology with Igs, it dismissed the assumption that all Ig-like structures had to play a role in immune functions. The Thy-1 antigen has already been referred to in this Chapter in the context of its expression as a surface antigens of rodent thymocytes (section 1.3.2) and brain (section 1.4.2). Thy-1 was first identified as the theta (θ) alloantigen of mouse thymocytes (Reif and Allen, 1964), and occurs as allotypic forms Thy-1.1 and Thy-1.2 (Reif and Allen, 1966). Only the Thy-1.1 allotype has been identified in the rat (Douglas, 1972). The curious brain/lymphoid distribution of Thy-1 was discussed in sections 1.3.3 and 1.3.4. In the rat, the antigen is the most abundant molecule on the thymocyte surface, with an estimated site number of 10^6 molecules per cell (Mason and Williams, 1980) and an equivalent amount on neuronal cells (Barclay and Hyden, 1978). Both the thymocyte (Letarte-Muirhead et al., 1975) and brain (Barclay et al., 1975) forms have been purified and have similar antigenicities and amino acid compositions (Barclay et al., 1975, 1976). Brain Thy-1 has a molecular weight of 17,500 and thymocyte Thy-1 of 18,700 (Kuchel et al., 1978). The sequence of brain Thy-1 has been determined (Campbell et al., 1981) and is identical to the thymocyte sequence deduced from a cDNA clone (Seki et al., 1985). The molecular weight difference between the two forms can be accounted for by substantial differences in carbohydrate (Barclay et al., 1976).

The complete amino acid sequence of rat Thy-1 (Campbell et al., 1981) shows that the molecule consists of 111 amino acids with 2 disulphide bonds from Cys⁹ to Cys¹¹¹ and Cys¹⁹ to Cys⁸⁵, the latter bond being in the position of the conserved disulphide bonds of Igs. The circular dichroism spectrum of the molecule indicates extensive β structure (Campbell et al., 1980) and secondary structure predictions locate β strands along the sequence, consistent with the Ig folding pattern (Cohen et al., 1981) and most like the V-type fold (Williams and Gagnon, 1982). Thy-1 has amino acid sequence homologies with both V and C type domains (Williams and Gagnon, 1982), and the significance of this is discussed in section 1.5.3.1.

1.5.2.3 Other members of the Ig superfamily

Several other cell surface molecules are now known to share sequence homology with Ig, and their structures are sketched in figure 1.3. These molecules comprise of different numbers of Ig-like domains plus other sequence. The homology extends to the gene arrangement as each Ig-like domain in the MHC Class I (Jordan et al., 1985) and Class II (Lee et al., 1982), Thy-1 (Seki et al., 1985), and T8 (Littman et al., 1985b) molecules is encoded by a separate exon. In the T cell receptor, the V domain comprises of separate V, D and J elements analogous to those of heavy chain Ig genes, which are brought together by DNA rearrangement (Chien et al., 1984a). The organisation of the Poly Ig receptor gene is not yet known.

The majority of these polypeptides participate in various aspects of the immune response, and their possible functions are listed in table 1.6. B and T cell receptors and MHC Class I and II antigens are directly involved with antigen recognition, while the CD8 molecule has an indirect role as discussed in section 1.3.1. The Poly Ig receptor does

Figure 1.3

MOLECULES IN THE Ig SUPERFAMILY

Key : V = V-like domain
 C = C-like domain
 $\begin{smallmatrix} S \\ | \\ S \end{smallmatrix}$ = intrachain disulphide bond
 (S-S) = interchain disulphide bond
 $\bullet$ = N-linked carbohydrate structure

References for sequence data :

Thy-1	Williams and Gagnon, 1982
CD8	Littman et al., 1985b, Sukhatme et al., 1985
T cell receptor	Hedrick et al., 1984b; Saito et al., 1984
MHC Class I	Ploegh et al., 1981
MHC Class II	Kaufman et al., 1984
IgM	Kehry et al., 1980
Poly Ig receptor	Mostov et al., 1984

The figure is modified from Williams et al., 1984 and Williams, 1985

THE IMMUNOGLOBULIN SUPERFAMILY

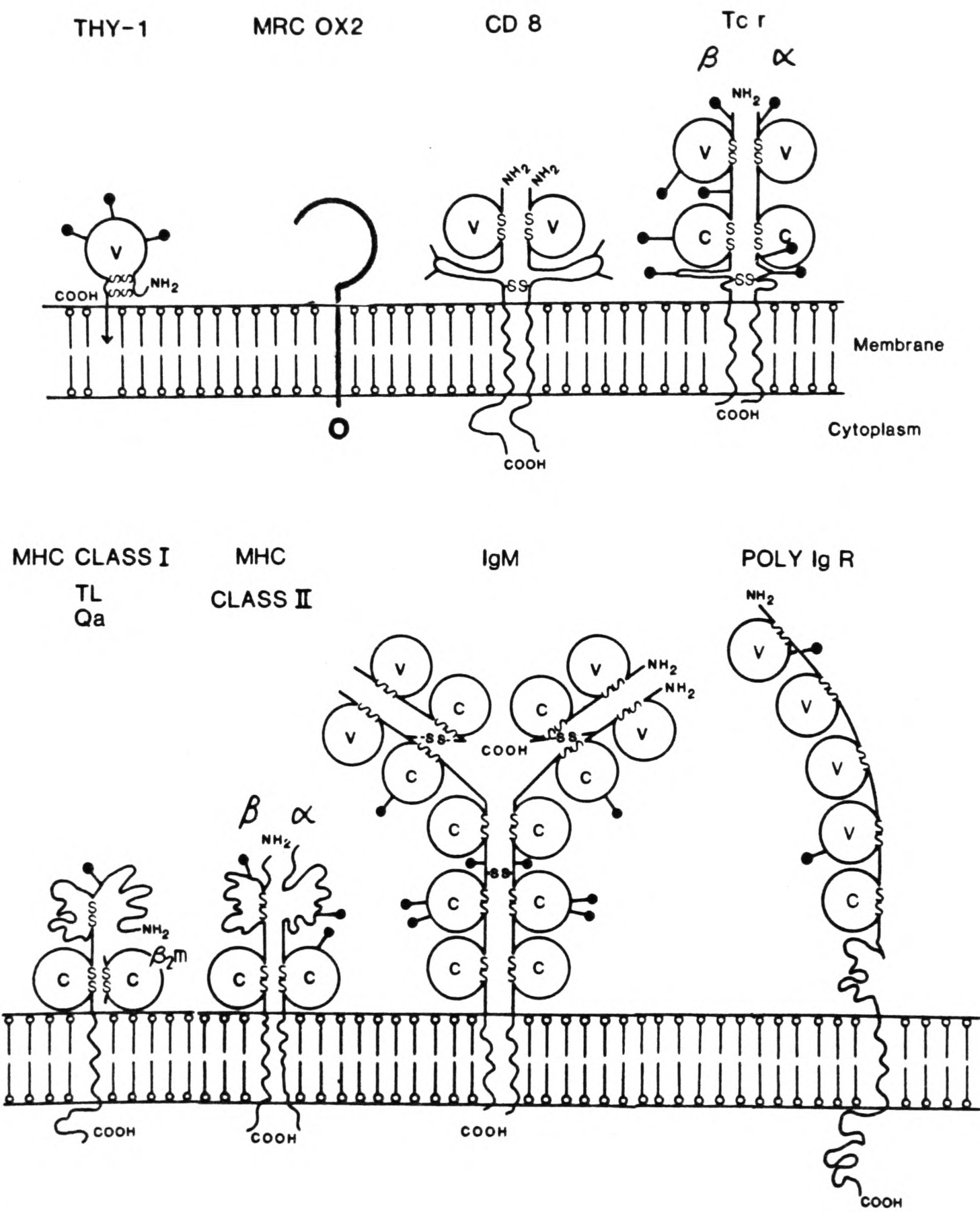


Table 1.6

FUNCTIONS OF MOLECULES WITH IMMUNOGLOBULIN-RELATED STRUCTURES

Molecule	Function	Reference
<u>INVOLVED IN ANTIGEN RECOGNITION</u>		
Immunoglobulin	Receptor for antigen on B cells	(1)
MHC class I	Major histocompatibility antigens recognised in graft rejection. MHC determinants are recognised in association with foreign antigen in T cell responses. MHC class I + Ag by T _{CYTOTOXIC} subset MHC class II + Ag by T _{HELPER} subset	(2)
MHC class II		
T cell receptor	Receptor on T cells for antigen in association with MHC class I or II	(3)
<u>ACCESSORY FUNCTION IN ANTIGEN RECOGNITION</u>		
CD8	Found on T _{CYTOTOXIC} subset of T cells. Involved in interaction between cytotoxic cells and targets bearing antigen with MHC Class I.	(4)
<u>TRANSPORT OF IMMUNOGLOBULIN</u>		
Poly Ig Receptor	Transports IgA and IgM across gut epithelia, then is released as "secretory component"	(5)
<u>UNKNOWN FUNCTION</u>		
TL, Qa	Differentiation antigens expressed on various cell types. TL antigens have been identified on thymus cells of certain mouse strains, and leukaemias of all mouse strains. Genetically linked to the MHC complex.	(6)
Thy-1	Lymphoid/brain antigens of rodents, conserved on neuronal tissue across species.	(7)

References :

- | | |
|----------------------------|--------------------------|
| (1) Warner, 1974 | (5) Poger and Lamm, 1974 |
| (2) Klein et al., 1981 | (6) Kindt, 1981 |
| (3) Collins and Owen, 1985 | (7) Williams, 1982 |
| (4) McDonald et al., 1982 | |

not interact with antigen, but has a role in immunity as it transports IgM and IgA across epithelia into the gut. The Qa and TL antigens in the mouse are of unknown function, but they have a similar structure to that of the Class I antigens and they form a multigene family genetically closely linked to the MHC. Thy-1 stands out as the only member of the Ig superfamily which is unlikely to be involved with immunity since it is conserved on neuronal but not on lymphoid tissues across species. (Campbell et al., 1981).

It could be argued that the sequence similarities seen between the molecules above reflect convergent evolution from different genes to produce a type of structure that is favourable for expression at cell surfaces. Evidence against their convergent evolution is the observation that a 3-dimensional structure resembling an Ig domain can occur without sequence homologies. This is seen for the enzyme superoxide dismutase, which has an Ig-like fold but no conserved disulphide bond or detectable sequence identities (Richardson et al., 1976), and hence would not be considered a member of the Ig superfamily. The assignment of a molecule as being Ig-related must be based on both structure and conserved sequence patterns over the entire length of the domain. Marchalonis et al. (1984), for example, rejected Thy-1 from the Ig superfamily after computer analyses of sequence alignments because they failed to take into account the key Ig-like sequence patterns throughout the molecule.

1.5.3 Possible evolutionary origins of the Ig superfamily

The Ig domain hypothesis (Hill et al., 1966) postulated that a single primordial Ig domain gave rise to all the Ig molecules. Gene duplication of the full Ig domain might produce an Ig light chain-like structure from which all the immune recognition molecules in figure 1.3 evolved. Others (Borgois, 1975; McLachlan, 1980) have suggested an

earlier evolutionary stage was a half domain structure corresponding to β strands A,B,C,C' in figure 1.2B, which may have formed a dimer before its structure was expressed as one gene corresponding to the single domain (McLachlan, 1980).

1.5.3.1 Thy-1 antigen looks most like the primordial domain

It has been argued (Cohen et al., 1981; Williams 1982) that Thy-1 is most like the molecule from which the whole family evolved. This conclusion was based on the following features anticipated for the primordial domain :

a) The primordial structure was a single domain. Both Thy-1 and $\beta_2\text{M}$ exist as single domains, but $\beta_2\text{M}$ is non-covalently associated with MHC Class I heavy chain.

b) The primordial domain should have V- and C- like characteristics. $\beta_2\text{M}$ fits better with the C domain pattern, whereas Thy-1 has a V-like folding pattern but contains sequences very characteristic of both V and C domains (Williams and Gagnon, 1982). In addition to the conserved disulphide bond of the Ig fold, Thy-1 contains a second disulphide bond between Cys 9 and Cys 111, in a similar position to that seen in some C domains (Cohen et al., 1981)

1.5.3.2 Possible functions of the primordial domain

If Thy-1 is directly descended from the primordial domain, then its present day role may reflect the functions of primitive Ig domains. Any functions proposed for Thy-1 must reconcile its unusual tissue distribution and lack of conservation of pattern of expression across species.

The functions of Igs can be divided into (i) antigen recognition by V domains, and (ii) mediation of effector functions by C domains. It is unlikely that Thy-1 is involved with antigen recognition as its

sequence does not vary within individuals of a species. In addition, the structure of the Poly Ig receptor (Mostov et al., 1984) illustrates that domains with V-like folding patterns are not necessarily involved in antigen recognition.

Instead, Cohen et al. (1981) have suggested that Thy-1 is functioning in a manner analogous to the effector functions of Ig C domains, examples of which are given in table 1.7. These interactions can be viewed as a set of ligand (i.e. Ig domain) : receptor systems each of which leads to a different effect as a result of antigen binding. This lead Cohen et al. to propose the hypothesis that Ig-related structures like Thy-1 mediate cell recognition in morphogenesis. The idea has been discussed further by Williams (1982, Williams et al., 1984) and the general points are as follows:

- 1) A family of Ig-related molecules exist which function to trigger cell-cell interactions.

- 2) A set of receptors has co-evolved along with the molecules they recognise, analogous to the ligand-receptor systems of Ig C domains.

- 3) The same ligand-receptor pair can be used in anatomically distinct tissues. This could account for the lack of conservation of tissue-specific antigen expression in different species, if the interaction is a consequence of the differentiated state of the cells and not the triggering system. Any pair of ligand-receptor molecules would potentially be suitable.

- 4) The involvement of Ig domains in cell interactions predates the immune response. Simple Ig-like molecules were first expressed at cell surfaces at an early stage in the evolution of multicellular organisms, and immune function molecules evolved from these recognition systems.

Table 1.7

EXAMPLES OF THE EFFECTOR FUNCTIONS OF IMMUNOGLOBULINS

Function	Interacting structures	Reference
Activation of the classical pathway of complement	i) IgG CH ₂ domain with C1q ii) Fd with C4	Colomb & Porter, 1975 Campbell et al., 1980
Transfer of IgG across placenta	IgG with Fc receptors	McNabb et al., 1976
Interaction of IgG with macrophages	IgG C _H 3 domain with Fc receptor on macrophages	Barnett Foster et al., 1980
Stimulation of mast cells to release mediators of anaphylaxis	IgE with mast cell receptor	Metzger, 1978
Transport of IgA across gut epithelia	IgA and IgM with the Poly Ig receptor	Poger & Lamm, 1974

This hypothesis anticipates a set of simple Ig-related structures at the cell surface, but up to this point Thy-1 was the only known member of this family. Thus it would be valuable to identify and characterise other membrane molecules which do not participate in the immune response but show sequence homology with immunoglobulin.

1.6 OX-2 ANTIGEN

A promising candidate for a Thy-1-like surface molecule is the antigen recognised by monoclonal antibody MRC OX-2, which also has an unusual brain/lymphoid distribution and shares similar biochemical properties with Thy-1.

The MRC OX-2 monoclonal antibody producing hybrid cell line was produced by fusion of mouse myeloma NS-1 cells with spleen cells from a mouse immunised with thymocyte membrane glycoproteins. The proteins had been separated by gel filtration and a fraction covering the molecular weight range 40,000-60,000 was used in the immunisation (McMaster and Williams, 1979).

The unusual tissue distribution of the antigen has already been commented on (section 1.3.3). FACS analysis and inhibition binding assays have detected OX-2 antigen on rat thymocytes, brain and B cells (weakly) (McMaster and Williams, 1979) and also on T cell blasts, produced by Concanavalin A stimulation of thymocytes (W. Jefferies, personal communication). On tissue sections, the antibody binds to follicular dendritic cells in spleen, lymph node and Peyer's patch, as well as to vascular endothelium and some smooth muscle (Barclay, 1981). In the rat ovary, the antigen is expressed mostly on structures which do not develop further, e.g. the granulosa of developing antral follicles (Bukovský et al., 1983). Figure 1.4A shows the pattern of antibody

Figure 1.4

A LOCALISATION OF OX-2 ANTIGEN IN THE SPLEEN

Indirect immunoperoxidase staining of rat spleen with MRC OX-2 antibody gives dendritic-like staining in the follicle (f) and also stains the smooth muscle of the atriole (a) and unidentified structures around the marginal sinus (↓).

p = periarteriolar lymphocyte sheath
r = red pulp

magnification = x 260

Reproduced from Barclay (1981)

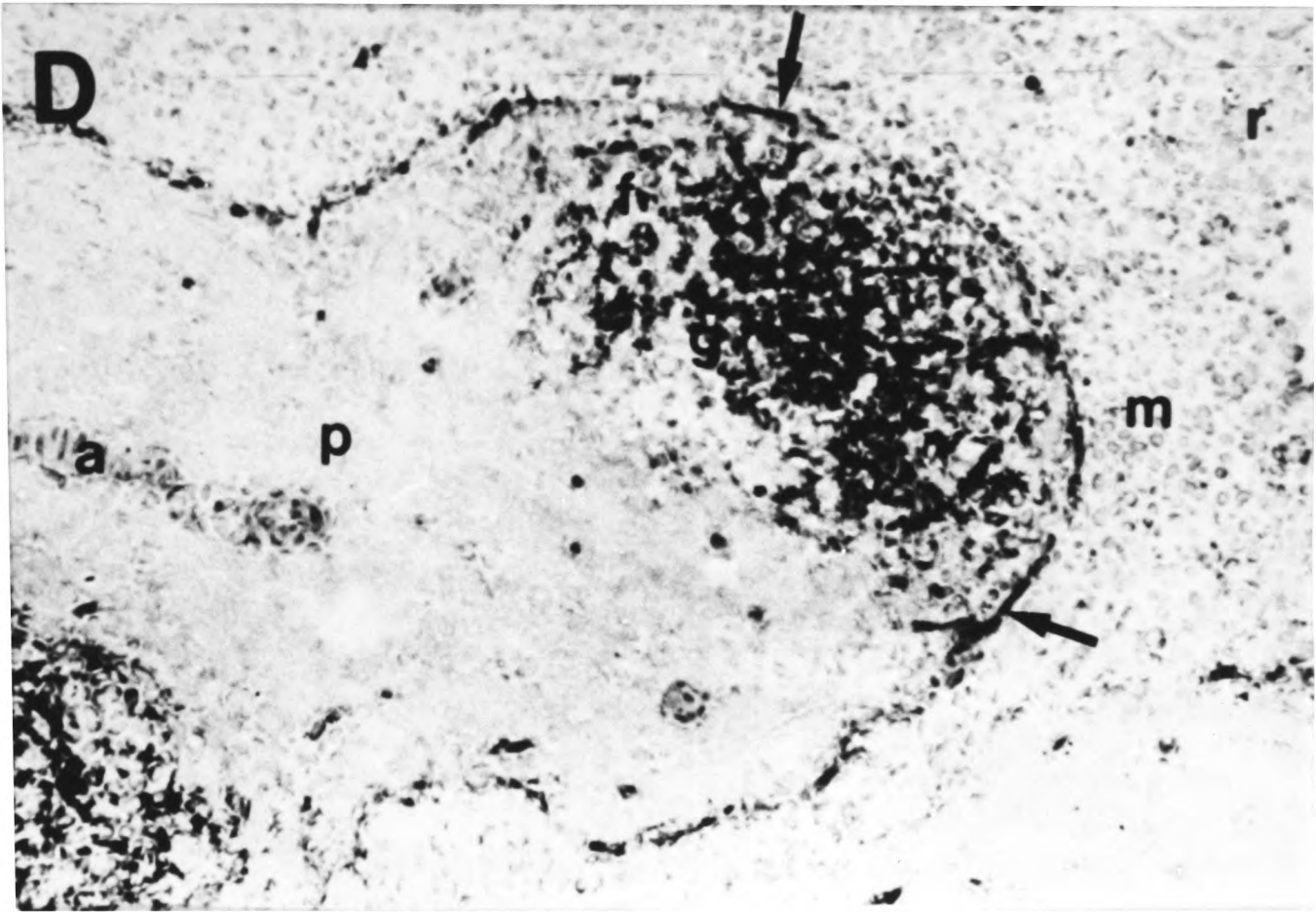
B SDS-PAGE ANALYSIS OF OX-2 AND THY-1 ANTIGENS

Brain and thymocyte forms of OX-2 and Thy-1 (reduced) were run on a 10% polyacrylamide gel and stained with silver stain.

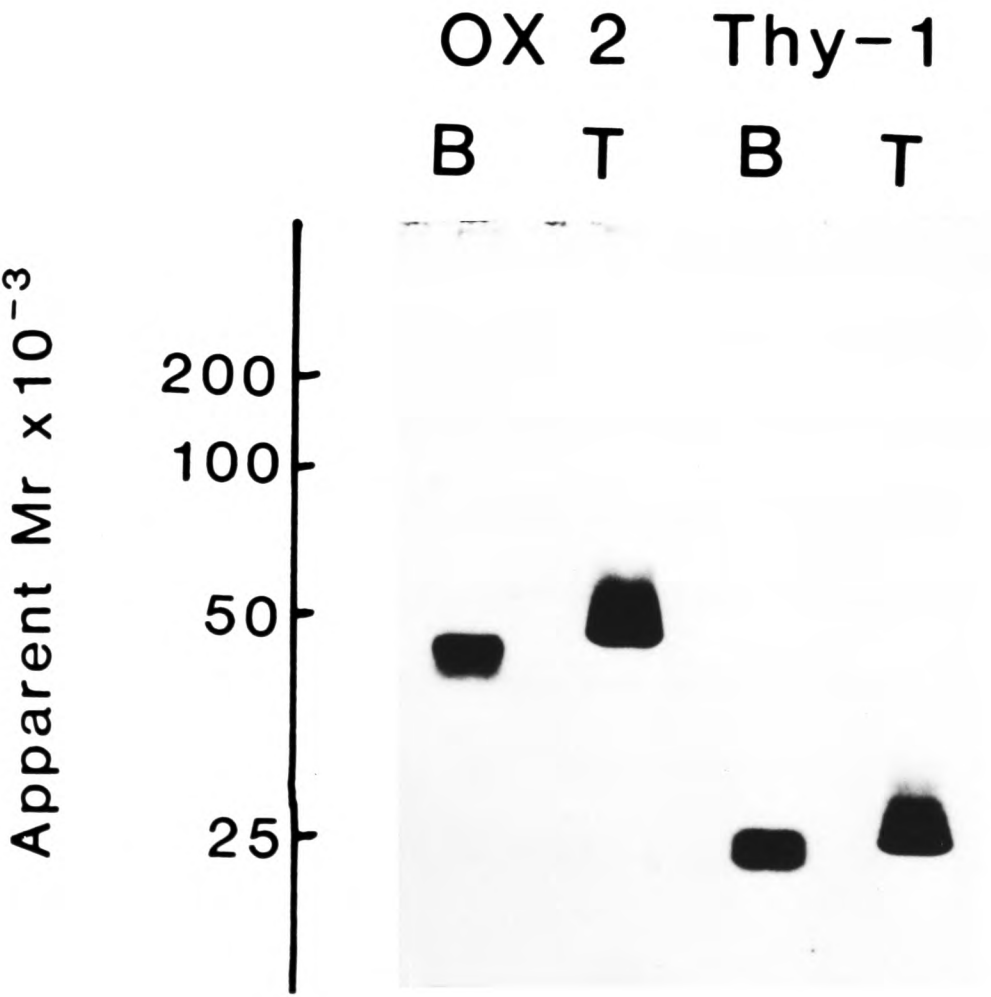
B = protein purified from brain
T = protein purified from thymocyte

Reproduced from Barclay and Ward, 1982

A



B



binding on a section of spleen, illustrating the staining of follicular dendritic cells. These cells are thought to be involved in the generation of B cell memory (Klaus et al., 1980).

The levels of expression of OX-2 antigen in the brain fluctuate during neonatal development from 95% of adult levels at birth, to 150% at day 11, reaching adult levels by day 40 (Webb and Barclay 1984). The postnatal expression of Thy-1 antigen also varies, but follows different patterns. The antigen is present at low levels at birth (7% of adult) and increases rapidly from 30% of adult levels at day 9 to reach adult levels at day 22 (Barclay, 1979).

The brain and thymocyte forms of OX-2 have been purified by solubilisation in DOC and affinity chromatography, and their molecular characteristics are compared with those of Thy-1 in table 1.8. The similarity in the migration of the antigens on SDS-PAGE is particularly striking (figure 1.4B). Brain OX-2 appears as a doublet (M_R 41,000) and thymocyte OX-2 as a trail (M_R 47,000), while brain Thy-1 also appears as a doublet (M_R 24,000) and thymocyte Thy-1 as a trail (M_R 25,000). The antigenicity and amino acid composition of the tissue forms of OX-2 are indistinguishable, so the apparent size difference is probably due to differences in the carbohydrate structure. OX-2 antigens are highly glycosylated, with brain OX-2 containing 24% and thymocyte OX-2 33% by weight carbohydrate, present as N-linked structures. The difference in the composition of the carbohydrate on thymocyte and brain forms of OX-2 is very similar to that seen for Thy-1. The thymocyte forms of both OX-2 and Thy-1 contain higher levels of galactose and sialic acid and lower levels of fucose than the brain forms, and this reflects overall differences in the patterns of glycosylation between the two tissues (Barclay and Ward, 1982).

Table 1.8

COMPARISON OF OX-2 AND THY-1 ANTIGENS

	RAT OX-2 ANTIGEN	RAT THY-1 ANTIGEN
TISSUE DISTRIBUTION	Neurons Thymocytes (not mature T cells) Endothelium of vessels Follicular dendritic cells, T blasts, B cells, smooth muscle.	Neurons Thymocytes (not mature T cells) Pericyte sheath around vessels. Haemopoietic stem cells, eosinophils, fibroblasts.
SDS-PAGE	Brain M_R : 41,000 doublet Thymocyte M_R : 47,000 trail	Brain M_R : 24,000 doublet Thymocyte M_R : 25,000 trail
MOLECULAR CHARACTERISTICS	1 chain glycoprotein Brain : 24% CHO Thymus : 33% CHO Brain forms have higher fucose, lower galactose and sialic acid.	1 chain glycoprotein Brain : 29% CHO Thymus : 33% CHO Brain forms have higher fucose, lower galactose and sialic acid.

From Barclay and Ward (1982)

The tissue distribution of OX-2 might suggest a role in cell surface events, but the function of OX-2 is as yet unknown as no in vivo or in vitro effects have been observed for the monoclonal MRC OX-2. An interaction between OX-2 antigen and mast cells was observed when tissue sections were incubated with purified OX-2 antigen, followed by localisation with MRC OX-2 antibody and immunoperoxidase staining (Barclay and Ward, 1982). The significance of this interaction is unclear, as a similar staining is seen with the leucocyte-common antigen.

The conservation of OX-2 expression on tissues of other species is also unknown. No cross-reactions are observed between rabbit antisera raised against purified OX-2 and mouse tissues, in contrast to the antiserum raised against Thy-1 (Williams et al., 1976) and against MHC Class II antigens (Fukamoto et al., 1982) which detect a murine homologue. However, a failure to detect an equivalent molecule by serology has been observed before, for the rat and mouse leucocyte-common antigen (Dr A.F.Williams, personal communication).

The possibility that OX-2 and Thy-1 are structurally related can be evaluated by determining the complete amino acid sequence of OX-2.

1.7 STRATEGIES FOR DETERMINING THE SEQUENCE OF CELL SURFACE MOLECULES

Monoclonal antibodies can be used to isolate pure antigen of a quality suitable for sequence studies, but only the most abundant cell surface molecules such as Thy-1 (Campbell et al., 1981) can be prepared in sufficient amounts to permit complete sequence determination by the traditional approach of overlapping peptides (Allen, 1981). Recent advances in recombinant DNA technology have created a range of strategies for isolating cDNA clones encoding cell surface antigens, and these strategies require only limited protein sequence data or none at

all. Examples are listed in table 1.9, and as discussed below, oligonucleotide hybridisation would appear to be the most straightforward approach for cloning the OX-2 antigen.

1.7.1 The oligonucleotide approach to sequencing cell surface antigens

The key points of this strategy are:

- i) Purify sufficient antigen from which to obtain reliable peptide sequence.
- ii) Synthesise an oligonucleotide corresponding to the protein sequence.
- iii) Isolate cDNA clones from an appropriate cDNA library which hybridise to the oligonucleotide.

Stage (i) only requires a short stretch of sequence (4-6 amino acids) but it must be unambiguous and avoid amino acids such as Ser and Thr which are encoded by a large number of codons. Thus several peptides need to be obtained and it is important to be able to isolate enough antigen of sufficient purity. For example, the calculations in Chapter 3 section 2.1 estimated that 100 nmoles of brain OX-2 antigen would be required if the peptides were to be sequenced on a spinning cup sequencer. For more sensitive gas-phase sequencers, 10 nmol of starting protein should be sufficient. Monoclonal antibody affinity chromatography followed by a gel filtration step is suitable for the isolation of most cell surface antigens, provided that the protein retains its antigenicity when extracted with detergent, and the monoclonal antibody has a reasonably high affinity for the solubilised monomeric antigen (reviewed by Williams and Barclay, 1985).

The sequence of the N-terminus of the protein, or of peptides separated by standard chromatographic techniques (Allen, 1981), is determined by automated Edman degradation. Spinning cup sequencers

Table 1.9

METHODS CURRENTLY IN USE FOR THE ISOLATION OF cDNA CLONES IN THE
DETERMINATION OF THE AMINO ACID SEQUENCE OF LYMPHOCYTE SURFACE ANTIGENS

Strategy	Example of cell surface antigen sequenced by this approach	Reference
Oligonucleotide hybridisation	L-CA Thy-1	Thomas et al. 1985 Moriuchi et al. 1984
Expression in eukaryotic systems	T8	Kavathas et al. 1984 Littman et al. 1985b
Subtractive hybridisation	T cell receptor	Hedrick et al. 1984b
Hybrid selection-translation	β_2 microglobulin	Parnes et al. 1981
Polysome immuno-precipitation	MHC Class II	Korman et al. 1982

require 1-20 nmol of peptide, while the more sensitive gas-phase sequencer needs only 0.1-1 nmole (Hunkapillar et al., 1984). From the protein data, a region is selected for the design of the oligonucleotide probe. Because of the redundancy of the genetic code, most amino acids are encoded by more than one codon and thus a variety of possible combinations of nucleotide sequence exist for each peptide. The following two alternative strategies to probe design have been employed:

1) Synthesis of a mixture of oligonucleotides which correspond to every possible combination of codons. To minimise the complexity of the oligonucleotide mix required, probe sequences must therefore be based on stretches of protein sequence rich in amino acids which have a restricted number of codons such as Trp, Met (1 codon), Gln, Phe etc. (2 codons). Oligonucleotides 14 long have been used successfully (Kornblihtt et al., 1983) but longer sequences have less chance of hybridising to an identical sequence occurring at random in the wrong clone, and oligonucleotides have been employed which are, for example, 16 bases long with 8 variations (Reyes et al., 1981) or 17 bases long with 64 sequence variations (Bentley and Porter, 1982; Thomas et al., 1985).

2) Synthesis of long single copy probe which has been designed taking the codon utilisation, G:T basepairing and dinucleotide frequencies of published sequences into account (reviewed by Lathe, 1985). The probe relies on increased length to confer specificity, even though it will not be an exact match. This approach has been used for the isolation of clones encoding the human coagulation factor IX using a probe 52 nt long (Jaye et al., 1983) and epidermal growth factor using a probe 51 nt long (Ullrich et al., 1984).

The oligonucleotide is then used to screen a cDNA library, made from the RNA of a tissue known to express the antigen. Synthesis of cDNA libraries has been discussed by Maniatis et al. (1982) and details of the construction of a thymocyte cDNA library are given in Chapter 2.3.16. Briefly, total RNA is prepared from the tissue, poly A⁺ RNA selected using oligo dT affinity chromatography, and double strand cDNA generated by copying the RNA template using reverse transcriptase for the first strand and either DNA Pol 1 or reverse transcriptase for the second strand. The cDNA may then be clonally isolated by insertion into a plasmid vector and transfection into a competent bacterial host. The resulting library is screened with the radiolabelled oligonucleotide and clones which hybridise are isolated and sequenced.

The disadvantage of the oligonucleotide approach is that it requires some initial protein chemistry. For molecules such as OX-2 for which there is a good source of starting material and a monoclonal with which to purify the antigen, this should not be a problem. Once a cDNA clone has been isolated, the protein sequence not chosen for the oligonucleotide design is valuable for confirming whether or not the correct clone has been identified.

1.7.2 Other strategies for cloning cell surface antigens

If it was not feasible to obtain sequence data on OX-2 protein, then the following strategies (table 1.9) would be considered:

1.7.2.1 Hybrid selection-translation

cDNA clones taken at random from a library are pooled and their DNA bound to nitrocellulose filters. The filters are hybridised with RNA from the tissue expressing the antigen and transcripts which specifically bind to the filters are eluted. The mRNA is translated in vitro and the products screened by immunoprecipitation with the

monoclonal. The pool of cDNA clones which selected the RNA encoding the antigen are then cloned out and the procedure repeated. The method has been used to clone β_2 microglobulin but is dependent on the monoclonal or a xenoantiserum being able to recognise an in vitro translation product. For OX-2, a rabbit antiserum was available, but had not yet been tested for its ability to immunoprecipitate in vitro translation products..

1.7.2.2 Immunoprecipitation of polysomes

Monoclonal or polyclonal antibodies may be used to immunoprecipitate polysomes that are in the process of translating the mRNA of the surface antigen, thus producing an RNA preparation enriched for specific transcripts, suitable for preparing or screening cDNA libraries. The HLA-DR heavy chain has been cloned by this strategy (Korman et al., 1982).

1.7.3.3 Expression in eukaryotic systems

Genomic DNA is introduced into a eukaryotic cell (usually of a different species) where it becomes integrated into the host genome. If the foreign DNA is expressed, those cells containing the DNA encoding the cell surface antigen can be identified with the monoclonal antibody and isolated by FACS sorting or rosetting. The transfection is normally repeated using DNA from the primary transformant to generate secondary transformants, and the foreign DNA is recovered by cloning, assuming that the gene encoding the antigen is closely linked to repetitive sequences characteristic of that species e.g. Alu repeats of human DNA. The human DNA can thence be retrieved by screening a DNA library from the secondary transformants with a repetitive sequence probe, made by nick-translating human genomic DNA. The gene for the human transferrin receptor has been cloned in this manner (Kühn et al., 1984). If the

foreign DNA is not closely linked to a repetitive sequence it will be necessary to isolate the gene by an alternative strategy, such as subtractive hybridisation in which a cDNA probe is constructed from transfectants expressing the antigen and made specific by absorption with mRNA from untransfected cells. The cDNA can be used to make a library or to screen a cDNA library constructed from cells known to express the antigen. This approach has been successful for the isolation of the human T cell antigen T8 (CD8) (Littman et al., 1985b).

1.7.3.4 Restricted cDNA libraries

In the absence of either protein data or specific antibodies, it has been possible to isolate cDNA clones for certain molecules which have a very restricted pattern of expression, for example the T cell receptor (Hedrick et al., 1984a). The strategy requires that the antigen be expressed on one cell type (i.e. T cell) but absent from a second very similar cell type (B cell). Cloned DNAs common to both cell types can be removed by hybridisation, selecting T cell-specific DNA. A further screen is then necessary to discriminate between these cDNAs. For the T cell receptor, the criteria was that the genes should be rearranged in T cells. This method is particularly suited to the T cell receptor, and would not, of course, be suitable for antigens expressed in a less specific fashion, for which there would be problems in choosing the correct clone for the antigen.

1.8 DISCUSSION TO CHAPTER 1

This thesis evaluates whether the similarities in the tissue distribution and molecular characteristics of OX-2 and Thy-1 antigens reflect a sequence and structural homology. The approach chosen was to purify protein and obtain peptide sequence with which to design an

oligonucleotide. cDNA clones encoding OX-2 antigen were isolated from a thymocyte cDNA library using the oligonucleotide. The complete sequence of OX-2 deduced from the clones is compared with the sequence of Ig-related molecules. The cDNA clones are used as hybridisation probes to investigate the structure of OX-2 mRNA, and to isolate a genomic clone encoding the human OX-2 gene. Preliminary data on the gene organisation and partial sequence is presented.

CHAPTER 2

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 Animals

Sprague Dawley rats, used for the bulk preparation of OX-2 antigen were either from Olac (1976) Ltd., Bicester, Oxford or Department of Psychology, University of Oxford.

AO and PVG rats used as a source of RNA and DNA were bred in the MRC Cellular Immunology Unit, Oxford, under specific pathogen free conditions.

(DBA/2 x Balb/c) F1 mice used for the preparation of ascites and DNA were from the MRC Cellular Immunology Unit animal house, Oxford.

2.1.2 Chemicals

<u>Autoradiography</u>	Kodak X-Omat X ray film	Kodak
<u>Bacterial media</u>	Bacto-tryptone	Difco Labs
	Agar No 1	Oxoid Ltd
	Yeast extract	Oxoid Ltd
<u>Enzymes</u>	Alakaline phosphatase	Boehringer
	DNA Polymerase I	Amersham
	(Klenow fragment)	
	Nick-translation kit	Amersham
	Polynucleotide kinase (T4)	NEN
	Restriction endonucleases	Amersham or
		New England Biolabs
	S1 nuclease	Amersham
	Trypsin TPCK	Worthington

<u>Gel filtration</u>	Biogel P30	Bio-Rad labs
	Sephacryl S300	Pharmacia
	Sephadex G50	
	Sepharose CL-4B	
<u>HPLC</u>	Acetonitrile	Rathburn
	Ammonium acetate AnalaR grade	BDH
	Methanol	Rathburn
<u>Radiochemicals</u>	Iodo 2- ³ H acetic acid	NEN
	[γ - ³² P] ATP (>5000Ci/mmol)	
	[α - ³² P]dNTP (3000 Ci/mmol)	Amersham
	Na ¹²⁵ I	
<u>Various chemicals</u>	Bovine serum albumin	BRL
	(for DNA work)	
	Deoxycholate	BDH

All other chemicals were AnalaR grade, purchased from BDH or Sigma.

Addresses of chemical suppliers

Amersham	Radiochemical Centre, Amersham, Bucks.
BDH	BDH, Broom Road, Poole, Dorset.
Bio Rad	Bio Rad Laboratories Ltd, Caxton Way, Watford, Herts.
Boehringer	BCL, Bell Lane, Lewes, East Sussex.
BRL	Bethesda Research Laboratories(UK) Ltd, Science Park, Cambridge.
Difco	Difco Ltd, Central Avenue, East Moseley, Surrey
Kodak	Kodak, Hemel Hempstead, Herts.
NEN	DuPont (UK) Ltd., Wedgewood Way, Stevenage, Herts.
Oxoid	Oxoid Ltd., Wade Road, Basingstoke, Hants.
Pharmacia	Pharmacia Ltd., Midsummer Boulevard, Milton Keynes.

Rathburn	Rathburn Chemicals Ltd., Walkerburn., Peeblesshire, Scotland.
Sigma	Sigma London Chemical Co., Fancy Road, Poole, Dorset.
Worthington	Worthington Biochemical Corporation, Freehold, New Jersey, USA

2.1.3 Antibodies

RAM	Kindly provided by Dr A.F. Williams. Prepared as in Jensenius and Williams, (1974 a)
MRC OX-2	For binding assays, MRC OX-2 IgG was in the form of spent tissue culture medium from cultures of MRC OX-2 hybridoma (McMaster and Williams, 1979). For affinity chromatography, MRC OX-2 IgG was prepared from ascites fluid of (DBA/2 x Balb/c) F1 mice growing the cell line as a tumour, as for MRC OX-1 (Sunderland et al., 1979).

2.1.4 Buffers

DAB	Made from tablets of Dulbecco's A and a concentrate of Dulbecco's B obtained from Oxoid Ltd. The final buffer is 0.136M NaCl, 0.0027M KCl, 0.0081M Na_2HPO_4 , 0.0015M KH_2PO_4 , 0.001M CaCl_2 , 0.001M MgCl_2 , 0.003M NaN_3 pH 7.3.
Denaturing loading dye	2 x = 80% deionised formamide, 50mM Tris-borate, 50mM borate, 2mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue.
Denhardt's solution	100 x Denhardt's = 2% Ficoll, 2% polyvinyl pyrrolidone, 2% BSA
0.5% DOC buffer	0.5% DOC (w/v), 0.02% NaN_3 , 10mM Tris-HCl pH8.0

L-Broth	Bacto-tryptone 10g/L, yeast extract 5g/L, NaCl 10g/L
Medium salt buffer	50mM NaCl, 10mM MgCl_2 , 1mM DTT, 10mM Tris-HCl, pH 7.5
MOPS	10 x MOPS = 0.2M MOPS, 0.05M Na acetate, 0.01M EDTA, pH 7.0
NET	6 x NET = 1M NaCl, 0.9M Tris-HCl, 6mM EDTA pH 7.5
Normal loading dye	3 x = 30% glycerol, 15mM EDTA, 0.5% bromophenol blue
SM	0.1M NaCl, 8mM MgSO_4 , 0.1% gelatin, 50mM Tris-HCl pH 7.4
SSC	20 x SSC = 3M NaCl, 0.3M Na citrate pH 7.0
SSPE	20 x SSPE = 3.6M NaCl, 0.2M NaH_2PO_4 , 0.02M EDTA, pH 7.4
T-broth	Bacto-tryptone 10g/L, NaCl 5g/L
TBE	10 x TBE = 0.89M Tris-borate, 0.89M boric acid, 0.002M EDTA
2 x TY	Bacto-tryptone 16g/L, Yeast extract 10g/L, NaCl 5g/L

2.2 METHODS : PROTEIN CHEMISTRY

2.2.1 Purification of OX-2 antigen from rat brain

The method follows the protocol of Barclay and Ward (1982). All stages were performed at 4°C.

2.2.1.1 Preparation of rat brain deoxycholate extract

Sprague-Dawley brains were removed, frozen in liquid nitrogen, and stored at -20°C. 150g wet weight of brains were thawed and homogenised in a Waring Blender with 4ml of 30mM Tris-HCl, 0.02% NaN₃ pH8.0 per gram of brain (i.e. 600ml of buffer). The homogenate was centrifuged at 18,000 x g for 60 minutes, and the pellet resuspended in 3% DOC, 30mM Tris-HCl, 0.02% NaN₃ pH8.0 to a ratio of 0.25g DOC per g starting weight of brain (i.e. 1.25L of DOC buffer) Iodoacetamide to 1 mM and PMSF to 0.1mM final concentration were added immediately to inhibit proteolysis. The mixture was returned briefly to the Waring Blender to disrupt large clumps of tissue, then homogenised finely in a hand-held Potter-Elvehjem homogeniser. The DOC suspension was centrifuged at 18,000 x g for 60 minutes and the supernatant decanted. The pellet was re-homogenised with 200ml of the 3% DOC buffer and re-centrifuged. The supernatants were pooled and referred to as "rat brain DOC extract".

2.2.1.2 The "Batch Method" MRC OX-2 affinity chromatography

MRC OX-2 antibody was coupled to sepharose CL-4B beads by the method of Porath (1974) at a ratio of 10mg IgG per ml beads. Before use, the beads were pre-eluted with 50mM diethylamine, 10mM Tris-HCl, 0.02% NaN₃ pH 11.5.

10 ml of the antibody-coupled beads were mixed with 1.1L of rat brain DOC extract in a screw capped plastic bottle and maintained in

suspension by attaching the bottle to a wheel revolving at 2 rpm. The bottle was filled up completely with the extract to avoid frothing during rotation. Most of the OX-2 antigenic activity was removed from the extract in 6 hours (see Chapter 3) but in practice it was convenient to leave the mixture rotating overnight (12-16 hours). The beads were recovered by gently drawing the DOC extract through a coarse grade (No 1) sinter funnel. The beads were washed whilst still on the funnel by resuspending them in 50 ml 0.5 % DOC, 10mM Tris-HCl, 0.02% NaN_3 pH8.0. The buffer was drawn through the sinter without allowing the beads to dry, and the procedure repeated about 5 times until the washings were colourless and the beads no longer formed viscous clumps.

2.2.1.3 Elution of antigen bound to MRC OX-2 sepharose beads

After the final wash on the sinter funnel, the slurry of MRC OX-2 Sepharose beads was poured into a conventional column arrangement flowing at approximately 30 ml/hr. The antibody column was washed further with the 0.5% DOC buffer, then with the 0.5% DOC buffer containing 0.15M NaCl to remove non-specifically bound material, and finally in the 0.5% DOC buffer again. In each case, the wash continued until the absorbance at 280nm of the eluate was the same as that of the applied buffer. The bound antigen was eluted with 50mM diethylamine, 0.5% DOC, 10mM Tris-HCl, 0.02% NaN_3 pH 11.5. The absorbance at 280nm of each eluted fraction was monitored, and those containing protein were immediately neutralised with solid glycine, and pooled. The eluate was dialysed against 3 changes of 0.25% DOC, 10mM Tris-HCl, 0.02% NaN_3 pH8.0.

2.2.1.4 Gel filtration of affinity purified OX-2 antigen

OX-2 antigen eluted from the MRC OX-2 affinity column was purified from high and low molecular weight contaminants by gel filtration on a Sephacryl S300 column.

3 batches of eluted rat brain OX-2 antigen were pooled in order to minimise non-specific loss of protein during gel filtration. The combined eluates were concentrated from approximately 200ml down to 5 ml by ultrafiltration on an Amicon P10 membrane and the concentrate was given a clarifying spin in a Beckman type 35 rotor at 30,000 rpm for 30 minutes. The supernatant was loaded onto a Sephacryl S300 column (2.5cm x 77cm) running in 0.5% DOC, 10mM Tris-HCl, 0.02% NaN₃ pH8.0. The column flow rate was 13ml/hr, and 4ml fractions were collected. The absorbance at 280nm of the column eluate was monitored, and samples of the peak fractions analysed on a 12% polyacrylamide gel. The fractions containing the highest purity OX-2 antigen were pooled and concentrated to 1ml by ultrafiltration as before. The preparation was dialysed extensively against 0.1M NH₄HCO₃ to remove DOC, and protein quantitated by amino acid analysis.

2.2.2 Chemical modification of OX-2 antigen

2.2.2.1 Reduction and alkylation of brain OX-2 antigen

1.8mg affinity purified OX-2 antigen was dried under vacuum and resuspended at 2.5 mg/ml in 7M guanidine-HCl, 0.5M Tris-HCl, 0.002M EDTA pH8.2. DTT was added to a final concentration of 8mM which gave a 15-fold molar excess of DTT over Cys residues in the protein. The solution was incubated for one hour at 37°C under nitrogen. 100uCi (2umoles) of 2³H-iodoacetic acid at 1mCi/ml and 50mCi/mmol was added per 100 nmoles of protein. The mixture was incubated at room temperature for 5 minutes.

Ice-cold unlabelled iodoacetic acid was added in 2.2 fold molar excess over the DTT and incubation continued for a further 30 minutes on ice. Unincorporated reagents were removed by extensive dialysis against 0.1M NH_4HCO_3 then against 10mM Tris-HCl pH8.0 prior to succinylation. Incorporation of ^3H -IAA into the reduced and alkylated protein was estimated by scintillation counting, and was typically 3-5% of that added.

2.2.2.2 Succinylation of reduced and alkylated OX-2

Reduced and alkylated OX-2 was dried under vacuum and resuspended at 1.8 mg/ml in 1M Tris-HCl pH10.0. Powdered succinic anhydride was added in 200-fold molar excess over free amino groups in the protein. The anhydride was added gradually over two hours and the solution shaken steadily at room temperature. Excess reagents were removed by dialysis against 0.1M NH_4HCO_3 . The reaction was repeated to ensure complete succinylation of lysine groups. Recovery of protein was estimated by scintillation counting.

2.2.3 Trypsin digestion of reduced, alkylated and succinylated OX-2

Reduced, alkylated and succinylated OX-2 at a concentration of 2 mg/ml in 0.1M NH_4HCO_3 was centrifuged at 9000 x g for 5 minutes to remove insoluble material. The supernatant was incubated with 2% (w/w) TPCK treated trypsin for 6 hours at 37°C and the peptides immediately separated by gel filtration (see below).

2.2.4 Purification of OX-2 tryptic peptides

2.2.4.1 Separation by gel filtration on Biogel P30

OX-2 tryptic peptides in a volume of approximately 1ml were loaded onto a Biogel P30 column (200-400 mesh) equilibrated in 0.1M NH_4HCO_3

pH8.0-8.5 at room temperature. The flow rate was 3 ml/hr and fraction volume 0.25 ml. Peptides were detected by their absorbance at 206 nm on a Uvicord S monitor and those containing SCMC identified by scintillation counting of a 10 μ l aliquot of each fraction. Peptide peaks were pooled and a 5% sample reserved for amino acid analysis.

2.2.4.2 Reverse phase HPLC

The OX-2 tryptic peptide pools collected from the P30 column were first analysed by reverse phase HPLC, then purified on a preparative scale if more than one peak was observed.

The sample in 0.1M NH_4HCO_3 was injected into a Waters μ Bondapak C_{18} HPLC column equilibrated in 10mM ammonium acetate (Buffer A). The absorbance at 206nm of the eluate was detected on a LKB Uvicord S monitor. When the salt front and any peptides which did not adhere to the column had come through, a linear gradient of buffer B was run from 0% B to 85% B (where B is 70% acetonitrile, 10mM ammonium acetate) over one hour. Eluted peaks were collected manually and analysed for amino acid content and ^3H counts. The preparative scale separation of a peptide sample was performed in two HPLC runs. 30% of the sample was separated in the first run, and if the resolution of the peaks was satisfactory, the remaining 70% was fractionated.

2.2.5 Lowry assay

The method of Lowry et al. (1951) was used to determine protein content during the purification of OX-2. Samples were diluted into 0.5% DOC, 10mM Tris-HCl pH8.0 and quantitated by comparison with standard dilutions of BSA in the same buffer.

2.2.6 Amino acid analysis

Approximately 10ug of protein or 1-2 nmole of peptide was dried down in an acid washed glass tube. The sample was hydrolysed in 200µl of constant boiling HCl (5.7M), 7mM β mercaptoethanol, 5mM phenol in the sealed tube under vacuum at 110°C for 24 hours. The tube was opened and the sample dried under vacuum then analysed by Mr T.Gascoyne of the MRC Immunochemistry unit, Oxford, on a LKB 4400 amino acid analyser.

2.2.7 Automated amino acid sequence determination

Amino acid sequences of peptides were determined by Dr Jean Gagnon of the MRC Immunochemistry Unit, Department of Biochemistry, Oxford. Peptides T1 to T5 (10-30 nmol) and intact OX-2 (20 nmol) were sequenced by automated Edman degradation on an upgraded Beckman 890C spinning cup sequencer. Polybrene (2mg) was added to the sequencer cup prior to the application of each sample. Peptide T6 (2nmol) was sequenced on an Applied Biosystems gas phase sequencer.

Amino phenylthiohydantoin (PTH) derivatives produced at each sequencing step were identified by reverse phase HPLC on an ODS-Hypersil 5µ column (5mm x 100mm), which was equilibrated with 0.01M sodium acetate pH 4.2 containing 10% v/v methanol and eluted with a linear gradient of ethanol/acetonitrile (1:1 v/v) from 5 to 45% (v/v) in 27 minutes at 1ml/min.

Histidine and arginine were identified on a µ Bondapak C₁₈ column with a linear gradient of methanol from 20 to 35% (v/v) in an aqueous buffer.

S (³H)-carboxymethylcysteine was confirmed by counting an aliquot of the PTH extract.

2.2.8 Indirect binding assay to quantify OX-2 antigen

Inhibition of indirect radioactive binding assays was used to follow OX-2 antigenic activity through the purification stages. (Morris and Williams, 1975; Barclay and Ward, 1982). To enable extracts to be assayed in detergent, the target cells were fixed with glutaraldehyde (Williams, 1973).

Serial dilutions of the test sample were made in 0.2% DOC, 0.5% BSA, 10mM Tris-HCl, 0.02% NaN_3 pH 8.0 and 150 μl of the dilution was incubated with 150 μl of MRC OX-2 tissue culture supernatant diluted 1/15 for 3-16 hours at 4 °C. Aggregated material was removed by centrifugation at 9000 x g for 5 minutes at 4 °C. Two 50 μl aliquots of the supernatant were incubated with 50 μl target cells (2×10^6 rat thymocytes, 5×10^6 sheep erythrocytes) for 1 hour at 4 °C. The cells were washed twice in 0.1% BSA, 10mM Tris-HCl, 140mM NaCl, 0.02% NaN_3 pH8.0 and resuspended in 50 μl 0.5% BSA, 10mM Tris-HCl, containing 2×10^5 cpm ^{125}I -labelled rabbit F(ab')_2 anti-(mouse IgG) at trace levels (0.13 $\mu\text{g/ml}$, 20 $\mu\text{Ci/ug}$). After incubation for 1 hour at 4 °C the cells were washed twice as before and counted on a LKB Rackgamma.

The data was plotted as a graph of cpm of ^{125}I RAM bound to target cells against $\log_{10}$ dilution of the test sample and one unit of antigenic activity defined as the amount of antigen required to give 50% inhibition of the assay.

High concentrations of deoxycholate had a slight inhibitory effect on the assay. Thus when samples in 3% DOC buffer were assayed, equivalent dilutions of the 3% DOC buffer itself were included, and the antigenic activity of the sample corrected for the effect of the detergent.

2.2.9 Polyacrylamide gel electrophoresis in SDS

The purity of brain OX-2 antigen in the latter stages of its purification was monitored by polyacrylamide gel electrophoresis in SDS according to the method of Laemmli (1970).

Samples were boiled in 20% glycerol, 5% SDS, 2% DTT, 0.005% bromophenol blue and electrophoresed on 12% or 15% 0.75mm thick polyacrylamide slab gels with 3% stacking gels. Protein bands were visualised by staining with Coomassie Brilliant blue.

2.3 METHODS : RECOMBINANT DNA TECHNIQUES

The majority of these protocols are standard laboratory methods and have been described elsewhere in detail, particularly in the laboratory manual by Maniatis et al. (1982). Therefore I have not given full experimental details for each technique, but have described the method in outline and indicated the points at which my version differs from the published protocol.

Buffer compositions are given in Section 2.1.4

2.3.1 Microbiological techniques

2.3.1.1 Storage and growth of transformed E.coli MC1061

Transformation of MC1061 with pAT153/PvuII/8 confers ampicillin resistance on the bacteria, which can then be grown in L-broth containing 100ug/ml ampicillin (LA broth).

Bacteria plated onto L-agar can be stored for a few weeks at 4°C.

a) Long term storage is in

- i) stab agar (L-broth with 7g/L agar) at 4°C. or
- ii) in 85% 2xTY, 15% glycerol at -20°C or -70°C.

b) Plating and isolation of single colonies

i) Plating an aliquot of the rat thymocyte cDNA library The library aliquot (stored at -70°C), was thawed rapidly at 37°C. 5-10 volumes of 2xTY were added to dilute the glycerol, and allowed to recover at 37°C for 1 hour. 300 ul of bacteria were spread over a large (14cm) L-agar plate and grown up overnight.

ii) Streaking A sterile platinum loop was dipped into a bacterial culture and streaked across the agar to produce a variation in colony density.

iii) Toothpicking The tip of a sterile wooden cocktail stick was dipped into a single well isolated bacterial colony then lightly pressed onto the surface of one or more agar plates in a regular grid array.

c) Growth of bacteria in liquid culture.

i) Overnight cultures of bacteria were grown by inoculating 10ml of L-broth with a scoop of bacteria from a plate, stab, or culture, and incubating at 37°C overnight.

ii) Chloramphenicol amplification of liquid cultures for plasmid preparation L-broth inoculated with bacteria was incubated at 37°C with vigorous shaking until the OD₆₀₀ of the culture had reached 0.8. Chloramphenicol was added to 200ug/ml and incubation continued overnight at 37°C with shaking.

2.3.1.2 Storage and growth of λ EMBL3

The bacterial host strain for λ EMBL3 is E.coli Q359. This bacterium grows in non-selective T-broth or on T-agar plates (15g/L agar). The bacteriophage is stored and plated in SM buffer.

a) Preparation of Q359. An overnight culture was grown in 10ml T-broth supplemented with 10mM MgSO₄ and 0.2% maltose. The bacteria were pelleted by centrifugation at 3000rpm for 10 minutes, resuspended in 10ml 10mM MgSO₄, 10mM Tris-HCl pH 7.4, and stored at 4°C for several hours.

b) Absorption and plating. Serial dilutions of phage stock in SM buffer were incubated with 50µl of Q359 at 37°C for 30 minutes, to allow the bacteriophage to absorb. The mixture was added to 4ml of melted T agarose (T broth + 0.7% agarose) at 47°C together with 100µl of Q359,

and poured onto small (9cm) T-agar plates. For 14 cm plates, the volume of T agar was 10ml and the volume of plating bacteria 250 μ l. When the agarose had set, the plates were inverted and incubated at 37 $^{\circ}$ C overnight.

c) Picking plaques and plaque purification. Well-isolated plaques were cut out with a sterile scalpel blade and the agar plug dropped into 500 μ l SM buffer and vortexed with 30 μ l of chloroform. The phage supernatant was plated at a dilution which would give well-separated plaques. The tip of a sterile toothpick was dipped into a single well-isolated plaque then lightly pressed onto a lawn of Q359 to produce a grid arrangement of plaques.

d) Storage of λ phage Phage supernatants in SM buffer are usually stable at 4 $^{\circ}$ C for several years, and the phage titre was typically 5×10^6 pfu/ml.

2.3.2 Preparation of phage and plasmid DNA

2.3.2.1 Small scale preparation of plasmid DNA

Plasmid DNA was prepared rapidly on a small scale by the alkaline lysis method described by Maniatis et al. (1982), the only modification being that plasmid copy number in the initial bacterial culture was increased by chloramphenicol amplification. Typical yields were 2-5 μ g of plasmid DNA per 1ml bacterial culture.

2.3.2.2 Large scale preparation of plasmid DNA

Plasmid DNA of sequencing quality was prepared by the alkaline lysis method of Birnboim and Doly (1979) as described by Maniatis et al. (1982), starting with a 500ml culture of bacteria and increasing plasmid copy number by chloramphenicol amplification.

Plasmid DNA was purified by centrifugation to equilibrium on CsCl-ethidium bromide gradients following the protocol described in Maniatis et al. (1982), and after butanol extraction to remove the ethidium bromide, the plasmid was recovered by precipitation with ethanol. Plasmid quality was assessed by analysis of digested and undigested samples on an agarose minigel, and DNA quantitated by absorbance at 260nm assuming an extinction coefficient of 1 OD unit = 50 µg/ml. Recoveries ranged between 100µg-300µg plasmid DNA per 500ml of bacterial culture.

2.3.2.3 Small scale preparation of λ DNA by the plate lysate method.

DNA from the human OX-2 clone in λ EMBL3 was isolated by the plate lysis method essentially as described in Maniatis et al. (1982). Phage were plated onto large T-agarose plates at a dilution which gave confluent plaques ($>2 \times 10^5$ pfus). The phage were recovered by adding 12.5ml SM buffer to each plate and swirling gently at 27°C over 2 hours. Bacterial debris was removed by centrifugation and the method of Maniatis et al. (1982) followed thereafter. A large confluent plate yielded around 10µg of λ DNA.

2.3.3 Preparation of genomic DNA

High molecular weight DNA for Southern blot analysis was prepared from rat and mouse tissues by a modification of the method of Kunkel et al. (1977) as described below.

i) Preparation of thymocytes The thymus was removed from a 7 week old PVG rat and thymocytes teased, The cells were filtered, washed twice in DAB, 0.2% BSA, and cell number determined on a Coulter counter. 5×10^8 cells were recovered by centrifugation and resuspended in 4.5ml of 75mM NaCl, 25mM EDTA, pH8.0.

ii) Preparation of liver. A lobe of liver weighing approximately 0.5g was removed from the rat (7 week old PVG) or mouse (DBA/2 x Balb/c) and diced finely. The tissue was homogenised in 4.5ml of 75mM NaCl, 25mM EDTA, pH8.0.

iii) DNA isolation. To each tissue suspension was added 250 μ l of 10% SDS and 100 μ l of 10mg/ml proteinase K, and the mixture incubated with shaking at 37 $^{\circ}$ C overnight. Detritus was removed by centrifugation, and the viscous supernatant gently extracted twice with an equal volume of phenol pH8.0, and twice with an equal volume of chloroform-butanol (4:1). The solution was made to 0.1M with respect to Na acetate, and two volumes of ethanol added. The DNA formed a gelatinous precipitate which was removed by spooling on to a glass rod and rinsed with 70% EtOH. The liver DNA was resuspended in 5ml and the thymocyte DNA in 30ml of 10mM Tris-HCl pH8.0, 0.1mM EDTA by gentle shaking overnight. DNA quality was assessed by electrophoresis on an agarose minigel and quantitated by absorbance at 260nm assuming an extinction coefficient of 1 OD unit = 50 μ g/ml. Recoveries were 1mg DNA from 5×10^8 thymocytes and 2mg DNA from 0.5g liver. However, the spectrophotometric reading tended to overestimate DNA content by a factor of two, based on ethidium bromide staining of agarose gels.

2.4.4 Digestion of DNA with restriction enzymes.

DNA was digested by the majority of restriction endonucleases in a medium ionic strength buffer. A typical digestion reaction contained 0.1-1 μ g DNA, 2-10 units of enzyme, and 1 μ l of a 10 x stock of medium salt buffer in a final volume of 10 μ l. Occasionally spermidine was added to 4mM final, and DNase-free ribonuclease to 100 μ g/ml. The mixture was incubated at 37 $^{\circ}$ C for 1-3 hours and if the restriction fragments

were not to be analysed directly, the enzyme was inactivated by the addition of 2 μ l of 0.2M EDTA or heating to 70°C for 15 minutes.

2.3.4.2 Digestion of genomic DNA.

Genomic DNA for Southern blot analysis was digested with restriction endonucleases under the buffer conditions suggested by the enzyme supplier. A typical digestion reaction contained 5-10 μ g DNA, 25-50 units of enzyme, 100 μ g/ml BSA, and the recommended salt conditions. The mixture was incubated overnight at 37°C and a 5 μ l aliquot analysed on an agarose minigel to check that digestion was complete. Further enzyme was added if necessary and the sample precipitated with 0.3M Na acetate and 2 volumes of ethanol prior to large scale electrophoresis.

2.3.5 ³²P radiolabelling of DNA

2.3.5.1 5' Labelling of oligonucleotides with T4 kinase

The OX-2 oligonucleotide was labelled with ³²P at its 5' end essentially as described in Maniatis et al. (1982) . 10 pmoles of oligonucleotide (60ng) were labelled with 10 pmoles (50 μ Ci) of [³²P]ATP using 3 units of T4 polynucleotide kinase, and separated from unincorporated nucleotide by electrophoresis on a 20% denaturing polyacrylamide gel, as described below. A typical labelling incorporated 4×10^7 counts, to a specific activity of 4×10^8 cpm/ μ g.

2.3.5.2 5' labelling cDNA with T4 polynucleotide kinase

DNA restriction fragments were labelled at their 5' ends for Maxam and Gilbert sequencing or for use as S1 mapping probes by the protocol described in Maniatis et al. (1982) . Approximately 200ng of a 250-900 bp long DNA fragment with blunt or protruding 5' termini was dephosphorylated by incubation with 5 μ g of alkaline phosphatase in 0.1M

Tris-HCl pH 9.5 for 30 minutes at 37°C. The DNA was extracted twice with phenol-chloroform, precipitated, then labelled with 10 pmoles (50 μ Ci) of [γ -³²P]ATP using T4 polynucleotide kinase. The fragment was separated from unincorporated nucleotides by electrophoresis through a 6% native polyacrylamide gel.

2.3.5.3 Labelling of cDNA with DNA Pol I (Klenow fragment)

DNA restriction fragments were labelled at their 3' ends for Maxam and Gilbert sequencing or for use as S1 mapping probes by the protocol described in Maniatis et al. (1982) .

Approximately 400ng of intact plasmid or 100ng of electroeluted DNA fragment see (section 2.3.7) was digested with a restriction endonuclease which generated protruding 5' termini. The recessed 3' end was then "filled in" using the appropriate [α -³²P]dNTP and 0.5 units of E.coli DNA polymerase I (Klenow fragment). Fragments for Maxam and Gilbert sequencing were made blunt-ended by a cold chase of all 4 dNTPs. The Klenow fragment was inactivated by heating to 70°C for 10 minutes and the DNA either digested with a second enzyme or loaded immediately onto a 6% native acrylamide gel to resolve the labelled fragments.

2.3.5.4 Nick-translation of DNA

Electroeluted fragments of DNA were labelled with ³²P for use as hybridisation probes using a nick-translation kit from Amersham International. The protocol was essentially as described by the manufacturers except that the reaction volume was scaled down to 20 μ l. The labelled DNA was separated from unincorporated nucleotides by chromatography on a column of Sephadex G50 (fine grade) in 10mM Tris-HCl pH7.5, 10mM EDTA. Typically, 3-5 x 10⁶ cpm were incorporated into 25ng DNA, corresponding to a specific activity of 1-2 x 10⁸ cpm/ μ g.

2.3.6 Analysis of DNA by gel electrophoresis

Polyacrylamide and agarose gel electrophoresis is described in detail by Maniatis et al. (1982).

2.3.6.1 Native polyacrylamide gel electrophoresis

End-labelled cDNA fragments were resolved by electrophoresis on 40cm x 0.3mm non-denaturing 6% polyacrylamide gels (acrylamide:bis ratio of 32:1) run in 1 x TBE buffer. Samples were loaded in 15% glycerol, 7mM EDTA, 0.25% bromophenol blue and electrophoresed until the tracking dye had migrated 4/5 of the way down the gel. These conditions resolved DNA fragments of 60-600bp. Radioactive bands were located by autoradiography and the DNA fragment recovered from the gel slice by elution into 0.4ml of 2M ammonium acetate at 37°C overnight.

2.3.6.2 Denaturing gel electrophoresis

a) Separation of labelled oligonucleotide from unincorporated phosphate

The 5' labelled oligonucleotide was boiled for 3 minutes in 1 x denaturing loading dye, and loaded in a 3cm slot of a 40cm x 0.3mm gel containing 8M urea, 20% acrylamide (ratio of acrylamide:bis of 19:1) and 1 x TBE. The gel was run in 1 x TBE until the fast blue marker had migrated 20-30 cm. The oligonucleotide was located by autoradiography and recovered from the excised gel slice by elution into 0.8 ml of 6 x NET buffer at 37°C overnight.

b) Gels for Maxam and Gilbert sequencing

An aliquot of each Maxam and Gilbert degradation reaction in 0.7 x denaturing loading dye, was boiled for 1 minute and loaded onto adjacent 5mm slots of a denaturing 8M urea, 1 x TBE gel of the dimensions and percentage acrylamide given in table 2.1. The gel was run for different lengths of time as indicated, depending on the region of the fragment to be read.

Table 2.1

ELECTROPHORESIS OF DENATURING GELS FOR MAXAM AND
GILBERT SEQUENCING

% gel	Length of gel	Marker dye	Distance dye run	Nucleotides read: position from labelling site
20%	40cm	BB	2/3 down	1 - 35
8%	40cm	BB	to base	25 - 110
8%	40cm	XC	to base	70 - 200
6%	80cm	XC	off base	110 - 250

BB = bromophenol blue marker (fast blue)

XC = xylene cyanol marker (slow blue)

Gels were fixed in 10% methanol, 10% acetic acid for 20 minutes and dried under vacuum before autoradiography.

2.3.6.3 Strand separating gels

Radiolabelled DNA fragments 90-400bp long were strand separated according to Maniatis et al. (1982) on 6% polyacrylamide gels (acrylamide:bis ratio 50:1) in 0.5 x TBE. Samples were boiled for 3 minutes in 30% DMSO, and electrophoresed overnight at 4°C until the fast blue tracking dye had migrated 4/5 down the gel. Radioactive fragments were located by autoradiography and recovered by excision of the gel slice and elution into 0.4 ml of 2M ammonium acetate at 37°C overnight.

2.3.6.4 Agarose gel electrophoresis

Fragments of DNA 0.4-20 kb long were resolved by horizontal agarose gel electrophoresis as described by Maniatis et al., (1982). Samples were resuspended in 1 x normal glycerol loading dye, electrophoresed through 0.7-1.2% agarose gels in 1 x TBE buffer, and visualised by ethidium bromide staining.

a) Agarose minigels (25ml) were used for the rapid analysis of 50-500ng DNA.

b) Large agarose gels (250ml) were used for:

- i) Resolving 1-2 ug digested λ DNA or
- ii) Fractionating a digest of 5-15 μ g genomic DNA for Southern blotting. The gel was run overnight until the bromophenol blue tracking dye had migrated 12-15cm.

2.3.7 Electroelution of DNA fragments from agarose gels

Restriction fragments 0.7-1.4 kb in length were isolated by electroelution from agarose gels following the method of Girvitz et al. (1980). A digest of 20-30ug of plasmid DNA was electroeluted from a wide

lane of a 0.9% agarose minigel according to the published method, with the amendment that the DNA fragment was recovered by precipitation with 0.25 vols of 3M Tris-HCl pH 8.5 and 2 vols of isopropanol in an effort to reduce coprecipitation of agarose contaminants.

2.3.8 Identification of recombinant clones

2.3.8.1 Screening the cDNA library with the OX-2 oligonucleotide probe

One copy of the rat thymocyte cDNA library (see section 2.3.16) was plated over 18 large L-Agar plates at a density of about 10^4 colonies per plate. The colonies were transferred to Whatman 541 filters, amplified on agar plates containing 10-100 μ g/ml chloramphenicol, and the DNA lysed in NaOH and fixed as described by Gergen et al. (1979). Filters were prehybridised in 0.5% NP40, 0.1mg/ml herring sperm DNA according to Wallace et al. (1981) at 38°C and hybridised in the same conditions with 4×10^5 cpm/ml kinase-labelled oligonucleotide specific activity 7×10^8 cpm/ μ g. Filters were washed in 6 x SSC, 0.1% SDS for 30 minutes at 42°C, and dried and autoradiographed overnight at -70°C with intensifying screen and pre-flashed film.

Positive colonies were rescreened by aligning the autoradiograph with the agar plate and touching the positive area with a sterile pipette tip. The bacteria were streaked onto small agar plates, transferred to Whatman 541 filters and screened as above. One copy of the grid was screened, the other remained untouched as a master copy.

2.3.8.2 Screening cDNA clones with a cDNA probe

The rat thymocyte cDNA library transferred to Whatman 541 filters (section 2.3.8.1, above) was subsequently screened with the OX-2 cDNA probe. Four filters were prehybridised in the buffer conditions shown in table 2.2 then hybridised overnight in the same buffer containing 1.5 x

Table 2.2

A SUMMARY OF PREHYBRIDISATION AND HYBRIDISATION CONDITIONS USED FOR SCREENING FILTERS WITH cDNA PROBES

Section	Hybridisation to:	Deionised formamide	SSC	Denhardt's solution	Denatured herring sperm DNA	SDS	Dextran sulphate	Temp	Time
2.3.8.2	cDNA library transferred to Whatman 541	P 50%	5 x	5 x	500µg/ml	1%	-	42 ^o C	>5 hr
		H 50%	5 x	5 x	500µg/ml	-	-	42 ^o C	O/N
2.3.8.3	λ plaques transferred to nitrocellulose	P 50%	3 x	10 x	20µg/ml	-	-	42 ^o C	>5 hr
		H 50%	3 x	1 x	20µg/ml	-	5%	42 ^o C	O/N
2.3.10.1	Southern blot of genomic DNA on nitrocellulose	P 50%	2 x	2 x	100µg/ml	-	-	42 ^o C	O/N
		H 50%	2 x	1 x	50µg/ml	-	5%	42 ^o C	O/N
2.3.10.2	Southern blot of λ DNA on nitrocellulose	P 50%	3 x	2 x	100µg/ml	-	-	42 ^o C	>5 hr
		H 50%	3 x	1 x	50µg/ml	-	5%	42 ^o C	O/N
2.3.1.11	Northern blot of RNA on Gene - Screen membrane	P 50%	5 x	2 x	500µg/ml	1%	-	42 ^o C	O/N
		H 50%	5 x	1 x	500µg/ml	1%	-	42 ^o C	O/N

KEY : P = Prehybridisation H = Hybridisation O/N = Overnight

10^5 cpm/ml of the nick-translated Bam-Xba 900bp coding region fragment from clone pX2/11. The filters were washed in 3 x SSC at 65°C for 20 minutes, then 0.2 x SSC for 1 hour, and autoradiographed overnight at -70°C with preflashed film and intensifying screen.

2.3.8.3 Screening λ plaques with a cDNA probe

Bacteriophage λ EMBL3 were plated on a lawn of E.coli Q359 and plaques transferred in replica copies to Schleicher and Schuell 0.45 μ m pore nitrocellulose membranes. The filters were fixed following the method of Benton and Davis (1977) in 0.5M NaOH, 1.5M NaCl for 60 s followed by 0.5M Tris-HCl pH7.4, 1.5 M NaCl for 5 minutes and 2 x SSPE for 5 minutes. After baking at 80°C for two hours the filters were pre-hybridised then hybridised with 3×10^5 cpm/ml nick-translated cDNA probe (Bam-Xba 900 bp coding region fragment of clone pX2/11) in the buffer conditions shown in table 2.2. The filters were washed to a stringency of 3 x SSC at 65°C for 30 minutes, then 1 x SSC at 65°C for 30 minutes, followed by autoradiography at -70°C with intensifying screen and preflashed film.

The autoradiograph was aligned with the agar plate and the corresponding plaque removed into SM buffer for replating at a lower dilution and rescreening. The positive colony was plaque-purified and toothpicked onto a grid arrangement for a final screening.

2.3.9. Hybridisation of OX-2 oligonucleotide with dried down agarose minigels

Restriction fragments bearing the OX-2 oligonucleotide site were identified and mapped by the method of Singer-Sam et al. (1983). Restriction enzyme digests of putative OX-2 clones were fractionated on a 0.8% agarose minigel. The gel was transferred onto Whatman paper and

dried down. After removal of the Whatman paper the gel was hybridised with 1×10^5 cpm/ml of radiolabelled OX-2 oligonucleotide (7×10^8 cpm/ μ g) employing the same hybridisation and washing conditions as for screening Whatman 541 filters, Section 2.3.8.1. The gel was autoradiographed at -70°C with pre-flashed film and intensifying screen.

2.3.10 Southern blot analysis of DNA

The transfer of size-fractionated DNA fragments from agarose gels to nitrocellulose, and subsequent hybridisation with ^{32}P labelled cDNA probes is essentially as described by Southern (1979) with the following slight modifications.

2.3.10.1 Southern transfer of genomic DNA

7-15 μ g of genomic DNA digested with restriction endonucleases was fractionated on a 250ml 0.8% agarose gel. The DNA was denatured in the gel by soaking in 0.2M NaOH, 0.6M NaCl for 2 hours and 1M Tris-HCl pH7.4, 1.5M NaCl for 2 hours then transferred to Millipore 0.45 μ m pore nitrocellulose according to Southern (1979) in 20 x SSC buffer. The filter was baked, prehybridised, then hybridised with 1×10^6 cpm/ml nick-translated cDNA fragment (Bam-Xba 900 bp coding region fragment of clone pX2/11) in the buffer conditions described in table 2.2. The blot was washed at 65°C at a stringency of 0.2 x SSC to detect identical sequences, or 2 x SSC to detect cross-hybridisation.

2.3.10.2 Mapping λ EMBL3 genomic clone by Southern blotting

0.5-1 μ g HX2/ λ EMBL3 DNA was digested by restriction endonucleases and fractionated on a 250ml 0.9% agarose gel. The DNA was denatured in the gel by soaking in 0.25M HCl for 5', then 0.5M NaOH for 15'. The DNA was transferred to Millipore 45 μ m pore nitrocellulose in 20 x SSC (Southern, 1979), and the baked filter prehybridised then hybridised

with nick-translated OX-2 probe (Bam-Xba 900 bp coding region fragment of clone pX2/11) in the buffer conditions in table 2.2. The filter was washed under the same conditions as for screening λ plaques on nitrocellulose.

2.3.11 Preparation of RNA

RNA was prepared from rat tissues by the guanidine-isothiocyanate method of Chirgwin et al. (1979).

2.3.11.1 Preparation of tissues

a) Brain One brain from a 7 week old PVG rat was homogenised in 15ml of guanidine lysis buffer (4M guanidine isothiocyanate, 25mM Na citrate, 5mM EDTA, 0.1M β -mercaptoethanol, 0.5% SDS, pH7.0) and centrifuged at 6000 x g for 20 minutes to pellet debris.

b) Leucocytes i) Total thymocytes Thymocytes were teased from the thymuses of 5 PVG rats, filtered, and washed once in DAB 0.1% BSA before counting and checking viability. 3.6×10^9 thymocytes were obtained from 5 rats.

ii) T cell blasts [prepared by G. Newton] A preparation of thymocytes from 8 PVG rats was prepared as above, but in RPMI 1640 medium with 0.5% fetal calf serum. Cells were resuspended at a density of 2×10^6 /ml and Concanavalin A added to a final concentration of 5 μ g/ml. The cells were cultured as a still suspension in 5 x 200ml flasks over 2 days then recovered by centrifugation. The cells were centrifuged through an isopaque-ficoll gradient and the interface cells collected and washed in PBS, 0.2% BSA prior to counting and checking viability. 1×10^9 cells were recovered.

iii) B cell [prepared by G. Newton] Cervical and mesenteric lymph nodes were removed from 200 Sprague-Dawley rats and the cells teased. Dead

cells were removed by treatment with sorbitol (von Boehmer and Shortman, 1973) and the cells washed as in (i) above. Cells were rosetted with W3/13 antibody-coupled SRBC (Mason, 1981) to remove all the T cells, and the B cells recovered by centrifugation. The supernatant was flash-lysed to remove the SRBC, the B cells washed and counted, and the purity of the preparation assayed by FACS analysis. 5.7×10^9 B cells were obtained from 200 rat lymph nodes.

3.7ml of guanidine lysis buffer was added per $2-8 \times 10^8$ cells, and the mixture vortexed and passed through a 19 gauge needle.

2.3.11.2 Preparation of RNA

Cesium chloride was added to the lysis mixture to a final concentration of 2.2M. RNA was purified by ultracentrifugation of 3.5ml aliquots of the mixture through a cushion of 1.2ml of 5.7M CsCl, 0.1M EDTA pH7.5 (Chirgwin et al., 1979) at 35,000 rpm in a Beckman SW 50.1 rotor for 12-16 hours at 20°C. The RNA was resuspended in 3 ml 10mM Tris-HCl pH7.5, 5mM EDTA, 0.1% SDS, extracted twice with chloroform:butanol (4:1), then precipitated twice in 0.3M Na acetate plus 2.5 volumes of ethanol. RNA quality was analysed on an agarose minigel and the recovery quantitated by absorbance at 260 nm, assuming an extinction coefficient of 1 OD unit = 40µg RNA. Typical recoveries were 400µg RNA per rat brain, 1.4mg RNA per 3×10^9 thymocytes, 2mg RNA per 5×10^9 B cells, and 7mg RNA per 1×10^9 T blasts. T blasts are more active cells than thymocytes, hence the increased yield of RNA.

2.3.12 Northern blot analysis of RNA

Size-fractionated RNA was transferred to Gene Screen membrane and analysed with cDNA probes by the technique of "northern blotting".

a) Electrophoresis of RNA through denaturing formaldehyde gels

The protocol is detailed in the instruction manual supplied with 'Gene Screen' and is an adaption of the method of Lehrach et al. (1977). 15µg RNA was resuspended in 50% formamide, 6% formaldehyde, 1 x MOPS buffer and heated at 60°C for 15 minutes. Bromophenol blue tracking dye was added and the sample was loaded onto a 250ml 1.2% agarose gel containing 1 x MOPS buffer, 6% formaldehyde. Bacterial and eukaryotic ribosomal RNA size markers were run in the outside lanes. The gel was electrophoresed overnight until the marker dye had migrated 12-14cm. The marker lanes were cut from the gel, stained with ethidium bromide, and analysed under UV light.

b) Transfer and hybridisation of RNA

The following protocol for the transfer and hybridisation of RNA onto Gene Screen is based on the instructions in the original version of manufacturer's handbook. The gel was washed in dH₂O and fixed in 0.05M NaOH for 30', followed by 0.1M Tris-HCl pH7.4 for 30', then 0.01M Na phosphate buffer (0.01M Na₂HPO₄/NaH₂PO₄ pH6.5) for 20'. RNA was transferred to Gene Screen membrane by capillary blotting in 0.01M Na phosphate buffer overnight, and the membrane baked for 2 hours at 80°C. Prehybridisation and hybridisation buffer conditions are shown in table 2.2. The filter was hybridised with 1 x 10⁶ cpm/ml nick-translated OX-2 cDNA probe (Bam-Xba 900 bp coding region fragment of clone pX2/11) and washed in 2 xSSC, 0.5% SDS at 65°C for 30' and 0.2 x SSC at room temperature for 1 hour. Autoradiography was at -70°C for 1-10 days with preflashed film and intensifying screen.

2.3.13 S1 nuclease mapping of RNA

S1 mapping was used to investigate the splice points within OX-2 mRNA from various rat tissues. The method was essentially as described by Favaloro et al. (1980) and Berk and Sharp (1977).

^{32}P labelled single- or double-stranded DNA probes were designed and labelled as described in the legends to figures 7.2 and 7.3 in Chapter 7. 15ug of total poly A+ RNA was hybridised with a small quantity of probe (5-15 cps on a bench minimonitor) at 43°C for single stranded probes or 51°C for double stranded probes. Unhybridised single stranded DNA was digested with 70 units of S1 nuclease and the products analysed on a 6% denaturing polyacrylamide gel together with restriction fragment size markers.

2.3.14 Sequencing DNA by the method of Maxam and Gilbert

End labelled DNA fragments were sequenced by the base-specific chemical cleavage method of Maxam and Gilbert (1980).

Labelled DNA fragments for sequencing were generated by digestion of 400ng of plasmid DNA or 100ng of electroeluted DNA with the appropriate restriction enzyme, followed by :

- i) 3' labelling by filling in recessed 3'ends using DNA pol 1 (Klenow fragment) and [α ^{32}P]dNTP.
- or ii) 5' labelling of overhanging or blunt 5' ends using [γ ^{32}P]ATP and T4 polynucleotide kinase.

To produce fragments labelled at one end only, the DNA was then either :

- i) Redigested with a second enzyme and fractionated on a 6% native polyacrylamide gel.
- or ii) strand separated.

Labelled fragments were ethanol precipitated and subjected to the following base-specific chemical modifications described by Maxam and Gilbert (1980).

T + C : Hydrazine

C only : Hydrazine in the presence of NaCl

G + A : Formic acid

G only : Dimethylsulphate

Modified bases were cleaved by piperidine and excess reagents removed by precipitation and vacuum drying. Each sequencing reaction was resuspended in 15µl of 0.7 x denaturing formamide loading buffer. 2µl aliquots were run in adjacent 5mm wells of a denaturing polyacrylamide gel as described in section 2.3.6 and summarised in table 2.1.

2.3.15 Oligonucleotide synthesis

The OX-2 oligodeoxyribonucleotide was prepared with the assistance of Keith Gould in Professor G. Brownlee's laboratory (Sir William Dunn School of Pathology, Oxford). Oligonucleotide synthesis was by the solid-phase phosphotriester method (Gait et al., 1980, Gait, 1984) using a continuous flow method on a Kieselguhr-polyamide support and Omnifit apparatus. Protected mononucleotides were coupled to the solid support in the presence of the coupling agent MSNT and the catalyst 1-methylimidazole. For mixed positions A/G/C/T, the ratio of A and G monomers were increased over C and T to compensate for the slightly slower reaction rate of purines.

The completed nucleotide was deprotected, desalted on a Biogel P2 column in 20% ethanol, and purified by preparative scale gel electrophoresis on a 40 cm x 1.5 mm thick 20% polyacrylamide 7M urea denaturing gel. The oligonucleotide was visualised under long-wave UV.

eluted, and deionised on a DEAE Sephadex A25 column. After desalting on the Biogel P2 column, the oligonucleotide was resuspended in dH_2O and stored in aliquots at -20°C . Quantitation was on the basis of absorbance at 260 nm assuming an extinction coefficient of 1 OD unit = 30 μg .

2.3.16 Construction of the rat thymocyte cDNA library

The cDNA library was made in our laboratory by Drs M. Thomas and A.N. Barclay, and details of its construction are given in Thomas et al. (1985). Briefly, RNA was prepared from Sprague-Dawley thymocytes, and the Poly A⁺ fraction isolated by chromatography on oligo-dT cellulose. The first strand of cDNA was synthesised using reverse transcriptase and oligo(dT)₁₂₋₁₆, and the second strand using 'loopback' self-priming and DNA polymerase 1. Successive treatments with S1 nuclease and the Klenow fragment of DNA polymerase 1 created blunt-ended cDNA. Fractions larger than 1kb were selected by centrifugation through a 10%-30% (w/v) sucrose gradient, then blunt-end ligated into the vector pAT153/Pvu II/8 (Anson et al., 1984, and shown in figure 2.1A). Ampicillin-resistant transformants were selected by growth in LA broth, and the library amplified approximately 50 times before aliquotting and storage in 15% glycerol at -70°C . There were approximately 160,000 transformants prior to amplification, and >90% of these contained cDNA inserts.

2.3.17 Construction of the human genomic library λ DA Mbo 1

The library was constructed by Colin Sharpe of Professor Brownlee's group (Sir William Dunn School of Pathology, Oxford) and details are described by Sharpe (1985). Total genomic DNA was obtained from human peripheral blood lymphocytes and digested partially with Mbo 1. DNA fragments between 15-25 kb were selected after centrifugation

Figure 2.1

A : Restriction map of cloning vector pAT153/Pvu II/8

The blunt-ended cDNA insert is ligated into the Pvu II site of pAT153/Pvu II/8 (Anson et al., 1984) and is excisable with [EcoR 1 or Cla 1 or Hind III] plus Bam H1.

B : Restriction map of cloning vector λ EMBL3

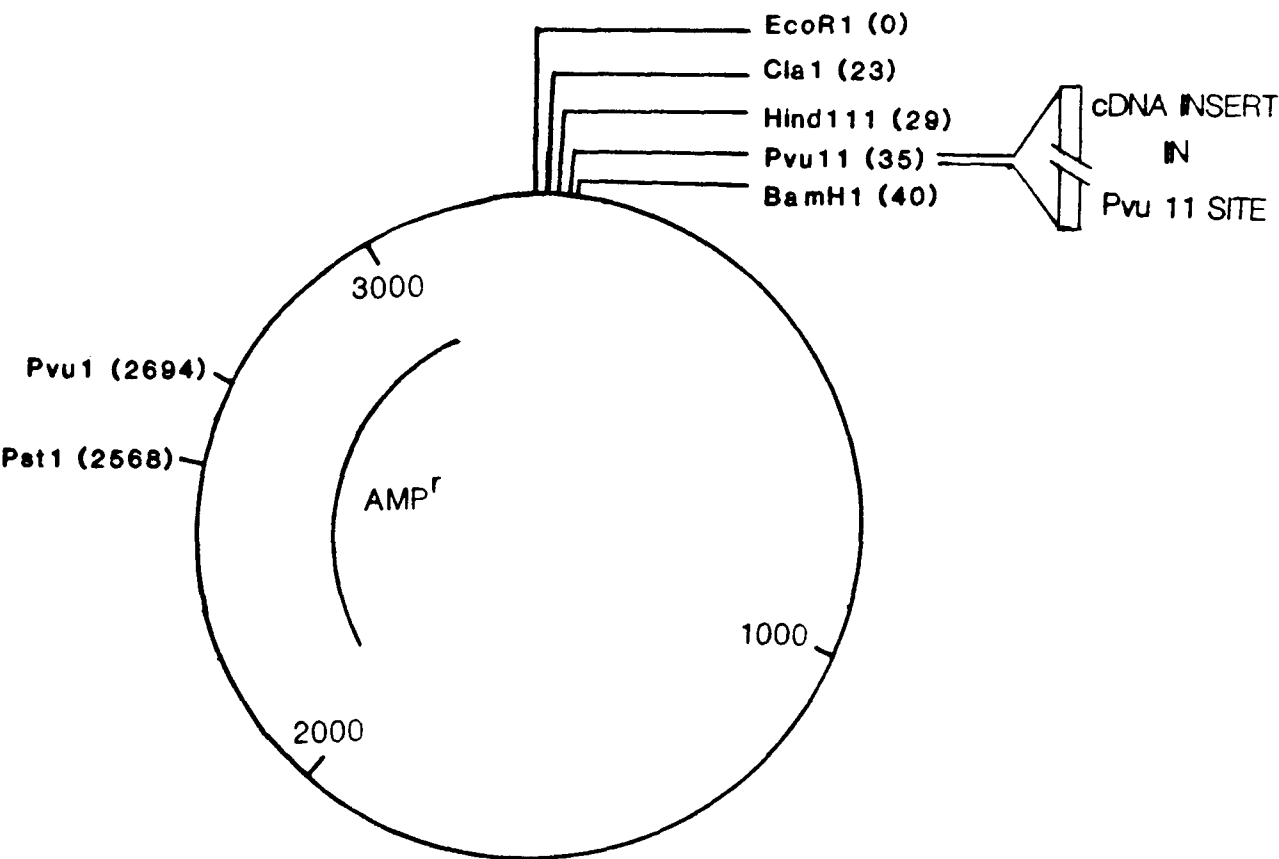
Partial Mbo 1 digested genomic DNA is ligated into the BamH1 cut vector λ EMBL3 (Frishauf et al., 1983) and is excisable with Sal 1.

L = Left end (long arm of lambda)

R = right end (short arm of lambda)

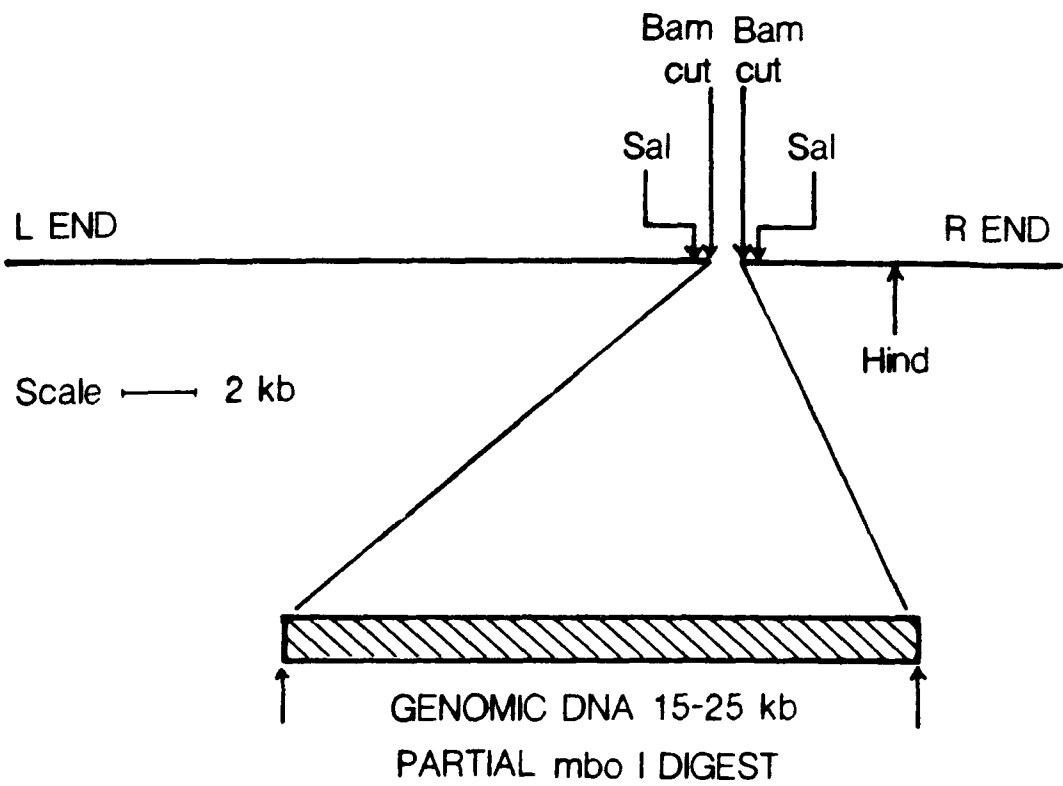
Figure 2.1

A RESTRICTION SITES IN THE CLONING VECTOR pAT153/Pvu11/8



TOTAL LENGTH 3322 NUCLEOTIDES

B RESTRICTION MAP OF λ VECTOR EMBL3



through a 10%-40% sucrose gradient, and ligated into Bam H1 cut vector λ EMBL3 (Frishauf et al. 1983, shown in figure 2.1B). Recombinant plaques were selected for the spi- phenotype on the bacterial strain Q359 which carries the P2 lysogen. The library was plated without amplification, and contained approximately 4×10^5 pfu. 1 genome has been estimated to be represented by approximately 3×10^5 pfu.

CHAPTER 3

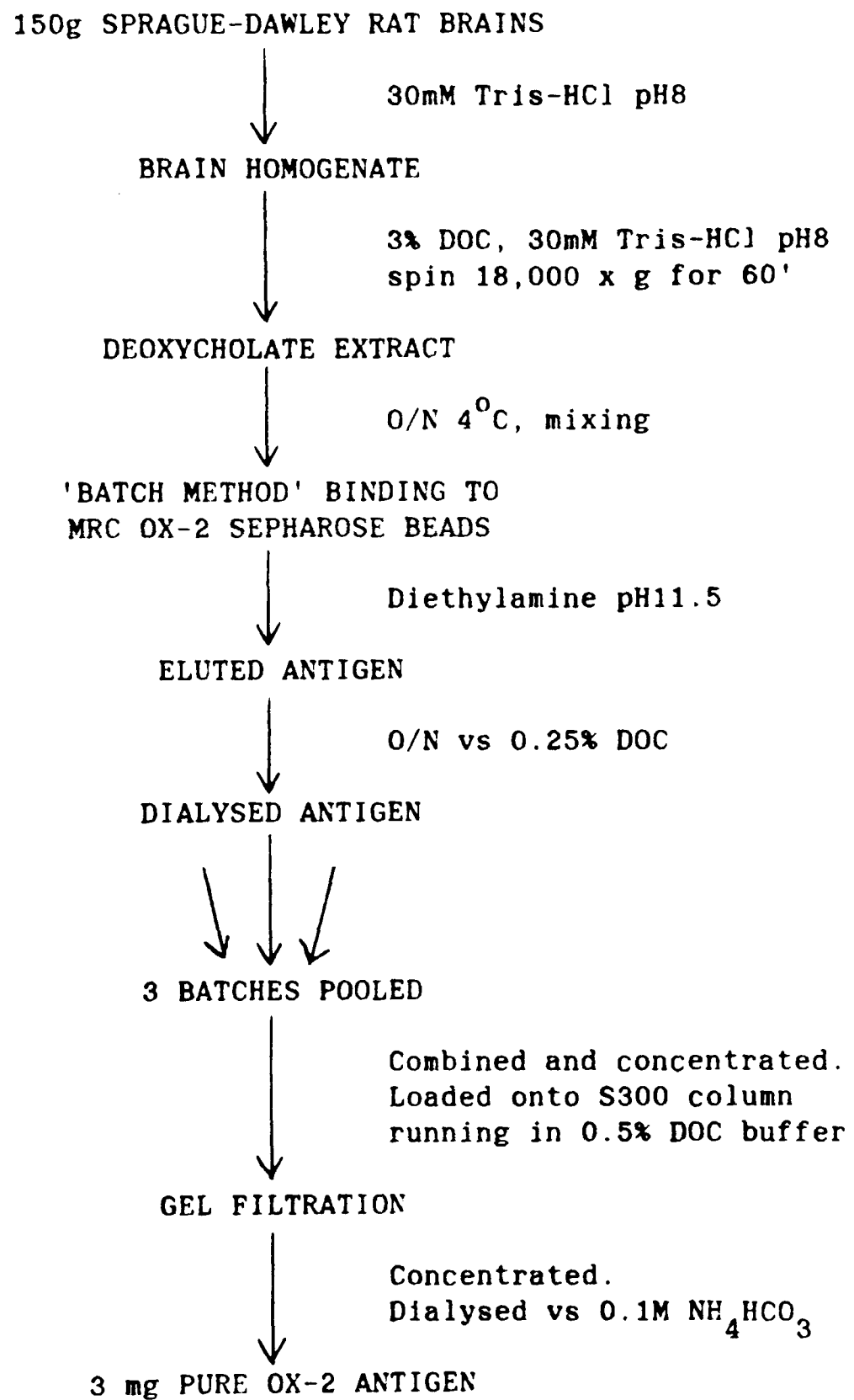
PARTIAL AMINO ACID SEQUENCE OF OX-2 ANTIGEN

3.1 INTRODUCTION TO CHAPTER 3

The mouse monoclonal antibody MRC OX-2 recognises a glycoprotein of approximate M_r 45,000, with an unusual pattern of expression on brain, lymphoid, and other tissues in the rat. As discussed in the Introduction section 1.6, the sequence of the OX-2 protein is of interest because similarities in tissue distribution and biochemical properties between OX-2 and Thy-1 antigens suggest an underlying structural homology between these two molecules. One strategy for determining the complete amino acid sequence of OX-2 is to use an oligonucleotide probe to isolate a cDNA clone encoding the protein, and the advantages of this approach have been discussed in the Introduction, section 1.7. The first step is to purify sufficient protein to generate reliable peptide sequence for designing the oligonucleotide, and the approach is feasible for OX-2 because the antigen is readily isolated by means of the monoclonal antibody (Barclay and Ward, 1982). This Chapter describes the preparation of OX-2 antigen from rat brain by standard procedures using solubilisation in DOC and antibody affinity chromatography, but on the scale required here a "batch" rather than a "column" step is introduced at the affinity chromatography stage to accomodate the large preparation needed. Tryptic peptides are isolated and sequenced, and their sequences examined for homology with Thy-1 and immunoglobulin.

Figure 3.1

PURIFICATION OF OX-2 ANTIGEN FROM RAT BRAIN



3.2 PURIFICATION OF OX-2 ANTIGEN FROM RAT BRAIN

The method, based on the protocol of Barclay and Ward (1982), is summarised as a flow chart in figure 3.1.

3.2.1 An estimation of the amount of OX-2 antigen required for sequence studies

A 7-8 week old AO rat has seven fold more OX-2 in the brain than in the thymus (Barclay and Ward, 1982) making rat brain a suitable source of OX-2 antigen. On the basis of the authors' data, 300 rat brains (450g) should yield about 3mg OX-2, equivalent to 100 nmoles of protein. Assuming peptide purification involves two chromatography steps each with a recovery of 50%, the overall yield of peptide would be 25 nmol, and this amount is adequate for the spinning cup sequencer.

3.2.2 Preparation of rat brain OX-2 antigen using the "batch" method of affinity chromatography

Sprague-Dawley rat brains were homogenised in 30mM Tris-HCl pH8, and membrane proteins solubilised by the addition of 3% deoxycholate (DOC). The supernatant after centrifugation, referred to as 'rat brain DOC extract' typically contained around 75% of the initial antigenic activity as shown in table 3.1A.

OX-2 antigen was then purified by affinity chromatography on MRC OX-2 antibody coupled to Sepharose beads. The conventional "column" approach, in which the extract is passed through a column of packed beads, is too slow to handle the one-to-two litre quantities of extract generated by this large scale preparation. Instead, a "batch" approach was employed, similar to the method used by Mellman and Unkeless (1980) for the purification of the Fc receptor. Beads and batches of extract were mixed gently together in bulk, then the beads recovered for washing

Table 3.1

PURIFICATION OF BRAIN OX-2 ANTIGEN

Antigenic activity was determined by inhibition of the binding of MRC OX-2 monoclonal antibody to fixed rat thymocyte target cells. The brain homogenate sample was disrupted in 3% DOC, 30mM Tris-HCl, pH8.0, prior to making serial dilutions for assay, and all assays included 0.2% DOC.

One unit of OX-2 antigenic activity is defined as the amount of antigen needed to give 50% inhibition of the assay, so the total units in a sample is the product of the dilution factor giving 50% inhibition and the sample volume.

High concentrations of DOC interfere slightly with the assay, so a titration of 3% DOC buffer alone was included as a control and samples with the same inhibition curve as the DOC buffer alone were considered to contain zero antigenic activity.

Protein content was determined by

- (a) Lowry assay
- or (b) Amino acid analysis

PURIFICATION OF BRAIN OX-2 ANTIGEN

A: RECOVERY OF ANTIGENIC ACTIVITY AFTER ELUTION OF MRC OX-2 BEADS

(ONE BATCH)

B: RECOVERY OF ANTIGENIC ACTIVITY AFTER GEL FILTRATION

(THREE BATCHES POOLED)

A:

STAGE	PROTEIN mg	UNITS OF ANTIGENIC ACTIVITY $\times 10^{-3}$	YIELD OF INITIAL ANTIGENIC ACTIVITY	RELATIVE SPECIFIC ACTIVITY
Brain homogenate	15,200 ^a	38.4	100 %	1.0
Brain homogenate + 3% DOC	8,680 ^a	32.2	83 %	1.5
DOC extract	7,500 ^a	22.4	58 %	1.2
After mixing with MRC OX-2 beads	N.D.	~ 0	~ 0	~ 0
Eluted from MRC OX-2 beads	3.9 ^a	13.7	36 %	1400

B:

Combined eluate from 3 preps	N.D	47.3	-	-
After gel filtration and dialysis	3.1 ^b	36.4	31 %	4650
OVERALL				

Figure 3.2

TIME COURSE OF THE BATCH MIXING METHOD

1.1 L of rat brain DOC extract was 'batch mixed' (methods, section 2.2.1) with 10ml MRC OX-2 IgG-Sepharose CL 4B beads. 1 ml aliquots were removed at regular intervals, and centrifuged at 9,000 x g for 1 minute to sediment the beads. The amount of OX-2 antigen remaining in the supernatant was determined by inhibition of binding assay.

One unit of antigenic activity is defined as the amount of antigen needed to give 50% inhibition of the assay.

Note that 3% DOC buffer has a slight effect on the assay, and a background of 3 units of activity is contributed by the DOC buffer alone.

TIME COURSE OF THE BATCH MIXING METHOD

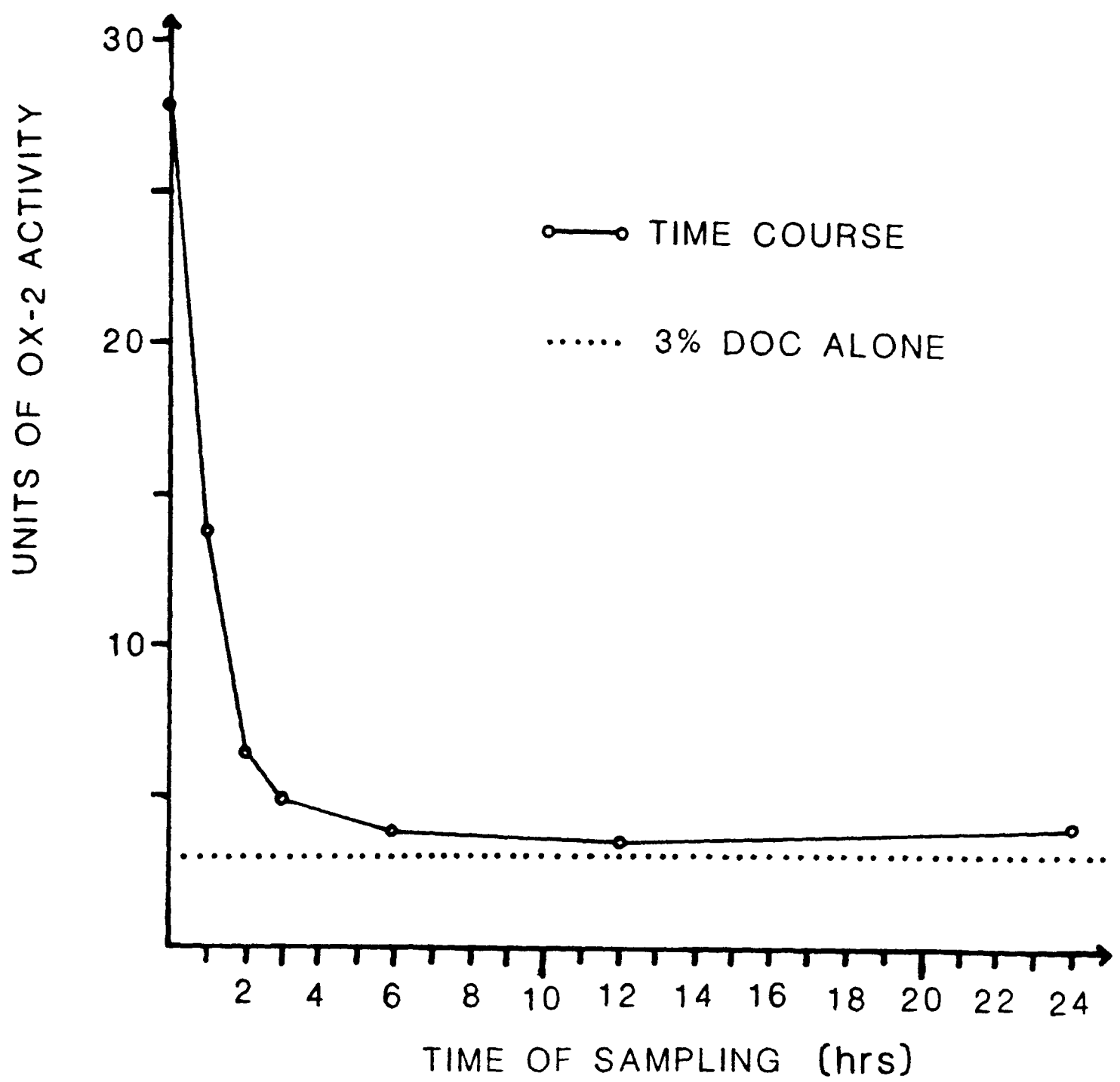


Figure 3.3

GEL FILTRATION AND SDS-PAGE OF AFFINITY PURIFIED OX-2 ANTIGEN

A : Gel filtration of OX-2 on Sephacryl S300

Approximately 3mg purified OX-2 antigen was separated on a Sephacryl S300 column (2.5 x 77cm) in 0.5% DOC buffer. 4ml fractions were collected and fractions 40 to 46 were pooled.

B : SDS PAGE analysis of the purification of OX-2 antigen

Samples from each stage in the purification of OX-2 antigen were reduced with DTT and analysed on a 15% SDS-polyacrylamide gel. The gel was stained with Coomassie Blue.

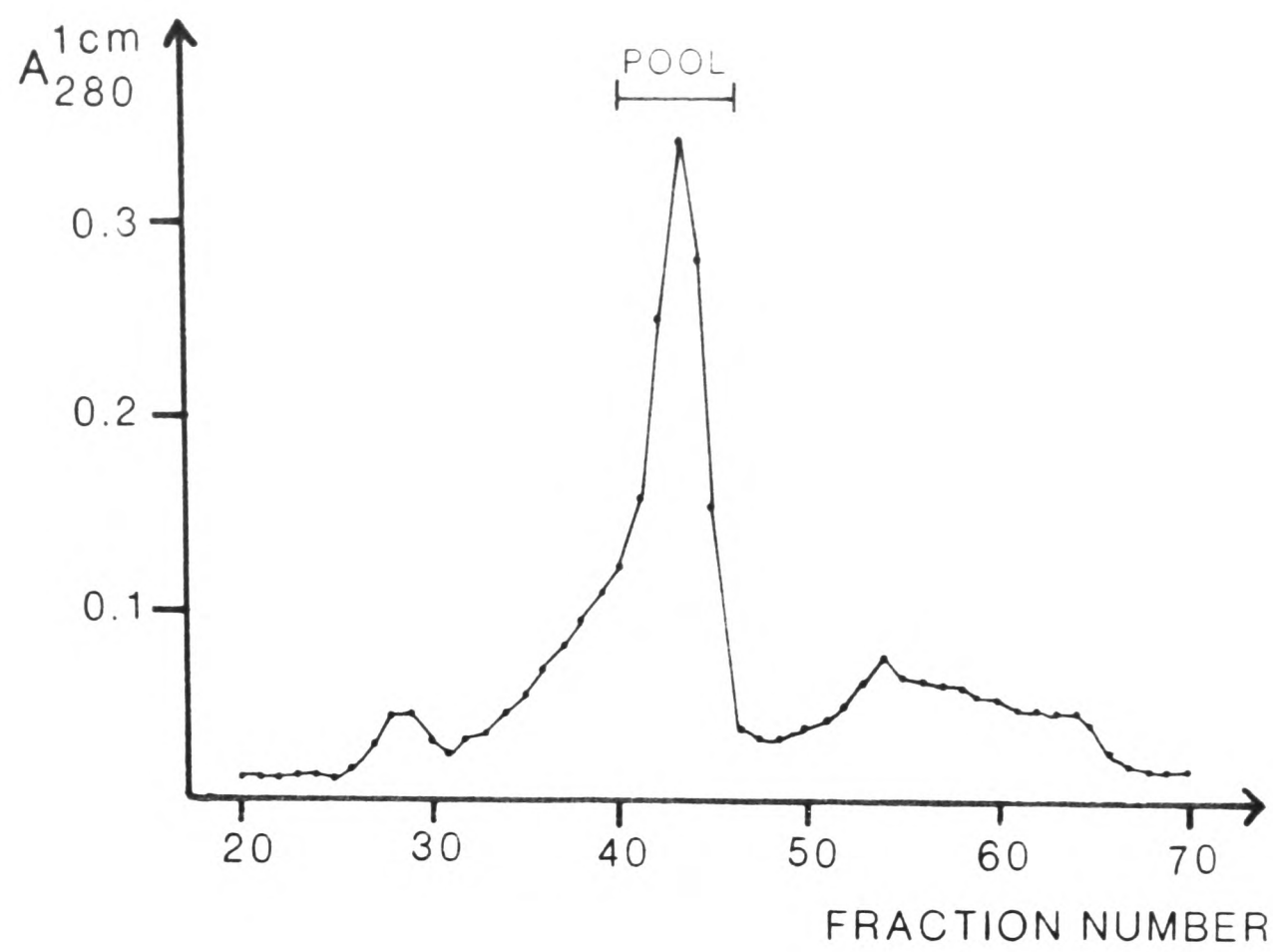
Lane A : DOC extract. Membrane proteins from 150g rat brains were solubilised in 1.1L 3% DOC, 30mM Tris-HCl, 0.02% NaN₃, pH8.0, and a 1ul aliquot run in lane A.

Lane B : Eluted OX-2 antigen. Three preparations of eluted OX-2 antigen purified by the 'batch method' of affinity chromatography were pooled and concentrated to 5.5ml. A 6ul aliquot was run in lane B.

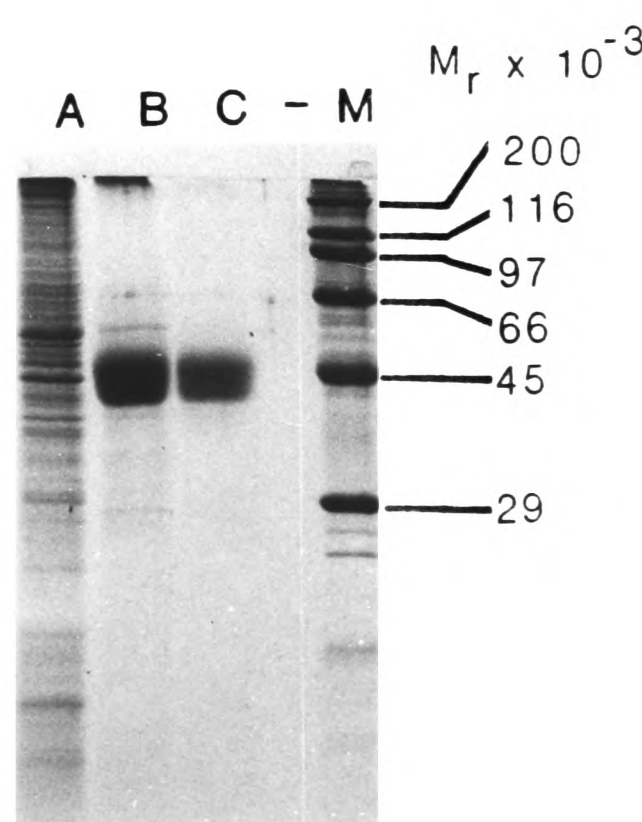
Lane C : Post gel filtration. Fractions 40-46 from gel filtration (figure 3.3A, above), were precipitated with 3 volumes of ethanol. The protein was resuspended in 1 ml dH₂O and a 1ul aliquot run in lane C.

M = standard protein size markers.

A GEL FILTRATION OF OX-2 ON SEPHACRYL S300



B SDS PAGE: PURIFICATION OF OX-2 ANTIGEN



and elution. The length of mixing time required was established by removing an aliquot of extract at regular intervals, and assaying for residual OX-2 antigenic activity. The time course plotted in figure 3.2 shows that 98% of the antigenic activity was depleted from 1L batch of extract in 6 hours.

The beads were returned to a conventional column arrangement for washing, and the antigen eluted with diethylamine pH11.5. The recovery of antigenic activity (table 3.1) compares favourably with the traditional column method - no OX-2 antigen remained in the extract and 60% of the activity was recovered in the eluate. The eluted protein is undegraded and free from major contaminants. Antibody-coupled beads which have been batch-mixed ten times show no signs of damage to the beads or loss of binding capacity.

3.2.3 Final purification of rat brain OX-2 antigen by gel filtration

Three batches of eluted OX-2 antigen were combined at this stage, and separated from minor low and high molecular weight contaminants by gel filtration. The elution profile of the S300 column is illustrated in figure 3.3A. The peak fractions were pooled and dialysed to remove DOC and shown to be essentially free from protein contaminants by SDS-PAGE analysis (figure 3.3B track C). The overall recovery of antigenic activity from the brain homogenate was 31% (table 3.1B) and about 1mg OX-2 was extracted per 150g brains, with an overall purification factor of 4600 which was similar to that obtained previously (Barclay and Ward, 1982).

3.3 PREPARATION AND PURIFICATION OF TRYPTIC PEPTIDES

The ideal length OX-2 peptide for sequence studies is 20-30 amino acids. Although the minimum sequence required for oligonucleotide

design is 5-6 amino acids (Introduction section 1.7), longer peptide sequences increase the choice of suitable probes, and the additional protein data subsequently confirms whether the correct cDNA clone has been isolated. In addition, the longer the peptide sequence the more reliable it is for assessing possible sequence homologies with immunoglobulin (Introduction section 1.5.1).

On the basis of the published amino acid composition of brain OX-2 (Barclay and Ward, 1982), if cleavage at lysine residues is prevented by succinylation, then digestion with trypsin at arginine residues (4 residues per 100) should generate peptides of average length 20 amino acids. Thus about 11 peptides should be generated from the OX-2 antigen, assuming it is about 250 amino acids in length, as estimated from the M_r of 41,000 and a carbohydrate content of 24% (Barclay and Ward, 1982).

3.3.1 Reduction, alkylation and succinylation of brain OX-2 antigen

Purified OX-2 antigen (1.8 mg, 70 nmoles) was reduced with DTT under denaturing conditions in guanidine buffer, and free sulphhydryls converted to S-carboxymethyl cysteine with 2-³H iodoacetic acid. Lysine groups were subsequently succinylated with succinic anhydride. A fine white precipitate observed after the reduction and alkylation stage contained no radioactivity and was removed by centrifugation. 85% of the ³H labelled reduced and alkylated protein was recovered after succinylation.

3.3.2 Digestion with trypsin and preliminary separation of peptides by gel filtration

Reduced, alkylated, succinylated OX-2 antigen was digested with 2% (w/w) trypsin and peptides separated by gel filtration on Biogel P30 in

0.1M NH_4HCO_3 (figure 3.4). Fractions within the horizontal bars were pooled and numbered T1 to T13.

A 5% aliquot of each pool was reserved for amino acid analysis, and for a typical peptide such as T5, the overall recovery after the gel filtration stage was 60% (40 nmole peptide).

The subsequent fate of each peptide pool is summarised in figure 3.5:-

- 1) Peptides T1-T8 analysed in Section 3.3.3
- 2) Peptides T9-T13 discarded because:
 - a) T9, T10 contained a complex mixture of long peptides.
 - b) T11 contained only short peptides
 - c) T12, T13 contained no peptides

3.3.3 Analysis of tryptic peptide pools T1-T8 by reverse phase HPLC

A 5% aliquot of each peptide pool T1 to T8 was further analysed by reverse phase HPLC on a Waters u Bondapak C_{18} column eluted with a linear acetonitrile gradient in 10mM NH_4 acetate as shown in figure 3.6. Major peaks were collected and amino acid composition and ^3H recovery determined. Peptide pools which were found to contain several components were then fractionated by preparative scale HPLC while those which contained only one major peptide were generally sequenced directly from the P30 pool as detailed below and summarised in figure 3.5.

- 1) Peptide T8 not analysed further as the HPLC profile was complex.
- 2) Peptides T5-T7. P30 pool sequenced directly.
 - a) T5, T6: P30 pools were judged to have only one major component.
 - b) T7: could not be eluted from the HPLC column (figure 3.6) so there was no alternative other than to sequence the unfractionated P30 pool.

Figure 3.4

GEL FILTRATION OF OX-2 TRYPTIC PEPTIDES ON BIOGEL P30

2mg (70 nmol) of reduced, alkylated and succinylated brain OX-2 antigen was digested with 2% w/w trypsin in 2 aliquots over 6 hours at 37⁰C. The peptides were loaded immediately onto a Biogel P30 column in 0.1M NH₄HCO₃ at room temperature. The flow rate was 3ml/hr and fraction volume³ 0.75ml.

Fractions corresponding to peptide peaks T1 to T13 were pooled as indicated by the horizontal bars.

GEL FILTRATION OF OX-2 TRYPTIC PEPTIDES ON BIOGEL P30

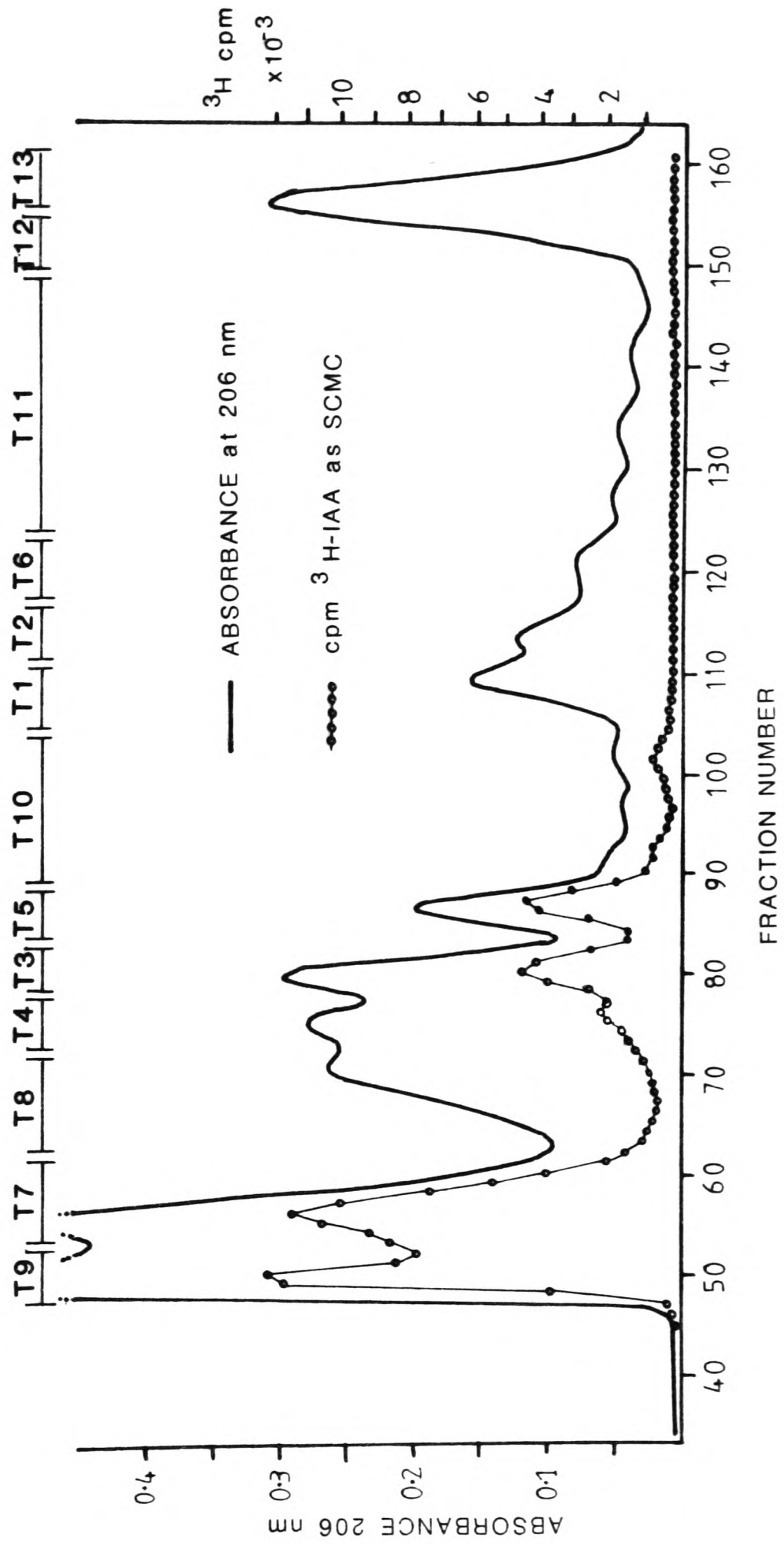
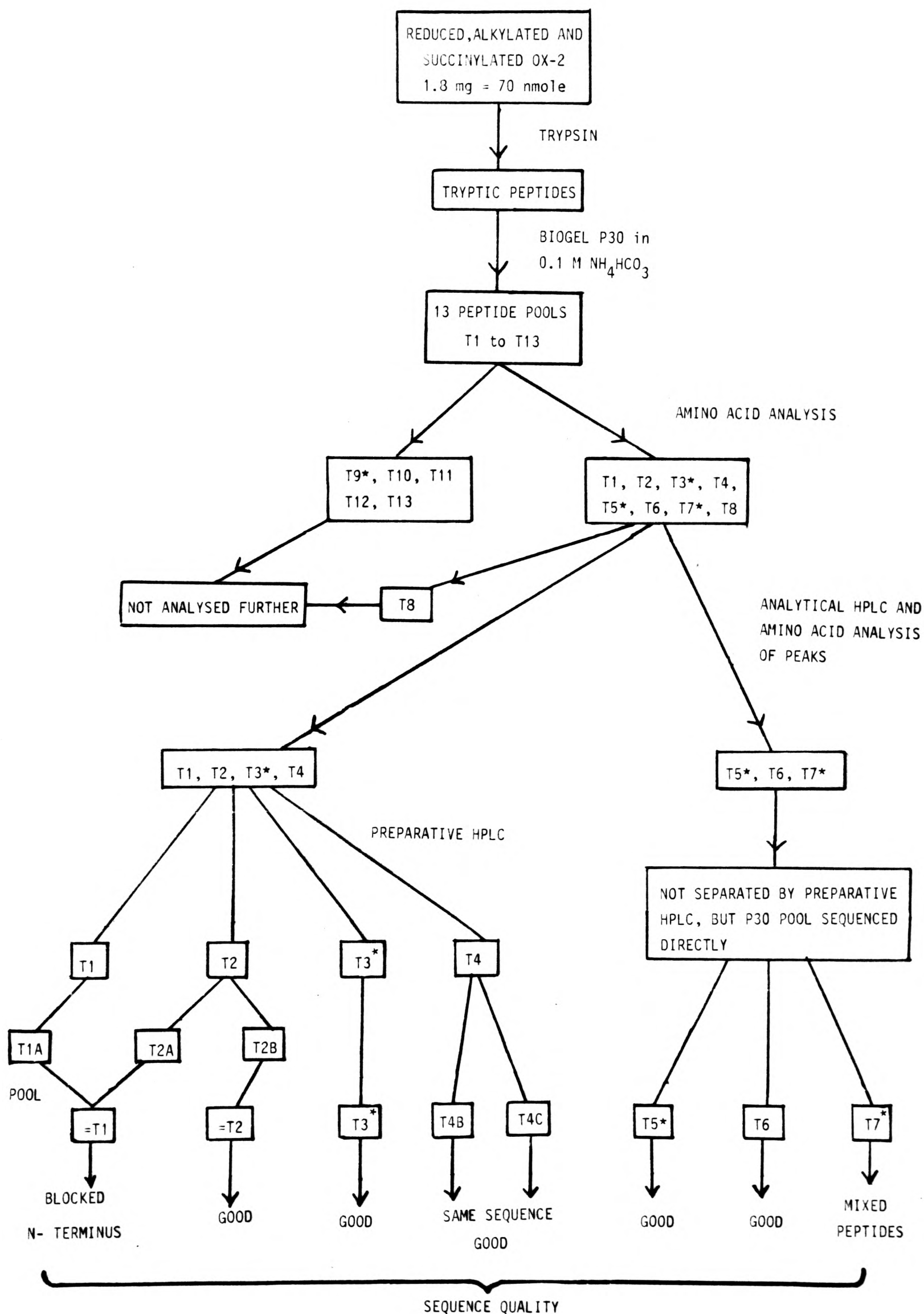


FIGURE 3.5

FLOW CHART INDICATING THE FATE OF OX-2 TRYPTIC PEPTIDE POOLS T1 to T13



KEY

* = ³H LABELLED CYS PEPTIDE

Figure 3.6

SEPARATION OF OX-2 TRYPTIC PEPTIDE POOLS BY REVERSE PHASE HPLC

5% of each P30 tryptic peptide pool T1 to T8 was loaded onto a Waters μ Bondapak C₁₈ HPLC column equilibrated in 10mM NH₄ acetate. Peptides were eluted with a 0-85% linear gradient of 70% (v/v) acetonitrile, 10mM NH₄ acetate, over 60 minutes.

The HPLC profiles are aligned vertically in the order that the peptide pools were collected from the the P30 column.

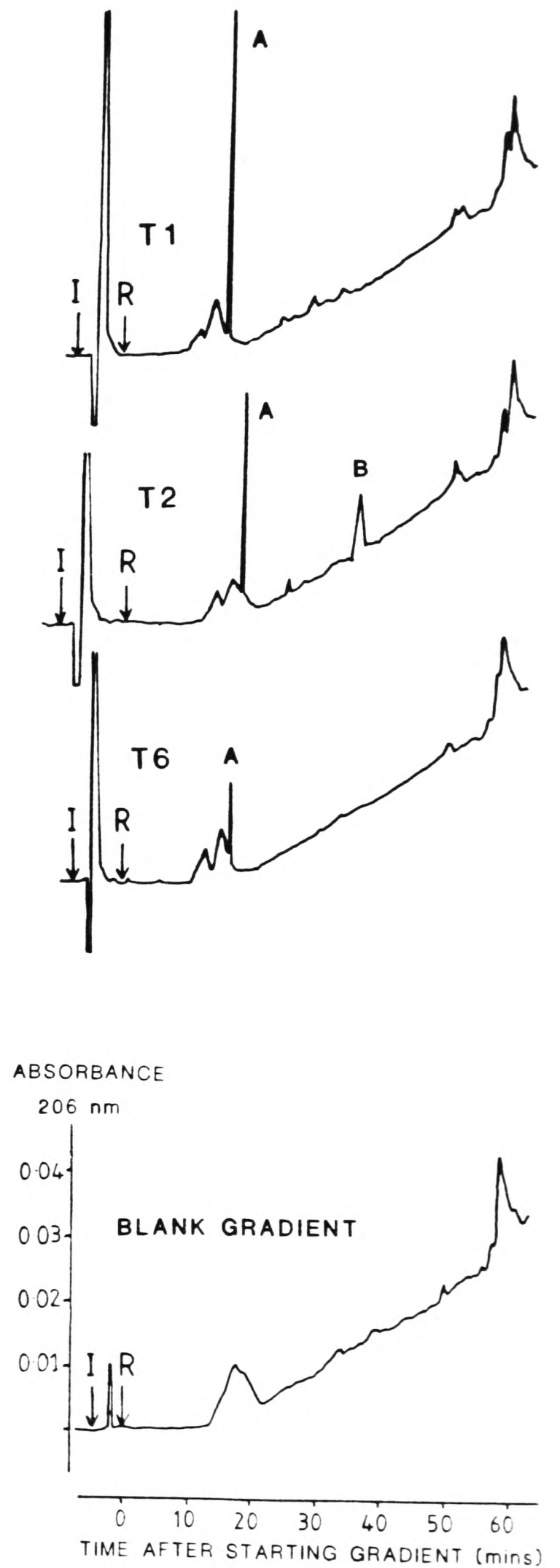
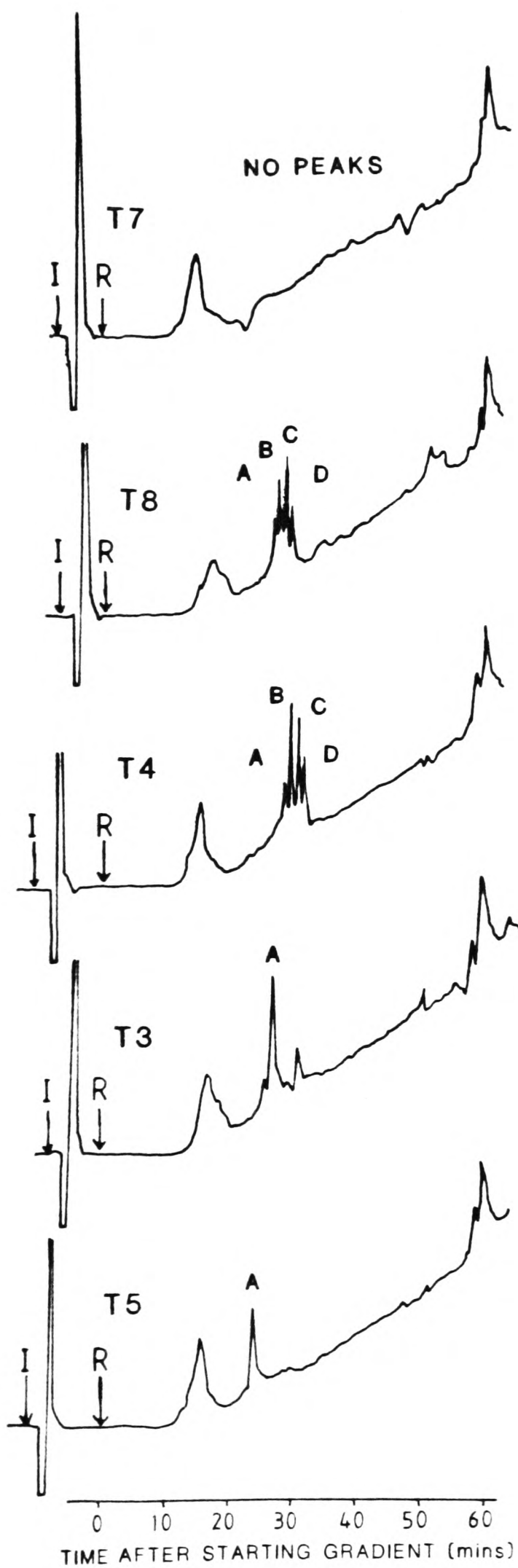
Key

Blank gradient = HPLC profile of the gradient alone

I = sample injected

R = gradient run commenced

SEPARATION OF OX-2 TRYPTIC PEPTIDE POOLS BY REVERSE PHASE HPLC



3) Peptides T1-T4. Preparative scale HPLC.

These peptide pools were freeze-dried to a small volume and each fractionated on two HPLC runs under the identical conditions used in the analytical runs.

a) Pool T1 contained only one major peak, T1A.

b) Pool T2 contained two peptides. T2A was pooled with T1A as they had identical amino acid compositions and elution positions, and this pool is called peptide T1.

Peak T2B eluted later in the gradient, had a different composition, and is referred to as peptide T2.

c) Pool T3 contained one major peptide, smaller contaminant peaks were discarded.

d) Pool T4 was resolved into four peaks. The peptide pool contained N-linked carbohydrate, so the peak splitting may have been due to carbohydrate heterogeneity on the same peptide chain. Peaks B and C were collected for sequencing.

The amino acid compositions of T1-T4 are shown in table 3.2A. The overall yield of peptide ranged from 16nmol (T4B) to 50nmol (T1 pooled) and a typical recovery of 60% was estimated for peptide T3:

	nmol peptide	³ H cpm as SCMC
P30 pool T3	42	1.1 x 10 ⁶
Peak T3 after HPLC	25	6.4 x 10 ⁵
% recovery after HPLC	60%	56%

3.4 THE SEQUENCES OF OX-2 TRYPTIC PEPTIDES

The sequences of tryptic peptides T1 to T7 were determined by automated Edman degradation and are shown in figure 3.8. The sequences are compared with the experimentally determined peptide compositions in

Table 3.2 A and 3.2 B

AMINO ACID COMPOSITIONS AND DETAILS OF THE AUTOMATED EDMAN DEGRADATION
OF OX-2 TRYPTIC PEPTIDES T1 TO T7

For each peptide:

The right hand column gives the experimentally determined compositions (mol of residue/mol of peptide) normalised to the residue (Arg or Tyr) indicated at the head of the column.

The left hand column gives the values obtained by amino acid sequence analysis.

A = nmol peptide recovered after purification

B = nmol peptide used in automated sequence analysis

C = Amount of peptide recovered on first sequencing cycle

D = Average repetitive yield

E = Total number of residues in peptide, or number of sequencing cycles run

TABLE 3.2 A

AMINO ACID COMPOSITIONS OF PEPTIDES T1 to T4

PURIFIED BY GEL FILTRATION AND REVERSE PHASE HPLC

		T1		T2		T3		T4	
AMINO ACID		aa ratio	peptide	aa ratio	peptide	aa ratio	peptide	aa ratio	peptide
		Arg=1	sequence	Arg=1	sequence	Tyr=1	sequence	Arg=1	sequence
SCMC		-	BLOCKED	-	0	0.65	1	-	(0)
ASX		1.0	∴	-	0	1.0	1	1.1	(1)
THR		0.9	NO	1.4	2	2.5	4	1.8	(2)
SER		0.9	SEQUENCE	1.1	1	1.2	1	2.8	(4)
GLX		3.0	↓	-	0	3.0	4	1.5	(2)
PRO		-		-	0	2.2	2	1.5	(2)
GLY		1.0		0.5	1	1.9	1	2.0	(3)
ALA		-		0.8	1	1.0	1	1.6	(2)
VAL		1.9		-	0	1.7	3	1.3	(0)
MET		-		-	0	0.6	1	-	(0)
ILE		-		-	0	0.6	1	1.0	(2)
LEU		-		1.6	3	1.6	2	0.5	(0)
TYR		-		-	0	1.0	1	-	(0)
PHE		-		-	0	-	0	-	(0)
HIS		-		0.75	1	-	0	1.0	(0)
LYS		-		0.8	0	3.0	4	1.0	(1)
ARG		1.0		1.0	1	-	0	1.0	(0)
TRP		N.D.		N.D.	0	N.D.	1	N.D.	(1)
A Total nmol peptide		50		30		25		16	
B nmol peptide for sequencing		42		25		20		13	
C nmol peptide from 1st sequencing cycle		BLOCKED		9 (GLY)		10 (SCMC)		2.6 (PRO)	
D average repetitive yield during sequencing		BLOCKED		97%		86%		86%	
E No. residues sequenced		BLOCKED		ALL=10		ALL=28		19 CYCLES NOT COMPLETE	

TABLE 3.2 B

AMINO ACID COMPOSITIONS OF PEPTIDES T5, T6 AND T7

PURIFIED BY GEL FILTRATION

		T5		T6		T7		
	AMINO ACID	aa ratio Tyr=1	peptide sequence	aa ratio Arg=1	peptide sequence	aa ratio Arg=1	peptide sequence	
	SCMC	0.9	1	-	0	1.7	MIXED PEPTIDES ↓	
	ASX	1.1	1	-	0	4.3		
	THR	1.1	1	} 2	0	3.9		
	SER	0.4	0		2	2.1		
	GLX	2.8	3	2.7	3	2.3		
	PRO	-	1	-	0	1.2		
	GLY	1.3	1	1.9	2	3.0		
	ALA	-	0	-	0	1.7		
	VAL	3.3	4	-	0	1.6		
	MET	-	0	0.8	1	0.8		
	ILE	0.5	1	-	0	2.3		
	LEU	0.9	1	-	0	4.4		
	TYR	1.0	1	-	0	2.0		
	PHE	-	0	-	0	1.8		
	HIS	-	0	-	0	1.4		
	LYS	2.2	3	-	0	1.0		
	ARG	-	0	1.0	1	1.0		
	TRP	N.D.	0	N.D.	0	N.D.		
A	Total nmol peptide	40		64		(~55)		
B	nmol peptide for sequencing	35		2		(~40)		
C	nmol peptide from 1st sequencing cycle	26 (VAL)		0.8 (GLY)		0.5 (SER) 2.1 (ASP)		
D	average repetitive yield during sequencing	87%		70%		N.D.		
E	No. residues sequenced	ALL = 18		ALL = 9		23 CYCLES NOT COMPLETE		

Figure 3.7

SEQUENCING OF OX-2 TRYPTIC PEPTIDE T3

A : Yields of amino acids during sequencing of OX-2 peptide T3

The yield in nmol for each residue during the sequencing of OX-2 tryptic peptide T3 was determined by comparison to standards and is plotted against the position of that residue in the peptide sequence.

B : Stepwise yields

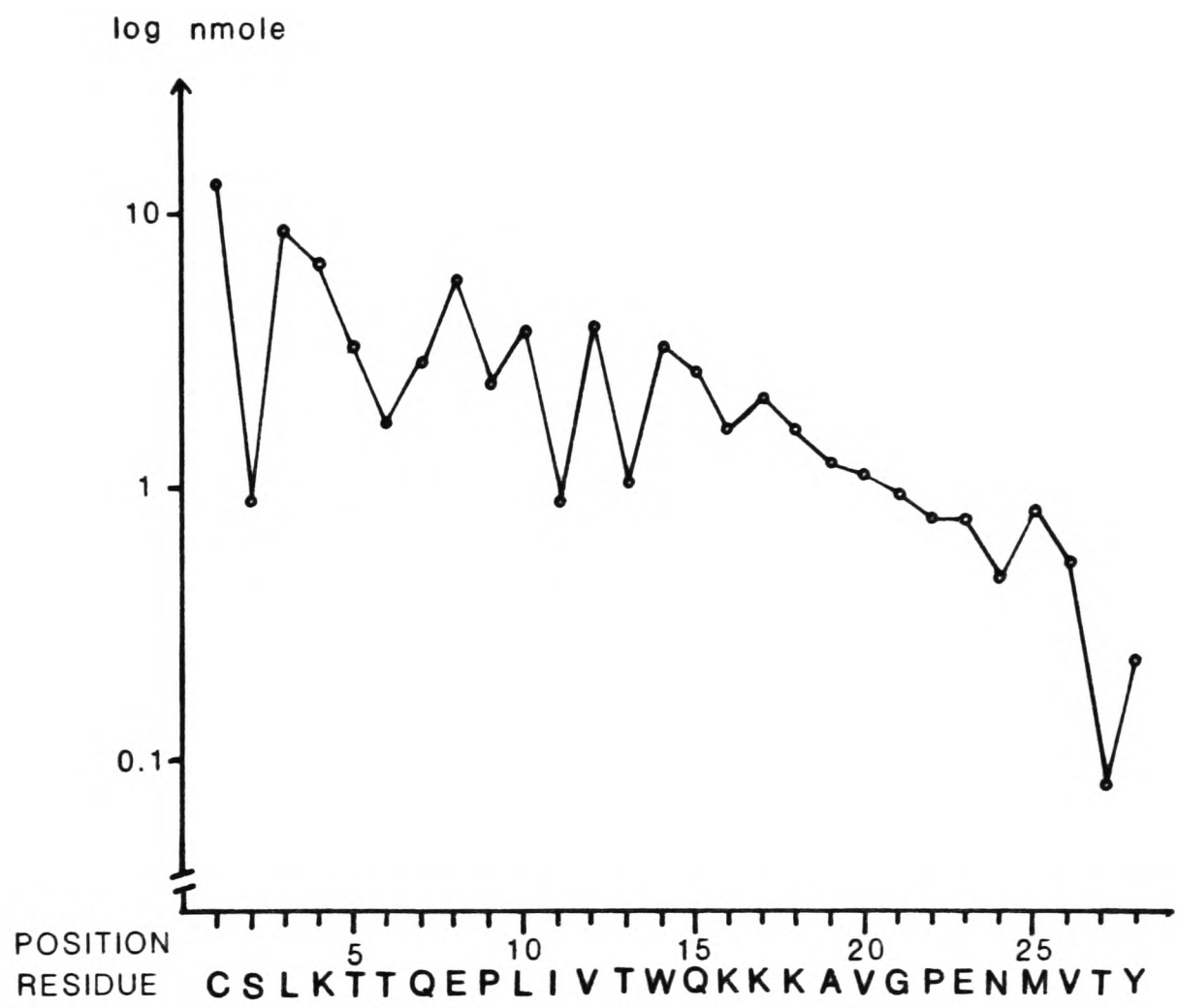
Stepwise yields for repeating residues may be calculated as follows:

If a residue has a yield of x nmol at position n
and " y " " " m

$$\text{the stepwise yield } n \rightarrow m = \left[\frac{n}{m} \right]^{\frac{1}{(y-x)}}$$

Stepwise yields are calculated for Lys, Thr and Val.

A YIELDS OF AMINO ACIDS DURING SEQUENCING
OF OX-2 PEPTIDE T3



B STEPWISE YIELDS

LYSINE		THREONINE		VALINE	
POSITION	%	POSITION	%	POSITION	%
4 → 16	89	5 → 6	54	12 → 20	85
16 → 17	125	6 → 13	93	20 → 26	88
17 → 18	76	13 → 27	83		
OVERALL		OVERALL		OVERALL	
4 → 18	90	5 → 27	85	12 → 26	87

Tables 3.2A and 3.2B and for the peptides which were sequenced to completion (T2, T3, T5 and T6) the amino acid sequences were in agreement with the analyses. Average repetitive yields were 70% - 97%, and a typical profile of recoveries at each cycle is shown in figure 3.7 for peptide T3.

Peptide T1 gave no sequence and thus is presumed to be the N-terminal peptide. As reduced and alkylated whole OX-2 protein gave no sequence either, the N-terminus of the intact protein is presumably blocked. Peptide T1 contains glutamate or glutamine residues, so the blockage could be due to cyclisation of glutamine to pyroglutamic acid as observed for Thy-1 (Campbell et al., 1981) and immunoglobulins (Wilkinson et al., 1966).

Peptides T2, T3, T5 and T6 yielded unambiguous sequence. There was no indication of more than one residue at any position except for the first position of T2 where small amounts of lysine, tyrosine and tryptophan accompanied the predominant glycine residue. Peptides T3 and T5 both end in a tyrosine residue instead of arginine, and a similar type of cleavage by trypsin has been noted before by Campbell et al. (1981) during the preparation of Thy-1 tryptic peptides.

Peptides T4B and T4C

Both peptides gave identical sequences but there was insufficient material to complete the sequencing run to the end of the peptide and the amino acid analysis (table 3.2A) predicts more residues than were sequenced. At position 16, no residue was obtained after the sequencing cycle, and is likely to be N-glycosylated asparagine which yields no phenylthiohydantoin amino acid derivative. This conclusion is supported by the fact that the peptide was glycosylated, and the sequence Asn-Ser-

Figure 3.8

AMINO ACID SEQUENCES OF OX-2 TRYPTIC PEPTIDES

Dotted lines (.....) indicate that peptides T4^B/C and T7 were not sequenced to completion.

The dash (-) in peptide T4^B/C indicates that no phenylthiohydantoin amino acid derivative was determined at this position, and the residue is believed to be glycosylated asparagine (see text).

Pool T7 consisted of a mixture of two peptides and at each sequencing cycle yielded the two residues within each pair of brackets. A dash (-) in this peptide indicates that only one residue was identified at that sequencer cycle.

AMINO ACID SEQUENCES OF OX-2 TRYPTIC PEPTIDES

PEPTIDE	SEQUENCE
T1	No sequence obtained. Thus is probably blocked at N-terminus Composition (Glx) ₃ (Val) ₂ (Arg) ₁ (Asx) ₁ (Thr) ₁ (Gly) ₁ (Ser) ₁
T2	G L L H T T A S L R
T3	C S L K T T Q E P L I V T W Q K K K A V G P E N M V T Y
T4B } T4C }	P A P A I S W K G T G S G I E - S T E S
T5	V K D P K T Q V G K E V I C Q V L Y
T6	G E S S Q G M Q R
T7	Mixed peptide sequence (S/D) (K/R) (A/I) (-/T) (G/I) (V/T) (V/E) (F/L) (Q/G) (P/L) (Y/L) (Y/-) (I/T) (D/S) (-/I) (I/T) (-/F) (I/W) (T/-) (E/L) (L/T) (G/L)

Thr would be compatible with the consensus signal Asn-x-Ser/Thr for N-linked glycosylation.

Peptide T7

The only peptide to give ambiguous sequence was T7 which contained a mixture of two peptides whose sequences could not be distinguished.

3.5 ANALYSIS OF THE SEQUENCE OF OX-2 TRYPTIC PEPTIDES

Comparison of the OX-2 tryptic peptide sequences with those of Ig-related molecules compiled by Kabat et al. (1983) showed that two OX-2 peptides shared sequence homology with immunoglobulin (figure 3.9). Peptide T3 can be aligned with residues 19 - 55 of rat Thy-1, and peptide T4 with residues 139 - 163 of rabbit κ chain (constant region). Identical residues are boxed, and particularly striking is the block of 7 identities "SGIENTE" between T4 and κ chain constant region. It is usually necessary to insert gaps to maximise homology between Ig-related sequences, and the number of gaps here is comparable to the alignments between some of the other members of the Ig superfamily.

3.6 DISCUSSION TO CHAPTER 3

This chapter has described the preparation of rat brain OX-2 antigen and the isolation and sequencing of seven tryptic peptides.

The preparation of milligram quantities of OX-2 antigen for sequence studies was facilitated by the "batch" approach of monoclonal antibody affinity chromatography first reported by Mellman and Unkeless (1980). The batch approach has recently been satisfactory for purifying the hematopoietic/endothelial antigen recognised by monoclonal OX-45 from spleen extract (Dr. J. Arvieux, manuscript in preparation). However, a preparation of leucocyte-common antigen from thymocytes using

Figure 3.9

ALIGNMENT OF OX-2 TRYPTIC PEPTIDES WITH IMMUNOGLOBULIN-RELATED SEQUENCES

A : ALIGNMENT OF PEPTIDE T3 WITH RAT THY-1

T3		C	S	L	K	-	T	T	Q	E	P	L	I	-	-	-	V	T	W	-	-	Q	K	K	K	A	V	-	-	-	-	G	-	P	E	N	M	V	T	Y	
THY-1		C	R	H	E	N	N	T	-	L	P	-	I	Q	H	E	F	S	L	T	R	E	K	K	K	H	V	L	S	G	T	L	G	V	P	E	-	-	H	T	Y

B : ALIGNMENT OF PEPTIDE T4 WITH RABBIT KAPPA CHAIN CONSTANT REGION

T4		(R)P	A	P	A	-	I	S	W	K	G	T	G	-	-	-	S	G	I	E	(N)	S	T	E	S		
XB9		R	P	D	D	I	T	V	T	W	K	V	D	D	E	I	Q	Q	S	G	I	E	N	S	T	T	P

Residues identical with OX-2 are boxed.
Dashes show gaps in the sequence inserted to maximise alignments.
The first arginine of peptide T4 is inferred from the cleavage specificity of trypsin.
The residue at position 16 of peptide T4 is believed to be glycosylated asparagine (see text).

References: Thy-1 sequence	Campbell et al. (1981)
Kappa B9	Kabat et al. (1983)

MRC OX-1 antibody was contaminated with other proteins (Dr M. Thomas, personal communication). If this latter problem is caused by non-specific binding to antibody-coupled beads then it may be circumvented by pre-mixing extract with bovine γ globulin coupled beads.

There were two objectives in determining the partial sequence of OX-2: (1) to design an oligonucleotide probe and thus complete the sequence by cDNA cloning, and (2) to search for sequence homologies which might indicate an evolutionary relationship with Thy-1. As the complete amino acid sequence was not required at this stage it was only necessary to isolate a few tryptic peptides, and those which could not be purified easily were discarded. Five out of seven peptides gave reliable, unambiguous sequence covering almost one-third of the OX-2 molecule, and the use of these for designing an oligonucleotide probe is discussed in the next chapter.

The sequence homology between two of the OX-2 peptides and members of the immunoglobulin superfamily is an encouraging indication that OX-2 antigen might have Ig-related structure. However, as discussed in the Introduction section 1.5, homology comparisons are best assessed over the entire length of the domain, taking into account the key conserved residues. Thus it is important to determine the arrangement of these peptides in relation to each other and to the rest of the OX-2 sequence. The next chapter describes the completion of the OX-2 sequence by recombinant DNA techniques, and the homology between the full OX-2 sequence and immunoglobulin-related molecules is discussed in Chapter 5.

CHAPTER 4

ISOLATION AND SEQUENCING OF cDNA CLONES FOR OX-2

4.1 INTRODUCTION TO CHAPTER 4

The partial amino acid sequence of OX-2 antigen determined in Chapter 3 provided data for the design of an oligonucleotide probe which would enable the OX-2 sequence to be completed by cDNA cloning (Introduction, section 1.7.1) The peptide data was derived from the brain form of OX-2, and as Barclay and Ward (1982) could detect no differences between the protein moieties of brain and thymus forms of the antigen, it seemed reasonable to use the oligonucleotide to screen a thymocyte cDNA library which was available in our laboratory.

This Chapter describes the design of the oligonucleotide probe, and the isolation and sequencing of full length cDNA clones for OX-2 antigen, following the strategy summarised in the flow chart in figure 4.1. The general features of the OX-2 cDNA and its translate are discussed.

4.2 DESIGN AND SYNTHESIS OF THE OLIGONUCLEOTIDE PROBE FOR OX-2 ANTIGEN

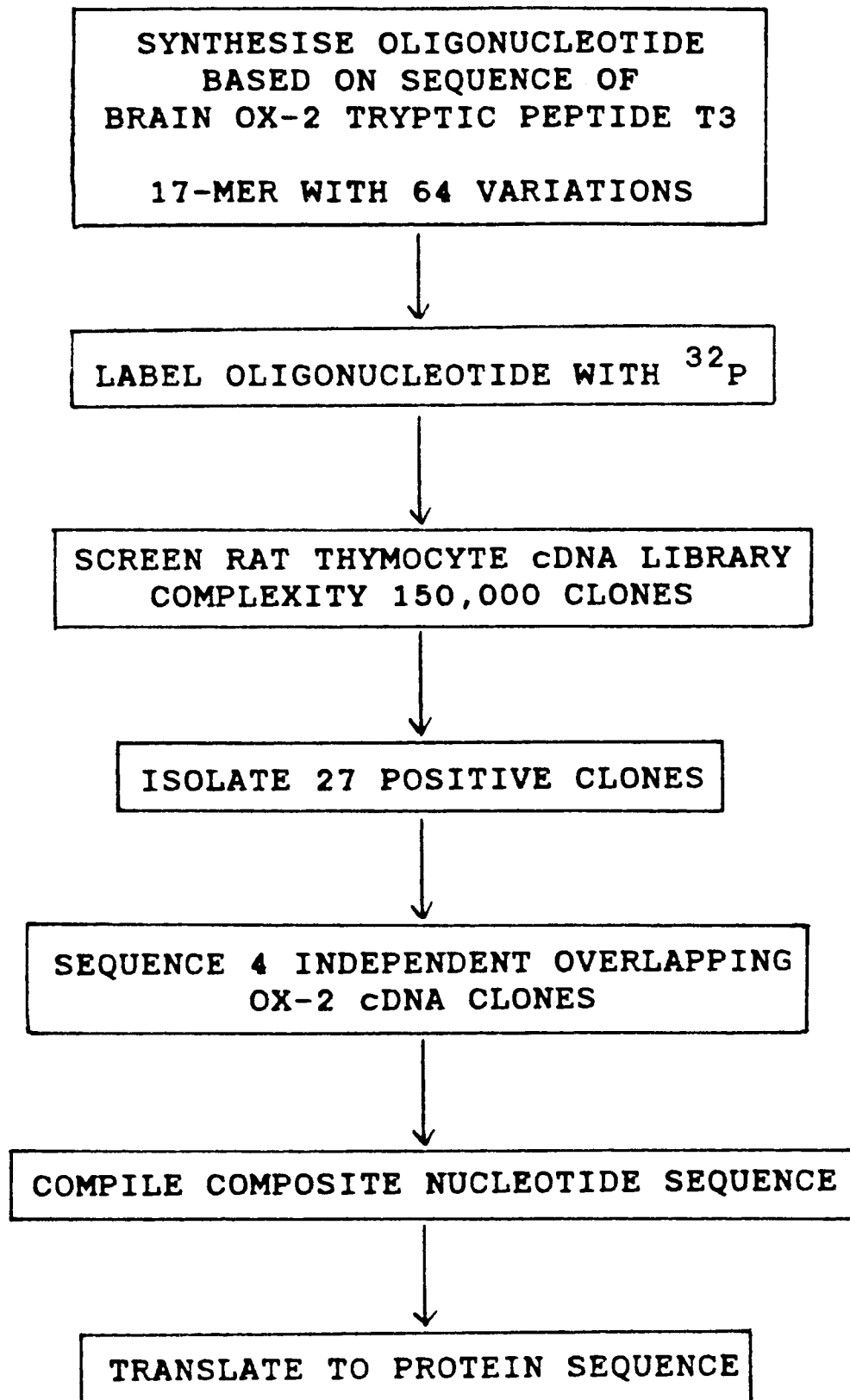
4.2.1 Design of the oligonucleotide

The factors influencing oligonucleotide probe design were discussed in the Introduction, section 1.7.1. The following points were considered when selecting the probe sequence from the peptide data in Chapter 3.

a) Probe length Mixed sequence oligonucleotides of length 14 nucleotides (Kornblihtt et al., 1983) to 17 nucleotides (Bentley and Porter, 1984) have been used successfully. The longer the probe, the

Figure 4.1

STRATEGY FOR ISOLATING THYMOCYTE cDNA CLONES FOR OX-2
ANTIGEN



lower the chance of the sequences in the mixture occurring at random in other cDNA clones.

b) Probe complexity The probe must be synthesised as a mixture of all possible codon combinations, so to minimise complexity the protein sequence should avoid amino acids with 6 codons and favour those with 1 or 2 codons. Two 17-long oligonucleotides can be made from the OX-2 sequence, both from peptide T3:

A : 17 long with 32 variations

amino acids	W	Q	K	K	K	A
nucleotides	UGG	CA ^A _G	AA ^A _G	AA ^A _G	AA ^A _G	GC

B : 17 long with 64 variations

amino acids	E	N	M	V	T	Y
nucleotides	GA ^A _G	AA ^U _C	AUG	GUN	ACN	AU

c) Unique protein sequence Neither of the sequences above are found in the Doolittle protein data bank.

d) Overall nucleotide composition Design A was discarded because the three lysine clones in the middle of the sequence would result in string of A residues in some of the oligonucleotide species in the mixture (the most extreme variant containing 11 consecutive As), and the probe might cross-hybridise with poly T from cloned poly A tails in the cDNA library.

Probe B was preferable as it has a moderate [G+C] content and an even distribution of G and C residues. The approximate hybridisation temperature of oligonucleotides can be calculated by the "Wallace Rule":

$$T_{\text{HYB}} = [(G+C) \times 4] + [(A+T) \times 2] \quad (\text{Gait, 1984})$$

For the component which has the lowest [G+C] content, T_{HYB} is 42°C.

e) Long single-copy oligonucleotide probes The alternative to a mixed-sequence oligo is a single copy probe about 50 nucleotides long (Introduction, 1.7.1), designed employing codon usage and dinucleotide frequencies, and avoiding amino acids encoded by 6 codons. Both peptides T3 (residues 11-28) and T5 (residues 1-16) contained suitable stretches of amino acid sequence. However, the mixed sequence oligonucleotide approach was preferred because of local expertise available for the synthesis and use of such probes. For example, the complement component C2 had recently been isolated using an oligonucleotide of the same length and [G+C] content as that proposed above (Bentley and Porter, 1984).

4.2.2 Synthesis of the oligonucleotide

The 17 residue oligonucleotide probe shown in figure 4.2 was synthesised as a mixture of all 64 possible sequences using the solid phase phosphotriester method of Gait et al.(1980). The probe was constructed as the complementary strand to the OX-2 message so that it could be used as a primer for cDNA synthesis if necessary, for example as employed by Shoulders and Baralle, (1982). The oligonucleotide was purified by preparative scale denaturing polyacrylamide gel electrophoresis. The yield was 1.8mg and overall recovery 12%, which is within the expected range.

4.3 THE RAT THYMOCYTE cDNA LIBRARY

4.3.1 Construction of the library

A cDNA library was constructed from Sprague-Dawley thymocyte mRNA by Drs N.Barclay and M.Thomas, as described in Methods section 2.3.16 and Thomas et al.(1985). Size-fractionated cDNA was blunt-end ligated

SEQUENCE CHOSEN FOR OX2 OLIGONUCLEOTIDE

C S L K T T Q E P L I V T W Q K K K A V G P E N M V T Y
 PROTEIN: E N M V T Y
 DNA: 5' G A G A A T C A T G G T C A C T A 3'
 T C C T

5' T A $\begin{matrix} \text{A} \\ \text{G} \\ \text{C} \\ \text{T} \end{matrix}$ G T $\begin{matrix} \text{A} \\ \text{G} \\ \text{C} \\ \text{T} \end{matrix}$ A C C A T $\begin{matrix} \text{A} \\ \text{G} \end{matrix}$ T T $\begin{matrix} \text{T} \\ \text{C} \end{matrix}$ T C 3'

17 nucleotides long, 64 variations

into the plasmid vector pAT153/Pvu 11/8 and the resulting cDNA library had a complexity of 160,000 clones.

4.3.2 An estimation of the abundance of OX-2 clones in the library

This cDNA library had not yet been screened for any thymocyte cell surface protein cDNAs. An estimate of the abundance of OX-2 clones was therefore based on the library of Moriuchi et al. (1983) who reported the frequency of Thy-1 clones as 1/5000. If cDNA frequency reflects cell surface abundance, then OX-2 clones should be at least 30 fold less frequent than Thy-1 clones. This would suggest about 1 OX-2 clone per 150,000, i.e. one clone per copy of the cDNA library.

4.4 SCREENING THE THYMOCYTE cDNA LIBRARY WITH THE OX-2 OLIGONUCLEOTIDE

4.4.1 The first screen

One copy of the library (~160,000 clones) was screened after transfer of the colonies to Whatman 541 paper and chloramphenicol amplification. The filters were fixed, prehybridised, and hybridised with radiolabelled OX-2 oligonucleotide probe. The hybridisation was performed at 4°C below T_{HYB} calculated in section 4.2, i.e. at 38°C. Filters were washed at T_{HYB} , i.e. at 42°C. After autoradiography overnight, colonies which aligned with 45 radioactive spots of varying intensity were selected for rescreen. A typical autoradiograph is shown in figure 4.3 filter A.

4.4.2 Rescreen of positive clones

Potential positives were taken and rescreened with the oligonucleotide at a lower colony density. 28 out of 45 gave a strong hybridisation signal (figure 4.3, filter B). Positive colonies were recloned by toothpicking individual clones onto a grid arrangement which was screened with the oligonucleotide (figure 4.3, filter C).

Figure 4.3

SCREENING THE THYMOCYTE cDNA LIBRARY WITH THE OX-2 OLIGONUCLEOTIDE PROBE

Bacterial colonies on LA agar plates were transferred to Whatman 541 filters and chloramphenicol amplified. The filters were fixed and hybridised with ^{32}P 5' labelled OX-2 oligonucleotide at 38°C , then washed in 6 x SSC at 42°C for 45 minutes.

The three filters illustrate the identification and isolation of the typical OX-2 cDNA clone pX2/25.

A : First screen

The thymocyte cDNA library was plated at a density of 10,000 colonies per 14cm plate. After transfer and hybridisation as above, the filters were autoradiographed overnight at -70°C with preflashed film and intensifying screen. The arrowed colony was taken for rescreen.

B : Second screen

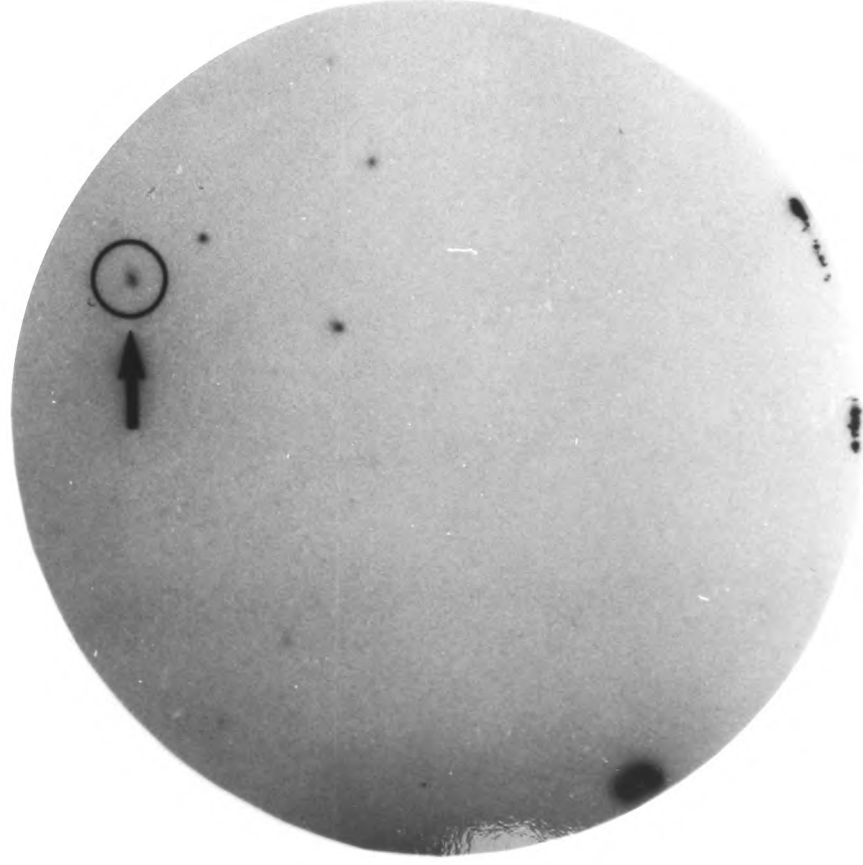
The colony from filter A was streaked across a 9cm LA agar plate. After transfer and hybridisation as above, the filter was autoradiographed for 4 hours at -70°C with preflashed film and intensifying screen.

C : Final rescreen

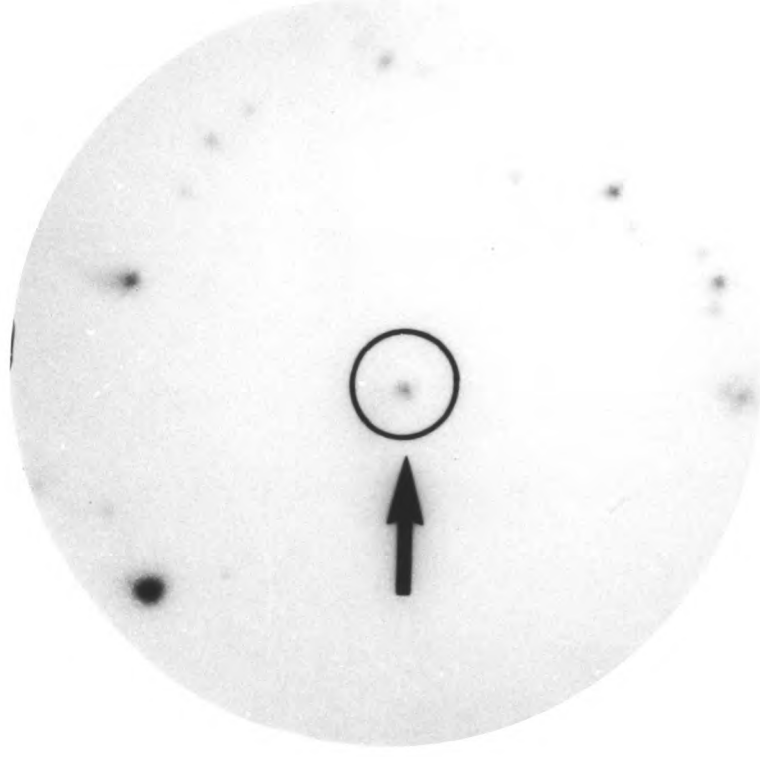
The positive colony from figure B was streaked out and 10 well-isolated colonies toothpicked into a row on an agar plate, together with putative OX-2 clones from other plates. After transfer and hybridisation as above, the filter was autoradiographed for 4 hours at -70°C with preflashed film and intensifying screen. Colonies of clone pX2/25 are in the row indicated by the arrow.

SCREENING THYMOCYTE cDNA LIBRARY WITH OX-2 OLIGONUCLEOTIDE PROBE

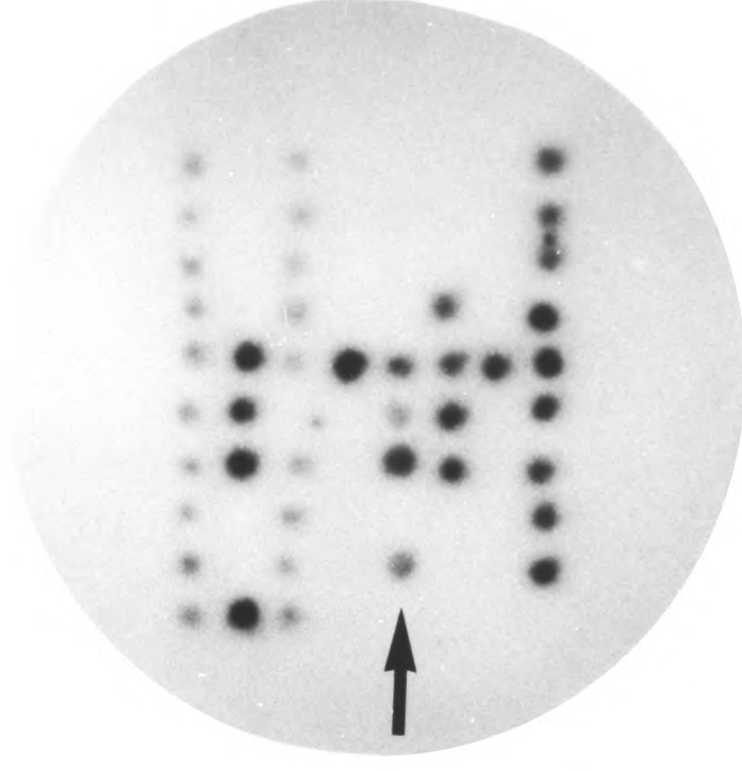
A FIRST SCREEN



B RESCREEN



C FINAL RESCREEN



27 out of 28 colonies were positive on the final screen. The intensity of signal from 5 clones, e.g. ^{top}~~arrowed~~ row, figure 4.3 C, was consistently less strong than the rest, and these clones were not pursued further.

4.5 IDENTIFICATION OF THE OLIGONUCLEOTIDE SITE IN POSTIVE CLONES

The OX-2 oligonucleotide would give an equally strong hybridisation signal whether it bound to the OX-2 cDNA, or to an identical nucleotide sequence occurring by chance in another cDNA clone. Thus it was important to identify the true OX-2 clones swiftly by localising the oligonucleotide site within the clones and sequencing the surrounding DNA.

4.5.1 Restriction mapping the positive clones to localise the oligonucleotide site

Plasmid DNA was isolated from 12 clones and the cDNA insert cleaved from the vector with restriction endonucleases Hind III and Bam HI (see figure 2.1 for a restriction map of the vector pAT153). Insert lengths were analysed on an agarose minigel and found to range from between 1.8 to 2.6 kb, and contained no internal Hind III or Bam HI sites. Four clones of varying insert sizes were mapped further by digestion with Hind III or Bam HI in combination with Xba I and Bgl II which have recognition sites of 6 nucleotides and do not cut the vector. Digests were analysed on the agarose minigel in figure 4.4A, from which the restriction map for clone pX2/25 could be derived. All 4 clones shared the characteristic Xba I and Bgl II sites, so presumably their sequences were related.

The agarose minigel was then dried down and hybridised directly with the radiolabelled OX-2 oligonucleotide to identify which fragment

Figure 4.4

LOCALISING THE OX-2 OLIGONUCLEOTIDE SITE

A : Agarose gel analysis of restriction enzyme digests of 3 OX-2 clones

Approximately 200ng of restriction enzyme digested plasmid DNA was run on a 0.8% agarose minigel with standard size markers.

Key: Lane a = clone pX2/25 (arrowed)
 Lane b = clone pX2/28
 Lane c = clone pX2/24
 Lane d = clone pX2/13

Digest 1 = Hind III + Bgl II

Digest 2 = Hind III + Xba I

Digest 3 = Bgl II + Xba I

Far right hand lane = clone pX2/9, linearised by digestion with Hind III.

B : Hybridisation of agarose gel with OX-2 oligonucleotide

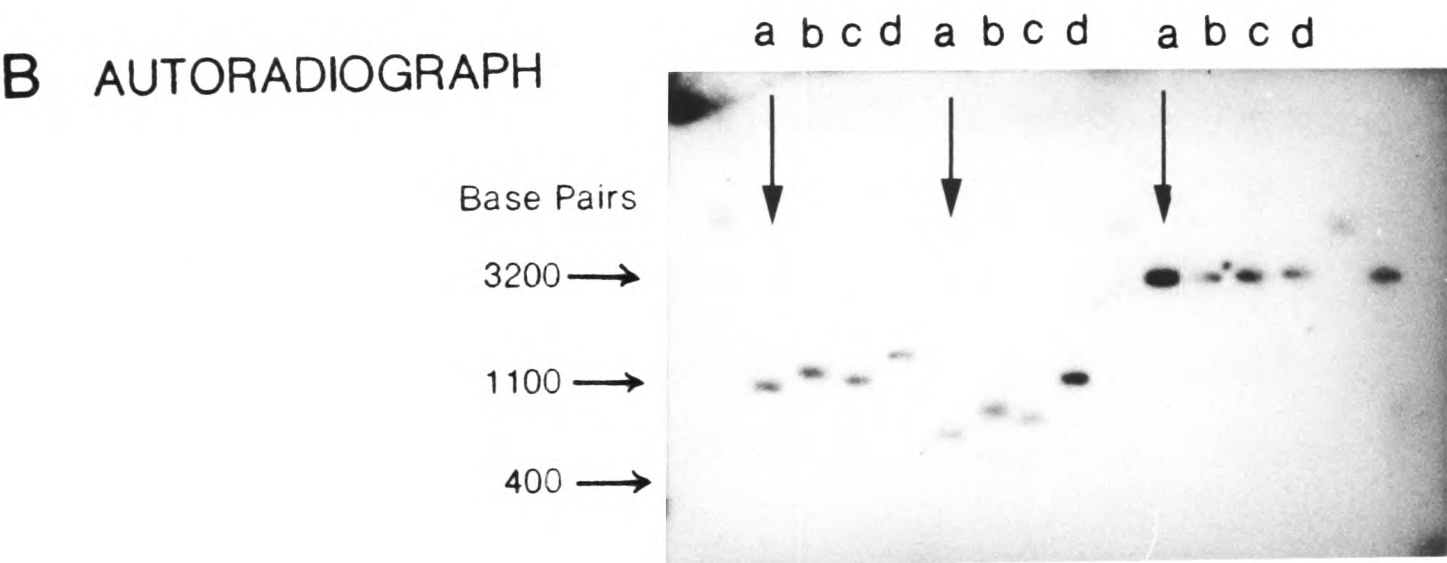
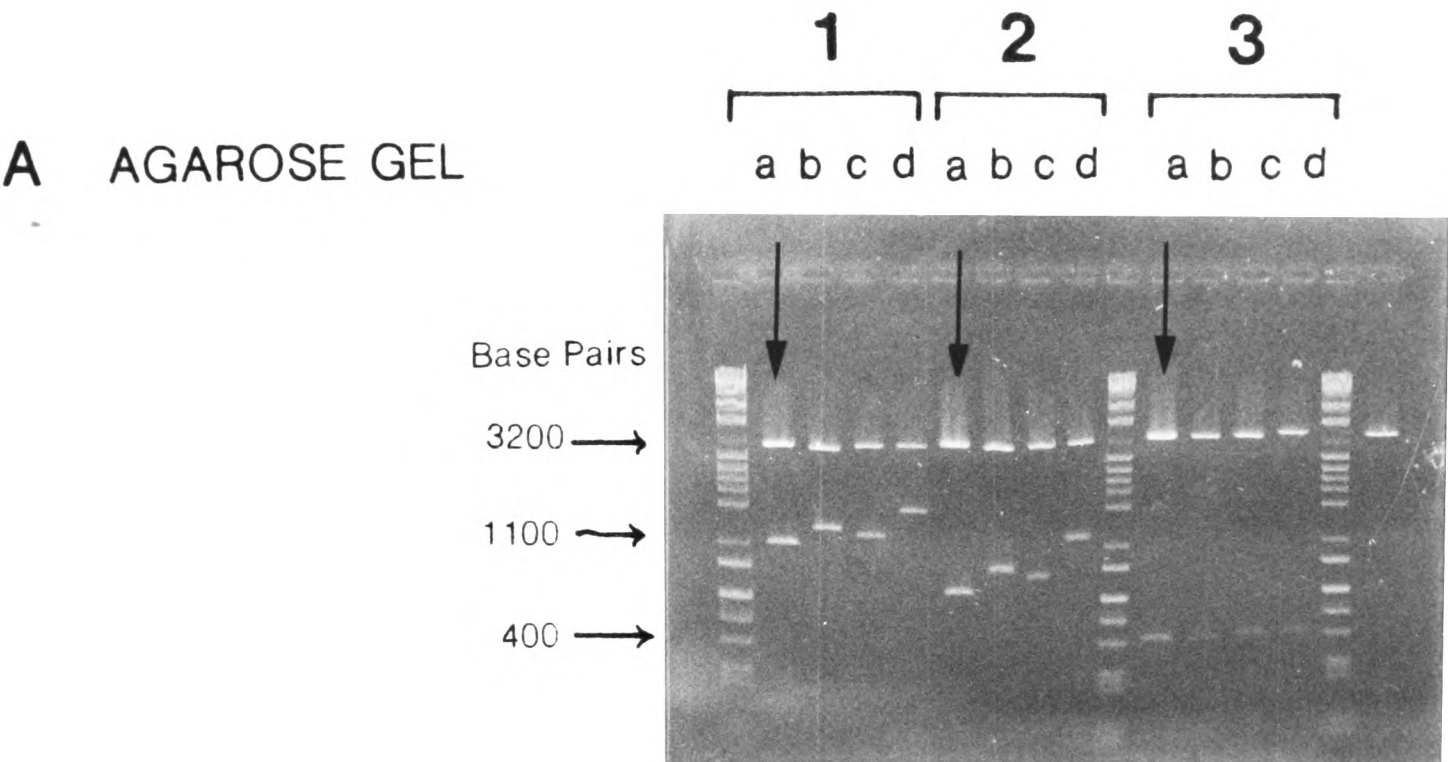
The agarose gel in figure A was dried down and hybridised with the ³²P 5' labelled OX-2 oligonucleotide at 38°C, and washed in 6 x SSC at 42°C. The gel was autoradiographed overnight at -70°C with preflashed film and intensifying screen.

The lanes correspond to figure A, above. Arrows indicate clone pX2/25.

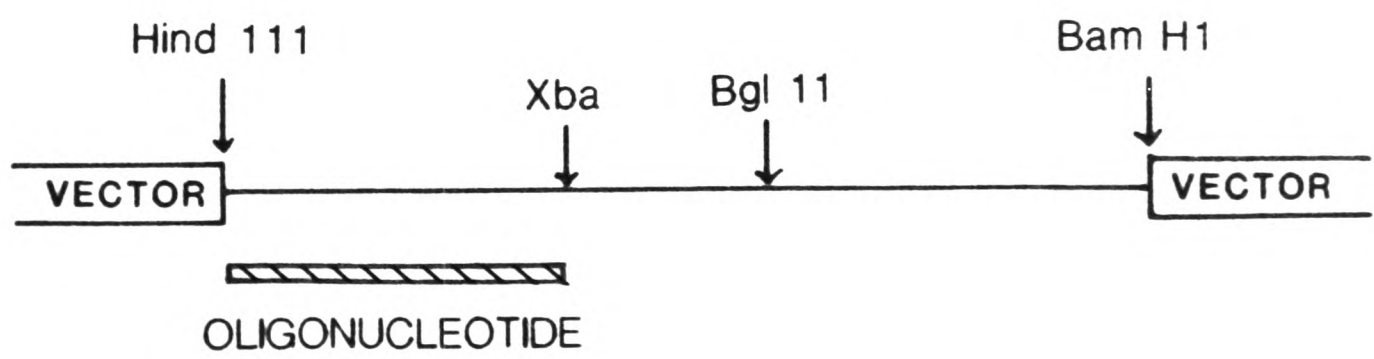
C : Restriction map for pX2/25 derived from (A) and (B)

The OX-2 oligonucleotide hybridises to a sequence somewhere within the 700bp fragment indicated by the hatched bar.

LOCALISING THE OX-2 OLIGONUCLEOTIDE SITE



C RESTRICTION MAP FOR pX2/25



Scale ⇐ 100 base pairs

carries the probe site. From the autoradiograph in figure 4.4B, it is apparent that the oligonucleotide lies within a 700 bp Hind III-Xba I fragment in clone pX2/25.

4.5.2 Preliminary sequencing of clone pX2/25

Plasmid pX2/25 was purified on a large scale by CsCl centrifugation. The 700 bp Hind III-Xba I fragment was labelled with ^{32}P at the Hind III site using DNA polymerase (Klenow fragment) and sequenced by the method of Maxam and Gilbert (1980). This region of clone pX2/25 was found to correspond to nucleotides 154 onwards (figures 4.6 and 4.7), and encodes the oligonucleotide sequence, peptide T3, and a portion of peptide T2. This confirmed the identity of clone pX2/25 as the cDNA copy of OX-2 mRNA.

4.6 STRATEGY FOR DETERMINING THE COMPLETE OX-2 cDNA SEQUENCE

Clone pX2/25 was apparently not full length at its 5' end because it lacked an initiator methionine residue and leader peptide. Therefore the sequences of other clones isolated with the OX-2 oligonucleotide were used to derive a composite full-length sequence for OX-2.

4.6.1 Preparation of cDNA fragments for sequencing

Plasmid DNA was purified from clones pX2/11, pX2/13 and pX2/24 by the CsCl gradient method. Fractionation of restriction fragments on agarose gels enabled a restriction map to be compiled for each clone as shown in figure 4.5A. Using this map, DNA fragments 0.7-1.4 kb long were prepared from each clone by preparative scale digestion and electroelution of the band from an agarose minigel. Typical recoveries ranged from 40%-90%, yielding 3-6 μg of a fragment which was of suitable quality for restriction mapping, sequencing, subcloning and nick-translation.

Figure 4.5

SUMMARY OF THE FEATURES OF OX-2 CLONES pX2/13, pX2/11, pX2/24 AND pX2/25, AND THE SEQUENCING STRATEGY INVOLVING THESE CLONES

A : Restriction map of clones pX2/13, pX2/11, pX2/24 and pX2/25, and the sequencing strategy used to derive the composite OX-2 nucleotide sequence

Solid arrows = regions sequenced by the method of Maxam and Gilbert
 → = noncoding strand labelled and sequenced
 ← = coding strand labelled and sequenced
Dotted arrow = sequenced by the dideoxy method.

B : Summary of rearrangements and abnormal structures in the cDNA clones used to determine the OX-2 sequence

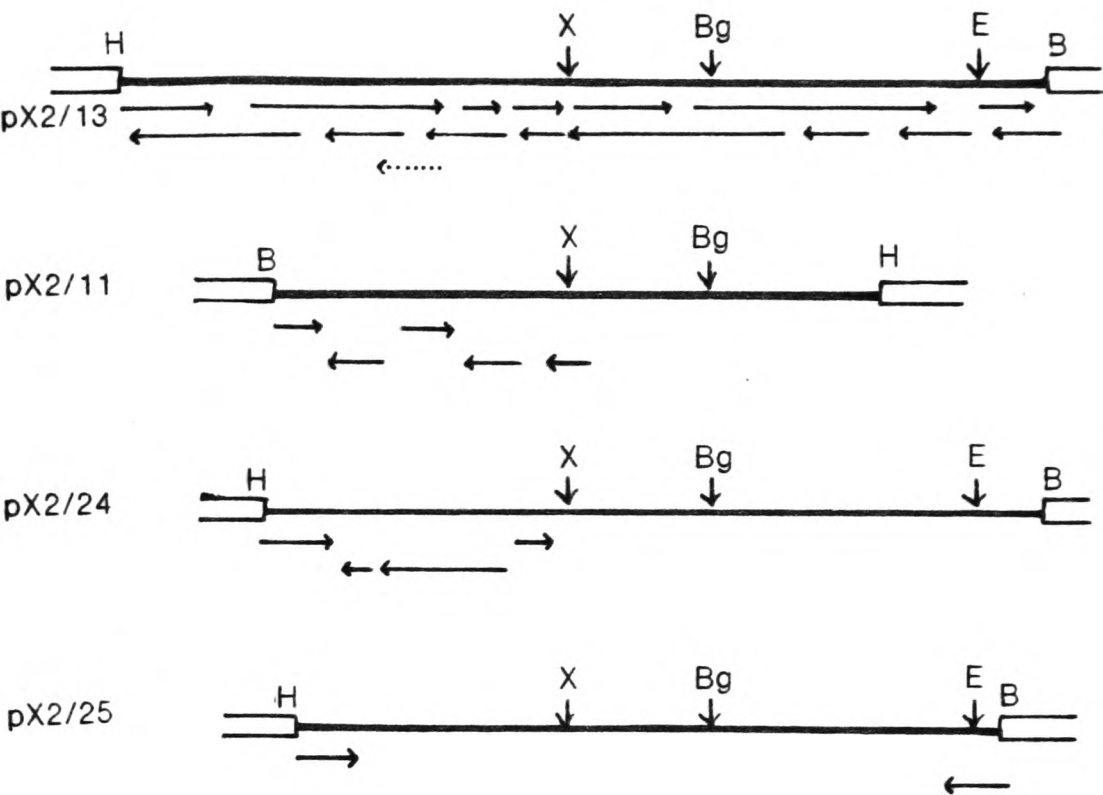
The numbers refer to the nucleotide position in the composite sequence (figure 4.6).

↵ = position of "loopback" inversion (see text)
∇ = position of deletion

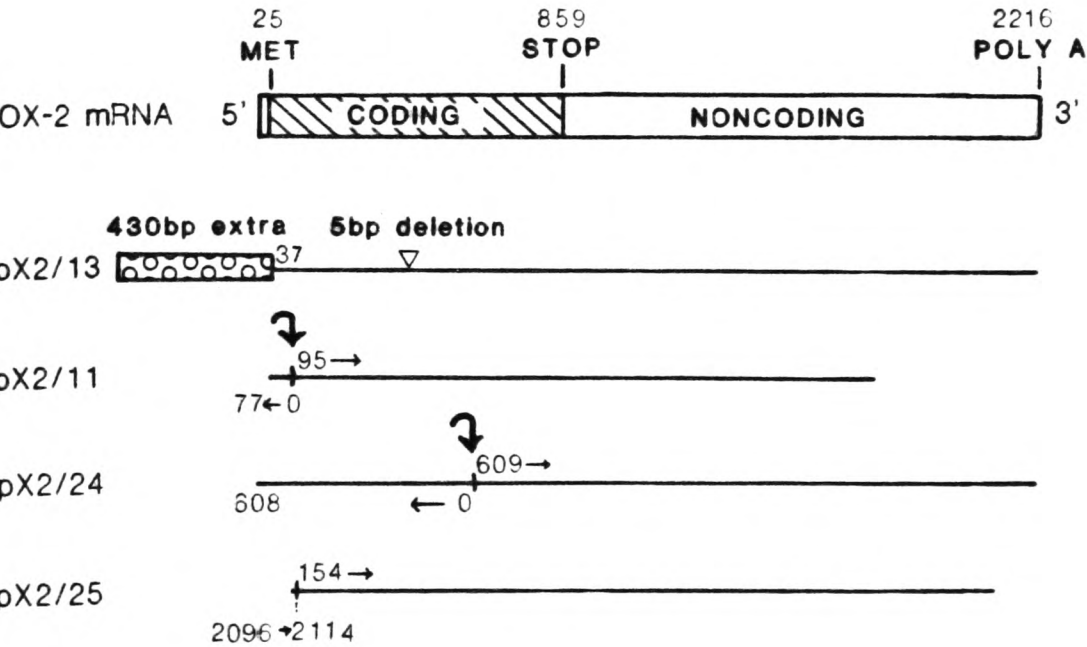
The thin lines represent sequence common to all three clones. The shaded portion in clone pX2/13 represents additional sequence unique to this clone.

H = vector Hind III site
B = vector Bam H I site
X = insert Xba I site
Bg = insert Bgl II site
E = insert EcoR I site

A : MAXAM AND GILBERT SEQUENCING STRATEGY



B : SUMMARY OF CLONE REARRANGEMENTS



SCALE $\longleftrightarrow$ 200 bp

4.6.2 Restriction mapping and sequencing OX-2 cDNA clones

Electroeluted fragments were mapped for the commonly occurring restriction enzyme sites such as Hinf I, Dde I and Sau 96I by digestion with one or more nucleases and labelling with DNA pol 1 (Klenow fragment). Fragments 50-500 bp long were analysed by electrophoresis on native 6% polyacrylamide gels with size markers. The detailed restriction maps derived in this manner indicated that the clones differed in length at both their 5' and 3' ends, and were rearranged towards their 5' ends.

The sequencing strategy for each clone is indicated by the arrows in figure 4.5A. The DNA was sequenced predominantly by the method of Maxam and Gilbert (1980) except for one region of clone pX2/13 (indicated by the dotted arrow) which was subcloned into M13 mp8 and sequenced by the dideoxy method (Sanger et al., 1977, Messing, 1983).

4.6.3 "Anomalies" in the cDNA clones

Unexpected features encountered during the mapping and sequencing of the OX-2 clones are summarised in figure 4.5B.

4.6.3.1 Inversions at the 5'ends

The construction of the cDNA library involved a "loopback" step to generate the second strand cDNA. It has been noted before that libraries made by this method contain inversions of various lengths at their 5' ends (Anson et al., 1984; Reid et al., 1984). Three inversions were identified in the OX-2 clones, involving bases 1-77 in pX2/11, and 1-608 in pX2/24, and a very extensive loopback in pX2/25 which placed nucleotides 2096-2114 at the 5'end of the clone.

Presumably as a consequence of the inversion, clone pX2/11 is missing nucleotides 78-94, and pX2/24 has lost a G residue at position 618.

4.6.3.2 Additions and deletions in clone pX2/13

a) The sequence of clone pX2/13 differs from that of clones pX/11 or pX/24 5' of nucleotide 37, and consists of 430 bp of different cDNA sequence which has not originated from the 3' end of the clone. Initially this was thought to have resulted from the ligation of two cDNAs into the same vector, but an alternative interpretation is suggested in Chapter 7.

b) pX2/13 also contains a deletion of 5 nucleotides (446-450) which disrupts the reading frame of the coding region. This deletion is discussed further in Chapter 7.

4.6.4 Compiling the composite OX-2 cDNA sequence

In view of the anomalies described in the previous section :

a) The sequence was determined on both strands of the cDNA, to resolve any artefacts or ambiguities intrinsic to the Maxam and Gilbert sequencing method.

b) Sequencing of the coding region was performed chiefly on clone pX2/13, then checked on pX2/24 and pX2/11.

c) Wherever a discrepancy occurred, the corresponding region was sequenced on all three clones.

The composite restriction map and nucleotide sequence is shown in figure 4.6.

4.7 NORTHERN BLOT ANALYSIS OF OX-2 THYMOCYTE MRNA

The length of the thymocyte message for OX-2 was determined by northern blotting of total thymocyte mRNA (Chapter 7, figure 7.4, track A). The OX-2 mRNA is a single band of apparent size 2.4kb, which is compatible with the value of 2.2kb for the cDNA in figure 4.6, excluding the poly A tail.

Figure 4.6

COMPOSITE RESTRICTION MAP AND cDNA SEQUENCE OF A THYMOCYTE OX-2 cDNA CLONE

A : Composite restriction map

Cleavage sites are shown for enzymes which have a recognition site of 6 nucleotides.

B : Composite nucleotide sequence

The sequence is compiled from data from clones pX2/13, pX2/11, pX2/24 and pX2/25.

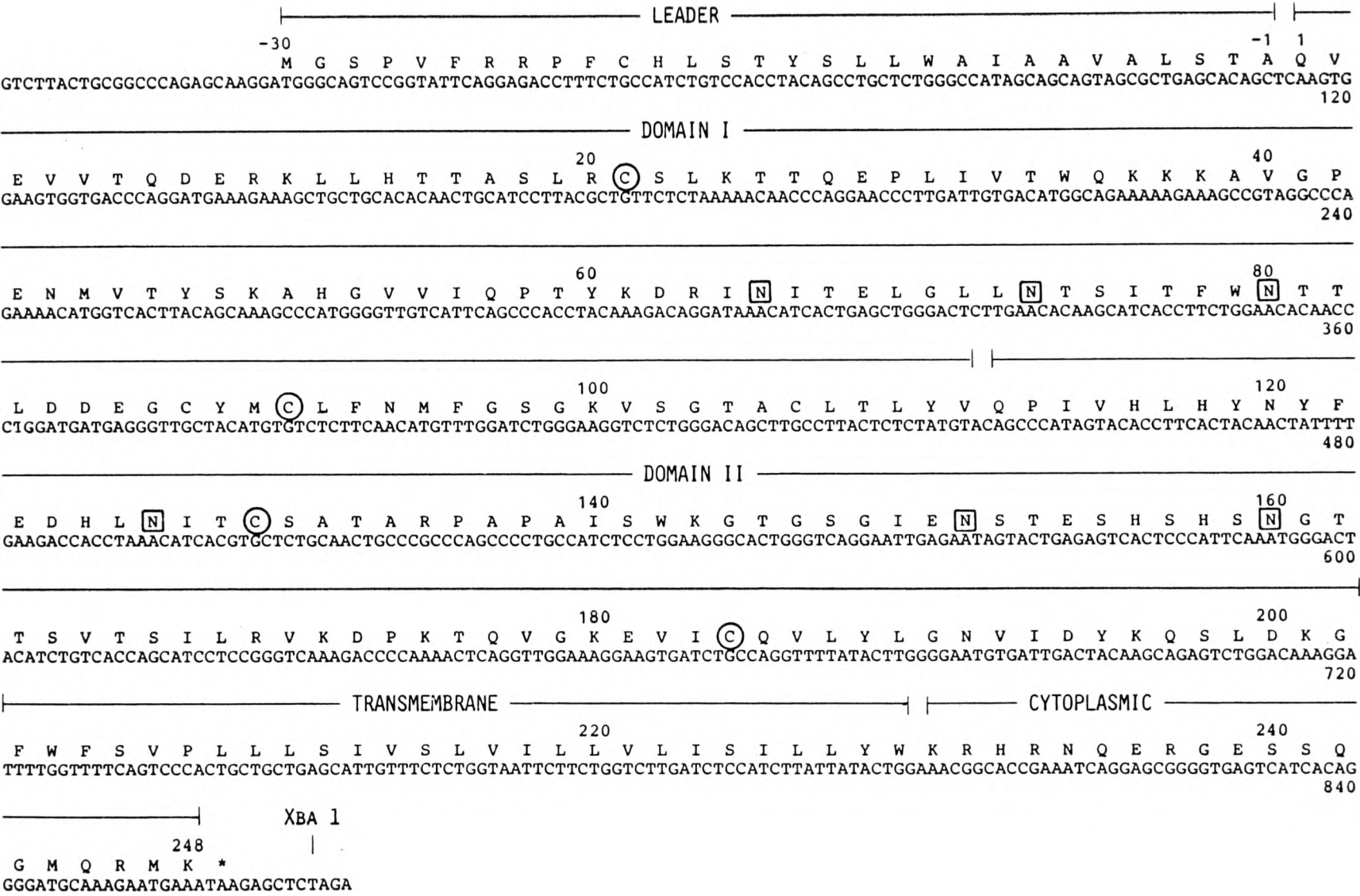
The coding region from the initiator methionine to the stop codon is underlined.

A double bar indicates the possible polyadenylation signal, and the sequence ends in a poly A tail.

GTCTTACTGCGGCCAGAGCAAGGATGGGCGAGTCCGGTATTTCAGGAGACCTTTCTGCCATCTGTCCACCTACAGCCTGCTCTGGGCCATA
90
GCAGCAGTAGCGCTGAGCACAGCTCAAGTGGGAAGTGGTGACCCAGGATGAAAGAAAGCTGCTGCACACAACCTGCATCCTTACGCTGTCT
180
CTAAAAACAACCCAGGAACCCCTTGATTGTGACATGGCAGAAAAAGAAAGCCGTAGGCCAGAAAAACATGGTCACTTACAGCAAAGCCCAT
270
GGGGTTGTCAATTCAGCCCACCTACAAAGACAGGATAAACATCACTGAGCTGGGACTCTTGAACACAAGCATCACCTTCTGGAAACACAACC
360
CTGGATGATGAGGGTTGCTACATGTGTCTCTTCAACATGTTTGGATCTGGGAAGGCTCTCTGGGACAGCTTGCCCTTACTCTCTATGTACAG
450
CCCATAGTACACCTTCACTACAACATATTTTGAAGACCACCTAACATCACGTGCTCTGCAACTGCCCGCCAGCCCTGCCATCTCCTGG
540
AAGGGCACTGGGTCAGGAATTGAGAATAGTACTGAGAGTCACTCCCATTCAAATGGGACTACATCTGTCAACAGCATCCTCCGGGTCAAA
630
GACCCCAAAACTCAGGTTGGAAAGGAAGTGTCTGCCAGGTTTTATACTTGGGGAATGTGATTGACTACAAGCAGAGTCTGGACAAGGGA
720
TTTTGGTTTTCAGTCCCACTGCTGCTGAGCATTGTTTCTCTGGTAATTCTTCTGGTCTTGATCTCCATCTTATTATACTGGAAACGGCAC
810
CGAAATCAGGAGCGGGGTGAGTCAACAGGGGATGCAAGAATGAAATAAGAGCTCTAGAGAAATTATACAGAACCTGAACGTGTTTT
900
CCTGGTCTACTTGAATCTGATGTTAAAGAAAAAGCAGGAGGGAAAAAGGTCACTTCTCATAGGACCTAAAAAGAGCAAAAGATGCABAAGG
990
AAGCCTGTGAGGGATTTGACTTTTTGCTGCTGTCCCAGGTCCTCAGTATTTACATTCGAAGAGGAAGTTGTGTGCCCTGGGTCTGTTGCA
1080
GGACTTGATTTTTGTGGAGCAATGCAGCGTCATTGCTGTTAGAAAGGCACCAGACTTAGAATCACCAGTGCCAAGCCGGCTCGCAGACCG
1170
ACTAGGGCTCCCACCGGAGGGACAAATCATAGTCAACTTACCTCACAGAGCTTCTGGTCTTTATACAAAGTAGAAAGGAGTGGGACCAG
1260
AAAATTTGCCATGTCTGAGATCTGATGGAATTTTAGGAAGAAAACCTGAAGAGCGATCAAAAGAAAGAGAGAGAACACAGAAGGGTCAAG
1350
AGCTTCGGAGAGTACCTTTTGCTTTCTGTTGGGGTCCACCTCTGGTTTTGTTCTTAGGTCCACCAGTTTGTTTCCCTGTTGTTGAGT
1440
ATCTAATTGACTACCTGCTACTGTTCCGCTGATTGTTGGCCTTGCTAAAATCCCTGATTCCCTGCCGTTCTCTATGTGCTTCTATGAGG
1530
GTTACTATGATGAAAATAGAGAAGAATTTAGTGTGAAGTAACATTGGCAACCGTAATGTGTCCATTTAACTTATTTTTATAGCAGGTAGG
1620
CAAATACCGTTAGTCTTAGCAAGTAGTTCACATCTTTACAAAAGCATGCTTTCCCTGTCCATTTGGCCAGGATATCACCTCTTTGAGG
1710
CCATCCTGAATCCTGTCTCGTGTAACGATAATATATTATGAAAACAGATGTGTTAAGAATTTCTGTACAGCAGTCAGTTGTATATTCTC
1800
TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAACTTCTTTTTCTGTGACTTTATTTTCAAAAGAGAAGGCCAAC
1890
TCTGGAATACAATCGCTTTGTTCTGAAGACATTTGTGGACATCTTAACCCTTTGACAGAACTAGTGACGTTGTTTTCTGTATCTTTTGTCT
1980
TCATCTGTCTCTGTAGAGTGACCTAGGAATTCAGTGTAAAGTTGTTTTTATTGTCAAACCTCGATTTTATATACTTGGTATGCTTTTCAT
2070
GTGTTATTTAATTCCGTATAATTTTCCTATATTTGTATTAAATACTGAGCAATTAAGAGTGTCAACTAAATAGTTGATGTGTGACATTC
2160
CCTTGAGAAATATAGAAATAAAGAAATAACACTTGGATTTTTTTTTTCTCTGTATTC (A)_n

Figure 4.7

COMPOSITE NUCLEOTIDE SEQUENCE OF THE CODING REGION OF A
THYMOCYTE OX-2 cDNA CLONE WITH ITS CORRESPONDING TRANSLATION



KEY

- = CYS RESIDUES BELIEVED TO PARTICIPATE IN THE CONSERVED DISULPHIDE BOND WITHIN EACH DOMAIN
- = ASN RESIDUES IN POTENTIAL GLYCOSYLATION SITES

4.8 TRANSLATION AND INTERPRETATION OF THE OX-2 CDNA SEQUENCE

4.8.1 Overall mRNA arrangement

The composite cDNA sequence in figure 4.6 contains one long open reading frame between bases 1-859 comprising of a 5' noncoding region of 25 nucleotides and a coding region of 834 nucleotides. This is followed by an unusually long 3' noncoding region of 1357 nucleotides. The cDNA sequence is full length at its 3' end as it finishes with a poly A tail located 36 nucleotides downstream from a conventional polyadenylation signal AATAAA (Proudfoot and Brownlee, 1974). However, the complete OX-2 mRNA might be longer at its 5' end than indicated here, as two other OX-2 clones (pX2/28 and pX2/9) extend a further 80bp 5' of nucleotide 1 (see Chapter 7).

4.8.2 Agreement between the predicted OX-2 sequence and the protein data

The translate of the open reading frame shown in figure 4.7 is strongly supported by the peptide data in Chapter 3. The first peptide sequence recognised is that of T2, commencing at amino acid 11. All the other peptides can be located in the translate through to residue 246 where peptide T6 ends (figure 4.8). Peptide T7, which yielded dual sequence, is a mixture of two overlapping peptides commencing at residues 49 and 61.

The only anomaly is the first residue of T2 which is Gly in the peptide sequence compared with Lys in the cDNA sequence. The amino acid analysis of T2 (table 3.2A) predicted a molar ratio of lysine in the peptide, but the first sequencing cycle yielded predominantly glycine. This might have been a result of contamination of the peptide with glycine prior to sequencing.

Figure 4.8

THE LOCATION OF OX-2 TRYPTIC PEPTIDES T1 TO T7 IN THE THYMOCYTE SEQUENCE

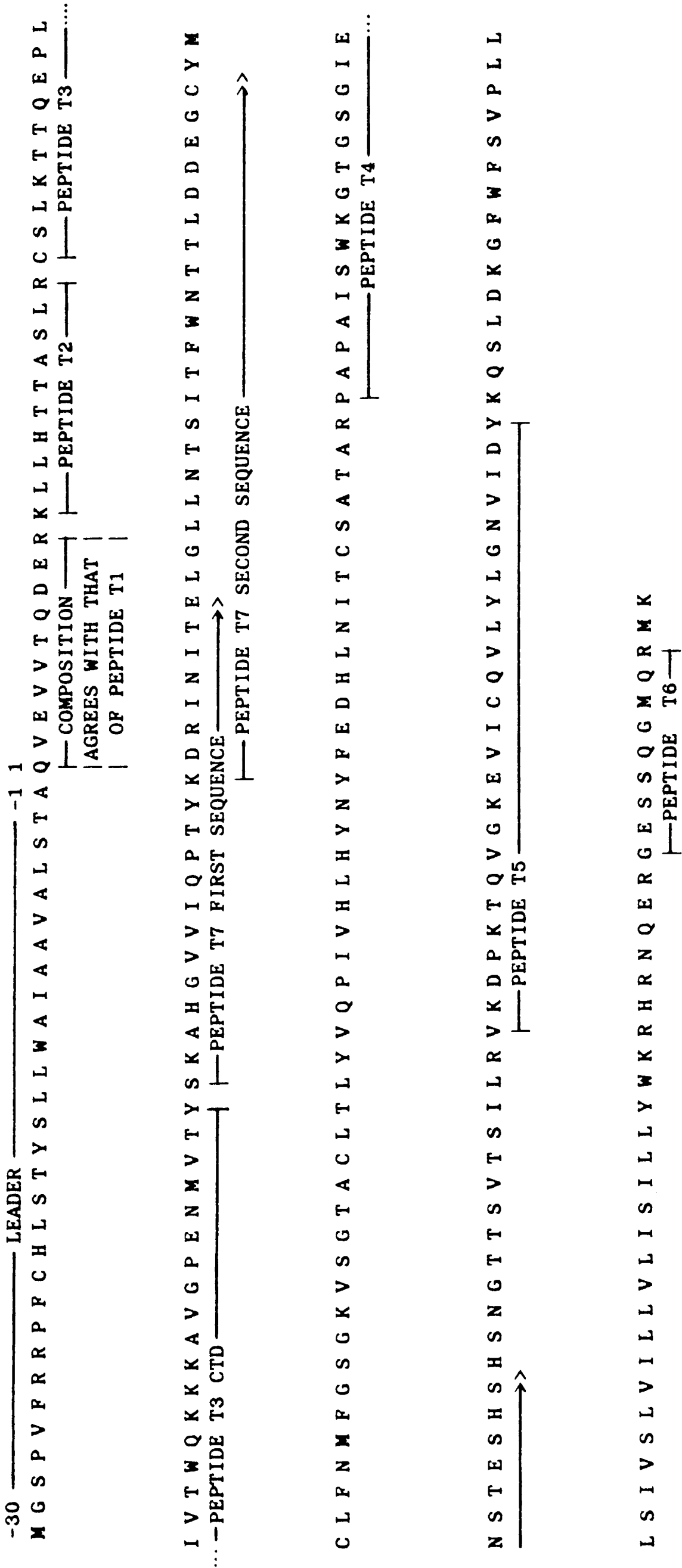


Table 4.1

COMPARISON OF EXPERIMENTALLY DETERMINED AMINO ACID COMPOSITION OF
THYMOCYTE OX-2 WITH THAT DEDUCED FROM THE NUCLEOTIDE SEQUENCE OF THE
THYMOCYTE cDNA CLONE

Residue	Determination	Prediction
Asx	21.3	19
Glx	26.0	24
His	8.0	8
Lys	15.0	16
Arg	10.0	9
Thr	22.8	23
Ser	20.8	22
Pro	8.7	8
Ala	10.9	8
Cys	5.2	6
Gly	18.4	16
Tyr	8.4	9
Val	18.6	21
Ile	13.1	16
Leu	25.0	27
Phe	6.7	6
Met	3.4	5
Trp	5.2	5

The figures for the experimentally determined composition are taken from Barclay and Ward (1982) and are expressed as residues per 248 amino acids.

Further confirmation of the coding region is shown by the good agreement between the experimentally determined amino acid composition and that deduced from the coding sequence minus the leader sequence (table 4.1).

The predicted molecular weight for the OX-2 polypeptide is 27,840. On the basis of carbohydrate contents of 33% for the thymus and 24% for the brain forms of OX-2, the calculated molecular weights for the glycoprotein are 41,500 for thymocyte and 36,600 for brain. These figures are fairly consistent with the values of 47,000 and 41,000 observed on SDS-PAGE, given that highly glycosylated proteins tend to run somewhat anomalously.

4.8.3 Features of the coding region

4.8.3.1 The leader sequence and the start of the mature protein

The potential initiator methionine codon at nucleotide 25 is contained within the consensus sequence A/G NNATG A/G found around most eukaryotic initiator methionine codons (Kosak, 1983). The 29 amino acids following this methionine have the characteristics of a leader peptide (von Heijne, 1983; Watson, 1984) i.e.:

a) A charged residue within the first few amino acids, seen in the OX-2 sequence as Arg at positions 7 and 8.

b) A core of uncharged, predominantly hydrophobic amino acids.

c) Small, neutral residues at positions -1 (Ala) and -3 (Thr).

The mature protein is predicted to start at Gln position 1. The N-terminus of the OX-2 protein was blocked (Chapter 3, section 4) and this could be accounted for if the Gln was cyclised to form pyroglutamic acid. Also, the composition of peptide T1 (table 3.2A), which is believed to be the N-terminal peptide, fits reasonably well with the tryptic peptide predicted by residues 1 to 10. The T1 composition cannot

be accounted for by tryptic peptides elsewhere in the sequence. The only discrepancy is the presence of Gly and Ser in the analysis, which may have originated from laboratory contaminants or a small amount of peptide T6 in the analysis.

4.8.3.2 The transmembrane segment and cytoplasmic region

The mature protein contains 248 amino acids and the only region likely to span the lipid bilayer is between residues 203-229. This sequence contains predominantly hydrophobic amino acids and lacks charged residues which are rarely seen in transmembrane segments (for example Kaufman et al., 1984).

Residues 230-248 include a high proportion of charged amino acids, typical of sequences on the cytoplasmic side of a transmembrane region. This implies that the C-terminal portion of the molecule is inside the cell, and the N-terminal residues 1-202 are outside.

4.8.3.3 The extracellular portion of the molecule

Carbohydrate structures have only been found on the extracellular portions of membrane proteins, so the presence of sites for N-glycosylation (recognition sequence Asn-X-Ser/Thr) at positions 65, 73, 80, 127, 151, and 160 confirms the orientation of the molecule. This is supported by the sequencing run on peptide T4 which indicated that residue 151 was Asn with N-linked carbohydrate (Chapter 3 section 4). As the OX-2 antigen has a high content of carbohydrate, it is likely that all six asparagines are glycosylated. The amino acid sequence of the extracellular portion of OX-2 is analysed in further detail in the next chapter.

4.9 DISCUSSION TO CHAPTER 4

Thymocyte cDNA clones encoding OX-2 antigen have been isolated from a cDNA library by means of an oligonucleotide probe, and their sequenced compiled to derive the complete primary structure of OX-2 protein.

All the tryptic peptides sequenced in Chapter 3 can be identified in the coding region of the cDNA clones, and no peptides are unaccounted for. Apart from a discrepancy at amino acid position 11 (see section 4.8.2), no sequence difference was detected between the brain form of OX-2 (peptide sequence) and the thymocyte form (cDNA clones).

Translation of the nucleotide sequence predicts a polypeptide 278 amino acids long with the characteristics of a membrane bound protein. The first 30 amino acids resemble a leader peptide which would be processed to leave a 248 residue mature protein. The N-terminal 202 residues are likely to be extracellular as they contain sites for N-linked glycosylation, and are followed by a hydrophobic membrane-spanning segment and a short hydrophilic cytoplasmic region.

At this point, it is interesting to note that cDNA clones for Thy-1 (Seki et al., 1985) also predicted a transmembrane segment yet the mature protein lacks this region and is anchored to the lymphocyte surface by covalent bonding to a non-protein hydrophobic tail (Tse et al., 1985). The OX-2 antigen, however, is unlikely to be modified in this manner because the isolation of peptide T6 encoded by residues 238-246 indicates that the transmembrane and cytoplasmic portions must be present in the mature OX-2 protein.

Two features of the 3' noncoding region of the OX-2 cDNA are somewhat unusual. Firstly, mRNAs encoding proteins the size of OX-2 generally contain 3' noncoding regions a few hundred

nucleotides in length, for example 175 bp in the case of the T cell receptor β chain (Yanagi et al., 1984). The OX-2 mRNA 3' region is distinctive in that it extends for 1357 bp and is 50% longer than the coding region. The significance of a longer-than-average 3' noncoding region is not apparent, but this feature has been reported for a few other cDNAs, notably Thy-1 (coding 483 bp, noncoding 1278 bp, Seki et al., 1985), the T8 antigen (coding 705 bp, noncoding 1182 bp., Littman et al., 1985b), and the transferrin receptor (coding 2544 bp, noncoding >2466 bp, Schneider et al., 1984).

Secondly, the 3' noncoding region includes the simple repetitive sequence (CT)₂₅ between residues 1798-1846. Similar repetitive sequences are sometimes encountered in introns, for example in the Thy-1 gene (Seki et al., 1985), but the function of such sequences in the noncoding regions of RNAs is not known.

The frequency of OX-2 cDNA clones in the thymocyte library was far higher than expected. Preliminary calculations (4.3.2) estimated about 1 OX-2 clone per copy of the library, yet 20 clones were identified with the oligonucleotide. When the library was rescreened with a coding region OX-2 probe, around 280 positive clones were counted, i.e. a frequency of 1 in 550. The data for OX-2 and other thymocyte membrane glycoproteins cloned from this library is summarised in table 4.2. The calculations in section 4.3.2 were based on the library of Moriuchi et al. (1983) which had not included a cDNA size-selection step prior to cloning, whereas the thymocyte library used here is enriched for messages longer than 1 kb. When this library was screened for Thy-1, a much higher frequency (1/330) was found than by Moriuchi (1/5000). From table 4.2, it can be seen that the abundance of cDNA clones does not correspond with cell surface abundance, as there is 30 times more Thy-1

Table 4.2

FREQUENCY OF cDNA CLONES FOR THYMOCYTE CELL SURFACE ANTIGENS CLONED FROM
THE RAT THYMOCYTE cDNA LIBRARY

Antigen	mRNA size	Frequency of cDNA clones	Site No. on thymocyte (AO rat)	Reference
OX-2	2.4 kb	1/550	2×10^4	This thesis
Thy-1	1.8 kb	1/330	1×10^6	M.J.Clark (unpublished)
L-CA	4.8 kb	1/5000	7×10^4	Thomas et al. 1985
OX-8	1.8 kb	1/500	4×10^4	Johnson et al. 1985
T cell receptor β chain	1.3 kb (major) 1.0 kb (minor)	1/200	n.k.	K.Morgan (unpublished)

n.k. = not known

The estimate of site numbers of each antigen on rat thymocytes is taken from Williams and Barclay (1985).

than OX-2 on a thymocyte, yet the Thy-1 clones are only 1.7 times more plentiful. Likewise, the leucocyte common antigen is 3 times more abundant than OX-2, but there are only about 1/10 the number of cDNA clones. Chapter 7 explores the possibility that not all OX-2 mRNA is capable of encoding a functional protein.

In the next Chapter, the translated OX-2 sequence is examined for homologies with the sequences of Ig-related molecules.

CHAPTER 5

SEQUENCE HOMOLOGY OF OX-2 ANTIGEN WITH IMMUNOGLOBULIN-RELATED MOLECULES

5.1 INTRODUCTION TO CHAPTER 5

The preliminary data from OX-2 tryptic peptides (Chapter 3) indicated sequence homology between OX-2 and Ig-related molecules. Now that the complete primary structure of OX-2 antigen has been determined by cDNA cloning (Chapter 4), the sequence comparisons can be extended to the entire molecule. This chapter discusses the evidence that the OX-2 antigen comprises of two Ig-like domains, and examines the sequence homology between OX-2 and other members of the Ig superfamily. A model for the structure of the OX-2 antigen is proposed.

5.2 WHAT TO LOOK FOR IN Ig HOMOLOGIES

The characteristics of Ig domains have already been discussed in the Introduction section 1.5.1, and illustrated in figure 1.2. The following features of the sequence of OX-2 antigen strongly indicate that the molecule has a structure related to Igs :

5.2.1 OX-2 antigen has the characteristic features of the Ig-type fold.

1) The size of the homology unit The Ig domain is about 100 amino acids long, and homology comparisons should be made over units of sequence the size of the Ig domain, i.e 100 residues. The extracellular portion of OX-2 is 202 amino acids long, sufficient for two domains designated I and II in figure 4.7.

2) The conserved disulphide bond All Ig domains contain the two Cys residues which form the conserved disulphide bond between β strands B and F (Figure 1.2). In OX-2 domain II there is one pair of correctly

positioned Cys at 130 and 184. In OX-2 domain I there is a candidate strand B Cys residue at position 20, and three further Cys at positions 88, 91 and 106.

3) β strand secondary structure In Ig domains, stretches of alternating hydrophobic and hydrophilic residues form β strands A to G which fold into two β sheets held together by the disulphide bond and hydrophobic interactions. The assignment of β strands in the OX-2 domains I and II is indicated in figure 5.1. In these regions of β strand are seen sequence patterns characteristic of Ig domains. notably:

i) The Cys residues in domains I and II are flanked by stretches of sequence which constitute β strands B and F:

Domain I strand B L R C S L

Domain I strand F Y M C L F

ii) β strand C is usually characterised by a tryptophan residue about 14 residues away from the Cys in β strand B. There is an appropriate Trp in at position 34 in domain I, and at position 142 in domain II.

iii) In β strand E, a patch of sequence like Leu Thr Leu is commonly found. The corresponding sequence in OX-2 is:

Domain I S I T F

Domain II I L R V

4) V or C type fold There are two main variants of folding pattern, V-type and C-type, which were illustrated in figure 1.2 B and C. The V-type folding pattern contains an extra loop of sequence between β strands C and D which is absent from the C-type fold. OX-2 domain I contains this extra sequence designated strand C' in figure 5.1 A, and thus fits a V-type fold. Domain II lacks the additional sequence (figure 5.1 B) and looks like a C domain.

Figure 5.1

ALIGNMENT OF OX-2 SEQUENCES WITH IMMUNOGLOBULIN DOMAINS

A : Alignment of OX-2 domain I sequence with V-type domains

Alignment with the sequence of:

Thy-1	Campbell et al., 1981
Poly Ig receptor	Mostov et al., 1984
Mouse MOPC 104E	Kabat et al., 1983
V _H New M	

B : Alignment of OX-2 domain II sequence with C-type domains

Alignment with the sequence of:

Rabbit kappa B9	
Mouse MOPC 315	Kabat et al., 1983
IgG EU C _H 3	
HLA-B7	

The numbers of identities and conservative substitutions between the sequences shown in A and B are given in figure 5.2

C : Alignment of OX-2 domain I with the V-type domains of T8 and the T cell receptor

Alignment with the sequence of:

Human T cell receptor β chain	Yanagi et al., 1984
Human T8 antigen (CD8)	Littman et al., 1985b

D : Alignment of OX-2 domain II with The C-type domain of the T cell receptor

Alignment with the sequence of the human T cell receptor β chain (Yanagi et al., 1984)

In the comparisons A-D, residues identical to OX-2 are boxed and dashes show the gaps in the sequence inserted to maximise the alignment. The bars with letters A,B,C,C',C'',D,E,F and G below the sequence indicate residues involved in forming the β strands shown in figure 1.2.

Figure 5.1

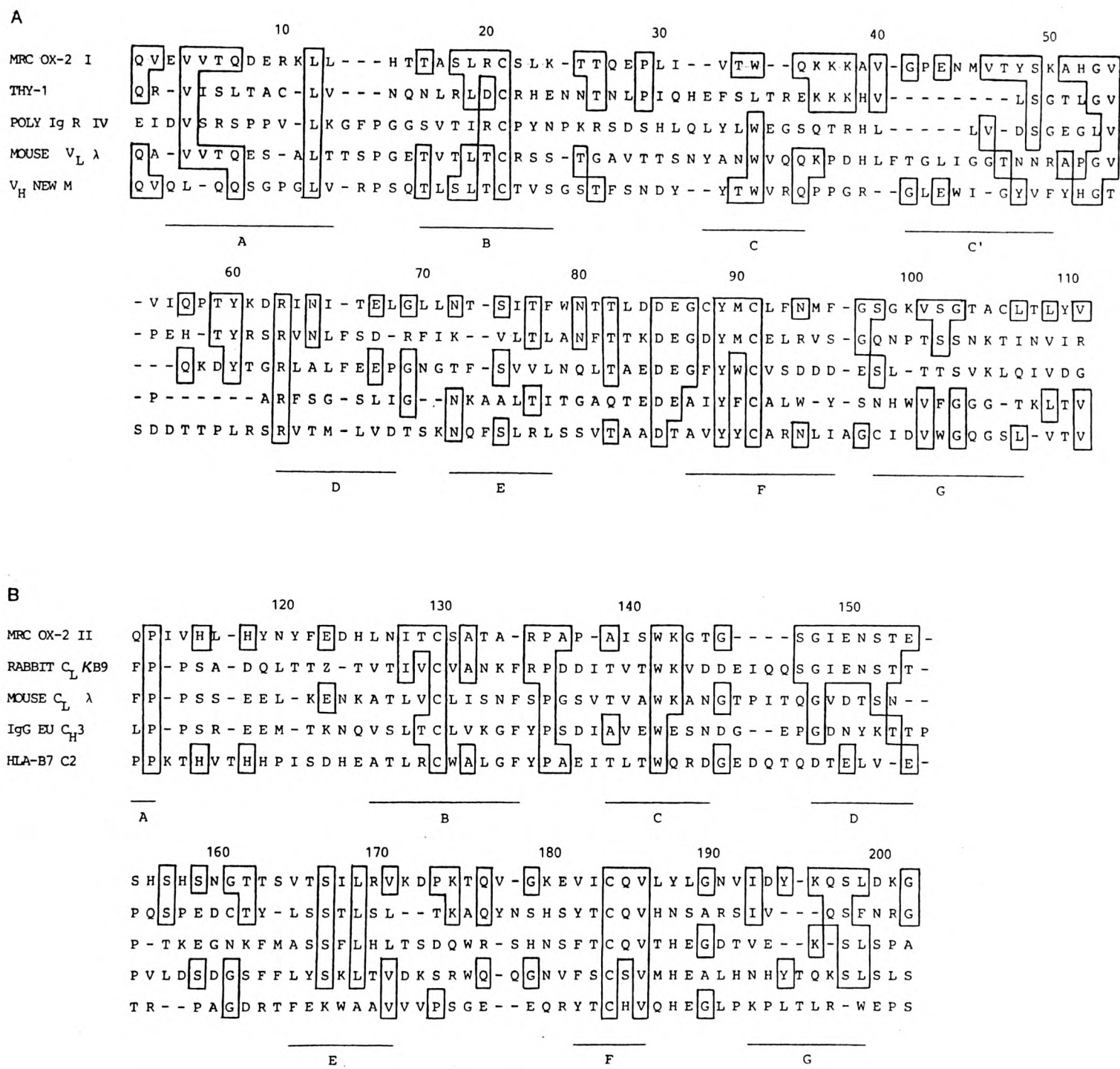
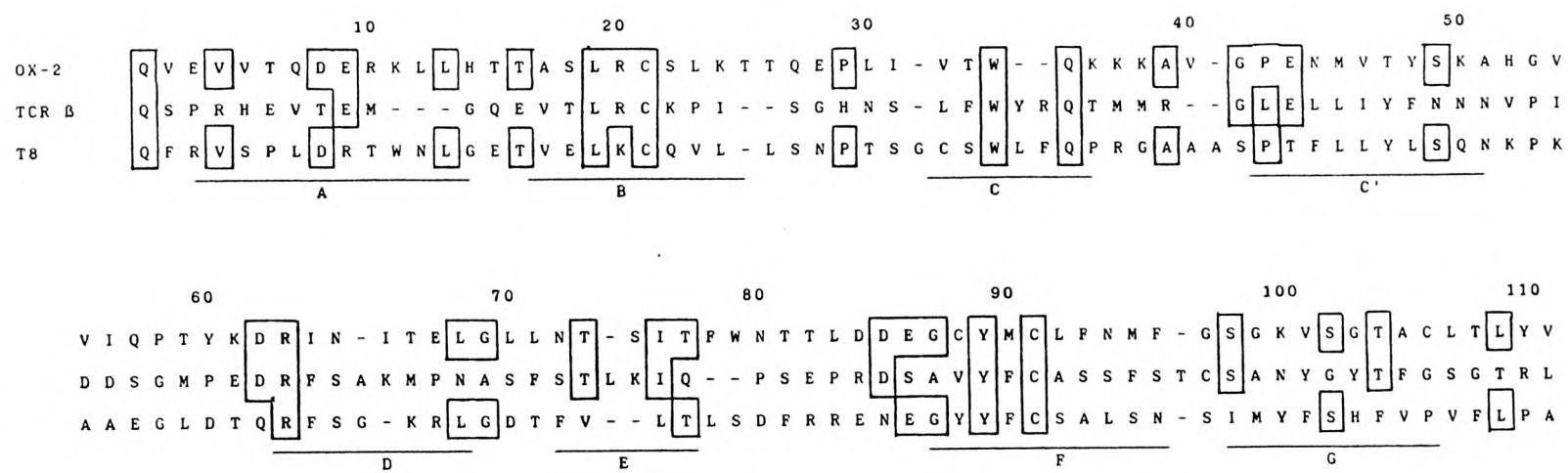
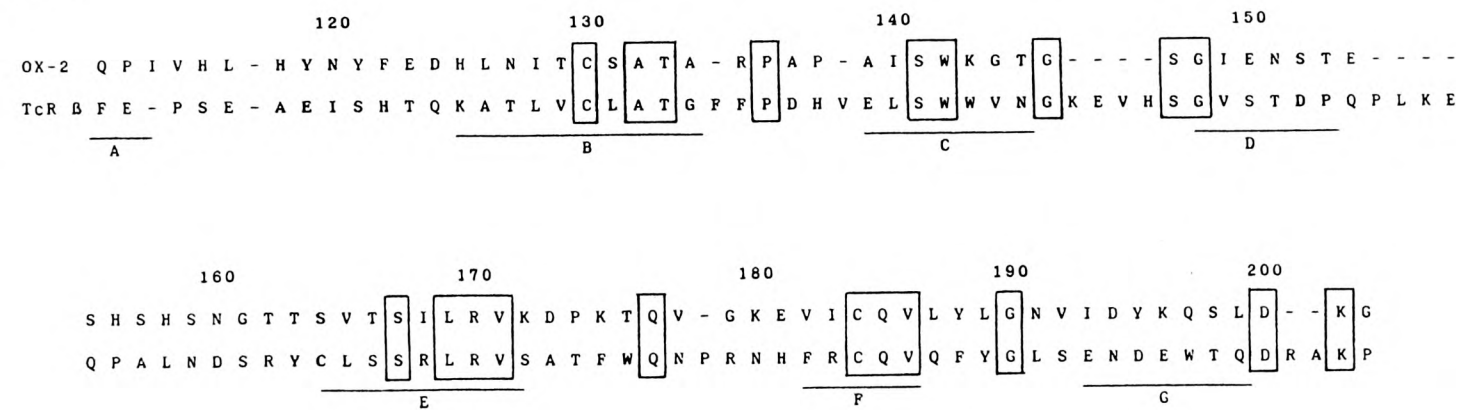


Figure 5.1 CTD.

C :



D :



5.2.2 Sequence homologies with Ig-related molecules

In addition to the folding pattern, some sequence patterns are specific for V- and C-type domains.

1) Alignment of Domain I with V regions

Figure 5.1 A and C illustrate sequences characteristic of V type domains aligned with OX-2 domain I, residues 1-111. The sequence around Cys 91 in β strand F is particularly distinctive. An Ala (A) or Gly (G) residue is almost always found 4 residues before the Cys in strand F, while the Asp (D) six residues before the Cys is highly conserved, probably because it forms a salt bridge with an Arg (R) residue (occasionally a Lys) near the base of β strand E (R.Poljak, personal communication). The domain I sequence has an arginine residue at this position.

The numbers of identities and conservative substitutions between the five sequences in 5.1 A are shown in figure 5.2. The scores for domain I with these sequences is at least as convincing as the scores within the sequences themselves. The OX-2 sequence also fits fairly well with the V-type folds of the T cell receptor β chain sequence (18/111 identities), and the T8 antigen (23/111 identities). Domain I appears to be slightly more homologous to Thy-1 (29 identities) than to the other V type domains in figure 5.1 (18-30 identities) especially in view of the fact that only 2 Thy-1 sequences are available for comparisons whereas there are many published immunoglobulin sequences. One point to note is that the V_H New M, T cell receptor β chain and OX-8 sequences are similar in length to OX-2, whereas the Thy-1, Poly Ig receptor and mouse V_L sequences are all shorter than OX-2 in the region of the β strand C' and consequently gaps extending across five or more residues were included in the alignments at this position.

Figure 5.2

IDENTITIES AND CONSERVATIVE SUBSTITUTIONS BETWEEN MRC OX-2 SEQUENCE AND
THE Ig V-AND C-DOMAIN SEQUENCES IN FIGURE 5.1

The two numbers in each square give the number of amino acid identities, and identities plus conservative substitutions respectively between the sequences indicated in figure 5.1 A and B. Conservative substitutions are those that occur more frequently than random chance would predict and are those with a score of 1 or more in Figure 84 in Dayhoff et al., (1978). For comparisons other than those involving the OX-2 sequence, the alignments were altered from those in figure 5. 1 A and B if adjustments, without excessive insertion of gaps, improved the scores. Thus in all cases the scores are for optimal alignments as far as could be judged.

C-DOMAINS - ALIGNMENTS WITH 91 RESIDUES OF
MRC OX-2 DOMAIN II

		C _K	C _λ	C _H ³	HLA B7	
						MRC OX-2
THY-1	29,54	28,46	18,43	20,47	17,38	C _K
POLY IgR (IV)	21,48	22,45		30,49	21,45	C _λ
V _λ	29,49	24,44	21,44		22,47	C _H ³
V _H	30,48	23,47	22,50	32,62		
	MRC OX-2	THY-1	POLY IgR	V _λ		

V-DOMAINS - ALIGNMENTS WITH 111 RESIDUES OF
MRC OX-2 DOMAIN I

2) Alignment between domain II and C type domains

Residues 112-202 from domain II align well with the C-type domains in figure 5.1B, and the scores amongst all the sequences are shown in figure 5.2. Fewer identities are seen between OX-2 and the Ig like domain of HLA B7, but the homology is as good as between the HLA domain and the other C-type domains. The homology between domain II and the T cell receptor β chain C domain is poorer (figure 5.1 D). The T cell receptor α chain C domain sequence is not included in this alignment because its sequence homology with Ig C regions is low (Saito et al., 1984).

5.3 COMPUTER SEARCH FOR SEQUENCE HOMOLOGIES

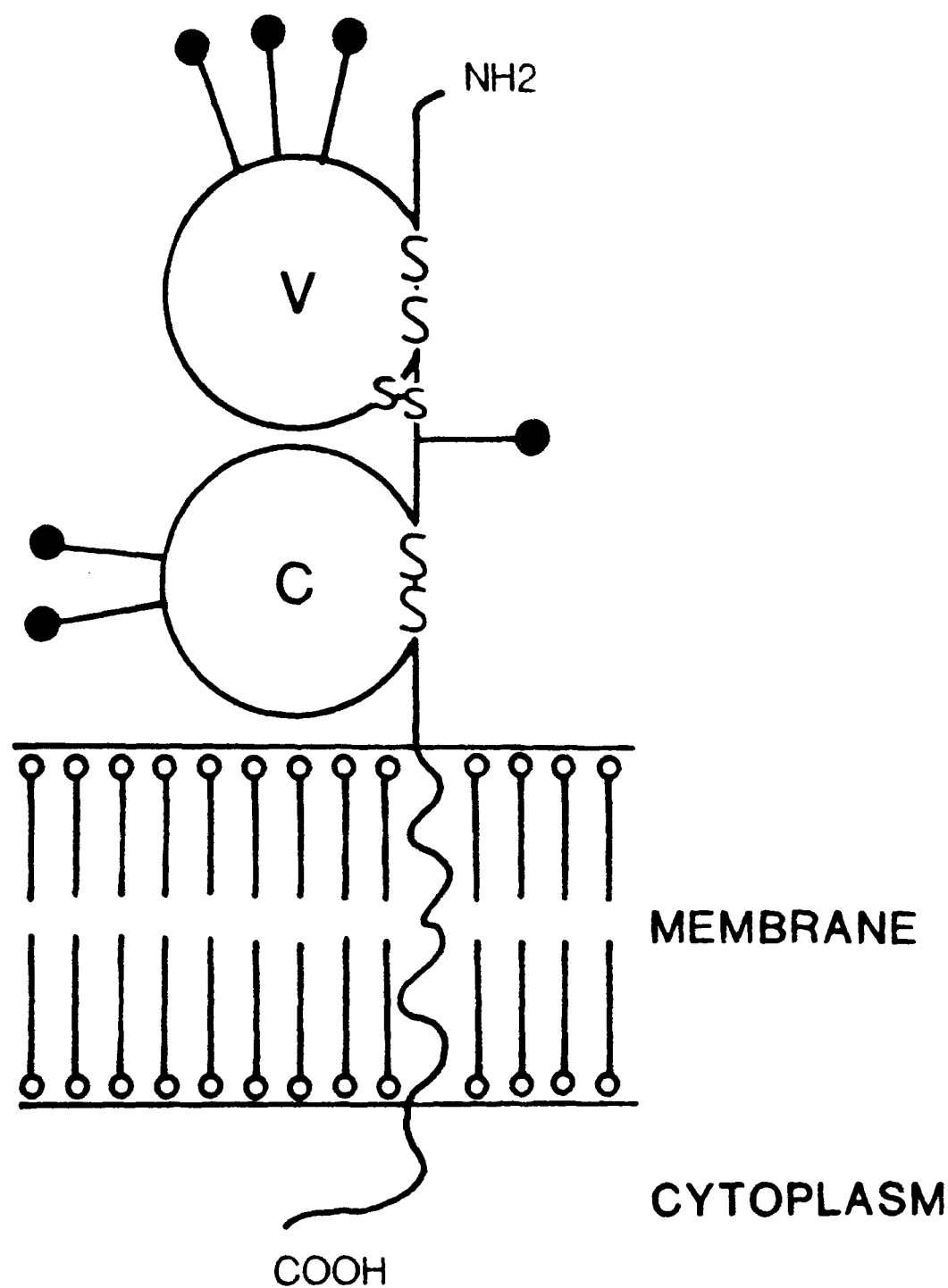
The complete OX-2 protein sequence was compared with 3061 sequences in the NIH PIR database using a basic protein database search program. The computer search did not detect any strong or consistent homologies with sequences other than those of immunoglobulins.

5.4 A MODEL FOR THE STRUCTURE OF OX-2 ANTIGEN

The sequence data from Chapter 4 and the Ig homologies discussed in this chapter are brought together in the model of the OX-2 molecule depicted in figure 5.3. Alignments with Ig-like sequences in figure 5.1 place the conserved disulphide link between Cys residues 20 and 91 in domain I and 130 and 184 in domain II. The 'extra' pair of Cys at positions 88 and 106 in domain I are in a suitable position to form an additional disulphide bond between strands F and G. These disulphide linkages await confirmation by protein chemistry techniques. The configuration of the molecule at the cell surface is discussed further in Chapter 8 section 2.

Figure 5.3

MODEL FOR THE STRUCTURE OF OX-2 ANTIGEN



KEY ● — = N-linked carbohydrate

§ = Ig fold conserved disulphide bond

§§ = Additional disulphide bond

} Positions not yet confirmed

5.6 DISCUSSION TO CHAPTER 5

In this Chapter, the sequence comparisons between OX-2 and Ig-related molecules have indicated that the extracellular portion of the OX-2 antigen comprises of two domains that show homology with immunoglobulin. The N-terminal domain fits best with Ig V domains and Thy-1, while the membrane proximal domain is like an Ig C domain. Thus the OX-2 molecule looks most like a single chain of the T cell receptor or an Ig light chain with a membrane spanning segment. The sequence of OX-2 antigen confirms that OX-2 and Thy-1 do indeed have similar simple Ig-related structures as anticipated at the outset of this thesis (Introduction section 1.6, and Barclay and Ward, 1982). The evolutionary relationship between OX-2 and the Ig superfamily will be discussed further in the Final Discussion, section 8.3.

Two other features of the OX-2 sequence are worth mentioning :

1) Homology between the beginning of the Thy-1 sequence (residues 2-10) and the junction of OX-2 domains I and II (residues 100-107).

Rat OX-2	K	V	-	S	G	T	A	C	L
Mouse Thy-1	K	V	I	S	L	T	A	C	L

A computer search of the NIH PIR database for S*TACL finds no other proteins with this sequence. The homology might reflect an ancient gene duplication event of the original Thy-1 sequence.

2) Residues 93 to 101 at the end of the OX-2 domain I correspond in position to the J region of immunoglobulin and T cell receptor variable regions. The OX-2 sequence shares some homology with these J segments, including the highly conserved FGXG sequence shown in figure 5.4. The sequence of different OX-2 clones show no variation in this region, so in OX-2 antigen this segment is not being used to generate diversity as it is in the immunoglobulin and T cell receptor genes.

Figure 5.4

HOMOLOGIES OF RAT OX-2 V-DOMAIN AND J REGIONS

RAT OX-2	M	C	L	F	N	M	F	G	S	G	-	K	V	S	G	T	A
TcR β JT3	S	A	E	T	L	Y	F	G	S	G	T	R	L	T	V	L	
TcR α	F	A	S	A	L	T	F	G	S	G	T	K	V	I	V	L	P
JH1	Y	W	Y	F	D	V	W	G	A	G	T	T	V	T	V	S	S
J κ 4					F	T	F	G	S	G	T	K	L	E	I	K	R
J λ 2					Y	V	F	G	G	G	T	K	V	T	V	L	G

Alignment of C-terminal region of OX-2 Domain I (residues 90 - 105) with typical J regions from the mouse T cell receptor chains, and heavy and light chains of mouse immunoglobulins.

Residues that are identical to OX-2 are boxed. A gap has been introduced into the OX-2 sequence to maximise homology.

References :

TcR β JT3	Chien et al., 1984b
TcR α	Saito et al., 1984
JH1	Sakano et al., 1980
J κ 4	Sakano et al., 1979a
J λ 2	Blomberg and Tonegawa, 1982

Thy-1, CD8 antigen, and the Poly Ig receptor molecules also contain V-like domains which are not involved in antigen recognition, yet these domains show no J sequence homology. The following chapter investigates the intron/exon boundaries in this region of the gene for the human homologue of OX-2.

CHAPTER 6

CLONING THE GENE FOR THE HUMAN HOMOLOGUE OF OX-2

6.1 INTRODUCTION TO CHAPTER 6

Members of the Ig superfamily have a common gene arrangement in which each Ig-like domain is encoded by a separate exon (Introduction 1.5.2.3). In the case of Igs (Tonegawa, 1983) and T cell receptors (Chien et al., 1984) the variable region comprises of a segment of DNA encoding the majority of the V domain, and separately encoded D segments (absent from Ig light chain genes) and J segments which are brought together by a sequential rearrangement during lymphocyte ontogeny to produce the continuous VDJ sequence. Different combinations of the germ-line repertoire of these segments, and variability in the position at which the segments join contributes to the diversity of antigen binding sites.

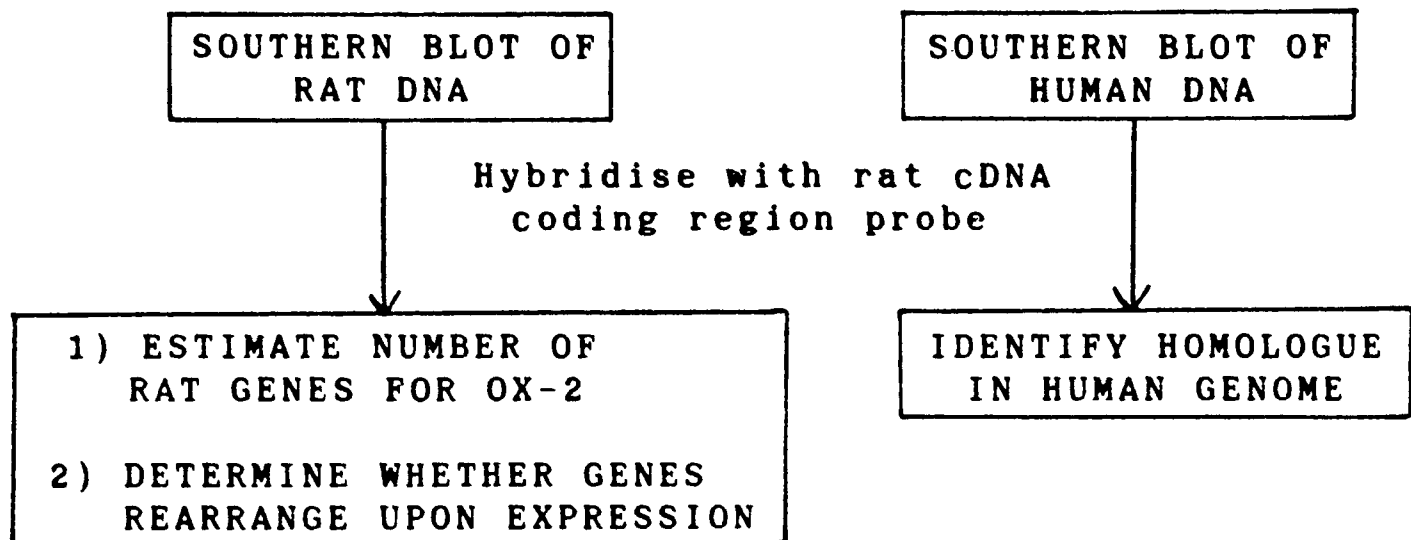
It was therefore of interest to see whether the gene for OX-2 has the typical structure of one exon per Ig-like domain, and in particular whether the region in domain 1 which shares sequence homology with Ig J regions (Chapter 5) is encoded by a separate J segment and becomes rearranged during expression. The gene structure of two other members of the Ig superfamily, the CD8 antigen and Thy-1, had recently been determined (Littman et al., 1985; Seki et al., 1985). These proteins have a domain with the sequence characteristics of a V-like folding pattern but no sequence homology with J regions, and a single exon encodes the whole V-like domain.

The extent to which the structure or expression of OX-2 antigen is conserved in species other than the rat is unknown because as yet no

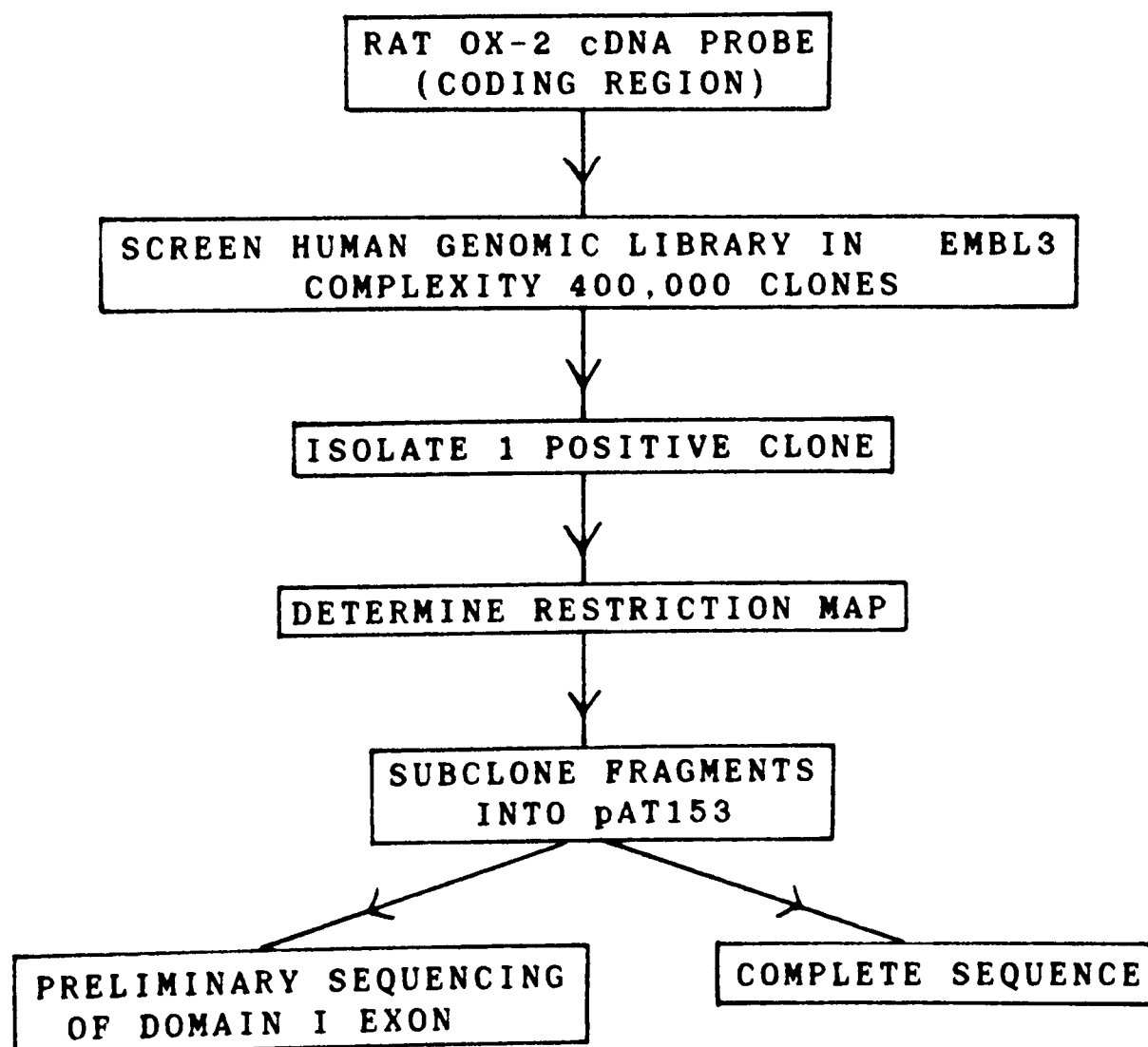
Figure 6.1

STRATEGY FOR ANALYSING THE GENE ENCODING THE HUMAN
HOMOLOGUE OF OX-2

I SOUTHERN BLOT ANALYSIS OF GENE ARRANGEMENT



II ISOLATION OF A GENOMIC CLONE FOR THE HUMAN OX-2 GENE



equivalent molecules have been identified serologically or biochemically (Introduction, 1.6). Thus it would be valuable to identify and characterise OX-2 homologues in other mammalian species to assess the degree of sequence conservation during evolution.

In this Chapter, the rat OX-2 cDNA probe is used in Southern blots to see whether the OX-2 gene becomes rearranged in tissues expressing the antigen. The homologue of the OX-2 gene in rodents and humans is analysed by Southern blotting, and a genomic clone for the human gene isolated in order to examine the gene structure of OX-2 and to characterise the human homologue (flow chart, figure 6.1).

6.2 SOUTHERN BLOT ANALYSIS OF THE GENE FOR OX-2

In this section, the gene(s) for OX-2 in rat and other species are examined using the technique of Southern blotting (Southern, 1979). The hybridisation probe is a nick-translated 900bp Bam-Xba coding region fragment of OX-2 clone px2/11 (figure 4.6, Chapter 4), which excludes the non-coding region in case the poly A tail or (CT)₂₅ tract bound to simple repetitive sequences in the genome.

6.2.1 The rat gene for OX-2

The size and number of genes coding for OX-2 was estimated by digesting rat liver DNA with a selection of nucleases which recognise hexanucleotide sequences and hence cleave the DNA into fragments averaging several kilobases in length. A Southern blot of rat genomic DNA probed with the OX-2 coding region fragment is shown in figure 6.2A. Kpn I and Xho I digests did not give any discernable bands, but in all the other tracks there are between 2 and 5 hybridising bands of various intensities. The limited number of bands indicate that OX-2 is not a member of a multigene family. In each track there is at least one very

weakly hybridising band which remains even after high stringency washing. This might represent an exon containing only a short stretch of OX-2 sequence, or a cross-hybridising sequence from elsewhere in the genome. The maximum size of the gene for the coding region of OX-2 is estimated to be about 8Kb, from the double enzyme digest in track 11.

To see whether the gene for OX-2 becomes rearranged in tissues which express the protein, DNA from rat liver (control) and rat thymocytes (express OX-2) was digested with 4 different restriction enzymes that were known to produce distinctive banding patterns. Major gene rearrangements should disrupt the size and number of the fragments hybridising in a Southern Blot, but as shown in figure 6.2B, the banding pattern is the same for both sources of the DNA. The slower migration of thymocyte DNA fragments in all tracks is probably due to the effect of salt in the DNA preparation. Thus no rearrangements in the OX-2 gene could be detected with the restriction nucleases described here.

6.2.2 The gene for OX-2 in other mammalian species

The rat OX-2 coding region probe was hybridised to Southern blots of mouse and human genomic DNA and washed under low stringency conditions (2 x SSC) to permit cross-hybridisation between homologous sequences. The autoradiograph is shown in figure 6.3. The rat probe hybridises strongly to a limited number of fragments in the mouse DNA (lane B) showing that the sequence of the gene for OX-2 is conserved between rodents. A Southern blot of human DNA gives a weaker signal with higher background (lane C) but there is a distinct band at 3.8⁷Kb indicative of an OX-2 homologue in the human genome.

Figure 6.2

A : SOUTHERN BLOT ANALYSIS OF THE OX-2 GENE

10µg of restriction enzyme digested DNA was run on a 0.8% agarose gel and blotted onto nitrocellulose. The filter was hybridised with the nick-translated Bam-Xba 900bp coding region fragment of clone pX2/11, and washed to a stringency of 0.2 x SSC. Autoradiography was overnight at -70°C with preflashed film and intensifying screen.

Restriction enzymes :

- 1 = Bam HI
- 2 = Bgl II
- 3 = BstE II
- 4 = EcoR I
- 5 = Hind III
- 6 = Kpn I
- 7 = Pst I
- 8 = Pvu II
- 9 = Xba I
- 10 = Xho I
- 11 = EcoR I + Hind III

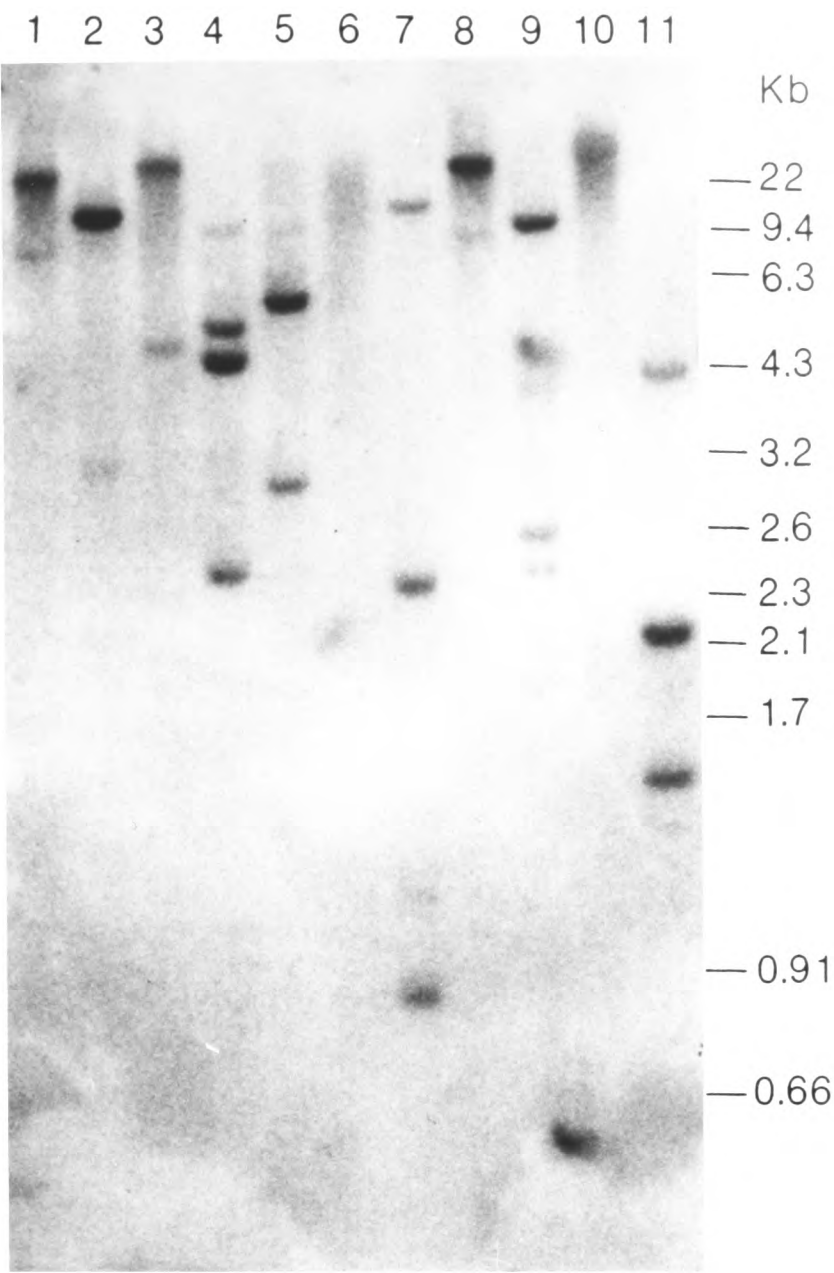
B : SOUTHERN BLOT ANALYSIS OF REARRANGEMENTS IN THE RAT OX-2 GENE

10µg of restriction enzyme digested DNA was run on a 0.8% agarose gel and blotted onto nitrocellulose. The filter was hybridised with the nick-translated Bam-Xba 900bp coding region fragment of clone pX2/11, and washed to a stringency of 0.2 x SSC. Autoradiography was overnight at -70°C with preflashed film and intensifying screen.

KEY T = DNA extracted from rat thymocytes
L = DNA extracted from rat liver

- Restriction enzymes :
- 1 = EcoR I
 - 2 = Hind III
 - 3 = Pst I
 - 4 = Xba I

A



B

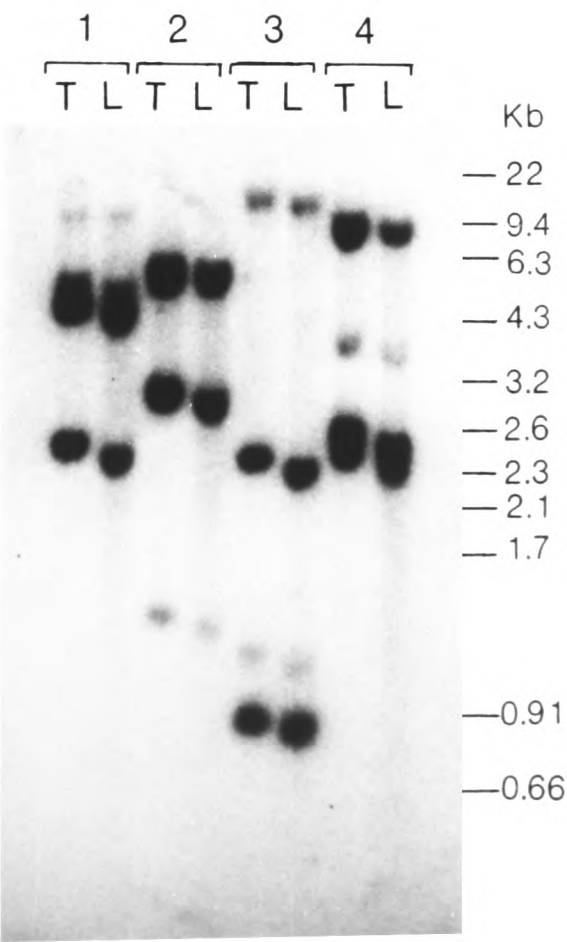
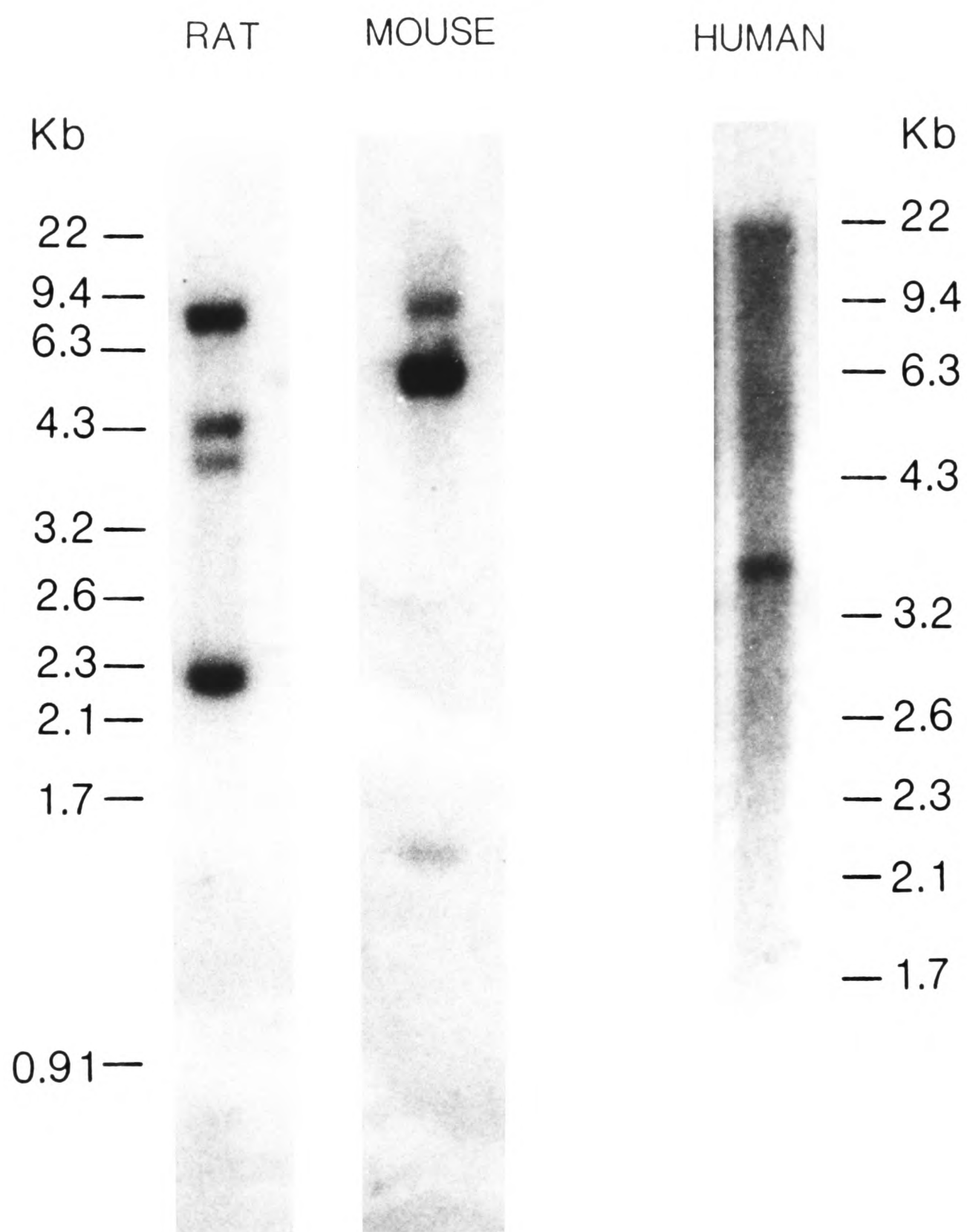


Figure 6.3

SOUTHERN BLOT ANALYSIS OF THE OX-2 GENE IN RAT, MOUSE AND HUMAN

5-10µg of restriction enzyme digested DNA was run on 0.9% agarose gels and blotted onto nitrocellulose. The filters were hybridised with the nick-translated Bam-Xba 900bp coding region fragment of clone pX2/11, and after washing (see below) were autoradiographed for 7 days at -70°C with preflashed film and intensifying screen.

Species	Restriction digest	Washing stringency
RAT	EcoR I	0.2 x SSC
MOUSE	EcoR I	2 x SSC
HUMAN	Hind III	2 x SSC



6.3 ISOLATION OF A GENOMIC CLONE FOR THE HUMAN HOMOLOGUE OF OX-2

Judging from the Southern blot hybridisation signal in figure 5.2C, it should be possible to use the rat OX-2 cDNA probe to isolate a genomic clone of the human OX-2 gene. The hybridising sequence is contained within a 3.7Kb Hind III band which provides a useful diagnostic restriction map for the correct OX-2 clone as well as being a conveniently sized DNA fragment for subsequent subcloning and sequencing.

6.3.1 The human genomic library in λ EMBL3

A human genomic library was constructed in the vector λ EMBL3 by Colin R. Sharpe of the Sir William Dunn School of Pathology, as described in the Methods Section 2.3.17. The library of complexity 400,000 pfu's was plated without amplification and prepared for screening by the method of Benton and Davis (1977). The filters had been screened successfully with cDNA probes for human Apolipoproteins AI and CIII and fibronectin, before being made available to search for the OX-2 gene.

6.3.2 Screening of the λ EMBL3 library

For the first screen of the library, the filters were hybridised with the nick-translated rat OX-2 coding region cDNA probe and washed to a stringency of 1 x SSC. 7 weakly positive spots could be distinguished after 48 hours autoradiography.

The positive clones were rescreened, and 1 out of the 7 gave a positive hybridisation signal. This clone was plaque purified, toothpicked onto a grid array and re-screened. The plaques gave a strong signal after 4 hours autoradiography, which is reasonable for a cross-species hybridisation. The positive clone is designated HX2/EMBL3.

6.3.3 Restriction mapping of the λ genomic clone

DNA was purified, digested with restriction enzymes, fractionated on 250ml agarose gels and transferred to nitrocellulose. Several nitrocellulose blots were made from each gel.

Figure 6.4 shows Southern blots of 3 restriction enzyme digests with or without the endonuclease Sal I which cleaves the genomic insert from the vector. Filter A is hybridised with the usual coding region probe while Filter B is hybridised with an M13 probe specific for domain I, containing nucleotides 1-445. Both probes hybridise to a 3.7Kb Hind III band characteristic of the human OX-2 gene as shown by Southern blotting in figure 6.3C.

The domain I probe recognises fewer bands in the Bgl II and EcoRI digests, than the total coding region probe and thus can be used to orientate the OX-2 gene. By hybridising the two probes with a blot of double digests (data not shown) it was possible to compile the restriction map of the OX-2 gene shown in figure 6.5.

The total length of the genomic insert in the clone is 18.6Kb, and the mapping places the OX-2 gene near the short arm (right end) of the vector. The 3.7Kb Hind III fragment contains an intron of at least 1.5Kb separating a 1.1Kb fragment homologous to the domain I probe from the remainder of the hybridising sequence. This Hind III band does not include any repetitive DNA as hybridisation of the Southern blots with nick-translated total human DNA detected the nearest repetitive sequence at least 9Kb downstream from the coding region of the OX-2 gene.

6.4 PRELIMINARY SEQUENCING OF THE GENE FOR THE HUMAN HOMOLOGUE OF OX-2

Confirmation that the clone coded for an OX-2 homologue required gene sequence data. The V-type domain was of particular interest

Figure 6.4

SOUTHERN BLOT ANALYSIS OF THE HUMAN OX-2 GENE IN THE λ GENOMIC CLONE
HX2/EMBL3

1 μ g of restriction enzyme digested λ genomic clone HX2/EMBL3 was run on a 0.9% agarose gel and transferred to nitrocellulose. Two copies were made from the same agarose gel.

Filter A was hybridised with the nick-translated Bam-Xba 900bp fragment of clone pX2/11 which covers the entire coding region.

Filter B was hybridised with a 32 P labelled M13 subclone containing the rat OX-2 domain I sequence only, corresponding to nucleotides 1 - 446 on the composite sequence, figure 4.6. (the probe was kindly provided by Dr G.McCaughan).

Both blots were washed to a stringency of 2 x SSC.

Blot A was autoradiographed overnight at -70°C with preflashed film and intensifying screen.

Blot B was autoradiographed for 6 days at -70°C with preflashed film and intensifying screen.

Track 1 = undigested

2 = Hind III

3 = Bgl II

4 = EcoR I

- = digest without Sal I

+ = digest with Sal I

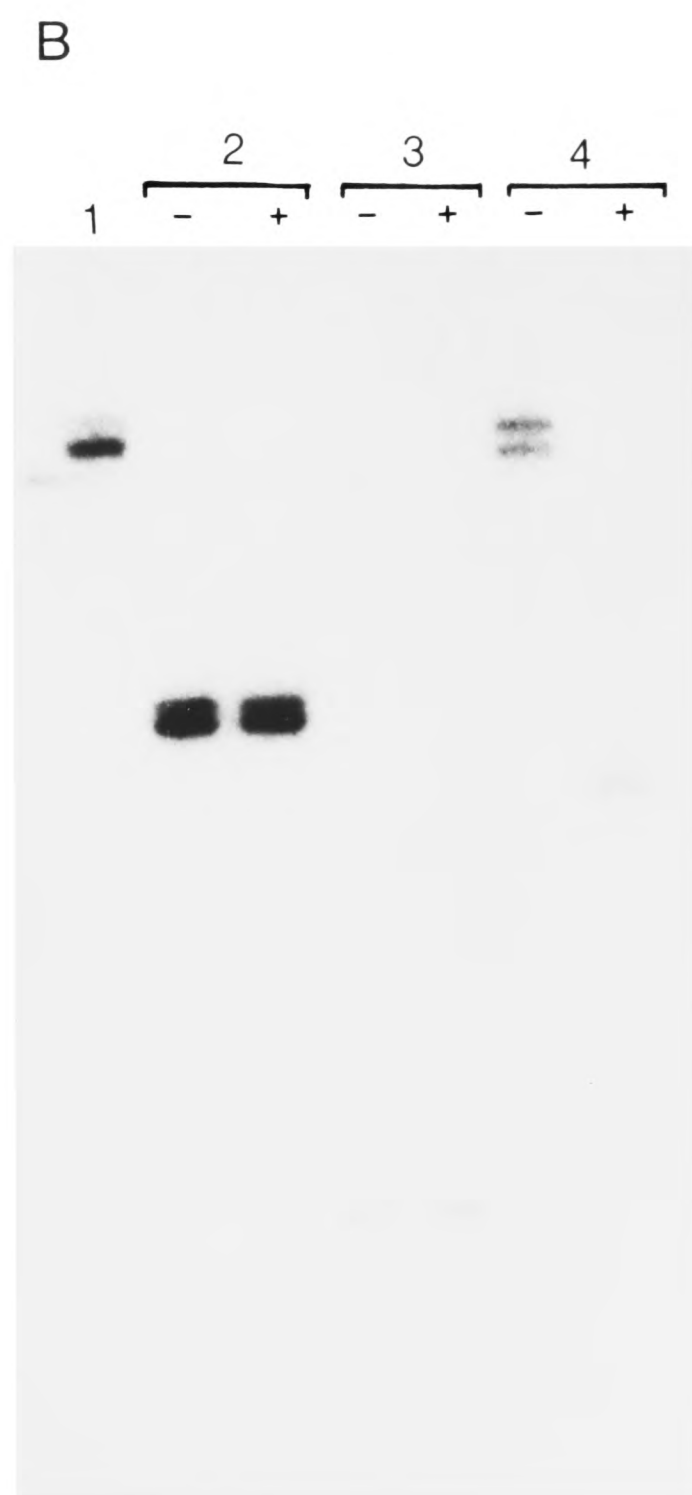
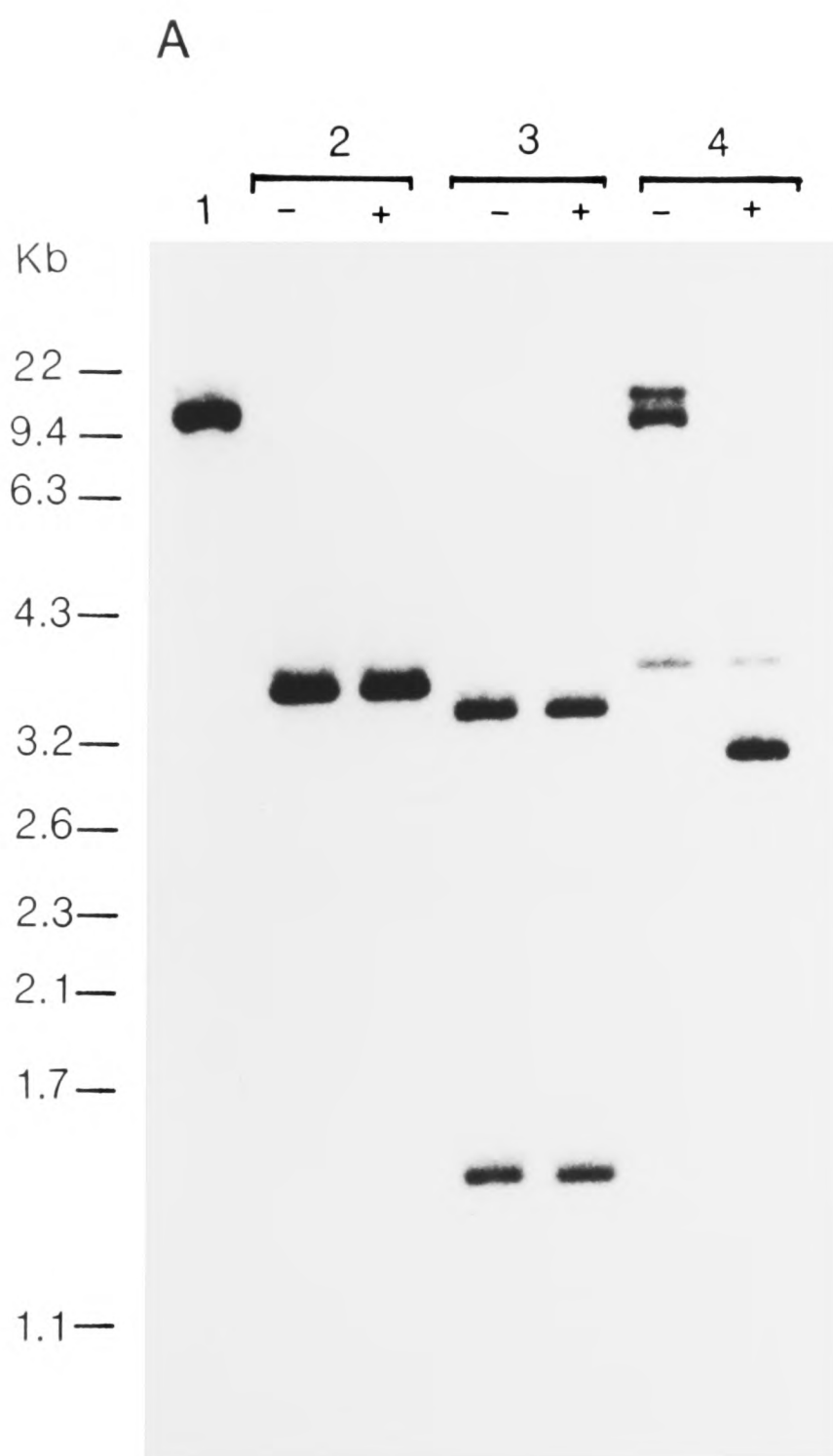


Figure 6.5

RESTRICTION MAP OF THE OX-2 GENE IN THE λ GENOMIC CLONE HX2/EMBL3

In each map, the thin line represents genomic DNA and the hatched bars vector DNA.

A : EcoR I map of the genomic insert of HX2/EMBL3

L END = left end of λ ("long arm")
R END = right end of λ ("short arm")

RI = EcoR I sites in the genomic insert.

B : Restriction map of the region around the human OX-2 gene

Sites for 5 restriction endonucleases were deduced from Southern blot analysis of the genomic clone HX2/EMBL3.

V = Fragment of DNA deduced to contain the gene for domain I because it hybridises with both the total OX-2 coding region probe and the domain I specific probe (see previous figure, 6.5)

C = Fragment of DNA deduced to contain the gene for domain II (+ transmembrane and cytoplasmic region) because it hybridises with the total coding region probe but not the domain I-specific probe (see previous figure, 6.5)

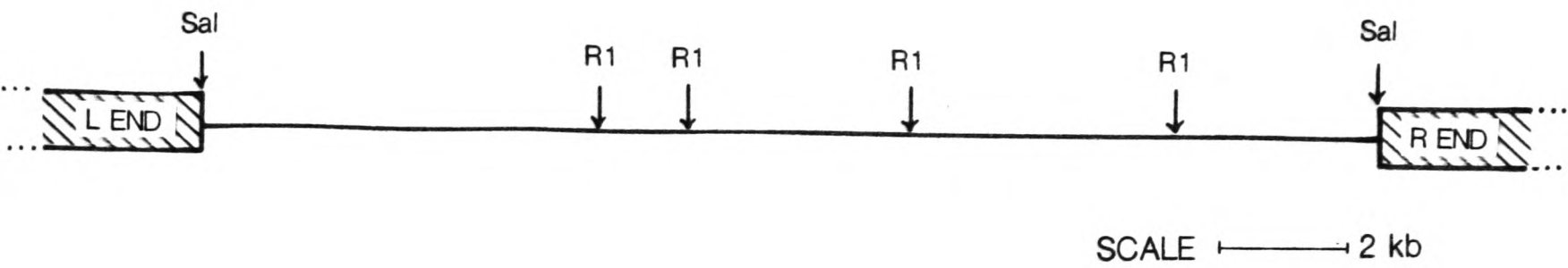
C : Restriction map of the 3.7kb Hind III fragment subcloned into pAT153/Pvu II/8

Restriction sites in the gene, and the flanking sites in the vector are indicated.

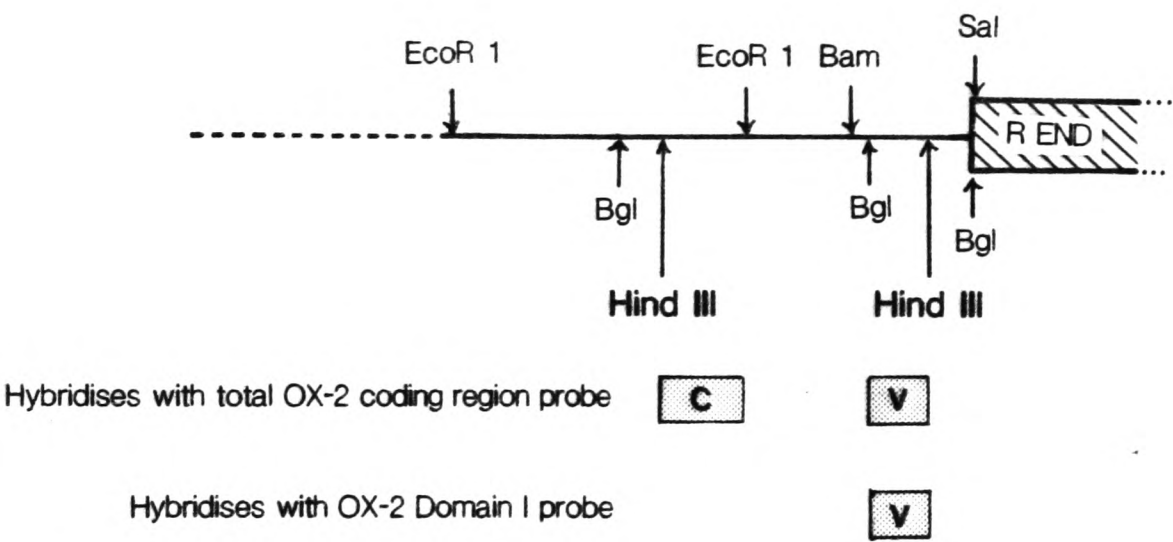
Note that the OX-2 gene is drawn in the opposite orientation to figures A and B.

Figure 6.5

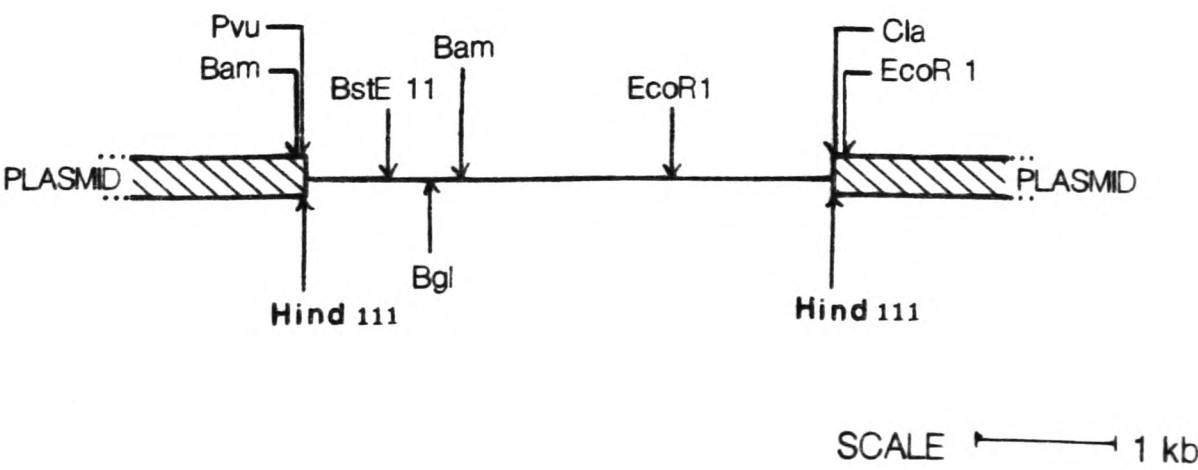
A : EcoR I MAP OF GENOMIC INSERT OF λ CLONE HX2/ EMBL3



B : RESTRICTION MAP OF THE REGION AROUND THE HUMAN OX-2 GENE



C : RESTRICTION MAP OF SUBCLONE pHX2/Hind 111/2



because of the possibility of a J exon (Introduction, 6.1) so this region was analysed first.

6.4.1 Subcloning the λ genomic clone

The 3.7Kb Hind III fragment was subcloned into the Hind III site of pAT153/Pvu11/8 by Dr G. McCaughan and transformants identified by screening with the nick-translated rat OX-2 coding region probe. Plasmid DNA was purified on CsCl gradients, and the restriction map of subclone pHX2/H III is shown in figure 6.5B.

6.4.2 The sequence of the human gene for OX-2 domain I

Preliminary Maxam and Gilbert sequencing around the Bgl II site in the subclone established that this portion of the gene coded for a protein homologous to domain I of OX-2. Figure 6.6A shows the nucleotide sequence and translate of this region. The sequence was confirmed and extended by Dr G. McCaughan by dideoxy sequencing of fragments subcloned into M13 mp8 (Sanger et al., 1977; Messing, 1983).

The protein sequence of rat and human OX-2 align very well along the entire length of domain I, as shown in figure 6.6B. The sequences are 78% identical at the protein level and 82% identical at the nucleotide level, and no gaps are needed to align the sequences. Conservative substitutions occur in nearly half the positions where the two sequences are not identical. All the key residues characteristic of Ig domains are conserved in the human OX-2 sequence, namely the conserved disulphide bond Cys residues, the Trp in β strand C, and the V-type sequence pattern in β strand F. The "additional" pair of Cys residues in the rat sequence is also present in the human, and all six N-glycosylation sites are conserved.

Domain I is encoded by 1 exon extending from the 2nd nucleotide of the codon for amino acid 2 to the 1st nucleotide of the codon for amino

Figure 6.6

A : NUCLEOTIDE SEQUENCE AND TRANSLATION OF THE EXON ENCODING DOMAIN I
OF THE HUMAN OX-2 GENE

The presumed intron splice sites AG and GT are underlined.

B : ALIGNMENT OF RAT AND HUMAN OX-2 DOMAIN I SEQUENCES

Identical residues are boxed. The numbering for the amino acids is from figure 4.7. The exon for domain I of the human gene does not include the codon for amino acid number 1, and only includes the last two nucleotides of the codon for amino acid 2.

A : NUCLEOTIDE SEQUENCE AND TRANSLATE OF THE EXON ENCODING
DOMAIN I OF THE HUMAN OX-2 GENE

Q V V T Q D E R E Q L Y T T
..TTATGCTTCCATAGTGCAAGTGGTGACCCAGGATGAAAGAGAGCAGCTGTACACAACT
A S L K C S L Q N A Q E A L I V T W Q K
GCTTCCTTAAAATGCTCTCTGCAAAATGCCCAGGAAGCCCTCATTGTGACATGGCAGAAA
K K A V S P E N M V T F S E N H G V V I
AAGAAAGCTGTAAGCCCAGAAAACATGGTCACCTTCAGCGAGAACCATGGGGTGGTGATC
Q P A Y K D K I N I T Q L G L Q N S T I
CAGCCTGCCTATAAGGACAAGATAAACATTACCCAGCTGGGACTCCAAAACCTCAAACATC
T F W N I T L E D E G C Y M C L F N T F
ACCTTCTGGAATATCACCTGGAGGATGAAGGGTGTTACATGTGTCTCTTCAATACCTTT
G F G K I S G T A C L T V Y
GGTTTTGGGAAGATCTCAGGAACGGCCTGCCTCACCGTCTATGGTGAGAATCTCGAG...

B : HOMOLGY BETWEEN RAT AND HUMAN DOMAIN I SEQUENCES

RAT	Q V E	V V T Q D E R	K L	L	H	T T A S L	R	C S L	K T T
HUMAN	..Q	V V T Q D E R	E Q	L	Y	T T A S L	K	C S L	Q N A

	Q E P	L I V T W Q K K K A V	G	P E N M V T	Y	S	K A	H
	Q E A	L I V T W Q K K K A V	S	P E N M V T	F	S	E N	H

	G V V I Q P	T	Y K D	R	I N I T	E	L G L	L	N	T S	I T F
	G V V I Q P	A	Y K D	K	I N I T	Q	L G L	Q	N	S T	I T F

	W N T T L	D	D E G C Y M C L F N	M	F G	S	G K	V	S G T
	W N I T L	E	D E G C Y N C L F N	T	F G	F	G K	I	S G T

	A C L T	L	Y
	A C L T	V	Y

acid 111. The exon is flanked by sequences corresponding to the consensus splice sites for introns (Breathnach and Chambon, 1981). At the 5' end of the exon is the sequence 5' (Py)₅ ATAG|TG 3' which corresponds with the intron acceptor sequence 5' (Py)_n XCAG|G^G/_T 3', while at the 3' end of the exon the sequence 5' ATG|GTGAG 3' corresponds to the intron donor sequence 5' ^A/_CAG|GT^A/_GAG 3'. Hence the amino acid residues 93 to 105 which have homology with the J sequence of immunoglobulins (Discussion, Chapter 5) are included in the exon for the V-type domain and are not encoded by a separate exon.

6.6 DISCUSSION TO CHAPTER 6

Homologues of the OX-2 gene have been identified by Southern blot analysis of mouse and human DNA, and the gene for the human homologue of OX-2 isolated and partially sequenced. The arrangement of the OX-2 gene resembles the organisation of the other Ig superfamily molecules i.e. the domain I forms a separate exon, from the leader sequence and domain II. The exon boundary at the 5' end of domain I is between residues 1 and 2 in the mature protein, in contrast to T8 (Littman et al., 1985b) and Thy-1 (Seki et al., 1985) where the corresponding splice site is a few amino acid residues into the leader sequence. At the 3' end of domain I, the sequence homologous to the J region of immunoglobulins is not contained with a separate exon but is continuous with the V-type domain sequence. No other J-like exons have been identified 200bp 3' of domain I or 400bp 5' of domain II (G. McCaughan, personal communication). Also, Southern blotting using four different restriction enzymes could not detect any rearrangement of the OX-2 gene in tissues which express the protein. Thus the OX-2 protein has a

sequence consistent with the folding pattern of a V-domain but lacks the mechanisms for creating diversity observed in Ig's and T cell receptors.

The low number of restriction fragments hybridising with the OX-2 cDNA probe on Southern blots indicates that there is probably only 1 copy (or at the most a few tightly linked copies) of the gene for OX-2. This is supported by the isolation of just one OX-2 clone from a genomic library which had a complexity approximately equivalent to 1 human genome (400,000 clones). However, the comparatively weak cross-species hybridisation signal and the background on the library filters may have obscured any other positive clones. It would be informative to screen another genomic library with the human probe to see if there were more clones for OX-2. The Thy-1 gene is also present at 1 copy per haploid genome (Seki et al., 1985) whereas T cell receptors (Chien et al., 1984) and Igs (Tonegawa, 1983) are multigene families consisting of large numbers of genes encoding V domains and several genes encoding C domains.

The OX-2 gene is conserved across several mammalian species, although Southern blotting cannot distinguish whether this gene is functional in the other species. From the sequence of the human gene completed so far, there is nothing to suggest that the gene is incapable of being expressed. Moreover, the high degree of sequence conservation between domain I of rat and human OX-2 (78% identity) implies a corresponding conservation of function. This degree of homology is higher than between human and rodent forms of some other Ig-related molecules, for example the V-like domains of the CD8 antigens T8 and OX-8 are only 42% homologous (Littman et al., 1985b; Johnson et al., 1985). The majority of the differences between the rat and human OX-2 sequence occur at the corners of the β strands (assigned in figure 5.1)

and this is also seen when mouse and rat Thy-1 sequences (82% homologous) are compared (Williams and Gagnon, 1982). Thus, the organisation and sequence of the human OX-2 gene provides additional evidence that OX-2 antigen is a descendent of the primordial Ig-like domain and a member of the Ig superfamily, and this is discussed further in Chapter 8. In the next Chapter, the information about the OX-2 gene arrangement is used to design S1 nuclease mapping experiments to investigate the structure of OX-2 mRNA.

CHAPTER 7

CHARACTERISATION OF ABNORMAL OX-2 TRANSCRIPTS IN DIFFERENT RAT TISSUES

7.1 INTRODUCTION TO CHAPTER 7

Not all OX-2 clones isolated from the rat thymocyte cDNA library in Chapter 4 appear to be capable of encoding a normal functional OX-2 protein. Insertions and/or deletions in several clones have removed portions of coding region and altered the reading frame so that the sequence can no longer be translated into the OX-2 polypeptide characterised by protein sequencing. The arrangement of the human OX-2 gene was determined in Chapter 6, and the observation below that deletions coincide with the intron-exon boundaries around domain I indicates that these cDNA clones might have been derived from incorrectly spliced thymocyte messenger RNA.

It is unusual for alternate splicing patterns to produce non-functional transcripts, and thus it was of interest to see whether species of OX-2 message exist which also have an abnormal structure. This chapter describes S1 nuclease mapping of OX-2 mRNA from different rat tissues.

7.2 THE ABNORMAL FEATURES OF THYMOCYTE CDNA CLONES COINCIDE WITH INTRON/EXON BOUNDARIES

Twelve independently isolated rat thymocyte OX-2 cDNA clones (Chapter 4) were mapped with restriction endonucleases. The clones fitted into six groups according to clone length and arrangement (table 7.1) and the members within each group were indistinguishable, suggesting that they are multiple copies of the same cDNA. Three out of

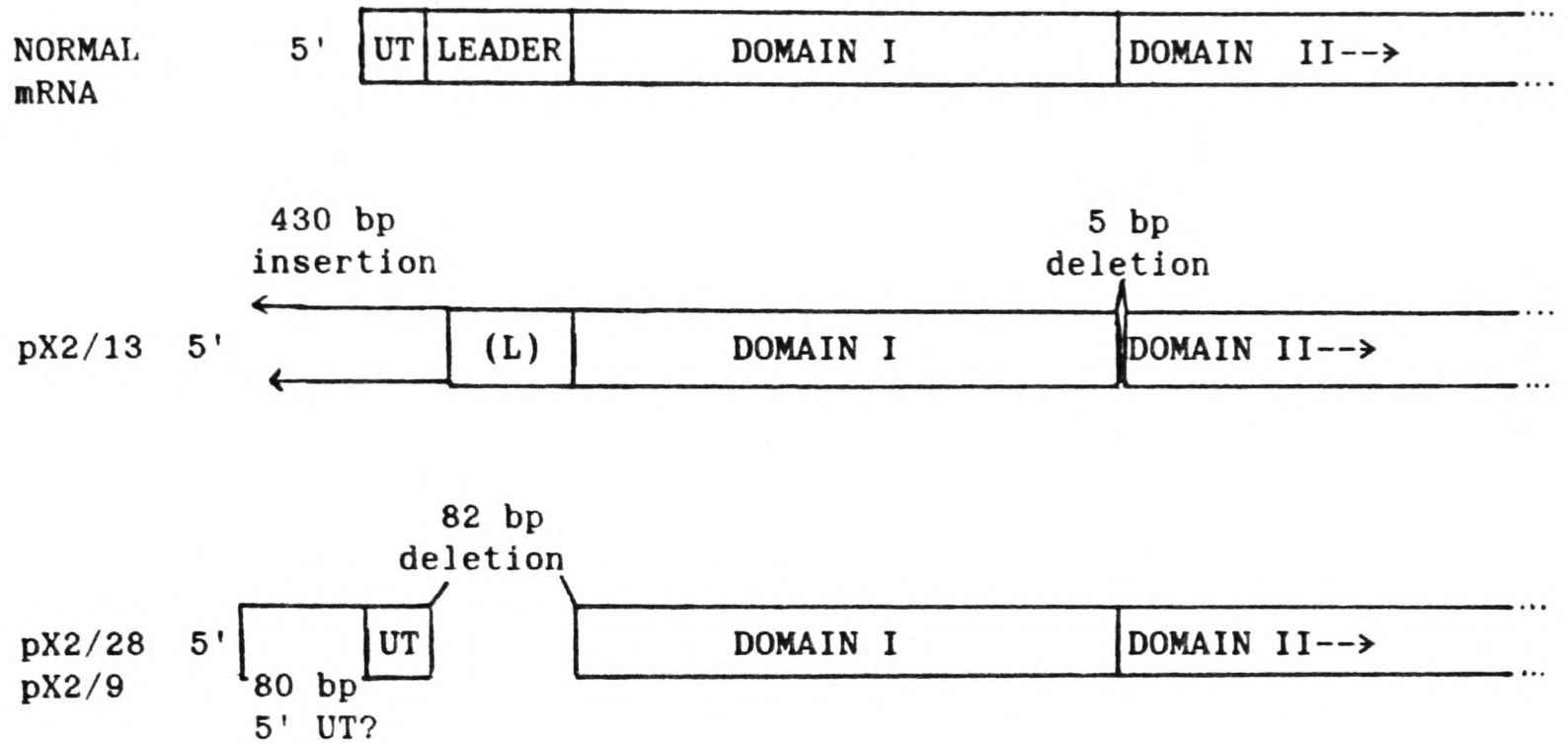
Table 7.1

CLASSES OF RAT THYMOCYTE cDNA CLONES

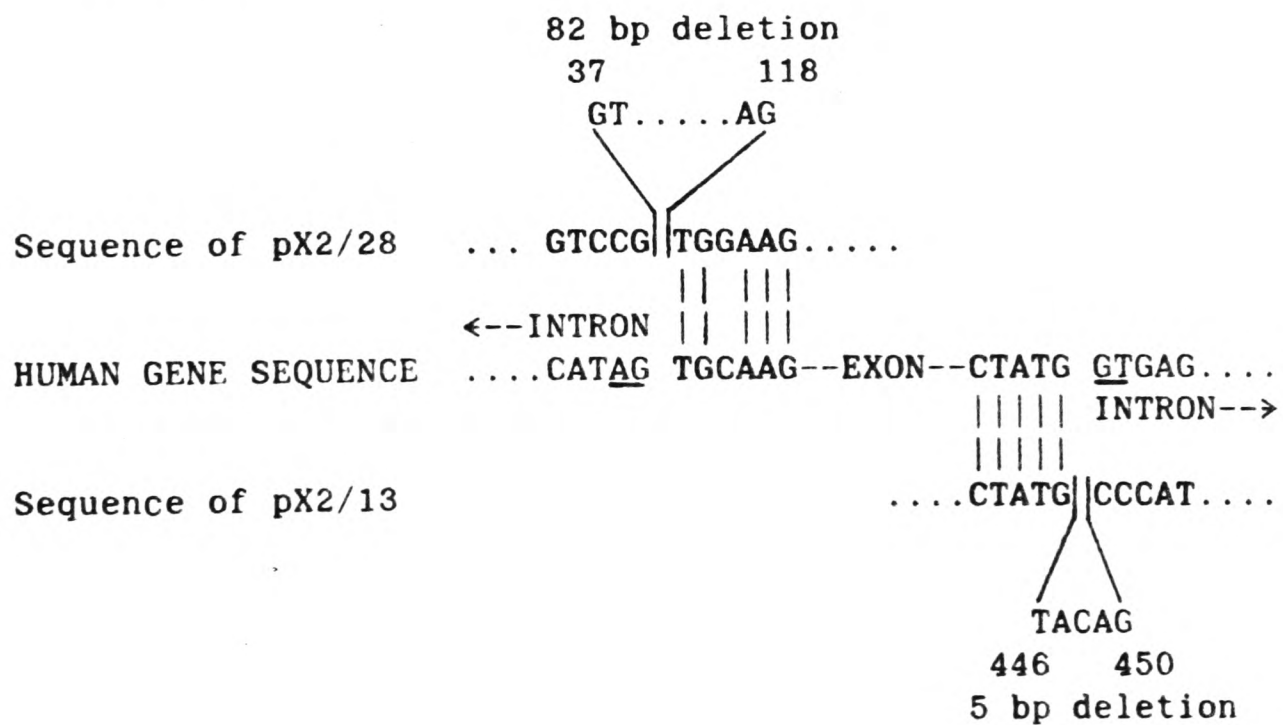
Group	Number of clones in this group	Representative clone	Clone length	Structural abnormality	Comments
A	2	pX2/13	2.6 kb	1) insertion 2) deletion	
B	2	pX2/28	2.1 kb	deletion	
C	3	pX2/9	2.1 kb	deletion	mapped only, not sequenced
D	3	pX2/24	2.2 kb	NONE	
E	1	pX2/11	1.7 kb	NONE	
F	1	pX2/25	1.9 kb	NONE	Only partially sequenced

Figure 7.1

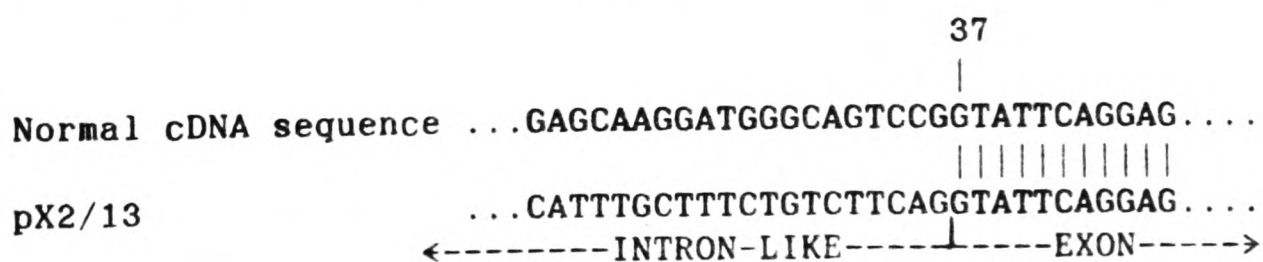
A : ANOMALOUS STRUCTURES OF OX-2 THYMOCYTE cDNA CLONES



B : DELETIONS COINCIDE WITH INTRON/EXON BOUNDARIES OF DOMAIN I OF THE HUMAN GENE



C : THE INSERTION IN CLONE pX2/13 RESEMBLES AN INTRON



six of these classes of clones contain structural abnormalities distinct from the "loopback" inversions (Chapter 4). Clone pX2/13 has been referred to earlier (section 4.6.3.2) while clones pX2/28 and pX2/9 are representative of groups of clones not previously examined. The positions of insertions and deletions identified in these clones are indicated in figure 7.1A. At first, these anomalous structures were thought to be cloning artefacts, but the data on the gene structure of OX-2 (Chapter 6) suggests that they can all be attributed to splicing events as described below:

1) Apparent deletion of the leader sequence exon Two separate groups of clones, pX2/28 and pX2/9, contain a deletion of 82bp (nucleotides 37-118, numbering as figure 4.7) which has removed the majority of the leader sequence and the first amino acid of the coding sequence, leaving the initiator methionine out of reading frame. The 3' end of the deletion coincides exactly with the splice junction between the leader sequence and domain I (figure 7.1B) The arrangement of the OX-2 gene at the 5' end of the deletion is not yet known, but in view of the sequence of clone pX2/13 in this region (see point (3) below), it is likely that this also coincides with a splice junction. It appears that a complete exon encoding the leader peptide has been removed. In fact, the deleted DNA sequence 5'GTATTC...CTCAAG 3' commences and ends with the characteristic donor and acceptor splice sites (Breathnach and Chambon, 1981) and could be recognised as an intron itself and excised.

Both pX2/28 and pX2/9 are 80 nucleotides longer at their 5' ends than the composite sequence in figure 4.7, and this extra DNA may represent additional 5' noncoding sequence absent from the cDNA clones examined in Chapter 4. When the extra sequence is translated in the same

reading frame as the rest of the coding sequence, the protein lacks an initiator methionine or leader peptide.

2) Deletion between domains I and II In clone pX2/13, a deletion of 5 bp (446-450) alters the reading frame of the coding sequence and introduces in-frame stop codons. The deletion coincides with the splice junction at the 3' end of domain I (figure 7.1B), and comprises of the first 5 nucleotides of the exon for domain II. Possibly the deleted sequence TACAG was mistaken for the end of the intron when domains I and II were spliced together.

3) Intron-like insertion in the leader sequence In clone pX2/13, the first 36 bp of the OX-2 cDNA sequence are missing, and replaced by 430bp of unidentified DNA (figure 7.1C). This removes the initiator methionine, and substitutes a sequence which does not resemble a leader peptide. The sequence 5' TCTTCAG|GT 3' at the junction between the extra DNA and the normal OX-2 cDNA agrees with the consensus intron acceptor sequence 5' (Py)_nXCAG|GG/T 3' (Breathnach and Chambon, 1981). The position of the junction (nucleotide 37) corresponds to the 5' end of the deletion in clones pX2/28 and pX2/9, and thus the additional DNA in clone pX2/13 could be an unspliced intron within the leader sequence.

7.3 S1 NUCLEASE MAPPING OX-2 MRNA FROM DIFFERENT TISSUES

To investigate whether species of OX-2 transcripts exist which also have an abnormal structure, S1 mapping of RNA was performed using cDNA probes which covered the splice junctions 5' and 3' of domain I.

7.3.1 S1 nuclease mapping the junction between leader sequence and domain I

Figure 7.2C shows the design of the S1 mapping probe, a 152 nucleotide single stranded Sau 961 fragment generated from a normal cDNA

Figure 7.2

S1 NUCLEASE MAPPING OF OX-2 mRNA ACROSS THE JUNCTION BETWEEN THE LEADER
SEQUENCE AND DOMAIN I

A : The S1 probe shown in figure C was hybridised with RNA from different rat tissues and the S1 nuclease products analysed on a 6% denaturing polyacrylamide gel. The gel was fixed and dried and autoradiographed for 10 days at -70°C with pre-flashed film and intensifying screen.

M = size markers, sizes as indicated on the right.

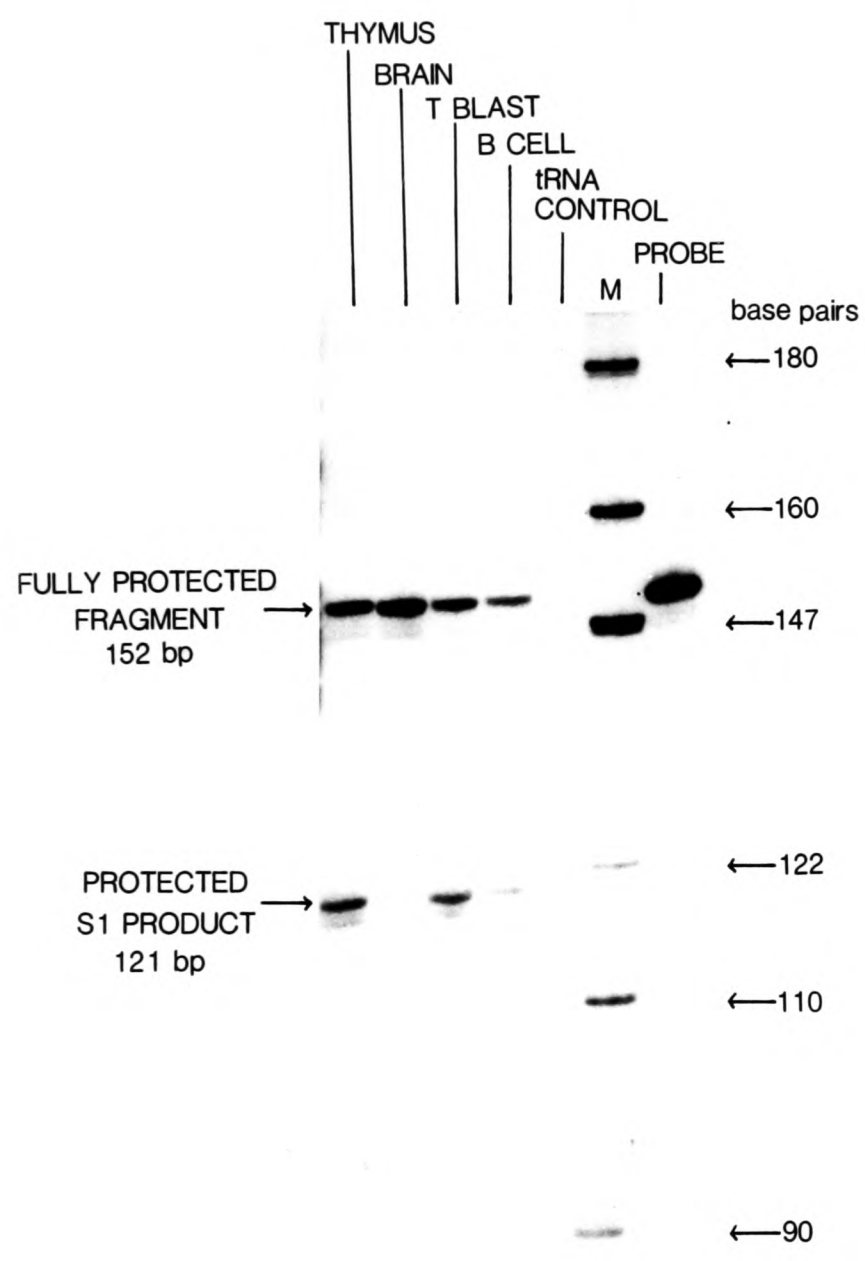
B : Densitometer scan of autoradiograph A.

$$\text{Ratio } \frac{\text{area of 121bp band}}{\text{area of 152bp band}} = \frac{\text{amount of 'abnormal' RNA}}{\text{amount of 'normal' RNA}}$$

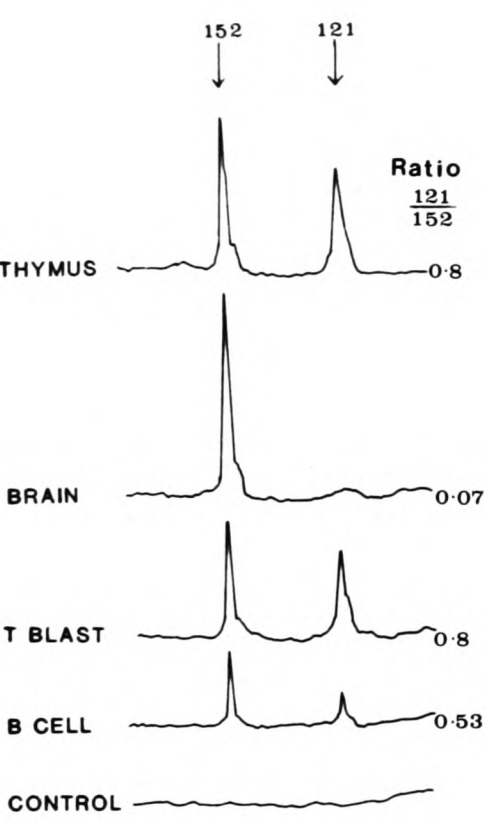
C : Design of the radiolabelled probe

The 152bp Sau 96 fragment (nucleotides 85-238) indicated by the heavy black arrow was prepared by digesting the 'normal' cDNA clone pX2/24 with Sau 96 and labelling the 5' ends with T4 polynucleotide kinase and [$\gamma^{32}\text{P}$]dATP. The single stranded antisense strand fragment labelled at nucleotide 238 was isolated by strand separation.

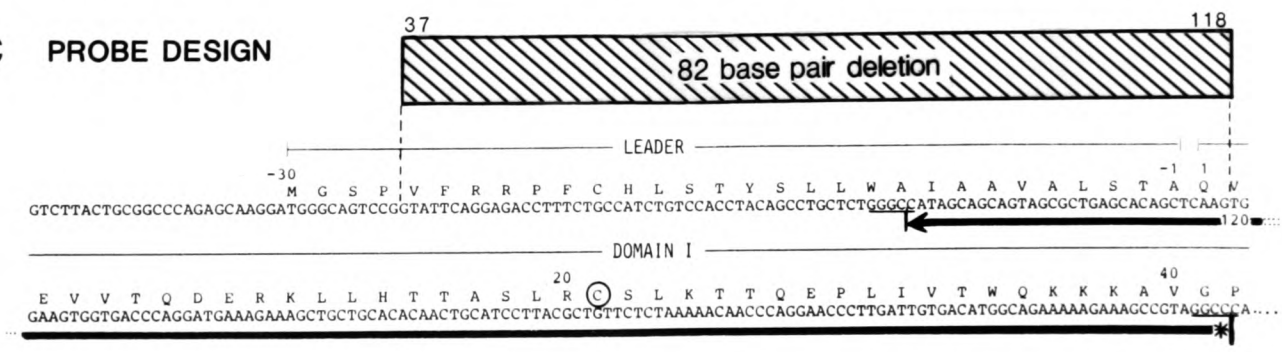
A S1 MAPPING OX-2 mRNA FROM DIFFERENT TISSUES



B DENSITOMETER SCAN



C PROBE DESIGN



clone. The probe is labelled at nucleotide 238 on the 5' end of the antisense strand, and extends from domain I back into the leader sequence, crossing the 3' end of the region missing from clone pX2/28. The probe was hybridised with total RNA from various rat tissues and S1 nuclease digestion products analysed (Figure 7.2A). The upper band (152 bp) represents fully protected probe, while the lower band (121 bp) corresponds to the S1 product expected from abnormal RNA whose sequence differs from the normal OX-2 mRNA 5' of nucleotide 119. An estimate of the relative quantities of the two mRNA species is shown in the densitometer scan in figure 7.2B.

It is apparent that the abnormal transcript accounts for a substantial proportion (30%-45%) of the total OX-2 message in thymocytes, T cell blasts and B cells, but is barely detectable in brain RNA.

7.3.2 S1 mapping the junction between domains I and II

The region corresponding to the deletion in clone pX2/13 was mapped from both directions using the Hinf probes A (antisense strand labelled at 3' end) and B (antisense strand labelled at 5' end), as indicated in figure 7.3C. These probes were double stranded because the DNA fragment could not be strand separated. Each probe was hybridised with total RNA from various rat tissues as before, and the S1 digestion products analysed in figures 7.3A and 7.3B.

For both probes, the upper band represents the hybridisation of probe to the normal message while the lower bands represent the S1 product from an abnormal structure transcript.

(a) Probe A is labelled at nucleotide 324, and the S1 product size is close to the 122bp predicted if the sequences of probe and transcript differ downstream of the splice site at the end of domain I.

Figure 7.3

S1 NUCLEASE MAPPING OF OX-2 mRNA ACROSS THE JUNCTION BETWEEN DOMAINS I AND II

A,B : The radiolabelled probe A or B shown in figure C was hybridised with RNA from different rat tissues and the S1 nuclease digestion products analysed on a 6% denaturing polyacrylamide gel. The gel was fixed and dried and autoradiographed at -70°C for 14 days with pre-flashed film and intensifying screen.

A = S1 nuclease mapping of RNA with probe A
B = S1 nuclease mapping of RNA with probe B

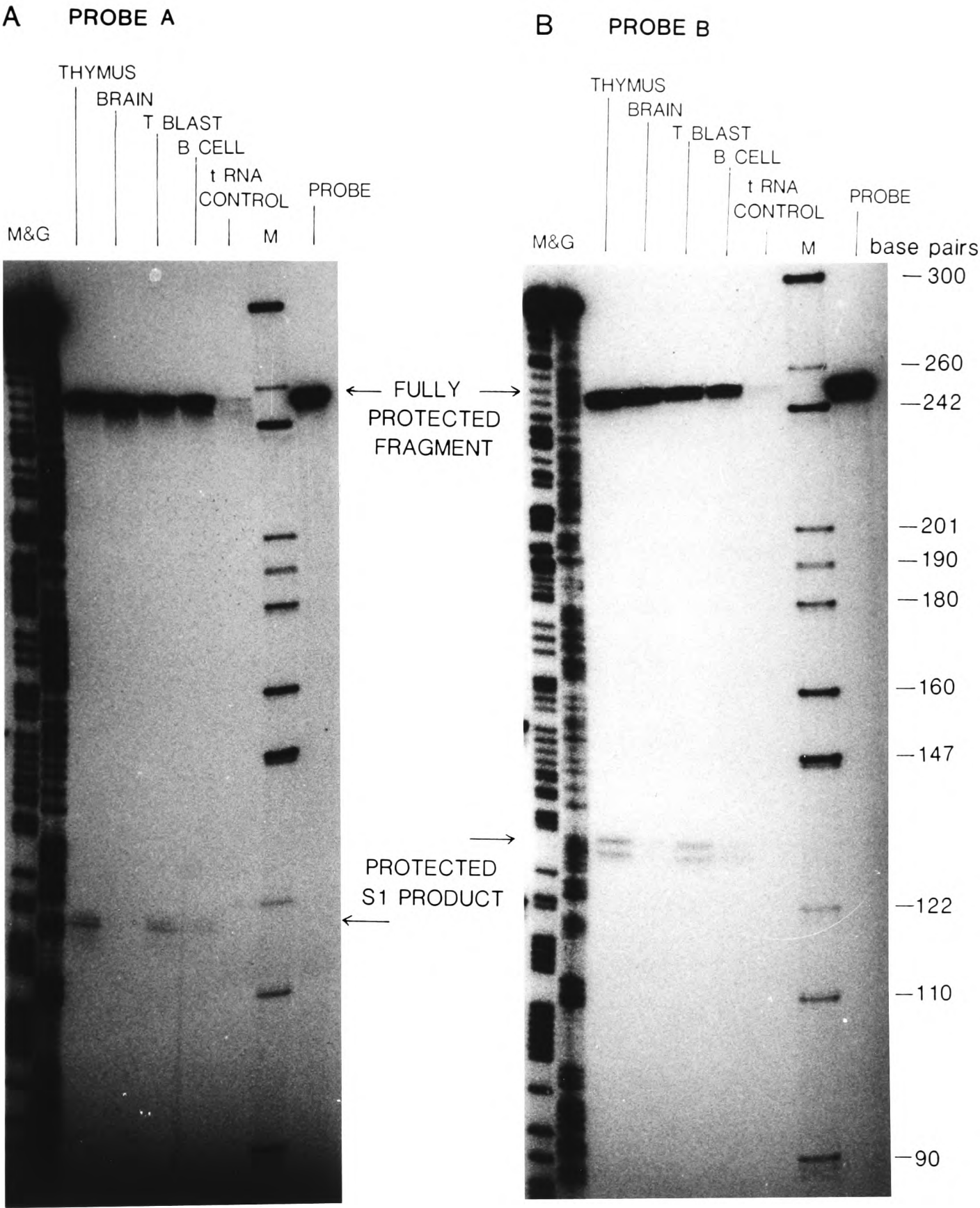
Key to gels A and B :

M&G = [A+G] tracks and [C+T] tracks of a Maxam and Gilbert sequencing degradation (provided by P. Winship)

M = radiolabelled size markers, sizes as indicated on right hand side of gel B

C : The arrows indicate the design and orientation of the radiolabelled probes. Clone pX2/24, which lacks the 5bp deletion, was digested with Hinf and labelled with ^{32}P as described below. The double stranded fragment was not strand separated.

Probe A : (black arrow)	Digestion with Hinf and 'filling in' with $[\alpha\text{-}^{32}\text{P}]\text{dATP}$ and DNA Pol 1 (Klenow fragment) generates a probe 256 bp long, labelled at nucleotide 324 at the 3' end of the antisense strand.
Probe B : (white arrow)	Digestion with Hinf and labelling with $[\gamma\text{-}^{32}\text{P}]\text{rATP}$ and T4 Polynucleotide kinase generates a probe 254 nucleotides long, labelled at nucleotide 579 at the 5' end of the antisense strand.



C PROBE DESIGN



(b) Probe B is labelled at nucleotide 579, and the S1 product is a doublet of 130 and 132 bp (restriction fragment size markers), or 133 and 135 bp (Maxam and Gilbert sequencing track). This is approximately the size expected if mRNA and probe sequence differ in the region around the junction between domains I and II, although it is a little larger than the 129 bp predicted if the message carried the same deletion as the cDNA clone pX2/13.

Both probes clearly indicate that the aberrant message is abundant in RNA from thymocytes, T cell blasts and B cells, but is almost entirely absent from brain RNA. This mapping experiment was not strictly quantitative because a small amount of the probe was resistant to S1 digestion (tRNA control lane).

7.4 NORTHERN BLOT ANALYSIS OF mRNA FROM DIFFERENT TISSUES.

The relative amounts of OX-2 message in the RNA used for S1 analysis was checked by Northern blotting as shown in figure 7.4. The blot should be quantitative as equal amounts of total RNA were loaded and all tracks hybridised with the same probe. Thymocyte, brain, and T blast RNA preparations contain roughly equivalent amounts of OX-2 message, while the B cell preparation contains much less. Thus the same quantity of OX-2 mRNA is found in brain despite the lack of abnormal structure transcripts.

All four tracks show just one hybridising band of approximate size 2.4 kb, and no distinct higher or lower molecular weight bands.

Figure 7.4

NORTHERN BLOT ANALYSIS OF OX-2 mRNA FROM DIFFERENT TISSUES

15µg total rat thymocyte RNA was fractionated on a 1.2% agarose gel in formaldehyde and transferred to Gene Screen membrane. The blot was hybridised with the nick-translated Bam-Xba 900bp coding region fragment of OX-2 clone pX2/11 and autoradiographed for 10 days at -70°C with pre-flashed film and intensifying screen. The size markers are bacterial (23S and 16S) and eukaryotic (28S and 18S) ribosomal RNA.

THYMUS

BRAIN

T BLAST

B CELL

rRNA
MARKERS

28S —

23S —

18S —

16S —



7.5 DISCUSSION TO CHAPTER 7

The results presented in this chapter, together with observations from Chapters 4 and 6, indicate that a high proportion (30%-45%) of OX-2 mRNA undergoes aberrant processing in lymphoid tissues but not in brain.

It is unlikely that the abnormal messages are the products of different OX-2 genes, as this would necessitate at least 3 separate genes, yet the Southern blot analysis of rat genomic DNA in Chapter 6 indicated that there was probably only one copy of the gene for OX-2. Although the library was prepared from outbred rats, the aberrant message is found in both outbred and inbred rats, as the B cell RNA was isolated from Sprague Dawely rats and the thymocyte, brain and T-cell blast RNA from PVGs. Thus the different mRNA species probably reflect alternative RNA processing.

The preliminary S1 analysis demonstrates that the abnormalities in the OX-2 mRNA occur at the same splice junctions as the deletions in the cDNA clones pX2/13 or pX2/28, although the mapping cannot distinguish whether the aberrant message carries the same deletion as the cDNA clones, or some other sequence difference. These deletions are not sufficiently large to affect the size of the message on a northern blot, and figure 7.4 confirms that all detectable OX-2 transcripts in brain and lymphoid tissues are approximately the normal length. The northern blot also shows that unspliced nuclear transcripts are too rare to account for the high proportion of abnormal message detected by S1 mapping.

Alternate splicing has been implicated in the control of gene expression for a variety of proteins, some of which are summarised in table 7.2. In general, the alternative transcripts encode functional proteins which are expressed either at different stages of development

Table 7.2 EXAMPLES OF ALTERNATELY SPLICED TRANSCRIPTS

Gene	Expression	Alternate transcription products	Reference
<u>A : TISSUE-SPECIFIC OR DEVELOPMENT-SPECIFIC PROCESSING</u>			
Ig μ chain	Development specific	Membrane bound or cytoplasmic forms of IgM or IgD	Maki et al., 1981
Calcitonin	Tissue specific	Different neural and thyroidal calcitonin-related peptides	Rosenfeld et al., 1983
Tachykinins	Tissue specific	mRNAs encode substance P alone, or substance P and substance K.	Nawa et al., 1984
α amylase	Tissue specific	Salivary gland and liver α amylase have different 5' noncoding regions	Young et al., 1981
Human fibronectin	Tissue specific	mRNA for liver fibronectin (plasma form) lacks an extra domain. mRNA for fibronectin produced by other cells (cellular form) includes species with and without the extra domain.	Kornblihtt et al., 1984
Troponin T	Tissue and development specific	Troponin T is a contractile protein. Multiple forms of rat troponin T mRNAs produced by combinations of exon splicing.	Breitbart et al., 1985
<u>B : FUNCTION OF ALTERNATE TRANSCRIPT UNKNOWN</u>			
γ fibrinogen	γ A and γ B forms	have different carboxy terminus and are produced concurrently in liver.	Crabtree and Kant, 1982
Q10	Q10 is a non-polymorphic secreted Class I molecule	of unknown function. Alternate transcript missing third exon produced concurrently in liver.	Lalanne et al., 1985
IL-2 receptor	cDNA clone identified	which contains a deletion and apparently encodes a protein which cannot bind IL-2.	Leonard et al., 1984

(e.g. the Ig μ chain), or in different tissues (e.g. the calcitonin gene-related protein) or concurrently in the same cell (e.g. γ fibrinogen). In some cases, such as the Q10 molecule, the alternative protein has been deduced from cDNA sequencing or RNA mapping, but its presence or function in vivo has not yet been proved. The IL-2 receptor is one of the few examples where the alternative product appears to be non-functional. In this case, a cDNA clone encoding the IL-2 receptor has been isolated and found to contain a deletion which leaves the protein still in reading frame. Transfection of this clone into mammalian cells does not result in the expression of functional IL-2 receptors (Leonard et al., 1984).

It is uncommon for alternate splicing to generate non-usable transcripts, yet it is hard to imagine how any of the structures corresponding to the abnormal OX-2 cDNA clones could encode a functional protein. The deletion of the leader sequence in clone pX2/28 creates a protein which presumably cannot be expressed at the cell surface. The 5bp deletion in clone pX2/13 shifts the reading frame so that the protein terminates 10 amino acids after the end of domain I. This could perhaps be a 'soluble' OX-2 molecule comprising of domain I only. However, the second abnormality found further upstream in this clone disrupts the leader peptide and would prevent secretion. It would be interesting to see whether other OX-2 cDNA clones carry the 5 bp deletion, but are otherwise normal.

cDNA clones containing unspliced introns resembling the structure in clone pX2/13 have occasionally been isolated from other cDNA libraries. For example, introns have been reported in cDNA clones coding for human fibroblast catalase (Korneluk et al., 1984), human erythropoietin (Jacobs et al., 1985), and apolipoprotein AII (Simon

Shelley, personal communication). Several OX-8 cDNA clones isolated from the same cDNA library as the OX-2 clones also contained intron-like insertions between coding region exons (Johnson et al., 1985), although this feature is not an artefact of the cDNA library as all the clones encoding the leucocyte-common antigen (Thomas et al., 1985) were normal. Clone pX2/13 may have been derived from an incompletely spliced nuclear transcript, but presumably this species of message is rare or an mRNA band of 2.6 kb would be visible in figure 7.4. Screening the northern blot or a copy of the cDNA library with a probe from the intron-like sequence would give an indication its frequency.

Truncated T cell receptor transcripts are abundant in thymocytes, but these differ from the examples discussed above because they are the products of genomic rearrangement rather than alternative transcriptional processing. The preliminary joining event between a D and a J segment results in the production of a DJC transcript from a conserved promoter upstream of the D segment (Siu et al., 1984). It is unlikely that this mechanism is operating for OX-2, because the gene apparently does not undergo any rearrangement in the thymus (Chapter 6).

The abnormal OX-2 message is abundant in lymphoid tissues i.e. thymocytes, T cell blasts and B cells, but barely detectable in brain. The lack of abnormal transcript in brain does not reflect a general lack of OX-2 mRNA in this tissue as the northern blot in figure 7.4 indicates comparable amounts of total OX-2 message in brain, thymocytes and T cell blasts.

If much of the OX-2 message is non-functional, then this might account for the high frequency of cDNA clones for OX-2 in the rat thymocyte library (Chapter 4) which was far greater than expected on the basis of the cell surface abundance of OX-2. Brain and thymocyte

membranes contain equivalent amounts of OX-2 antigen, so it is a little surprising that whilst brain and thymus apparently contain the same amount of OX-2 mRNA at least 40% of the thymocyte message is not usable. However, nothing is known about the turnover rates of OX-2 protein in brain and thymus.

The reason why lymphoid tissues should produce so much abnormal OX-2 message is not clear. The majority of thymocytes die in the thymus, so inaccurate splicing could perhaps be attributed to cell degeneration. However, the presence of abnormal transcripts in T cell blasts and B cells argues against this idea. It remains to be seen whether lymphoid tissues in general have higher levels of abnormal mRNA than other tissues.

CHAPTER 8

FINAL DISCUSSION

8.1 INTRODUCTION TO THE FINAL DISCUSSION

The MRC OX-2 antigen is a membrane glycoprotein of about 45,000 M_R present on lymphoid, neuronal and other cells of the rat. The complete primary structure of thymocyte OX-2 has been determined in this thesis, and the sequence shows that OX-2 is a member of the Ig superfamily along with the other molecules in figure 8.1.

This final Chapter brings together data presented in this thesis to propose a model for the structure of OX-2 at the cell surface, and to speculate on the evolutionary relationship between OX-2 and the Ig superfamily. Future experiments to investigate the expression and role of OX-2 are discussed.

8.2 THE STRUCTURE OF OX-2 AT THE CELL SURFACE

In Chapter 5, sequence homologies between OX-2 and Ig-related molecules indicated that OX-2 antigen has the two domain structure illustrated in figure 8.1. Although OX-2 is drawn here as an extended molecule, it is possible that interactions occur between the domains. In figure 8.2A, positions of Asn-linked carbohydrate are indicated on the predicted β strand folding pattern for OX-2. It can be seen that in both domains I and II, the three glycosylation sites are in similar positions and are found on the same face of the Ig fold. The domains could interact with each other via the non-glycosylated β sheets, so that the faces which are not exposed are mostly covered with carbohydrate as illustrated in figure 8.2B. Note that a single carbohydrate residue is

Figure 8.1

MODELS OF MOLECULES IN THE Ig SUPERFAMILY

This figure is essentially the same as figure 1.3 with the addition of the structure of the OX-2 molecule, as drawn in figure 5.3.

V = V-like domain

C = C-like domain

$\begin{array}{c} \text{S} \\ | \\ \text{S} \end{array}$ = Intrachain disulphide bond

S-S = Interchain disulphide bond

$\begin{array}{c} \bullet \\ | \end{array}$ = N-linked carbohydrate

References as in figure 1.3

THE IMMUNOGLOBULIN SUPERFAMILY

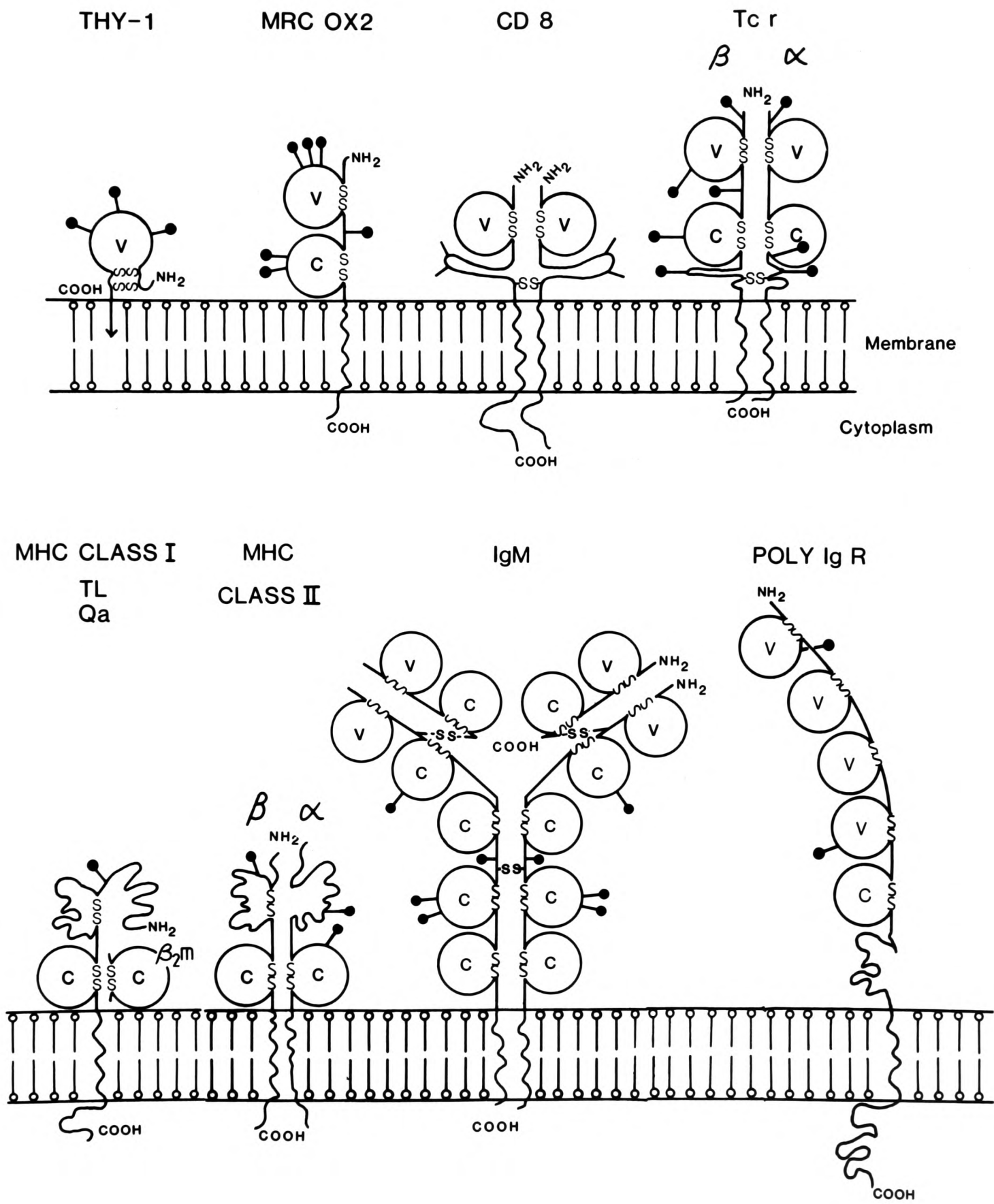


Figure 8.2

A : THE POSITION OF N-LINKED CARBOHYDRATE STRUCTURES IN RELATION TO THE PREDICTED Ig-LIKE FOLDING PATTERN FOR THE OX-2 DOMAIN

The β strand folding pattern of the Ig domain is drawn as in figure 1.2 with the positions of asparagine-linked carbohydrate indicated as circles (●) for OX-2 domain I and squares (■) for OX-2 domain II.

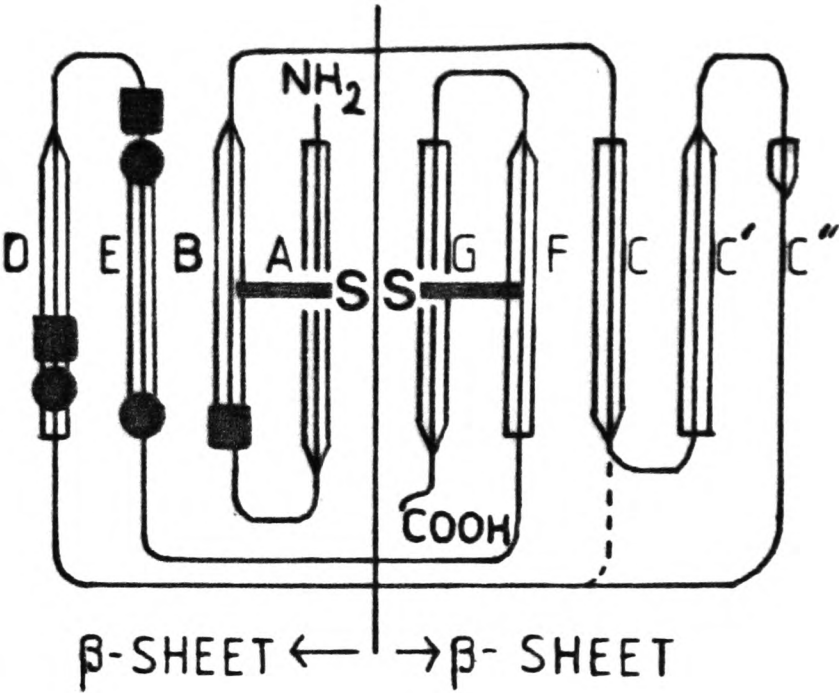
B : A MODEL FOR OX-2 AT THE CELL SURFACE

The stippled areas show how carbohydrate may cover most of the exterior surfaces of the molecule.

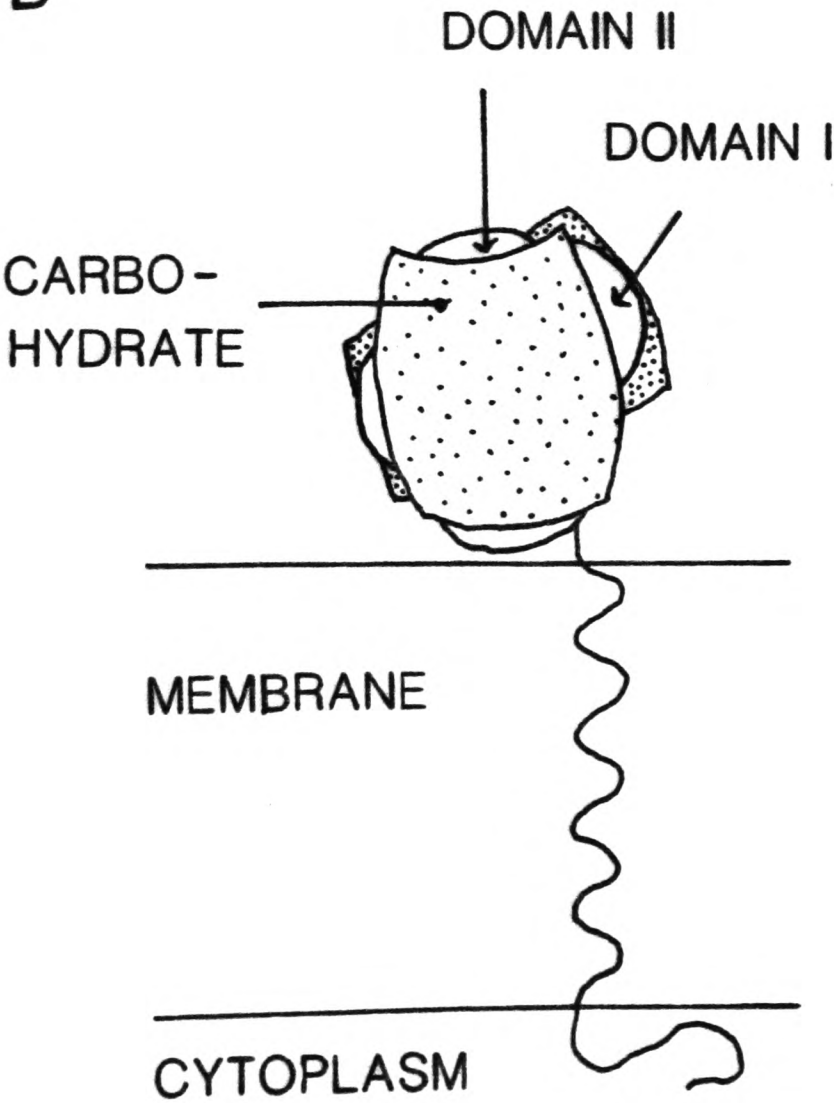
C : A TYPICAL ASIALO CARBOHYDRATE STRUCTURE drawn roughly to scale as figure B. There are estimated to be six of these structures attached to the OX-2 molecule, and hence the carbohydrate layer would be much thicker than indicated in B.

Figure 8.2

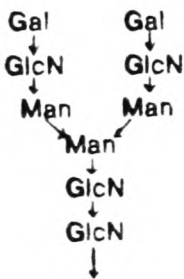
A



B



C



as large as the domain itself (figure 8.2C) so the carbohydrate layer would be a predominant feature of the molecule. However, some protein must be accessible to the exterior as the MRC OX-2 antibody specifically recognises the brain and thymocyte forms of the antigen which have considerable differences in the composition of the carbohydrate chains (Barclay and Ward, 1982).

It could be argued that the configuration proposed for the OX-2 molecule in figure 8.2B is unlikely because the associations so far found between Ig domains are either V-V or C-C interactions and V-C pairing has never been observed. However, most of the molecules in figure 8.1 are two chain structures and the associations occur between domains of the opposite chains. OX-2 and the Poly Ig receptor are the only known single chain structures containing more than one domain, and in the absence of a second chain with which to associate, it is possible that the Ig like domains interact with each other in this unconventional manner. The quaternary structure of the Poly Ig receptor will be of considerable interest in this respect.

The predicted location of the potential disulphide bonds in the OX-2 protein (Chapter 5 section 4) also supports the proposed folded configuration. In the immunoglobulin and T cell receptors, the C terminal conserved Cys of the V domain is separated by about 50 residues from the N-terminal Cys of the C domain, for example the separation is 48 residues in V_H EU (Kabat et al., 1983) and 55 residues in the T cell receptor β chain (Hedrick et al., 1985b). This separation is sufficient for the structure to be extended as indicated in figure 8.1. But in the OX-2 molecule, the corresponding distance is only 39 residues, and thus the two domains must be closer together. In addition, the 'extra' disulphide link between β strands F and G in domain I would keep β

strand G firmly within the Ig fold. The Poly Ig receptor is the only other molecule in the Ig superfamily which is thought to have such a short interdomain sequence (Mostov et al., 1984).

The disulphide linkages in the OX-2 protein were predicted on the basis of the position of the Cys residues and the sequence patterns of the flanking amino acids, but have not yet been established by protein chemistry. Clearly it is important that this disulphide bond arrangement is confirmed, as it would strengthen the arguments for the Ig-like structure of OX-2 and for the suggested configuration at the cell surface.

8.3 THE EVOLUTIONARY RELATIONSHIP BETWEEN OX-2 AND MEMBERS OF THE IG SUPERFAMILY

As discussed in the Introduction section 1.5, it seems likely that the starting point for the evolution of the Ig superfamily was the primordial Ig domain-like molecule suggested by Hill et al. (1966). Gene duplication, limited divergence and deletion gave a 2 domain light chain-like structure and the larger members of the Ig superfamily developed from this by further duplication, divergence, and the acquisition of non-Ig sequence exons. The final steps were the production of separate J and D exons, the duplication of germline gene segments, and the development of mechanisms for generating sequence diversity to increase the repertoire of antigen receptor sequences.

Where does OX-2 antigen belong in this evolutionary scheme? The following features of its structure and gene arrangement are all consistent with the idea that OX-2 looks most like the 2-domain light chain-like structure from which the molecules involved in antigen recognition evolved.

1) OX-2 antigen is a single chain polypeptide. Only one polypeptide chain is recognised by the monoclonal MRC OX-2, and the polypeptide does not appear to associate to form a homodimer during gel filtration in DOC (Chapter 3) as it elutes around the same position as bovine serum albumin (M_R of about 68,000). Given that gel filtration tends to overestimate the M_R of membrane proteins if calibrated with globular proteins, OX-2 antigen is eluting at the position expected for a monomeric polypeptide, not a dimer.

2) Sequence homologies suggest that the polypeptide comprises of an N-terminal V-like domain, followed by a C-like domain (Chapter 5), and hence the molecule looks like an Ig light chain but is attached to the membrane.

3) There is only one (or a limited number of) genes encoding OX-2 in the rat genome (Chapter 6). The molecules which have apparently undergone extensive duplication and divergence. i.e. Igs and T cell receptors, occur as clusters of closely related genes.

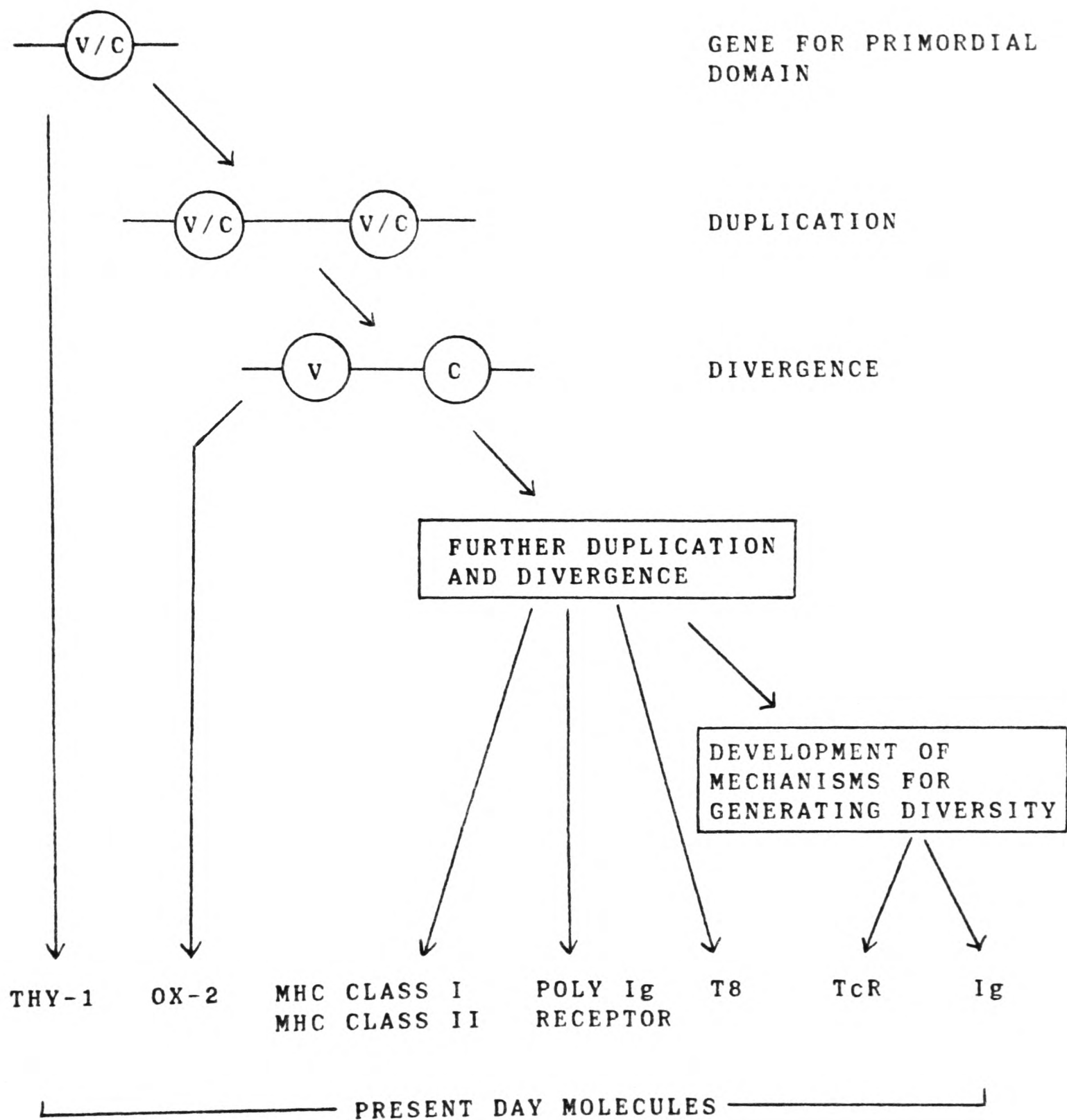
4) The OX-2 gene comprises of one exon for the V like domain (Chapter 6), and one exon for the C-like domain (Dr G. McCaughan, unpublished results), consistent with the molecule having evolved by duplication of the primordial gene encoding one domain.

5) Towards the end of the V-like domain, there is a sequence which has homology with J regions but this sequence is part of the V exon and not a separate J region exon as in the T cell receptors and Igs (Chapter 6).

A scheme incorporating OX-2 in the evolution of the Ig superfamily is illustrated in figure 8.3. The Thy-1 antigen is the present day molecule which looks most like the primordial domain (discussed in the

Figure 8.3

A POSSIBLE SCHEME FOR THE EVOLUTION OF THE IMMUNOGLOBULIN SUPERFAMILY



Introduction, section 1.5.3.1), while OX-2 looks most like the 2-domain structure proposed as the next evolutionary step.

It is important to note that since proteins and DNA molecules are not preserved in fossil remains, schemes of this type can only be deduced from the sequences of contemporary molecules and hence are essentially unprovable, although they can be modified as new sequences come to light. There is no such thing as a "primitive" molecule in present day organisms - all molecules are contemporary and all differ from their primordial genes to some extent. If the function of the primordial domain has been conserved through evolution, then OX-2 and Thy-1 will resemble their primitive ancestors more closely than the molecules involved in functions which developed later, e.g. antigen recognition, which required rapid gene duplication and divergence.

It could be argued that OX-2 is an offshoot of the Ig light chains or T cell receptor genes, its role and tissue distribution having changed drastically and its sequence having evolved comparatively recently. This is less easily reconciled with the data on the OX-2 gene structure because it requires the loss of the intervening sequence between the V and the J exons. Intron-less genes have been reported previously, for example in the mouse α globin pseudogene (Lueders et al., 1982), and it has been proposed that mRNA has been incorporated into the RNA genome of a retrovirus which has then integrated back into the host genome (reviewed by Flavell, 1982). Such a model demands that the newly integrated gene is flanked by retrovirus-like sequences, and these elements should be looked for when the sequence of the OX-2 gene is complete. However, rearrangement and transcription of Ig and T cell receptor genes only occurs in somatic cells, whereas if a new gene is to be transmitted to offspring, it must be present in the germline DNA.

Antibody molecules have not, as yet, been found in species more primitive than vertebrates (Nisonoff et al., 1975). If simple Ig-like molecules were found in invertebrates, this would support the theory that the primordial domain predated the immune response. The search for a Thy-1 homologue in the brain and nervous system of the squid yielded a molecule 84 amino acids long which has regions of sequence homology with immunoglobulin (Williams and Gagnon, 1982; Williams et al., 1984). The sequence does not fit with the conventional Ig domain, but may represent a half-domain structure which has been postulated by Borgois (1975) and McLachlan (1980) as a step in evolution prior to the formation of the single domain by a gene duplication event. In view of the large evolutionary distance between the species, serological cross-reactions between homologues would not be expected. It is surprising, then, that a protein has been identified in tunicates which cross-reacts with MRC OX-7 monoclonal antibody against Thy-1.1 (Mansour et al., 1985). The protein has biochemical similarities with Thy-1, but sequence data is required to confirm its identity.

No OX-2-like molecules have yet been found in invertebrates. One way to investigate this might be to search for genes which cross-hybridise with the rat cDNA probe on a Southern blots and cDNA libraries of invertebrate genomic DNA. However, the hybridisation signal between such distantly related species might be very weak.

8.4 POSSIBLE FUNCTIONS OF OX-2 AND THY-1

8.4.1 The role of the primordial Ig-like domain

How can knowledge of the structure of OX-2 contribute to thinking about the original function of the primordial domain?. As discussed in the Introduction 1.5.3.2, Cohen et al. (1981) and Williams (1982, 1984)

have suggested that the primitive domain mediated interactions at cell surfaces. If the functions of present day Thy-1 and OX-2 do indeed reflect those of their primitive ancestors, then the finding that both OX-2 and Thy-1 are expressed on neuronal tissue fits in with the idea that Ig-like structures first arose when primitive nervous systems evolved, and mediated recognition events in these tissues (Williams et al., 1984). Thy-1 is present on neuronal tissue of all species examined so far, although its expression on other tissues is variable. A key question now is whether OX-2 expression is also conserved in the nervous system of other species. This could be determined by probing northern blots of RNA from tissues of other species with a rat or human OX-2 DNA probe.

8.4.2 Investigation of the function of Thy-1 and OX-2

There is only limited functional data on Thy-1 and none at all on OX-2. Antibodies against Thy-1 have been demonstrated to have effects in some in vitro assays, for example monoclonals against Thy-1 coated onto glass enhance the growth of neural processes of rat retinal ganglion cells (Leifer et al., 1984). Rabbit anti-mouse-brain antisera (Jones, 1983) and one monoclonal antibody (Gunter, 1984) against Thy-1 have been found to stimulate T cell proliferation, while other monoclonals (Hollander, 1985) have the opposite effect and inhibit cell proliferation. To what extent these effects are due to the Thy-1 molecule per se or to the steric interference of a different functional molecule located in the cell membrane close to Thy-1, is not clear.

There are no functional data on OX-2, as the monoclonal MRC OX-2 has no demonstrable effect on in vitro or in vivo assays for immune function. As pure OX-2 antigen can be easily prepared, it might be worth

raising more monoclonals against different OX-2 determinants and testing these in systems such as the following :

- 1) The effect of antibodies against OX-2 on the growth of nerve processes in vitro in the assay system of Leifer et al. (1984).
- 2) The effect of in vivo administration of antibody to neonatal rats. The levels of OX-2 expression have been found to alter considerably during brain development in neonates, although fluctuations in other tissues have not been investigated. The antibody may therefore have a more profound effect on developing rats than on adults.
- 3) The effect on B cell responses. OX-2 is expressed on follicular dendritic cells, which are believed to be involved in the generation of B cell memory (Klaus et al., 1980). The effect of in vivo administration of MRC OX-2 antibody on the primary and secondary antibody responses against a specific antigen could be investigated.

8.5 POSSIBLE OX-2 COUNTERPARTS ON HUMAN THYMOCYTES

The tissue distribution of the human OX-2 homologue is not yet known, and if its pattern of expression parallels that of Thy-1 then it is possible that the molecule may not be present on lymphoid tissues at all. Antigens of comparable size to OX-2 are found on human thymocytes, and table 8.1 lists some of the 45 KDa glycoproteins which (like OX-2) are present on thymocytes but are absent from mature human T cells.

Sequence data are not available for these antigens, but on the basis of their biochemical characteristics and tissue distribution it is not immediately apparent that any of them correspond to the human homologue. None of the antigens are reported to be expressed on brain, and this is the one tissue in which OX-2 expression is likely to be conserved (section 8.4.1). The closest potential homologue is the T10

Table 8.1

HUMAN THYMOCYTE SURFACE GLYCOPROTEINS OF EQUIVALENT SIZE TO RAT OX-2 ANTIGEN

Species	M Ab	M _R glyco protein (SDS-PAGE)	+ β ₂ M	M _R [*] protein backbone (SDS-PAGE)	Distribution	Notes	Ref
RAT	OX-2	41K (brain) 47K (thymocyte)	NO	33K	Thymocytes (not mature T cells). Activated T and B cells. Brain, endothelium, follicular dendritic cells, B cells, muscle.		(1)
HUMAN	OKT6 NA1/34	49K	YES	33K	85% of thymocytes(not mature T cells). Langerhans cells, tonsil crypt epithelia, interfollicular areas in lymph node. Not present on brain or mitogen activated thymus derived lymph nodes.	T6 is thought to be associated with T8 in the thymus (6)	(2) (3) (4) (5)
	M241	43K	YES	33K	Thymocytes,some T leukemias, (not peripheral lymphocytes)	Antigenically distinct from T6 But some peptides in common	(7) (4)
	BL6	45K	N.K.	N.K.	Thymocytes, activated T and B Not peripheral T of humans, but of New world monkeys.	Thought to be Qa-like	(8) (9)
	T10	45K (reduced) 37K (non-reduced)	NO	N.K.	Thymocytes (not peripheral T) 5-10% B cells. Mitogen-activated thymus-derived lymphocytes.		(3)

KEY : N.K = not known

* = Determined by deglycosylation of antigen and analysis on SDS-PAGE

References :

(1) Barclay and Ward, 1982	(6) Snow et al., 1985
(2) McMichael et al., 1979	(7) Knowles and Bodmer, 1982
(3) Terhorst et al., 1981	(8) Clark et al., 1983
(4) van de Rijn et al., 1983	(9) Hansen et al., 1983
(5) McMillan et al., 1983	

molecule (Terhorst et al., 1981), although upon reduction the apparent molecular weight of T10 shifts to a lower value than that of OX-2.

The CD1 antigens (T6 or NA1/34 and M241) are unlikely to be OX-2 homologues because they appear to be noncovalently associated with β_2 microglobulin whereas OX-2 has never been found associated with other molecules, either when immunoprecipitated with MRC OX-2 monoclonal (McMaster and Williams, 1979), or purified by monoclonal antibody affinity chromatography (figure 3.3, this thesis). The folding configuration proposed for OX-2 in section 8.2 would leave no surfaces of the molecule free from carbohydrate and available to interact with β_2 m in the same manner as the MHC Class I molecule. The MRC OX-2 monoclonal might possibly cause dissociation of molecules noncovalently associated with OX-2, so the immunoprecipitation studies should be repeated with different monoclonals against OX-2 if these become available. The CD1 and OX-2 antigens are similar in that they are highly glycosylated, and when the carbohydrate is removed the protein runs with a M_R of 33,000 on SDS PAGE (van der Rijn et al., 1983 and Dr A.N. Barclay personal communication). This property distinguishes them from the murine TL antigens which are also associated with β_2 m but have a larger protein backbone (van der Rijn et al., 1983).

The antigen recognised by monoclonal BL6 has been reported to resemble the murine Qa antigens (Clark et al., 1983) but details of its biochemical properties or association with β_2 m have not been reported and there is insufficient data to assess whether or not it resembles OX-2 or not.

One approach for determining the identity and distribution of the human OX-2 homologue would be to incorporate the cloned human gene into the genome of mouse cells by DNA-mediated gene transfer and produce

transformants expressing the OX-2 protein on their surface. This technique has been used for the Thy-1 gene (Evans et al., 1984; see section 8.6). The transformed cells could be screened with all available monoclonals against human thymocyte antigens, and if none recognise the expressed protein then antisera or monoclonal antibodies raised against the transformed cells could be used for immunohistochemical localisation and immunoprecipitation studies.

8.6 FUTURE APPLICATIONS FOR THE HUMAN GENOMIC CLONE

The rat cDNA and human genomic clones for OX-2 are useful tools with which to investigate the structure and expression of OX-2, and their potential for analysing tissue distribution and conservation in evolution has already been mentioned.

OX-2 and Thy-1 are examples of membrane molecules expressed on just a few cell types, and their expression on some tissues, e.g. brain, fluctuates dramatically during neonatal development (discussed in Chapter 1.6). Therefore the OX-2 and Thy-1 genes are good models for the study of tissue-specific and developmental stage-specific gene regulation. Evans et al. (1984) have introduced a cloned Thy-1 gene into the genome of different cell lines and demonstrated that cells which would normally express Thy-1 (lymphoma and neuronal cell lines) express 50 fold more cell surface Thy-1 than cells which do not normally express the protein (L cells). These results suggest that the 8.1 kb fragment of the Thy-1 gene introduced into the cells contains the elements necessary for specific regulation.

The sequence of the mouse (Giguère et al., 1985) and rat (Seki et al., 1985) Thy-1 gene has been determined. The data of Giguère et al indicate that the 5' end of the gene contains a fourth exon of 5'

untranslated sequence. The Thy-1 promoter appears to be unusual in that it lacks the conserved TATA and CAAT boxes characteristic of most eukaryotic promoters (Corden et al., 1980) and instead has a high [G+C] content. It will be interesting to see whether the 5' regions of the OX-2 gene contain homologous nucleotide sequences which represent conserved controlling elements shared by the two genes. In view of the large (2.6kb) intron downstream of the 5' untranslated exon in the Thy-1 gene, it is likely that the OX-2 genomic clone λ HX2/EMBL3 is not full length at the 5' end, and further genomic clones will have to be isolated for sequence and expression studies.

The chromosomal location of the human OX-2 gene can be determined by Southern blotting of DNA from a panel of cell hybrids covering the full range of human chromosomes. These studies are in progress, and indicate that the OX-2 gene is not on the same chromosome as any of the other members of the Ig superfamily mapped so far (Dr F.Grosveld, unpublished results).

8.7 FAMILIES OF STRUCTURALLY RELATED CELL SURFACE MOLECULES

Many of the molecules on thymocytes have now been shown to have Ig-related structures, reinforcing the idea that the Ig domain is a structure well suited to mediating recognition events. Although the Ig fold is a popular theme, not all the thymocyte antigens are Ig-related as can be seen from table 8.2. The leucocyte-common antigen, for example, has no sequence homology with Ig at all (Thomas et al., 1985). Likewise, not all surface antigens with an unusual tissue distribution will have Ig-like structures. For example, the LGSP antigen (W3/13) completely lacks Cys residues and contains 60% carbohydrate, all of which is O-linked (Brown et al., 1981). The LGSPs may represent another

Table 8.2

STRUCTURAL RELATIONSHIPS OF SURFACE MOLECULES OF RAT LYMPHOCYTESIg-RELATED

IgM etc	Kehry et al., 1980
T antigen receptor(β)	Hedrick et al., 1984 *
(α)	Saito et al., 1984 *
MHC Class I and β_2M	Ploegh et al., 1981 *
MHC Class II	Kaufman et al., 1984 *
OX-8 (CD8)	Johnson et al., 1985
W3/25 (CD4)	Littman et al., 1985 a ⁺
Thy-1	Campbell et al., 1981
OX-2	Clark et al., 1985 and this thesis

NOT Ig-RELATED

L-CA	Thomas et al., 1985
W3/13	Brown et al., 1981
IL-2 receptor	Leonard et al., 1984 *
Transferrin receptor	Schneider et al., 1984 *

RELATIONSHIP TO Ig UNKNOWN

OX-19 (CD5)	Mason et al., 1983
-------------	--------------------

The rat lymphoid surface molecules described in table 1.1 are categorised according to whether they or their counterparts in other species have an Ig-related structure

* = Reference given is for the structure of the homologous molecule in a species other than the rat.

⁺ The complete sequence of CD4 has recently been reported (Maddon et al., 1985)

family of cell surface interaction molecules related to the glycophorin of human erythrocytes (Furthmayr et al., 1978).

8.8 CONCLUSIONS

The brain/lymphoid antigen OX-2 has been shown to have an Ig-related structure, but is more likely to have a role in recognition events at cell surfaces than in immunity. It is likely that further members of the Ig superfamily exist which have a function involved with cell interactions, and these will be revealed as the sequences of other cell surface proteins are determined.

REFERENCES

- Allen, G. (1981) in Sequencing of proteins and peptides'. Laboratory Techniques in Biochemistry and Molecular Biology Series, Elsevier/North-Holland Biomedical Press.
- Amzel, L.M. and Poljak, R.J. (1979) *Ann. Rev. Biochem.*, 48, 961.
- Anson, D.S., Choo, K.H., Rees, D.J.G., Gianelli, F., Gould, K., Huddleston, J.A. and Brownlee, G.G. (1984) *EMBO J.*, 3, 1053-1060.
- Arvieux, J., Gagnon, J. and Williams, A.F. (1985) Submitted to *Molecular Immunology*.
- Barclay, A.N. (1979) *J. Neurochem.* 32, 1249.
- Barclay, A.N. (1981) *Immunology* 44, 727-736.
- Barclay, A.N. and Hydén, H. (1978) *J. Neurochem.* 31, 1375.
- Barclay, A.N., Letarte-Muirhead, M. and Williams, A.F. (1975) *Biochem. J.* 151, 699.
- Barclay, A.N., Letarte-Muirhead, M., Williams, A.F., and Faulkes, R. (1976) *Nature* 263, 563.
- Barclay, A.N. and Ward, H.A. (1982) *Eur. J. Biochem.* 129, 447.
- Barnett Foster, D.E., Dorrington, K.J. and Painter, R.H. (1980) *J. Immunol.* 124, 2186.
- Bayer, J. and Reske, K. (1983) *Eur. J. Immunol.* 13, 18.
- Bentley, D.R. and Porter, R.R. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1212.
- Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180.
- Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721.
- Birnboim, H.C. and Doly, J. (1979) *Nucl. Acids Res.* 7, 1513
- Blake, C. (1983) *Nature* 306, 535.
- Blomberg, B. and Tonegawa, S. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 530.
- Bourgois, A. (1975) *Immunochemistry* 12, 873.
- Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50, 349.
- Breitbart, R.E., Nguyen, H.T., Medford, R.M., Destree, A.T., Mahdavi, V. and Nadal-Ginard, B. (1985) *Cell* 41, 67.
- Brockes, J. (1984) *Brit. Med. Bull.* 40, 229.

Brown, W.R.A., Barclay, A.N., Sunderland, C.A. and Williams, A.F. (1981) *Nature* 289, 456.

Bukovsky, A., Presel, J., Zidovsky, J. and Mancal, P. (1983) *Immunology* 48, 587.

Campbell, D.G., Gagnon, J., Reid, K.B.M. and Williams, A.F. (1981) *Biochem. J.* 195, 15.

Campbell, D.G., Williams, A.F., Bayley, P.M. and Reid, K.B.M. (1980) *Biochem. Soc. Symp.* 45, 45.

Campbell, R.D., Dodds, A.W. and Porter, R.R. (1980) *Biochem. J.* 189, 67.

Chickaraishi, D.M. (1979) *Biochemistry* 18, 3249.

Chien, Y.-H., Gascoine, N.R.J., Kavalier, J., Lee, N.E. and Davis, M.M. (1984a) *Nature* 309, 322.

Chien, Y., Becker, T.L., Okamura, M., Cohen, D.I. and Davis, M.M. (1984b) *Nature* 312, 31.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry* 18, 5294.

Christie, D.L. and Gagnon, J. (1982) *Biochem. J.* 201, 555.

Clark, E.A., Martin, P., Hansen, J.A. and Ledbetter, J.A. (1983) *Immunogenetics* 18, 599.

Clark, M.J., Gagnon, J., Williams, A.F., and Barclay, A.N. (1985) *EMBO J.* 4, 113.

Cohen, F.E., Novotny, J., Sternberg, M.J.E., Campbell, D.G. and Williams, A.F. (1981) *Biochem. J.* 195, 31.

Collins, M.K.L. and Owen, M.J. (1985) *Biochem. J.* 230, 281.

Colomb, M. and Porter, R.R. (1975) *Biochem. J.* 145, 177.

Conti-Tronconi, B.H. and Raftery, M.A. (1982) *Ann. Rev. Biochem.* 51, 491.

Corden, J., Wasylyk, A., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980) *Science* 209, 1406.

Cotmore, S.F., Crowhurst, S.A. and Waterfield, M.D. (1980) *Eur. J. Immunol.* 11, 597.

Crabtree, G.R. and Kant, J.A. (1982) *Cell* 31, 159.

Dalchau, D., Kirley, J. and Fabre, J.W. (1980) *Eur. J. Immunol.* 10, 745.

Dayhoff M.O., Hunt, T.L., McLaughlin, W., Barker, C. (1972) in 'Atlas of protein sequence and structure' M.O. Dayhoff, Ed. (National Biomedical Research Foundation, Georgetown University, Washington D.C.) Vol 5.

Dayhoff, M.O., Hunt, L.T., Barker, W.C., Schwartz, R.M. and Orcutt, B.C. eds., (1978) Protein Segment Dictionary, The National Biomedical Research Foundation, P.O. Box 629, Silver Spring, Maryland 20901, U.S.A.

Douglas, T.C. (1972) *J. Exp. Med.* 136, 1054.

Dulbecco, R., Unger, M., Bologna, M., Battifora, H., Syka, P. and Okada, S. (1981) *Nature* 292, 772.

Edelman, G.M., Hoffman, S., Chuong, G.M., Thiery, J.P., Brackenbury, R., Gallin, W.J., Grumet, M., Greenberg, M.E., Hemperly, J.J., Cohen, C. and Cunningham, B.A. (1983) *Cold Spring Harb. Symp. Quant. Biol.* 48, 515.

Evans, G.A., Ingraham, H.A., Lewis, K., Cunningham, K., Seki, T., Moriuchi, T., Chang, H.C., Silver, J. and Hyman, R. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 5532.

Fairclough, R.M., Finer, Moore, J., Love, R.A., Kristofferson, D., Desmeules, P.J. and Stroud, R.M. (1983) *Cold Spring Harb. Symp. Quant. Biol.* 48, 9.

Favaloro, J., Treisman, R. and Kamen, R. (1980) *Methods Enzymol.* 65, 718.

Flavell, R.A. (1982) *Nature* 295, 370.

Ford, W.L. (1975) *Prog. in Allergy* 19, 1.

Frischauf, A.-M., Lehrach, H., Poustka, A. and Murry, N. (1983) *J. Mol. Biol.* 170, 83.

Fukumoto, T., McMaster, W.R. and Williams, A.F. (1982) *Eur. J. Immunol.* 12, 237.

Furthmayr, H. (1978) *Nature* 271, 519.

Gait, M.J., Singh, M., Sheppard, R.C., Edge, M.D., Greene, A.R., Heathcliffe, G.R., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1980) *Nucleic Acids Res.* 8, 1081.

Gait, M.J. (1984) "Oligonucleotide synthesis: a practical approach" IRL Press Limited

Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) *Nucl. Acids Res.*, 7, 2115.

Giguère, V., Isobe, K.-I. and Grosveld, F. (1985) *EMBO J.* 4, 2017.

Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.L. (1980) *Anal. Biochem.* 106, 492.

Goridis, C., Hirn, M., Santoni, M.-J., Gennarini, G., Deagostini-Bazin, H., Jordan, B.R., Keifer, M. and Steinmetz, M. (1985) *EMBO J.* 4, 631.

Gowans, J.L. and Knight, E.J. (1964) *Proc. R. Soc. Lond. [Biol]* 159, 257.

- Gunter, K.C., Malek, T.D. and Shevach, E.M. (1984) *J. Exp. Med.* 159, 716.
- Hansen, J.A., Martin, P.J., Beatty, P.G., Clark, E.A. and Ledbetter, J.A. (1983) in *Human Leucocyte Markers Detected by Monoclonal Antibodies*. J. Dausset, C. Milstein and S.F. Schlossman (eds.) Springer-Verlag, Berlin.
- Hart, D.N.J. and Fabre, J.W. (1981) *Transplantation* 27, 110.
- Hart, D.N., Fuggle, S.V., Williams, K.A., Fabre, J.W., Ting, A. and Morris, P.J. (1981) *Transplantation* 31, 428.
- Hedrick, S.M., Cohen, D.I., Nielsen, E.A. and Davis, M.M. (1984a) *Nature* 308, 149.
- Hedrick, S.M., Nielsen, E.A., Kavalier, J., Cohen, D.I. and Davis, M.M. (1984b) *Nature* 308, 153.
- Hendry, S.H.C., Hockfield, S., Jones, E.G. and McKay, R. (1984) *Nature* 370, 267.
- Hill, R.L., Delany, R., Fellows Jnr, R.E., and Lebovitz, H.E. (1966) *Proc. Natl. Acad. Sci USA* 56, 1762.
- Hockfield, S. and McKay, R.D.G. (1983) *Proc. Natl. Acad. Sci USA* 80, 5758.
- Hollander, N. (1985) *J. Immunol.* 134, 2916.
- Hunkapillar, M., Kent, S., Caruthers, M., Dreyer, W., Firca, J., Giffin, C., Morvath, S., Hunkapiller, T., Tempst, P. and Hood, L. (1984) *Nature* 310, 105.
- Jacobs, K., Shoemaker, C., Rudersdorf, R., Niell, S.D., Kaufman, R.J., Mufson, A., Seehra, J., Jones, S.S., Hewick, R., Fritsch, E.F., Kawakita, M., Shimizu, T. and Miyake, T. (1985) *Nature* 313, 806.
- Jaye, M., De la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. and Lecocq, J.P. (1983) *Nucl. Acids Res.* 11, 2325.
- Jefferies, W.A., Brandon, M.R., Hunt, S.V., Williams, A.F., Gatter, K.C. and Mason, D.Y. (1984) *Nature* 312, 162.
- Jefferies, W.A., Brandon, M.R., Williams, A.F. and Hunt, S.V. (1985a) *Immunology* 54, 333.
- Jefferies, W.A., Green, J.R. and Williams, A.F. (1985b) *J. Exp. Med.* 162, 117.
- Jensenius, J.C. and Williams, A.F. (1974a) *Eur. J. Immunol.* 4, 91.
- Jensenius, J.C. and Williams, A.F. (1974b) *Eur. J. Immunol.* 4, 98.
- Johnson, P., Gagnon, J., Barclay, A.N. and Williams, A.F. (1985) *EMBO J.*, in press.

- Jones, B. (1983) *Eur. J. Immunol.* 13, 678.
- Jones, P.P., Murphy, D.B. and McDevitt, H.O. (1981) *Immunogenetics* 12, 321.
- Jordan, R., Caillol, D., Damotte, M., Delovitch, T., Ferrier, P., Kahn-Perles, B., Kourilsky, F., Layet, C., LeBouteller, P., Lemonnier, F.A., Malissen, M., N'Guyen, C., Sire, J., Sodoyer, R., Strachan, T. and Trucy, J. (1985) *Immunol. Rev.* 84, 73.
- Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M. and Perry, H. eds. (1983) *Sequences of Proteins of Immunological Interest*. U.S. Department of Health and Human Services, National Institutes of Health, Bethesda, U.S.A.
- Kaufman, J.F., Auffray, C., Korman, A.J., Shackelford, D.A. and Strominger, J. (1984) *Cell* 36, 1.
- Kavathas, P., Sukhatme, V.P., Herzenberg, L.A. and Parnes, J.R. (1984) *Proc. Natl. Acad. Sci. USA* 81, 7688.
- Kehry, M., Ewald, S., Douglas, R., Sibley, C., Raschke, W., Fambrough, D. and Hood, L. (1980) *Cell* 21, 393.
- Kindt, T.J. (1981) *Immunol. Today* 2, 169.
- Klareskog, L., Forsum, U. and Peterson, P.A. (1980) *Eur. J. Immunol.* 10, 958.
- Klaus, G.G.B., Humphrey, J.H., Kunke, A. and Dongworth, D.W. (1980) *Immunol. Rev.* 53, 3.
- Klein, J., Juretic, A., Baxevanis, C.N. and Nagy, Z.A. (1981) *Nature* 291, 455.
- Knowles, R.W. and Bodmer, W.F. (1982) *Eur. J. Immunol.* 12, 676.
- Knowles, D.M., Tolidjian, B., Marboe, C., D'Agati, V., Grimes, M. and Chess, L. (1984) *J. Immunol.* 132, 2170.
- Kohler, G. and Milstein, C. (1975) *Nature* 256, 495.
- Korman, A.J., Knudsen, P.J., Kaufman, J.F. and Strominger, J.L. (1982) *Proc. Natl. Acad. Sci. USA* 79, 1844.
- Kornblihtt, A.R., Vibe Pedersen, K. and Baralle, F.E. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3218.
- Kornblihtt, A.R., Vibe Pedersen, K. and Baralle, F.E. (1984) *EMBO J.* 3, 221.
- Korneluk, R.G., Quan, F., Lewis, W.H., Guise, K.S., Willard, H.F., Holmes, M.T. and Gravel, R.A. (1984) *J. Biol. Chem.*, 259 13819.
- Kozak, M. (1983) *Microbiol. Rev.* 47, 1.

- Kuchel, P.W., Campbell, D.G., Barclay, A.N. and Williams, A.F. (1978) *Biochem. J.* 169, 411.
- Kühn, L.C., McClelland, A. and Ruddle, F. (1984) *Cell* 37, 95.
- Kunkel, L.M., Smith, K.D., Boyer, S.H., Borgaonkar, D.S., Watchel, S.S., Miller, O.J., Breg, W.R., Jones, H.W. and Rary, J.M. (1977) *Proc. Natl. Acad. Sci USA* 74, 1245.
- Laemmli, U.K. (1970) *Nature* 227, 680.
- Lakin, K.H., Allen, A.K. and Fabre, J.W. (1983) *J. Neurochem.* 41, 385.
- Lalanne, J.L., Transy, C., Guerin, S., Darche, S., Meulien, P. and Kourilsky, P. (1985) *Cell* 41, 469.
- Lathe, R. (1985) *J. Mol. Biol.* 183, 1.
- Lee, J.S., Trowsdale, J., Travers, P.J., Carey, J., Grosveld, F., Jenkins, J. and Bodmer, W.F. (1982) *Nature* 299, 750.
- Lehrach, H., Diamond, D., Wozney, J.M. and Boedtker, H. (1977) *Biochemistry* 16, 4743.
- Leifer, D.A., Lipton, S.A., Barnstable, C.J. and Masland, R.H., (1984) *Science* 224, 303.
- Leonard, W.J., Depper, J.M., Crabtree, G.R., Rudikoff, S., Pumphrey, J., Robb, R.J., Kronke, M., Svetlik, P.B., Pfeffer, N.J., Waldman, T.A. and Greene, W.C. (1984) *Nature* 311, 626.
- Leonard, W.J., Pepper, J.W., Uchiyama, T., Smith, K.A., Waldmann, T.A. and Greene, W.C. (1982) *Nature* 300, 267.
- Letarte-Muirhead, M., Barclay, A.N. and Williams, A.F. (1975) *Biochem. J.* 151, 685.
- Littman, D.R., Maddon, P.J., Godfrey, M., Chess, L. and Axel, R. (1985a) Published as an abstract in *Advances in Gene Technology: Molecular Biology. Proceedings of the 17th Miami Winter Symposium, Miami, Florida, U.S.A. 11-15th February 1985 (ICSU Short Reports; U2)*.
- Littman, D.R., Thomas, Y., Maddon, P.J., Chess, L. and Axel, R. (1985b) *Cell* 40, 237.
- Lowry, O.H., Rosenbrough, N.J., Farr, A.L. and Randall, R.J. (1951) *J. Biol. Chem.* 193, 265.
- Lueders, K., Leder, A., Leder, P. and Kuff, E. (1982) *Nature* 295, 426.
- MacDonald, H.R., Glasebrook, A.L., Bron, C., Kelso, A. and Cerottini, J.C. (1982) *Immunol. Rev.* 68, 89.
- McGarry, R.C., Riopelle, R.J. and Roder, J.C. (1982) *Neuroscience Letters* 56, 95.

- McKay, R.D.G., Hockfield, S., Thomson, I., Johansen, J. and Fredericson, K. (1983) Cold Spring Harbor Symp. Quant. Biol. 48, 599.
- McKenzie, J.L., Allen, A.K. and Fabre, J.W. (1981) Biochem. J. 197, 629.
- McKenzie, J.L., Dalchau, R. and Fabre, J.W. (1982) J. Neurochem. 39, 1461.
- McLachlan, A.D. (1980) Early evolution of the antibody domain in "Proteins and Related Subjects". Protides of the Biological Fluids (Ed. H. Peeters) 28, 29. Pergamon Press, Oxford.
- McMaster, W.R. and Williams, A.F. (1979) Eur. J. Immunol. 9, 426.
- McMichael, A.J., Pilch, J.R., Galfrè, G., Mason, D.Y., Fabre, J. and Milstein, C. (1979) J. Immunol. 9, 205.
- McMillan, E.M., Brubaker, D.B., Peters, S., Jackson, I., Beeman, K., Wasik, R., Stoneking, L.E. and Pesler, D.R. (1983) Cancer Immunol. Immunother. 15, 221.
- McNabb, T., Koh, T.Y., Dorrington, K.J. and Painter, R.H. (1976) J. Immunol. 117, 882.
- Maddon, P.J., Littman, D.R., Godfrey, M., Maddon, D.E., Chess, L. and Axel, R. (1985) Cell 42, 93.
- Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabl, M., Raschke, W. and Tonegawa, S. (1981) Cell 24, 353.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. eds. (1982) Molecular Cloning, a Laboratory Manual, published by Cold Spring Harbor Laboratory Press, N.Y.
- Mansour, M.H. and Cooper, E.L. (1984) J. Immunol. 132, 2515.
- Mansour, M.H., De Lange, R. and Cooper, E.L. (1985) J. Biol. Chem. 260, 2681.
- Marchalonis, J.J., Vasta, G.R., Warr, G.W. and Barker, W.C. (1984) Immunol. Today 5, 133.
- Mason, D.W. (1981) Transplantation 32, 222.
- Mason, D.W., Arthur, R.P., Dallman, M.J., Green, J.R., Spickett, G.P. and Thomas, M.L. (1983) Immunol. Rev. 74, 57.
- Mason, D.W. and Williams, A.F. (1980) Biochem. J. 187, 1.
- Maxam, A.M. and Gilbert, W. (1980) Meths. Enzymol. 65, 499.
- Mellman, I.S. and Unkeless, J.C. (1980) J. Exp. Med. 152, 1048.
- Messing, J. (1984) Meths. Enzymol. 101, 20.
- Metzger, H. (1978) Immunol. Rev. 41, 186.

- Meuer, S.C., Cooper, D.A., Hodgdon, J.C., Hussey, R.E., Fitzgerald, K.A., Schlossman, S.F. and Reinhertz, E.L. (1983) *Science* 222, 1239.
- Moriuchi, T., Chang, H-C, Denome, R., and Silver, J. (1983) *Nature* 301, 80.
- Morris, R.J. and Williams, A.F. (1975) *Eur. J. Immunol.* 5, 274.
- Mostov, K.E., Friedlander, M. and Blobel, G. (1984) *Nature (Lond.)* 308, 37.
- Nawa, H., Kotani, H. and Nakanishi, S. (1984) *Nature* 312, 729.
- Nisonoff, A., Hopper, J.E. and Spring, S.B. (1975) "The Antibody Molecule" Academic Press, New York and London.
- Osawa, H. and Diamantstein, T. (1983) *J. Immunol.* 130, 51.
- Parham, P., Alpert, B.N., Orr, H.T. and Strominger, J.L. (1977) Carbohydrate moiety of HLA antigens. *J. Biol. Chem.* 252, 7555.
- Parnes, J.R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Apella, E. and Seidman, J.G. (1981) *Proc. Natl. Acad. Sci. USA* 78, 2253.
- Paul, W.E. (1984) *Fundamental Immunology*. Raven Press, N.Y.
- Peterson, P.A., Cunningham, I.B. and Geriald, M.E. (1972) *Proc. Natl. Acad. Sci. USA* 69, 1697.
- Ploegh, H.L., Orr, T. and Strominger, J.L. (1981) *Cell* 24, 287.
- Poger, M.E. and Lamm, M.E. (1974) *J. Exp. Med.* 139, 629.
- Porath, J. (1974) *Meths. Enzymol.* 34, 13.
- Proudfoot, N.J. and Brownlee, G.G. (1974) *Nature* 252, 359.
- Quakenbush, E.J., Cruz, T.F., Moscarello, A. and Letarte, M. (1985) *Biochem. J.* 225, 291.
- Reichardt, L.F. (1984) *Science* 225, 1294.
- Reid, K.B.M., Bentley, D.R. and Wood, K.J. (1984) *Phil. Trans. R. Soc. Lond.* 306, 345.
- Reif, A.E. and Allen, J.M.V. (1964) *J. Exp. Med.* 120, 413.
- Reif, A.E. and Allen, J.M.V. (1966) *Nature* 209, 521.
- Reyes, A.A., Johnson, M.J., Schöld, M., Ito, H., Ike, Y., Morin, C., Itakura, K. and Wallace, R.B. (1981) *Immunogenetics* 14, 383.
- Rhodin, J.A.G. (1974) in *Histology a text and atlas*. Oxford University Press.
- Richardson, J.S., Richardson, D.C., Thamas, K.A., Silverton, E.W. and Davies, D.R. (1976) *J. Mol. Biol.* 102, 221.

- Robinson, A.P., White, T.M. and Mason, D.W. (1985) Submitted to Immunology.
- Romanes, G.J. (1981) Cunningham's Textbook of Anatomy. 12th Edition. Oxford University Press.
- Rosenfeld, M.G., Mermod, J.-J., Amara, S.G., Swanson, L.W., Sawchenko, P.E., Rivier, J., Vale, W.W. and Evans, R.M. (1983) Nature 304, 129.
- Rostas, J.A., Shevenan, T.A., Sinclair, C.M. and Jeffrey, P.L. (1983) Biochem. J. 213, 143.
- Rutishauser, U. (1983) Cold Spring Harb. Symp. Quant. Biol. 48, 501.
- Saito, H.S., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N. and Tonegawa, S. (1984) Nature 312, 36.
- Sakano, H., Hüppi, K., Heinrich, G. and Tonegawa, S. (1979a) Nature 280, 288.
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. and Tonegawa, S. (1980) Nature 286, 676.
- Sakano, H., Rogers, J.H., Hüppi, K., Brock, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. (1979b) Nature 277, 627.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463.
- Schachner, M., Faissner, J., Kruse, J., Lindner, D.H., Meier, D.H., Rathjen, F.G. and Wernecke, H. (1983) Cold Spring Harb. Symp. Quant. Biol. 48, 557.
- Schneider, C., Owen, M.J., Banville, D. and Williams, J.G. (1984) Nature 311, 675.
- Seki, T., Moriuchi, T., Chang, H.-C., Denome, R. and Silver, J. (1985) Nature 313, 485.
- Sharpe, C.R. (1985) D. Phil. Thesis (Oxford).
- Shoulders, C.C. and Baralle, F.E. (1982) Nucl. Acids Res. 10, 4873.
- Siu, G., Kornenberg, M., Strauiss, E., Haars, R., Mak, T.W. and Hood, L. (1984) Nature 311, 344.
- Singer-Sam, J., Simmer, R.L., Keith, D.H., Shively, L., Teplitz, M., Itakura, K., Gartler, S.M. and Riggs, A.D. (1983) Proc. Natl. Acad. Sci. USA 80, 802.
- Siodak, A.W. and Nowinski, R.C. (1981) Immunogenetics 12, 45.
- Snow, P.M. and Terhorst, C. (1983) J. Biol. Chem. 258, 14675.
- Southern, E. (1979) Meths. Enzymol. 68, 152.

- Springer, T.A., Davignon, D., Ho, M.-K., Kürzinger, K., Martz, E. and Sanchez-Madrid, F. (1982) *Immunol. Rev.* 68, 171.
- Standring, R., McMaster, W.R., Sunderland, C.A. and Williams, A.F. (1978) *Eur. J. Immunol.* 8, 832.
- Sukhatme, V.P., Sizer, K.C., Vollmer, A.C., Hunkapiller, T. and Parnes, J.. (1985) *Cell* 40, 591.
- Sunderland, C.A., McMaster, W.R. and Williams, A.F. (1979) *Eur. J. Immunol.* 9, 155.
- Sutcliffe, J.G., Melner, R.J., Shinnick, T.M. and Bloom, F.E. (1983) *Cell* 33, 671.
- Terhorst, C., Agthoven, A., Le Clair, K., Snow, P., Reinherz, E. and Schlossman, S. (1981) *Cell* 23, 771.
- Thomas, M.L., Barclay, A.N., Gagnon, J., and Williams, A.F., (1985) *Cell* 41, 83.
- Tonegawa, S. (1983) *Nature* 302, 575.
- Trisler, G.D., Schneider, M.D. and Nirenberg, M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 2145.
- Tse, A.G.D., Barclay, A.N., Watts, A. and Williams, A.F. (1985) *Science*, in press.
- Tzartos, S.A., Rand, D.E., Einarson, B.L. and Lindstrom, J.M. (1981) *J. Biol. Chem.* 256, 8635.
- Tzartos, S.J., Seybold, M.E. and Lindstrom, J.M. (1982) *Proc. Natl. Acad. Sci. USA* 79, 188.
- Ullrich, A., Coussens, J., Hayflick, S., Dull, T.J., Gray, A., Tam, W., Lee, J., Yarden, Y., Liberman, T.A., Sclessinger, J., Downward, J., Mayes, E.L.V., Whittle, N., Waterfield, M.J. and Seeburg, P.H. (1984) *Nature* 309, 418.
- Van de Rijn, M., Lerch, P.G., Knowles, R.W. and Terhorst, C. (1983) *J. Immunol.* 131, 851.
- Von Boehmer, H. and Shortman, K. (1973) *J. Immunol. Meth.* 2, 293.
- Von Heijne, G. (1983) *Eur. J. Biochem.* 133, 17.
- Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981) *Nucl. Acids. Res.* 9, 879.
- Warner, N.L., (1974) *Adv. Immunol.* 19, 67.
- Watson, M.E.E. (1984) *Nucl. Acids Res.* 12, 5145.
- Watts, T.H., Brian, A.A., Kappler, J.W., Marrak, P. and McConnell, H.M. (1984) *Proc. Natl. Acad. Sci. USA* 81, 7564.

- Webb, M. and Barclay, A.N. (1984) *J. Neurochem.* 43, 1061.
- Wilkinson, J.M., Press, E.M. and Porter, R.R. (1966) *Biochem. J.* 100, 303.
- Williams, A.F. (1973) *Eur. J. Immunol.* 3, 628.
- Williams, A.F. (1977) *Comtemp. Topics Mol. Immunol.* 6, 83.
- Williams, A.F. (1980) *Biochem. Soc. Symp.* 45, 27.
- Williams, A.F. (1982) *J. Theoret. Biol.* 98, 221.
- Williams, A.F. (1985) *Nature* 114, 579.
- Williams, A.F., Barclay, A.N., Letarte-Muirhead, M. and Morris, R.J. (1976) *Cold spring Harbor Symposia on Quantitative Biology* 41, 51.
- Williams, A.F., Barclay, A.N., Clark, M.J. and Gagnon, J. (1984) Andersson, L.C., Gahmberg, C.G. and Ekblom, P. (eds.) *Proceedings of the Sigrid Juselius Symposium: Gene expression during normal and malignant differentiation.* Academic Press.
- Williams, A.F. and Barclay, A.N. (1985) Weir, D.M. and Herzenberg, L.A. (eds) in *"Handbook of Experimental Immunology"* 4th Edn. Chapter 25, Blackwell Scientific Publications Limited, Oxford.
- Williams, A.F. and Gagnon, J. (1982) *Science* 216, 696.
- Williams, A.F., Galfrè, G. and Milstein, C. (1977) *Cell* 12, 663.
- Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S.P., Aleksander, I. and Mak, T.W. (1984) *Nature* 308, 145.
- Young, R.A., Hagenbüchle, O. and Schilber, U. (1981) *Cell* 23, 451.
- Zipser, B., Stewart, R., Flanagan, T., Flaster, M. and Macagno, E.R. (1983) *Cold Spring Harb. Symp. Quant. Biol.* 48, 551.

MRC OX-2 antigen: a lymphoid/neuronal membrane glycoprotein with a structure like a single immunoglobulin light chain

Melanie J. Clark, Jean Gagnon¹, Alan F. Williams and A. Neil Barclay

MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, and ¹MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

Communicated by A. F. Williams

The MRC OX-2 antigen is a rat cell surface glycoprotein of mol. wt. 41 000–47 000 found on neurones, thymocytes, B cells, follicular dendritic cells and endothelium. We now report the amino sequence for this antigen as deduced from the nucleotide sequence of cDNA clones detected by use of an oligonucleotide probe. The sequence contains 248 amino acid residues of which 202 residues are likely to be outside the cell with two domains that show homology with immunoglobulins. The N-terminal domain fits best with Ig V domains and Thy-1 antigen while the C-terminal part is like an Ig C domain. Thus the structure overall is similar to an Ig light chain or the T cell receptor β chain. Three glycosylation sites are identified on each of the MRC OX-2 antigen domains.

Key words: MRC OX-2 antigen/thymic antigen/neuronal antigen/membrane glycoprotein/immunoglobulin superfamily

Introduction

The MRC OX-2 mouse monoclonal antibody was raised against rat thymocyte membrane glycoproteins (McMaster and Williams, 1979) and binds to rat thymocytes, neurones, follicular dendritic cells of lymphoid organs, vascular endothelium, some smooth muscle and B lymphocytes (Barclay, 1981; Webb and Barclay, 1984). The thymocytes and brain forms of the antigen have been purified by monoclonal antibody affinity chromatography and are highly glycosylated proteins of apparent mol. wts. 47 000 (thymocyte OX-2) and 41 000 (brain OX-2). The antigenicity and amino acid compositions of these forms were indistinguishable and the size difference is probably due to known differences in the carbohydrate structures (Barclay and Ward, 1982).

The finding of OX-2 antigen on neurones and thymocytes was reminiscent of Thy-1 antigen which is also found on these cell types. Furthermore, both Thy-1 and OX-2 antigens are found on diverse cell types without an apparent functional correlation. The function of neither antigen is known but the structure of Thy-1 antigen is of particular interest because it is homologous to Ig V domains (Williams and Gagnon, 1982). Rat Thy-1 antigen has an apparent mol. wt. of 25 000 on SDS gels and contains a polypeptide sequence of 111 amino acids (Campbell *et al.*, 1981).

On the basis of the tissue distributions and molecular characteristics it seemed possible that OX-2 and Thy-1 antigens may be related in evolution and to evaluate this the OX-2 antigen has been sequenced. Tryptic peptides were prepared from the pure glycoprotein and their sequences used to identify a region suitable for the synthesis of an oligonucleotide probe which was used to screen a cDNA library prepared

from rat thymocyte mRNA. From a number of cDNA clones the full sequence of the coding region for the OX-2 glycoprotein has been determined.

Results

Partial amino acid sequence of OX-2 antigen

OX-2 antigen was purified from rat brains and after reduction, alkylation and succinylation, was digested with trypsin. The peptides were fractionated and sequenced as described in Materials and methods. The sequences of five peptides are shown in Table I. They were all unambiguous except at the first position of T2 where small amounts of Lys, Tyr and Trp

Table I. Amino acid sequences of OX-2 tryptic peptides

Peptide	Sequence	Position in cDNA
T1	No sequence obtained. This is presumably blocked at N terminus. Composition (Glx) ₃ (Val) ₂ (Arg) ₁ (Asx) ₁ (Thr) ₁ (Gly) ₁ (Ser) ₁	1–10?
T2	G L L H T T A S L R	11–20
T3	C S L K T T Q E P L I V T W Q K K K A V G P E N M V T Y	21–48
T4	P A P A I S W K G T G S G I E (–) S T E S ...	136–155
T5	V K D P K T Q V G K E V I C Q V L Y	171–188
T6	G E S S Q G M Q R	238–246

A dash (–) indicated no phenylthiohydantoin amino acid derivative was determined at this position. The residue is likely to be glycosylated asparagine as this peptide contained carbohydrate. The dotted line (...) indicates that T4 was not sequenced to the end of the peptide. The region of peptide T3 used to design the oligonucleotide probe is underlined. Peptides T3 and T5 ended at a tyrosine residue. This type of cleavage has been observed before (Campbell *et al.*, 1981).

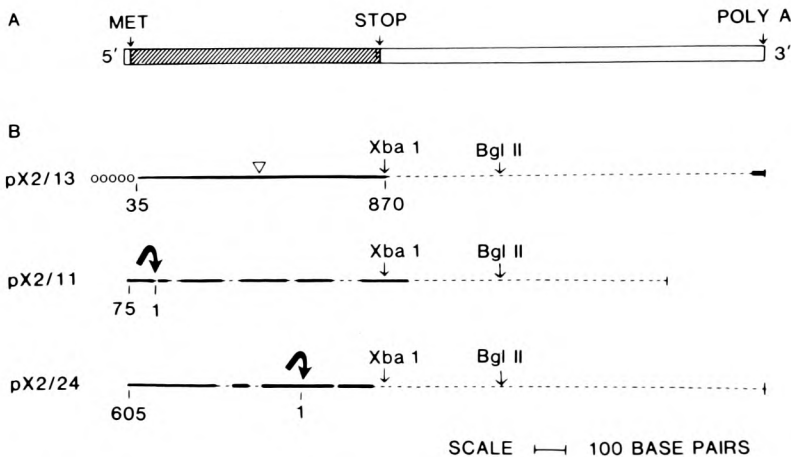


Fig. 1. Composite OX-2 mRNA structure and the three clones used to derive the sequence. **(A)** Structure of OX-2 mRNA. The hatched portion of the bar represents protein coding region, and the open portions represent non-coding regions. MET shows position of the initiator methionine and POLY A the beginning of the poly(A) region. **(B)** The three overlapping cDNA clones used to determine the sequence. The numbers refer to the nucleotide position in the composite sequence (Figure 2). Solid lines indicate those regions of the clones that have been sequenced, dashed lines those not sequenced. Cloning artifacts are marked 00000: double ligated insert, ∇: position of deletion, ↻: position of inversion.

were identified together with the predominant Gly residue. The N terminus of OX-2 is blocked because intact OX-2 gave no sequence when analysed by automated Edman degradation. Peptide T1 also gave no sequence and is probably the corresponding N-terminal tryptic peptide. Its composition is given in Table I.

Isolation and characterization of OX-2 cDNA clones

The amino acid sequence of the C-terminal portion of peptide T3 was used to design a 17 residue long oligonucleotide probe consisting of all 64 possible sequences complementary to the OX-2 mRNA as shown below:



A rat thymocyte cDNA library was constructed and screened with this probe and 30 positive cDNA clones isolated from a total of 150 000 clones. Twelve clones examined in detail all contained characteristic *Xba*I and *Bgl*II sites (see above, Figure 1). The nucleotide sequence was determined from three cDNA clones as shown in Figure 1 and all of these showed some cloning artifacts. Clone pX2/13 has apparently undergone a double ligation as 360 nucleotides of unidentified cDNA have become joined to the 5' end of the OX-2 cDNA. This clone also has a 5-bp deletion in the coding region. The other two clones both have an inversion at their 5' ends, involving the first 75 nucleotides of pX2/11, and the first 605 nucleotides of pX2/24. Similar inversions and deletions have been reported before (Anson *et al.*, 1984; Belt *et al.*, 1984). In view of these artifacts, the sequence of the coding region has been determined on both strands of the DNA and on more than one clone.

The coding sequence (see below) contained 834 nucleotides and this was followed by a long 3'-non-coding region of ~1300 nucleotides as determined by sizing restriction enzyme fragments on agarose gels. After this a poly(A) tail was iden-



Fig. 2. Northern blot analysis of thymocyte RNA. 20 μ g total rat thymocyte RNA was fractionated on an agarose gel in formaldehyde and transferred to Gene Screen membrane. The blot shown here was hybridised with the coding region probe of clone pX2/13 (5' end of clone to *Xba*I site). This fragment contains ~360 bases of extraneous DNA other than OX-2 coding sequence (see text) but an identical result was obtained when the same filter was stripped and probed with a fragment from the coding region only of clone pX2/11 which does not contain this extra sequence. Size markers are rat ribosomal 18S and 28S RNA.

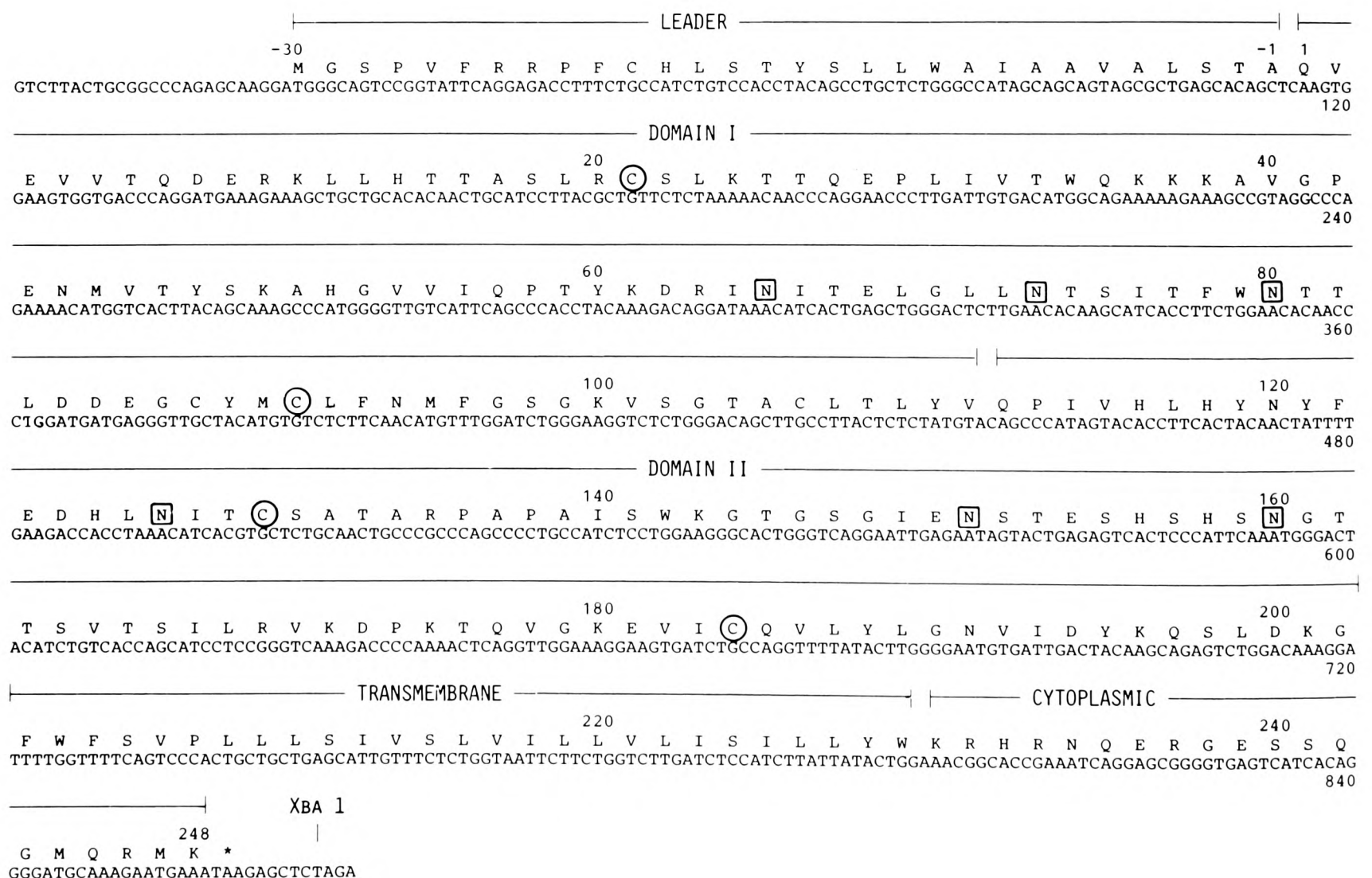


Fig. 3. Composite nucleotide sequence of the coding region of a thymocyte cDNA clone with its corresponding amino acid sequence. ○ = Cys residues believed to participate in the conserved disulphide bond within each domain. □ = Asn residues in potential glycosylation sites.

tified by sequencing clone pX2/13 from the 3' end. Thus the coding and non-coding regions total ~2200 nucleotides and this is consistent with a size of ~2500 nucleotides determined for the mRNA by Northern blot analysis (Figure 2).

Sequence of the coding region

The sequence shown in Figure 3 is a composite from the clones in Figure 1 and is strongly supported by the peptide data. One large open reading frame extends from a potential initiator methionine codon at nucleotide 25 contained within the consensus sequence A/G NNATG A/G found around most eukaryotic initiator methionine codons (Kozak, 1983). The first sequence established by peptide data is at amino residue 11 (peptide T2 Table I) and the other peptides occur throughout the sequence until residue 246 out of 248 where peptide T6 ends (Table I). The end of the protein is defined by a stop codon after residue 248. In the peptide data there is only one discrepancy, namely the first residue of peptide T2 which is a Gly in the peptide sequence compared with a predicted Lys in the cDNA sequence. Small amounts of Lys were found with the Gly in peptide T2 (see above) but the data were most consistent with the Gly designation shown in Table I. The other nine residues of T2 fitted with the nucleotide sequence and the inconsistency could thus be due to a polymorphism in the antigen; a difference between OX-2 of brain (peptides) and thymus (cDNA) or an artifact in the protein sequencing. Further confirmation of the coding region is shown by the good agreement in the amino composition determined for the glycoprotein compared with that predicted from the coding sequence minus leader sequence (Table II).

In Figure 3 a leader sequence extending to 30 amino acids is suggested with the protein sequence thus beginning at a Gln residue. If this Gln existed as a pyroglutamic acid residue then the blocked N terminus of OX-2 glycoprotein would be accounted for. Also the composition of peptide T1 which is believed to be the N-terminal peptide (Table I) fits reasonably

Table II. Comparison of experimentally determined amino acid composition of thymocyte OX-2 with that deduced from the nucleotide sequence of the thymocyte cDNA

Residue	Determination	Prediction
Asx	21.3	19
Glx	26.0	24
His	8.0	8
Lys	15.0	16
Arg	10.0	9
Thr	22.8	23
Ser	20.8	22
Pro	8.7	8
Ala	10.9	8
Cys	5.2	6
Gly	18.4	16
Tyr	8.4	9
Val	18.6	21
Ile	13.1	16
Leu	25.0	27
Phe	6.7	6
Met	3.4	5
Trp	5.2	5

The figures for the experimentally determined composition are taken from Barclay and Ward (1982) and are expressed as residues per 248 amino acids.

well with a tryptic peptide predicted by residues 1–10. There is a discrepancy in the presence of Gly and Ser in T1 composition but this could be due to small amounts of contamination with peptide T6 and also the T1 composition cannot be accounted for elsewhere in the OX-2 sequence. The putative leader sequence is similar to leader sequences of other glycoproteins with the only unusual point that it is longer than the average (von Heijne, 1983; Watson, 1984).

The mature protein as shown in Figure 3 contains 248 amino acids and amongst these only one segment is seen that seems likely to span the lipid bilayer. This is found at residues 203–229 and contains predominantly hydrophobic amino acids with no charged residues or amides. The latter residues are virtually never seen in transmembrane segments (Komaromy *et al.*, 1983; Kaufman *et al.*, 1984). The moderately hydrophobic amino acids Thr and Ser are commonly found in membrane-spanning segments and the sequence 203–229 contains four Ser residues. To the C-terminal side of residue 229 is found a stretch of basic residues and this is a typical feature of sequences on the cytoplasmic side of a membrane spanning piece. Thus it seems likely that residues 230–248 are inside the cell and that the sequence 1–202 is outside.

Carbohydrate structures have only been found on the extracellular parts of membrane glycoproteins and OX-2 sequences suitable for N-glycosylation are found at residues 65, 73, 80, 127, 151 and 160. At position 151 the peptide data (Table I) shows that glycosylation has occurred since no amino residue was detected in the sequencer at this position in peptide T4 (the expected result for Asn residues that have carbohydrate attached).

Homology of OX-2 glycoprotein with immunoglobulins, Thy-1 glycoprotein and the poly Ig receptor

Inspection of the sequence for OX-2 glycoprotein showed obvious homologies with the immunoglobulin superfamily with the N-terminal and C-terminal parts looking like Ig V and C domains respectively. This is shown in Figure 4a for V domain homologies and Figure 4b for the Ig C domains. In both cases the segments that form β -strands in the Ig domains are identified with letters and can be seen in the folding pattern for Ig domains shown in Figure 7a. In both domains I and II of the OX-2 sequence are found Cys residues that could form the conserved disulphide bonds between β -strands B and F and the tryptophan residue that is highly characteristic of β -strand C in Ig domains is also present in both OX-2 domains. Most Ig sequences are also characterised by a pattern of alternating hydrophobic residues in β -strands E and the sequence Ser Ile Thr Phe in OX-2 domain I fits well with this as does the Ile Leu Arg Val of domain II.

With regard to the V-like domains, identities are particularly seen around the Cys of β -strand F (residue 91) and this is a region that shows strong homologies in Ig V domains, Thy-1 antigen and the poly Ig receptor. At positions 6 and 4 residues before the Cys in β -strand F almost all Ig V domains have Asp and Ala or Gly, respectively, and OX-2 fits with this. The Asp residue may be conserved because it forms a salt bridge with a conserved Arg (more rarely Lys) residue near the base of β -strand D (position 63) (R. Poljak, personal communication; Williams *et al.*, 1984) and in this position an Arg is seen in the OX-2 domain I sequence.

The numbers of identities and conservative substitutions

were identified together with the predominant Gly residue. The N terminus of OX-2 is blocked because intact OX-2 gave no sequence when analysed by automated Edman degradation. Peptide T1 also gave no sequence and is probably the corresponding N-terminal tryptic peptide. Its composition is given in Table I.



Fig. 2. Northern blot analysis of thymocyte RNA. 20 µg total rat thymocyte RNA was fractionated on an agarose gel in formaldehyde and transferred to Gene Screen membrane. The blot shown here was hybridised with the coding region probe of clone pX2/13 (5' end of clone to *Xba*I site). This fragment contains ~360 bases of extraneous DNA other than OX-2 coding sequence (see text) but an identical result was obtained when the same filter was stripped and probed with a fragment from the coding region only of clone pX2/11 which does not contain this extra sequence. Size markers are rat ribosomal 18S and 28S RNA.

Isolation and characterization of OX-2 cDNA clones

The amino acid sequence of the C-terminal portion of peptide T3 was used to design a 17 residue long oligonucleotide probe consisting of all 64 possible sequences complementary to the OX-2 mRNA as shown below:



A rat thymocyte cDNA library was constructed and screened with this probe and 30 positive cDNA clones isolated from a total of 150 000 clones. Twelve clones examined in detail all contained characteristic *Xba*I and *Bgl*II sites (see above, Figure 1). The nucleotide sequence was determined from three cDNA clones as shown in Figure 1 and all of these showed some cloning artifacts. Clone pX2/13 has apparently undergone a double ligation as 360 nucleotides of unidentified cDNA have become joined to the 5' end of the OX-2 cDNA. This clone also has a 5-bp deletion in the coding region. The other two clones both have an inversion at their 5' ends, involving the first 75 nucleotides of pX2/11, and the first 605 nucleotides of pX2/24. Similar inversions and deletions have been reported before (Anson *et al.*, 1984; Belt *et al.*, 1984). In view of these artifacts, the sequence of the coding region has been determined on both strands of the DNA and on more than one clone.

The coding sequence (see below) contained 834 nucleotides and this was followed by a long 3'-non-coding region of ~1300 nucleotides as determined by sizing restriction enzyme fragments on agarose gels. After this a poly(A) tail was iden-

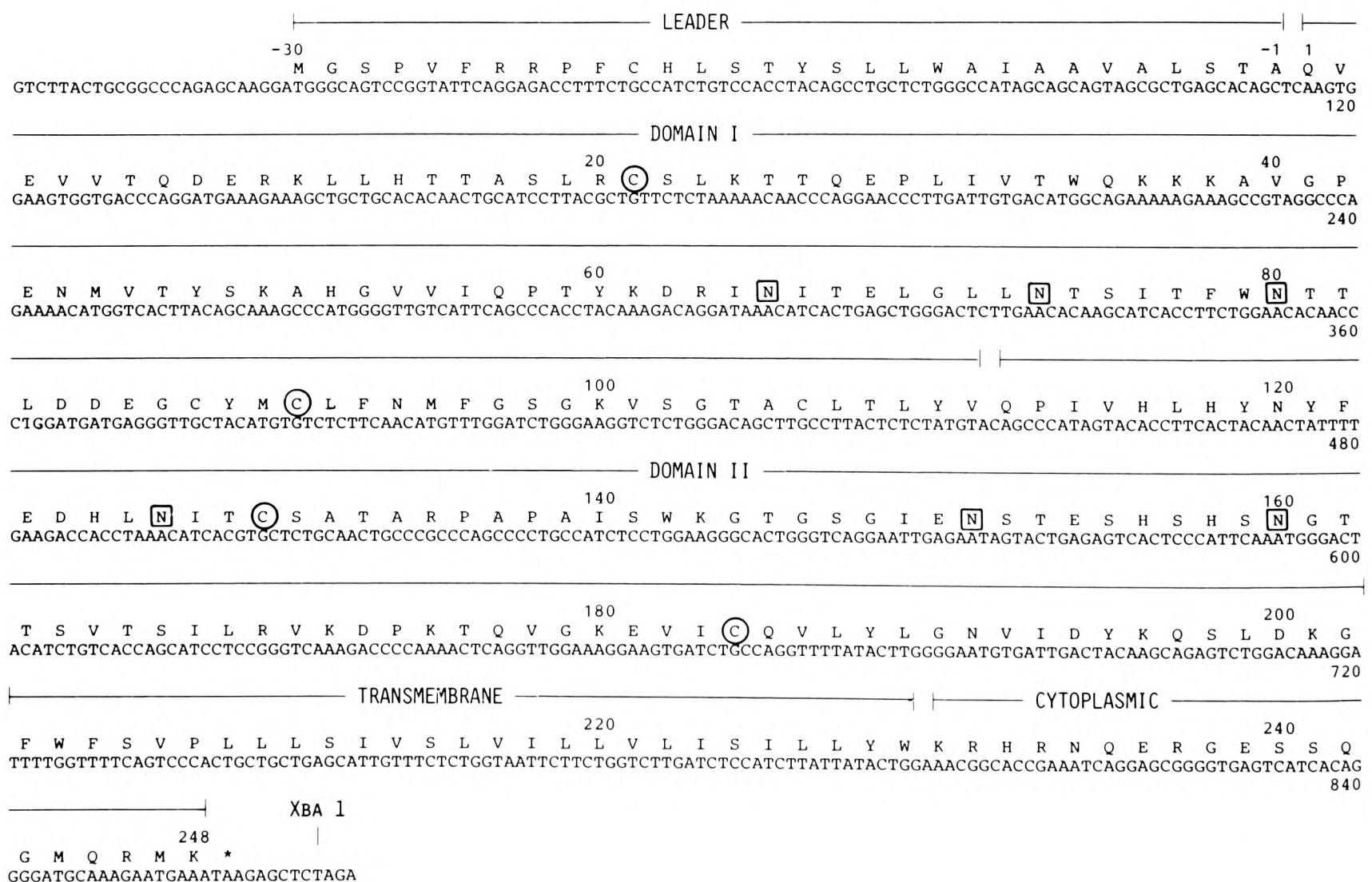


Fig. 3. Composite nucleotide sequence of the coding region of a thymocyte cDNA clone with its corresponding amino acid sequence. ○ = Cys residues believed to participate in the conserved disulphide bond within each domain. □ = Asn residues in potential glycosylation sites.

tified by sequencing clone pX2/13 from the 3' end. Thus the coding and non-coding regions total ~2200 nucleotides and this is consistent with a size of ~2500 nucleotides determined for the mRNA by Northern blot analysis (Figure 2).

Sequence of the coding region

The sequence shown in Figure 3 is a composite from the clones in Figure 1 and is strongly supported by the peptide data. One large open reading frame extends from a potential initiator methionine codon at nucleotide 25 contained within the consensus sequence A/G NNATG A/G found around most eukaryotic initiator methionine codons (Kozak, 1983). The first sequence established by peptide data is at amino residue 11 (peptide T2 Table I) and the other peptides occur throughout the sequence until residue 246 out of 248 where peptide T6 ends (Table I). The end of the protein is defined by a stop codon after residue 248. In the peptide data there is only one discrepancy, namely the first residue of peptide T2 which is a Gly in the peptide sequence compared with a predicted Lys in the cDNA sequence. Small amounts of Lys were found with the Gly in peptide T2 (see above) but the data were most consistent with the Gly designation shown in Table I. The other nine residues of T2 fitted with the nucleotide sequence and the inconsistency could thus be due to a polymorphism in the antigen; a difference between OX-2 of brain (peptides) and thymus (cDNA) or an artifact in the protein sequencing. Further confirmation of the coding region is shown by the good agreement in the amino composition determined for the glycoprotein compared with that predicted from the coding sequence minus leader sequence (Table II).

In Figure 3 a leader sequence extending to 30 amino acids is suggested with the protein sequence thus beginning at a Gln residue. If this Gln existed as a pyroglutamic acid residue then the blocked N terminus of OX-2 glycoprotein would be accounted for. Also the composition of peptide T1 which is believed to be the N-terminal peptide (Table I) fits reasonably

Table II. Comparison of experimentally determined amino acid composition of thymocyte OX-2 with that deduced from the nucleotide sequence of the thymocyte cDNA

Residue	Determination	Prediction
Asx	21.3	19
Glx	26.0	24
His	8.0	8
Lys	15.0	16
Arg	10.0	9
Thr	22.8	23
Ser	20.8	22
Pro	8.7	8
Ala	10.9	8
Cys	5.2	6
Gly	18.4	16
Tyr	8.4	9
Val	18.6	21
Ile	13.1	16
Leu	25.0	27
Phe	6.7	6
Met	3.4	5
Trp	5.2	5

The figures for the experimentally determined composition are taken from Barclay and Ward (1982) and are expressed as residues per 248 amino acids.

well with a tryptic peptide predicted by residues 1–10. There is a discrepancy in the presence of Gly and Ser in T1 composition but this could be due to small amounts of contamination with peptide T6 and also the T1 composition cannot be accounted for elsewhere in the OX-2 sequence. The putative leader sequence is similar to leader sequences of other glycoproteins with the only unusual point that it is longer than the average (von Heijne, 1983; Watson, 1984).

The mature protein as shown in Figure 3 contains 248 amino acids and amongst these only one segment is seen that seems likely to span the lipid bilayer. This is found at residues 203–229 and contains predominantly hydrophobic amino acids with no charged residues or amides. The latter residues are virtually never seen in transmembrane segments (Komaromy *et al.*, 1983; Kaufman *et al.*, 1984). The moderately hydrophobic amino acids Thr and Ser are commonly found in membrane-spanning segments and the sequence 203–229 contains four Ser residues. To the C-terminal side of residue 229 is found a stretch of basic residues and this is a typical feature of sequences on the cytoplasmic side of a membrane spanning piece. Thus it seems likely that residues 230–248 are inside the cell and that the sequence 1–202 is outside.

Carbohydrate structures have only been found on the extracellular parts of membrane glycoproteins and OX-2 sequences suitable for N-glycosylation are found at residues 65, 73, 80, 127, 151 and 160. At position 151 the peptide data (Table I) shows that glycosylation has occurred since no amino residue was detected in the sequencer at this position in peptide T4 (the expected result for Asn residues that have carbohydrate attached).

Homology of OX-2 glycoprotein with immunoglobulins, Thy-1 glycoprotein and the poly Ig receptor

Inspection of the sequence for OX-2 glycoprotein showed obvious homologies with the immunoglobulin superfamily with the N-terminal and C-terminal parts looking like Ig V and C domains respectively. This is shown in Figure 4a for V domain homologies and Figure 4b for the Ig C domains. In both cases the segments that form β -strands in the Ig domains are identified with letters and can be seen in the folding pattern for Ig domains shown in Figure 7a. In both domains I and II of the OX-2 sequence are found Cys residues that could form the conserved disulphide bonds between β -strands B and F and the tryptophan residue that is highly characteristic of β -strand C in Ig domains is also present in both OX-2 domains. Most Ig sequences are also characterised by a pattern of alternating hydrophobic residues in β -strands E and the sequence Ser Ile Thr Phe in OX-2 domain I fits well with this as does the Ile Leu Arg Val of domain II.

With regard to the V-like domains, identities are particularly seen around the Cys of β -strand F (residue 91) and this is a region that shows strong homologies in Ig V domains, Thy-1 antigen and the poly Ig receptor. At positions 6 and 4 residues before the Cys in β -strand F almost all Ig V domains have Asp and Ala or Gly, respectively, and OX-2 fits with this. The Asp residue may be conserved because it forms a salt bridge with a conserved Arg (more rarely Lys) residue near the base of β -strand D (position 63) (R. Poljak, personal communication; Williams *et al.*, 1984) and in this position an Arg is seen in the OX-2 domain I sequence.

The numbers of identities and conservative substitutions

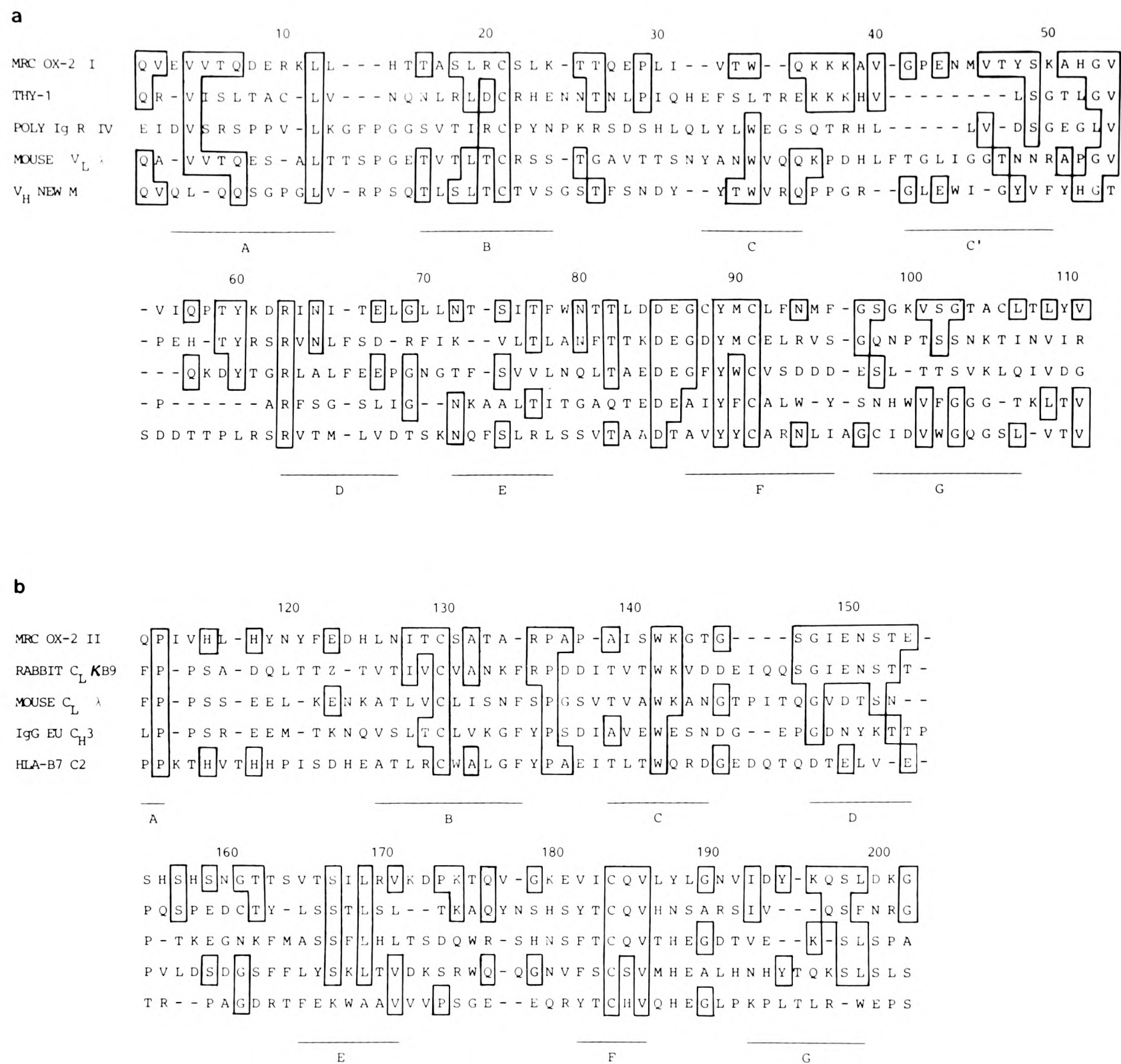


Fig. 4. Alignment of OX-2 sequence with immunoglobulin domains. **(A)** Alignment of OX-2 domain I sequence with rat Thy-1 (Campbell *et al.*, 1981), poly Ig receptor domain IV (Mostov *et al.*, 1984) mouse MOPC 104E V_λ, V_H NEW M (Kabat *et al.*, 1983). **(B)** Alignment of OX-2 domain II sequence with rabbit α B9, mouse MOPC 315_λ, IgG EU C_H3 and HLA-B7 C2 (Kabat *et al.*, 1983). In the comparisons with V-type domains **(A)** or C-type domains **(B)**, residues identical with OX-2 are boxed and dashes show gaps in the sequence inserted to maximize the alignment. The bars with letters A, B, C, C', D, E, F and G below the sequences indicate residues involved in forming β strands as shown in Figure 7a.

between all the sequences shown in Figure 4a are given in Figure 5 where it can be seen that the scores for OX-2 with the other sequences are at least as convincing as are the scores within the Thy-1, poly Ig R and V domain sequences.

The OX-2 domain II sequence clearly aligns well along the length of the Ig light chain C domain and heavy chain C_H3 domain sequences shown in Figure 4b. Fewer identities are seen with the Ig-like domain of HLA-B7. In Figure 5 the scores amongst all the sequences shown in Figure 4b are given and it can be seen that OX-2 domain II fits well with these Ig C domain sequences as does the HLA-B7 sequence.

One other homology of interest is that between mouse Thy-1 residues 2 – 10 and OX-2 residues, 100 – 107, namely:

rat OX-2	K V	-	S	G	T A C L
mouse Thy-1	K V	I	S	L	T A C L

The full Thy-1 sequence fits best with the OX-2 domain I sequence but the above identities are between the beginning of the Thy-1 sequence and the junction of the OX-2 domain I and II sequences. In the book of short sequences (Dayhoff *et al.*, 1978b) there are 3708 sequences containing Cys residues but only one TACL sequence (in bacteriophage ϕX174 gene D protein). Thus the above patches of identity may not have occurred by chance but might indicate a gene duplication event of a Thy-1 like sequence occurring long ago.

Discussion

The sequence data on the OX-2 glycoprotein shows that this molecule is a member of the Ig superfamily along with the other structures shown in Figure 6. The OX-2 molecule looks like a T receptor β chain or an Ig light chain with a membrane

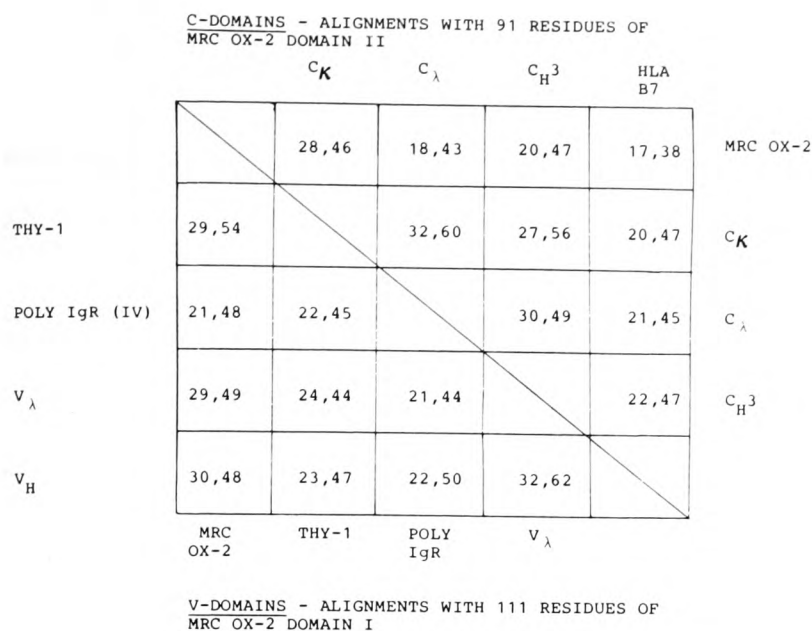


Fig. 5. Identities and conservative substitutions between MRC OX-2 sequence and Ig V and C domain sequences. The two numbers in each square give the number of amino acid identities and identities plus conservative substitutions respectively between the sequences indicated in Figure 4. Conservative substitutions are those that occur more frequently than random chance would predict and are those with a score of 1 or more in Figure 84 in Dayhoff *et al.* (1978a). For comparisons other than those involving the OX-2 sequence the alignments were altered from those in Figure 4 if adjustments, without excessive insertion of gaps, improved the scores. Thus in all cases the scores are for optimal adjustments as far as could be judged.

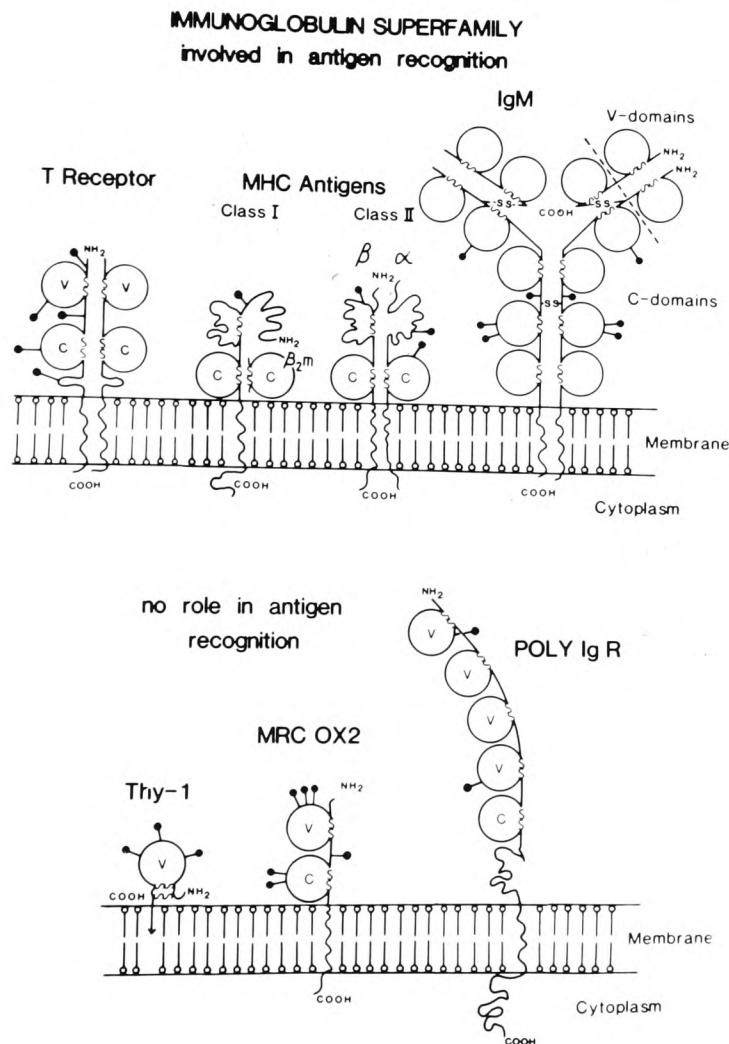


Fig. 6. Shows models for molecules in the Ig superfamily. Circles marked V are like an Ig V domain while those marked C are more like Ig C domains. Intrachain disulphide bonds are shown by symbols (S-S) and N-linked carbohydrate structures by (↑). The figure is modified from Williams *et al.* (1984). Data are from: mouse T cell receptor (Hedrick *et al.*, 1984; Saito *et al.*, 1984). Class I and Class II MHC (Ploegh *et al.*, 1981; Kaufman *et al.*, 1984). IgM (Kehry *et al.*, 1980). Thy-1 (Williams and Gagnon, 1982). Poly Ig R (Mostov *et al.*, 1984).

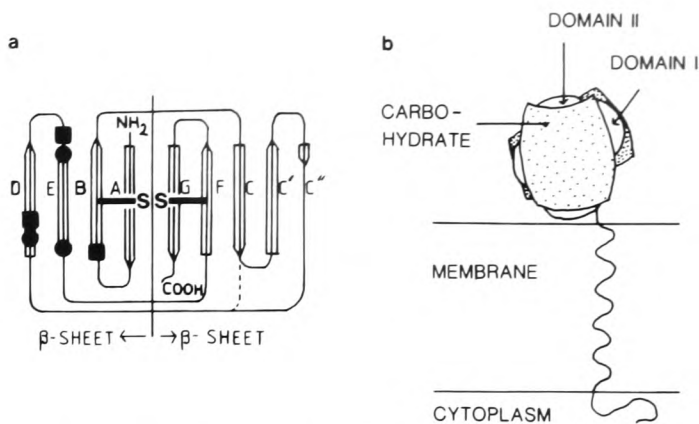


Fig. 7. (A) β strand folding pattern of Ig domains, modified from Amzel and Poljak (1979). The pattern for a C domain is shown by a dashed line directly connecting strands C and D. V domains have an extra loop of sequence in the middle of the domain, indicated here as strand C'. Positions of asparagine-linked carbohydrate are indicated as circles (●) for OX-2 domain I and squares (■) for OX-2 domain II. (B) A model for OX-2 at the cell surface. The stippled areas show how carbohydrate might cover most of the exterior surfaces of the molecule. The layer of carbohydrate would be thicker than drawn here.

integration segment. In Figure 6 we divide the Ig-related molecules into two categories based on whether or not they are involved in antigen recognition. The T receptors and immunoglobulins directly mediate this function and the MHC antigens are somehow co-recognised with foreign antigen by the T cell receptor. The poly Ig receptor clearly has no role in antigen recognition and the functions of Thy-1 and OX-2 glycoproteins are unknown. However it is very difficult to imagine that these latter molecules are in any way directly involved in the specificity of antigen recognition. A more likely possibility is that they are involved in mediating cell-cell interactions on various cell types (Cohen *et al.*, 1981; Williams, 1982; Williams *et al.*, 1984).

The Thy-1 and OX-2 glycoproteins are of considerable interest with regard to the evolution of the Ig superfamily since these are the two simplest structures shown in Figure 6. All of the structures shown are ultimately believed to have come from a primordial single domain structure and Thy-1 is the only known contemporary molecule that exists in this way (β_2 M is a single domain structure but is found associated with the large chain of Class I MHC antigens). Furthermore all of the molecules involved in antigen recognition were probably derived from a primordial precursor like an Ig L-chain and OX-2 glycoprotein is a contemporary molecule with these properties. It is interesting that both Thy-1 and OX-2 are found on neurones and in the case of Thy-1 this is the tissue on which this molecule is conserved between various species. It is thus possible that the Ig-related molecules first evolved to function in cell-cell interactions on sensory cells and this idea is discussed in detail elsewhere (Williams, 1982; Williams *et al.*, 1984).

In Figure 6 we show the OX-2 glycoprotein as an extended two domain structure at a cell surface but in reality interactions may occur between the domains. A relevant point here is the position of the N-linked carbohydrate structures in relation to the predicted Ig-like folding pattern for the OX-2 domains. This is shown in Figure 7a where it can be seen that in both domains the three glycosylation sites are in very similar positions. Furthermore they are found on the same face of the Ig-fold in both domains I and II. This can be contrasted with Thy-1 glycoprotein which has three N-linked glycosylation sites predicted to be on β -strands B, E and G and thus on the faces of both β -sheets of the domain. The proposed

restriction of carbohydrates to one side of each OX-2 domain suggests that the domains may interact *via* the non-glycosylated β -sheets. The faces that would be exposed might then in large part be covered by carbohydrate. A possible model for OX-2 glycoprotein taking these ideas into account is shown in Figure 7b.

Materials and methods

Purification of OX-2 antigen and sequencing of tryptic peptides

5 mg OX-2 antigen was purified from 750 g wet weight of frozen Sprague-Dawley rat brains by solubilisation with sodium deoxycholate, affinity chromatography with MRC OX-2 antibody and gel filtration as described previously (Barclay and Ward, 1982). After removal of sodium deoxycholate by dialysis against 0.1 M NH₄HCO₃, OX-2 antigen (2 mg protein; ~70 nmol) was reduced, alkylated, succinylated and digested with trypsin as described by Campbell *et al.* (1981). Peptides were purified by gel filtration and h.p.l.c. and sequenced on a Beckman spinning cup sequencer as in Campbell *et al.* (1981) and Williams and Gagnon (1982).

Screening the rat thymocyte cDNA library

The oligonucleotide mixture (see Results) was synthesised using the solid phase phosphotriester method developed by Gait *et al.* (1980) and purified by polyacrylamide gel electrophoresis. This probe was radiolabelled by phosphorylating the 5' ends using T4 polynucleotide kinase (Maniatis *et al.*, 1982) and [γ -³²P]ATP and used to screen a rat thymocyte cDNA library. This library which will be described in details elsewhere (M. Thomas and A. N. Barclay, in preparation) was prepared from cDNA made by the loopback procedure and fractionated on sucrose gradients to give a >1 kb fraction which was blunt end-ligated into the pAT 153/PvuII/8 vector (Anson *et al.*, 1984). A complexity of ~160 000 clones was achieved.

The cDNA library was transferred to Whatman 541 filter papers and chloramphenicol amplified as described by Gergen *et al.* (1979). The filters were screened with the oligonucleotide probe essentially as described by Wallace *et al.* (1981) employing a hybridisation temperature of 38°C and a washing temperature of 42°C.

DNA sequencing

Large mol. wt. fragments spanning the coding region of the cDNA clones were prepared by restriction enzyme digestion and purification by agarose gel electrophoresis and electro-elution (Girvitz *et al.*, 1980). The majority of these fragments were further digested with restriction enzymes and labelled at their 3' termini by the Klenow fragment of *Escherichia coli* DNA polymerase I and the appropriate ³²P-labelled deoxynucleotide triphosphate and then fractionated on polyacrylamide gels or strand separated (Maniatis *et al.*, 1982). The sequence was determined by the chemical degradation method of Maxam and Gilbert (1980). A small proportion of the sequence was established by dideoxy sequencing of cDNA fragments subcloned into M13mp8 (Sanger *et al.*, 1977; Messing, 1983).

Northern blot analysis

Total thymocyte RNA was fractionated on a 1.5% agarose gel containing formaldehyde (Lehrach *et al.*, 1977). Transfer to Gene Screen membrane, hybridisation and washing followed the procedure recommended by the manufacturer (New England Nuclear, Boston, USA). Probes from the pX2/11 and pX2/13 cDNA clones were prepared by purification of the fragments extending from the 5' end of the cDNA to the *Xba*I site (Figure 1). These were labelled with ³²P using a nick translation kit from Amersham International.

Acknowledgements

We are indebted to members of Professor G. G. Brownlee's group and to Dr Duncan Campbell for helping us establish the recombinant DNA techniques, to Keith Gould for assistance with oligonucleotide synthesis and Dr Matt Thomas for invaluable discussions. Expert technical assistance was provided by Ed Davey, Tony Gascoyne and Tony Willis, and Mrs Caroline Griffin typed the manuscript.

References

Amzel, L.M. and Poljak, R.J. (1979) *Annu. Rev. Biochem.*, **48**, 961-997.
Anson, D.S., Choo, K.H., Rees, D.J.G., Gianelli, F., Gould, K., Huddleston, J.A. and Brownlee, G.G. (1984) *EMBO J.*, **3**, 1053-1060.
Barclay, A.N. (1981) *Immunology*, **44**, 727-736.
Barclay, A.N. and Ward, H.A. (1982) *Eur. J. Biochem.*, **129**, 447-458.
Belt, K.T., Carroll, M.C. and Porter, R.R. (1984) *Cell*, **36**, 907-914.

Campbell, D.G., Gagnon, J., Reid, K.B.M. and Williams, A.F. (1981) *Biochem. J.*, **195**, 15-30.
Cohen, F.E., Novotny, J., Sternberg, M.J.E., Campbell, D.G. and Williams, A.F. (1981) *Biochem. J.*, **195**, 31-40.
Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978a) in Dayhoff, M.O. (ed.), *Atlas of Protein Sequence and Structure*, Vol. 5, National Biomedical Research Foundation, Baltimore, pp. 345-352.
Dayhoff, M.O., Hunt, L.T., Barker, W.C., Schwartz, R.M. and Orcutt, B.C., eds. (1978b) *Protein Segment Dictionary*, published by The National Biomedical Research Foundation, MD 20901, USA.
Gait, M.J., Singh, M., Sheppard, R.C., Edge, M.D., Greene, A.R., Heathcliffe, G.R., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1980) *Nucleic Acids Res.*, **8**, 1081-1096.
Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) *Nucleic Acids Res.*, **7**, 2115-2136.
Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.L. (1980) *Anal. Biochem.*, **106**, 492-496.
Hedrick, S.M., Nielsen, E.A., Kavalier, J., Cohen, D.I. and Davis, M.M. (1984) *Nature*, **308**, 153-158.
Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M. and Perry, H., eds. (1983) *Sequences of Proteins of Immunological Interest*, published by U.S. Department of Health, Bethesda, USA.
Kaufman, J.F., Auffray, C., Korman, A.J., Shackelford, D.A. and Strominger, J. (1984) *Cell*, **36**, 1-13.
Kehry, M., Ewald, S., Douglas, R., Sibley, C., Raschke, W., Fambrough, D. and Hood, L. (1980) *Cell*, **21**, 393-406.
Komaromy, M., Clayton, L., Rogers, J., Robertson, S., Kettman, J. and Wall, R. (1983) *Nucleic Acids Res.*, **11**, 6775-6785.
Kozak, M. (1983) *Microbiol. Rev.*, **47**, 1-45.
Lehrach, H., Diamond, D., Wozney, J.M. and Boedtker, H. (1977) *Biochemistry (Wash.)*, **16**, 4743-4751.
Maniatis, T., Fritsch, E.F. and Sambrook, J., eds. (1982) *Molecular Cloning, A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499-560.
McMaster, W.R. and Williams, A.F. (1979) *Eur. J. Immunol.*, **9**, 426-433.
Messing, J. (1983) *Methods Enzymol.*, **101**, 20-78.
Mostov, K.E., Friedlander, M. and Blobel, G. (1984) *Nature*, **308**, 37-43.
Ploegh, H.L., Orr, H.T. and Strominger, J.L. (1981) *Cell*, **24**, 287-299.
Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N. and Tonegawa, S. (1984) *Nature*, **309**, 757-762.
Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
von Heijne, G. (1983) *Eur. J. Biochem.*, **133**, 17-21.
Wallace, R.B., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981) *Nucleic Acids Res.*, **9**, 879-894.
Watson, M.E.E. (1984) *Nucleic Acids Res.*, **12**, 5145-5164.
Webb, M. and Barclay, A.N. (1984) *J. Neurochem.*, **43**, 1061-1067.
Williams, A.F. (1982) *J. Theor. Biol.*, **98**, 221-234.
Williams, A.F. and Gagnon, J. (1982) *Science (Wash.)*, **216**, 696-703.
Williams, A.F., Barclay, A.N., Clark, M.J. and Gagnon, J. (1984) in Anderson, L.C., Gahmberg, C.G. and Ekblom, P. (eds.), *Proceedings of the Sigrid Juselius Symposium; Gene Expression during Normal and Malignant Differentiation*, Academic Press, in press.

Received on 22 October 1984

Monoclonal Antibodies in the Study of Differentiation

Biochemical Immunology Group Colloquium organized and edited by M. A. Kerr (Dundee), R. M. E. Parkhouse (London) and J. F. Greally (Galway)

Neutrophil differentiation antigens

THOMAS G. COTTER

Department of Biology, St. Patrick's College, Maynooth, Co. Kildare, Ireland

The human peripheral blood neutrophil plays an important role in the defence of the body against pathogenic micro-organisms. While a substantial amount of information is known about the biological functions of this cell type, little is available on the sequence of events occurring during differentiation which lead to its production and release from the bone marrow. This has been due largely to the lack of suitable material for research purposes. With the recent availability of a number of stable human myeloid cell lines which can be maintained in continuous culture, it is now possible to begin to probe the nature of myeloid differentiation. The human promyelocytic leukaemia cell line (HL-60) is particularly useful in this respect (Collins *et al.*, 1977), because it can be induced to differentiate into neutrophil- or monocyte-like cells by culturing in the presence of a variety of agents. For example, retinoic acid, dimethylsulphoxide and dimethylformamide induce HL-60 cells to differentiate into neutrophils, whereas 1,25-dihydroxyvitamin D₃ and phorbol myristate acetate cause induction to monocytes (Rovera *et al.*, 1979; Miyaura *et al.*, 1981). On differentiation into neutrophils, HL-60 cells acquire the ability to respond to chemotactic stimuli, secrete lysosomal enzymes and perform other biological functions associated with the mature neutrophil (Collins *et al.*, 1978, 1979). When HL-60 cells are induced to differentiate into monocytes, they acquire the characteristics of the blood monocytes (Rovera *et al.*, 1979). Thus with the availability of these cell lines and the introduction of techniques for the production of monoclonal antibodies (Kohler & Milstein, 1975) analysis of neutrophil differentiation antigens has become possible.

In previous studies we have produced a number of monoclonal antibodies to the cell surface of human peripheral blood neutrophils (Cotter *et al.*, 1981a,b). These antibodies were selected on the basis of their ability to interfere with specific cell functions associated with the mature neutrophil. For example, the monoclonal antibody NCD 1 binds to peripheral blood neutrophils and inhibits chemotaxis and lysosomal enzyme secretion in response to stimulation with a variety of agents including C5a and fMet-Leu-Phe. HL-60 cells which do not possess these functions fail to bind this antibody (Cotter & Henson, 1983). When HL-60 cells are induced to differentiate into mature neutrophils they acquire the functions of chemotaxis and degranulation and now bind the antibody NCD 1 (Cotter & Henson, 1983). The antigen binding NCD 1 is a 100000 dalton glycoprotein, with approx. 25000 copies per cell (Cotter & Henson, 1984). A similar pattern of data is seen for the anti-neutrophil monoclonal antibody NCD 3, which

inhibits fMet-Leu-Phe-induced chemotaxis (Cotter *et al.*, 1981b). Neither of these two antibodies inhibit neutrophil phagocytosis or superoxide ion production. A third monoclonal antibody in this series, NCD 2, which has no effect on a variety of cell functions, is selectively internalized during phagocytosis of opsonized sheep erythrocytes (T. G. Cotter & P. M. Henson, unpublished work). Thus, all three monoclonal antibodies bind to mature neutrophils and are associated with cellular functions which are characteristic of these cells. Current studies in our laboratory are attempting to unravel the mechanism of action of the two inhibitory antibodies. This may shed some light on how the cellular machinery responsible for neutrophil chemotaxis/degranulation is assembled during differentiation.

Recently, we have attempted to produce monoclonal antibodies which would bind to the cell surface of HL-60 cells, with the intention of using these antibodies as probes to study differentiation-induced alterations in cell-surface antigens. Both immunofluorescent and immunoperoxidase assays have been used to detect such monoclonal antibodies. These attempts have, however, been hampered by the presence of a highly immunogenic antigen on the HL-60 cell surface, which is also present on the peripheral blood neutrophils. The vast majority of antibodies produced thus far appear to be against the same antigen if not against the same determinant. Preliminary data suggest that the carbohydrate sequence 3-fucosyl-N-acetyl-lactosamine may be responsible. This contention is supported by similar data reported by other investigators (Huang *et al.*, 1982; Gooi *et al.*, 1983). The trisaccharide appears to be a marker for both normal and malignant myeloid cells. Studies by Stockinger *et al.* (1984) have demonstrated that the trisaccharide is also present on certain acute lymphoblastic leukaemia cells but in a cryptic state, requiring prior treatment with neuraminidase for expression. This structure is not present on mature lymphocytes even after neuraminidase treatment. The potential function of this highly immunogenic structure remains unknown at present. In fact this statement applies to a large number of cell-surface antigens.

Fisher *et al.* (1982) produced an interesting monoclonal antibody which bound to an antigen termed AGF4.36 expressed on HL-60 cells, but not on peripheral blood neutrophils. These investigators also produced a variant HL-60 cell line which lacks the ability to differentiate when cultured in the presence of inducing agents. This variant also had a reduced expression of the antigen AGF4.36, indicating that this antigen may be involved in regulating myelopoiesis. Two antibodies S3.13 and S5.7, raised against acute myelogenous leukaemia by Pessano *et al.* (1983), identify antigens which are present on the surface of a large number of leukaemic cell lines. These monoclonal antibodies also bind to subsets of monocytes and lymphocytes and

with blast precursor cells of the myelomonocytic and erythroid lineage. Both antigens disappear during normal myeloid and erythroid differentiation.

In conclusion, the availability of both myeloid cell lines, particularly those which can be induced to differentiate into more mature cells, and monoclonal antibodies to these cells, may aid in our understanding of myelopoiesis. However, the presence of highly immunogenic antigens on a number of these cells make it that much more difficult to produce monoclonal antibodies against weakly immunogenic and potentially more interesting antigens. In this respect more sensitive screening assays and/or better immunization protocols are desirable.

This work was supported in part by a grant from the Medical Research Council of Ireland.

Collins, S. J., Gallo, R. C. & Gallagher, R. E. (1977) *Nature (London)* **270**, 347–349

Collins, S. J., Ruscetti, F. W., Gallagher, R. E. & Gallo, R. C. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2458–2462

Collins, S. J., Ruscetti, F. W., Gallagher, R. E. & Gallo, R. C. (1979) *J. Exp. Med.* **149**, 969–974

Cotter, T. G. & Henson, P. M. (1983) *Clin. Exp. Immunol.* **53**, 249–253

Cotter, T. G. & Henson, P. M. (1984) *Eur. J. Immunol.* **14**, 605–609

Cotter, T. G., Spears, P. & Henson, P. M. (1981a) *J. Immunol.* **127**, 1355–1360

Cotter, T. G., Keeling, P. J. & Henson, P. M. (1981b) *J. Immunol.* **127**, 2241–2245

Fisher, A. G., Bunce, C. M., Deniztoksoz, Stone, P. C. & Brown, G. (1982) *Clin. Exp. Immunol.* **50**, 374–381

Gooi, H. C., Thorpe, S. J., Hounsell, E. F., Rumpold, H., Kraft, D., Foster, O. & Feizi, T. (1983) *Eur. J. Immunol.* **13**, 306–312

Huang, L. C., Civin, C. I., Magnani, J. L., Shaper, J. H. & Ginsburg, V. (1982) *Blood* **61**, 1020–1023

Kohler, G. & Milstein, C. (1975) *Nature (London)* **256**, 495–497

Miyaura, C., Abe, E., Kuribayashi, T., Tanaka, H., Konno, K., Nishii, Y. & Suda, T. (1981) *Biochem. Biophys. Res. Commun.* **102**, 937–941

Pessano, S., Bottero, L., Faust, J., Trucco, M., Palumbo, A., Pegoraro, L., Lange, B., Brezin, C., Borst, J., Terhorst, C. & Rovera, G. (1983) *Cancer Res.* **43**, 4812–4815

Rovera, G., Santoli, D. & Damsky, C. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2779–2783

Stockinger, H., Majdic, O., Lischka, K., Aberer, W., Bettelheim, P., Lutz, D. & Knapp, W. (1984) *J. Natl. Can. Inst.* in the press

Amino acid sequence of lymphocyte-surface antigens

A. NEIL BARCLAY, MELANIE J. CLARK and MATTHEW L. THOMAS

M.R.C. Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, U.K.

The biochemical analysis of lymphocyte-surface antigens has, until recently, been hampered by difficulties in recognizing these molecules and purifying them in sufficient quantities for biochemical analysis. The introduction of the monoclonal antibody technique provided the means to recognize and purify lymphocyte-membrane antigens whilst the recent advances in peptide separation and sequencing and in recombinant DNA technology should allow rapid progress in sequencing of lymphocyte-membrane proteins. Fig. 1 illustrates a strategy using these techniques to sequence lymphocyte-membrane proteins and is of course applicable to other systems. We will discuss the basic strategy from the starting point of the monoclonal antibody of interest. Some key references are given for the more detailed methods.

The strategy outlined in Fig. 1 is one of a number of strategies that are currently being applied to lymphocyte-membrane proteins. The method involves the purification by monoclonal antibody affinity chromatography of sufficient antigen to prepare and sequence peptides. From the amino acid sequence a mixture of oligonucleotides containing all the sequences that code for a peptide sequence (see below) is synthesized, labelled with ^{32}P and used to probe a cDNA library. cDNA clones containing sequences that hybridize to the probe are selected and sequenced to give the protein sequence of the antigen. This method has advantages over other methods in that the peptide data confirm that the cDNA sequence obtained is correct (see Discussion). A key aspect of the strategy in Fig. 1 is the ability to purify sufficient antigen to allow peptide preparation and sequence and this is discussed below together with an outline of the nucleic acid techniques.

Abbreviations used: cDNA, complementary DNA; Ig, immunoglobulin.

Purification of lymphocyte-membrane antigens

How much pure antigen is needed? Current gas-phase sequencers will give good sequences with 1 nmol of peptide and even down to the 10–100 pmol range. If one assumes an overall yield of 10% from starting tissue to purified peptide as typical, then about 10 nmol of antigen in the starting material should be adequate. The amount of antigen per lymphocyte can be estimated by radioactive binding assay under saturating antibody conditions (see Mason & Williams, 1985). Thus if an antigen is present at 20000 molecules/cell, a total of 3×10^{11} cells would contain 10 nmol. Several lymphocyte antigens are present at these levels but the selection of cell lines, tumours or leukaemias with higher site numbers may be useful.

The main steps in the purification of membrane antigens are selection of monoclonal antibody, solubilization of

Production of monoclonal antibody to antigens of interest

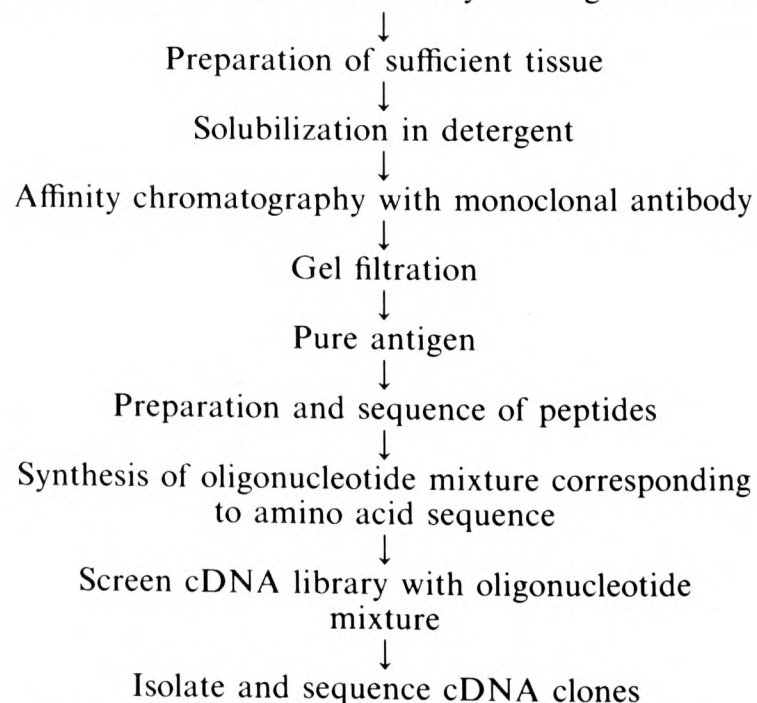


Fig. 1. Sequencing lymphocyte-membrane antigens

antigen and monoclonal antibody affinity chromatography. The monoclonal antibody should have a reasonably high affinity for solubilized monomeric antigen, preferably be IgG and not IgM, and its binding to antigen should be unaffected by detergent (see Mason & Williams, 1985). Sodium deoxycholate is probably the most effective detergent for solubilizing membrane antigens and use in affinity chromatography (e.g. Sunderland *et al.*, 1979) and methods for solubilization and affinity chromatography have recently been discussed in detail by Williams & Barclay (1985). After affinity chromatography a second cycle of affinity chromatography sometimes gives further purification and a final gel-filtration step often gives pure antigen (Sunderland *et al.*, 1979; Brown *et al.*, 1981).

Preparation and sequence of peptides

In order to construct an oligonucleotide probe (Fig. 1) one requires a peptide sequence that includes a stretch of five or six amino acids which have a low redundancy of codons. This could be yielded by a single sequencer run to determine the *N*-terminal sequence. The arguments against this are that the *N*-terminus may be blocked, that a suitable sequence may not be obtained and that the *N*-terminal sequence is the least suited to identifying cDNA clones if full length cDNA is not present. Thus it is preferable to prepare peptides initially and this yields considerably more sequence data. Common methods to prepare peptides are by cleavage by CNBr, trypsin, V8 proteinase or other enzymes followed by separation by gel filtration and high-performance liquid chromatography. For small proteins the gel filtration may not be necessary but for larger proteins this step allows large peptides to be separated and these may then be re-digested with a different enzyme to yield more peptides if necessary. The peptides are sequenced by spinning cup or the more sensitive gas-phase sequencer (Hunkapiller *et al.*, 1984).

Oligonucleotide probe synthesis

A mixture of oligonucleotides containing all the possible combinations of codons that correspond to the chosen peptide sequence is synthesized and used to probe, by hybridization, a cDNA library. Ideally the probe should be at least 14 nucleotides long and preferably longer and contain the minimum number of different sequences to maximize signal to noise and to minimize binding to the wrong clones. Oligonucleotides 17 long have commonly been used in mixtures of typically 128, 64 or less. Thus a peptide sequence lacking in amino acids with large numbers of codons such as Leu, Arg, Ser (six each) but not in Trp (one), Met (one) or Gln, Glu, Asp, etc. (two each) is selected and the oligonucleotide synthesized. After labelling with ³²P the oligonucleotide is used to probe the cDNA library.

Construction and screening of cDNA library

mRNA [poly(A)⁺ RNA] is isolated from a tissue known to produce the antigen of interest and cDNA synthesized using reverse transcriptase and oligo(dT) as a primer. The cDNA is ligated into a suitable bacterial plasmid or phage vector containing a drug-resistance marker to obtain a cDNA library whereby each cDNA is present in its vector in a clone of bacteria. The library is plated on to agar containing the drug, transferred to nitrocellulose or Whatman filters and after fixation the filters are hybridized with the ³²P-labelled synthetic oligonucleotide. The filters are washed, exposed for autoradiography and the positive

clones picked from the original plates. The cDNA inserts are analysed by restriction enzyme digestion and sequencing (Gergen *et al.*, 1979; Wallace *et al.*, 1981; Maniatis *et al.*, 1982).

Discussion

The strategy outlined above provides a general method for sequencing lymphocyte (or other) membrane antigens. It does require the use of biochemical separation techniques and access to a sensitive protein sequencer but it has the great advantage that when a clone is isolated and sequenced it can be readily identified from the protein sequence data and incorrect clones rejected. This is not necessarily the case with other strategies currently in use, e.g. the use of restricted cDNA libraries in the isolation of the T-cell receptor (Hedrick *et al.*, 1984), expression in eukaryotic systems (Kavathas & Herzenberg, 1983), hybrid arrest translation whereby putative clones are tested for their ability to bind the mRNA to the antigen of interest (Parnes *et al.*, 1981) and immunoprecipitation of polysomes to prepare or screen cDNA libraries (Korman *et al.*, 1982).

The strategy outlined in Fig. 1 is a general widely used method to sequence membrane or other antigens (e.g. Reyes *et al.*, 1981; Choo *et al.*, 1982; Noda *et al.*, 1982). The strategy has the added advantage that the availability of pure antigen may be used in other studies such as examining post-translation modification and functions of the protein.

- Brown, W. R. A., Barclay, A. N., Sunderland, C. A. & Williams, A. F. (1981) *Nature (London)* **289**, 456–460
- Choo, K. H., Gould, K. G., Rees, D. J. G. & Brownlee, G. G. (1982) *Nature (London)* **299**, 178–180
- Gergen, J. P., Stern, R. H. & Wensink, P. C. (1979) *Nucleic Acids Res.* **7**, 2115–2136
- Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. (1984) *Nature (London)*, **308**, 149–153
- Hunkapiller, M., Kent, S., Caruthers, M., Dreyer, W., Firca, J., Giffin, C., Horvath, S., Hunkapiller, T., Tempst, P. & Hood, L. (1984) *Nature (London)*, **310**, 105–111
- Kavathas, P. & Herzenberg, L. A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 524–528
- Korman, A. J., Knudsen, P. J., Kaufman, J. F. & Strominger, J. L. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1844–1848
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York
- Mason, D. W. & Williams, A. F. (1985) in *Handbook of Experimental Immunology* (Weir, D. M. & Herzenberg, L. A. eds.), 4th edn., Blackwell Scientific Publications Limited, Oxford, in the press
- Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Furutani, Y., Hirose, T., Asai, M., Inayama, S., Miyata, T. & Numa, S. (1982) *Nature (London)*, **299**, 793–797
- Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. & Seidman, J. G. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2253–2257
- Reyes, A. A., Johnson, M. J., Schödl, M., Ito, H., Ike, Y., Morin, C., Itakura, K. & Wallace, R. B. (1981) *Immunogenetics* **14**, 383–392
- Sunderland, C. A., McMaster, W. R. & Williams, A. F. (1979) *Eur. J. Immunol.* **9**, 155–159
- Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. & Itakura, K. (1981) *Nucleic Acids Res.* **9**, 879–894
- Williams, A. F. & Barclay, A. N. (1985) in *Handbook of Experimental Immunology* (Weir, D. M. & Herzenberg, L. A. eds.), 4th edn., chapter 25, Blackwell Scientific Publications Limited, Oxford

Differentiation antigens in acute leukaemia

DENIS J. REEN

Children's Research Centre, Our Lady's Hospital for Sick Children, Crumlin, Dublin 12, Ireland

The heterogeneity of the haemopoietic cell system has been recognized for many years because of the diversity of functional and morphological differences between various cell types. This is particularly well exemplified in the clinical expression of various leukaemias where the morphological appearance, clinical presentation and response to therapy are a reflection of the leukaemic cell type. Cells of the haemopoietic system and their malignant offspring may be classified into different types and subtypes based on their expression of specific or selective cell markers. These markers include cell-surface receptors such as those for sheep erythrocytes, complement receptors or insulin. A wide range of enzyme analyses describing a variety of enzyme abnormalities or alterations has also been used to distinguish subsets of leukaemia (Drexler *et al.*, 1981). The most extensive use of cell-specific markers for the study of leukaemia, however, has been the application of monoclonal antibodies for the identification and characterization of maturation and differentiation-associated cell antigens of the haemopoietic system. The advent of monoclonal antibody technology has had a major impact on our understanding of the immunobiology of human leukaemia since it is now recognized that leukaemias, especially those of the lymphoid system, can be subdivided into clinically relevant subgroups based on the expression of normal cell differentiation antigens (Thierfelder *et al.*, 1976). This immunological approach to the classification of normal and leukaemia cell differentiation and maturation antigens is based on the concept that as a cell differentiates it must change its molecular content so as to acquire new functional capabilities and discard old ones. Thus, each cell type has a fixed differentiation pathway plotted out along which it loses or acquires various structural antigens which accurately mirror the genetically determined commitment of the cell. Marker studies in normal haemopoiesis and in leukaemia are therefore used to gain some insight into the development biology and cell lineage relationship of normal haemopoietic tissue and by using this knowledge to gain an understanding of the biology, evolution and clinical heterogeneity of leukaemic disease.

The phenotypic profile and cellular characteristics of 'pluripotent stem cells', which are capable of differentiating to all blood cell lineages, are not known. The first recognizable stem cell which is already lineage committed expresses Ia-like (HLA-DR) antigens. These HLA-DR antigens are lost on all cell lineages of the haemopoietic system after maturation, with the exception of B-lymphocytes and a subset of the monocytes/macrophage series. The appearance of myeloid-specific cell-surface antigens corresponds to the loss of HLA-DR antigens on maturing myeloblast cells. A number of monoclonal antibodies recognizing a strongly immunogenic myeloid-specific cell-surface antigen expressed on all cells of the myeloid lineage have been characterized (Majdic *et al.*, 1981). This antigen which is recognized by the monoclonal antibody VIM-D5 consists of two polypeptide chains of molecular mass 150 and 105 kDa respectively (Knapp, 1982).

Much more detailed knowledge has been acquired on the expression of differentiation- and maturation-phase-specific antigens of lymphoid cells than for cells of other lineages through the use of monoclonal antibodies. The

most immature thymocyte precursor cells which originate in the bone marrow express the sheep erythrocyte receptor (E) and its analogous T11 differentiation antigen. Reinhertz has described four stages of lymphocyte maturation, each stage being identified by a restricted phenotype expression (Reinhertz *et al.*, 1980). The most immature thymocytes express the OKT9 and OKT10 antigens while more mature lymphocytes are present as two separate populations expressing either the T4 or T8 mutually exclusive antigens of the functionally distinct helper and suppressor lymphocyte subsets. Cells of the B-lymphocyte lineage have been identified by using a combination of heterologous and monoclonal antibodies as well as by DNA probes for the analysis of immunoglobulin gene re-arrangement as a marker of lineage commitment to the B-cell line. A number of monoclonal antibodies, e.g. BA-1, BA-2, VIB-C5 and B1, are available which recognize early or pre-B cells in advance of their expression of surface immunoglobulin (Stashenko *et al.*, 1980; Abramson *et al.*, 1981; Knapp, 1982). The use of monoclonal antibodies for the identification of differentiation antigens is most valuable when the functional role of the identified antigen is known, and when it is intimately related to the function of the cell type. OKT9 which reacts with immature thymocytes and with transformed cells recognizes the transferrin receptor on the cell surface (Sutherland *et al.*, 1981). The OKT11 monoclonal antibody which reacts with a 40 kDa glycosylated polypeptide recently identified as the sheep erythrocyte receptor on T-lymphocytes (Verbi *et al.*, 1982). However, the function of the majority of differentiation antigens presently recognized by monoclonal antibodies on leukaemic cells is not known.

The main observation arising from the investigation of antigen expression on leukaemic cells is the absolute reflection of normal differentiation cell antigens on leukaemia-transformed cells. This arises because of the restricted or clonal origin of these cells, the imposition of maturation arrest and the fairly broad conversion of a qualitatively normal phenotype on leukaemic cells. It is now accepted that ALL consists of two broad cell subtypes along the lines of normal lymphocyte differentiation, both of which originate in lymphocyte progenitors of the B-cell and T-cell lines. Within these two categories, subtypes can be defined which broadly reflect sequential stages of maturation within the early compartments of these two distinct cell lineages.

The subtypes of ALL may be divided into four clinically relevant groups based on the expression of a restricted or dominant phenotype. These are: 'common'-ALL (cALL), null-cell ALL, T-cell ALL and B-cell ALL, which have an increasing favourable clinical prognosis in the order B-cell ALL < T-cell ALL < null-cell ALL < cALL. In a study of 39 children with ALL we found that 73% of cases express the dominant cALL phenotype, cALLA⁺, Ia⁺, TdT⁺, E⁻, SmIg⁻. This 'common' subtype is defined by at least two monoclonal antibodies, J5 and VIL-A1, which recognize a 98 kDa polypeptide (cALLA) on the cell surface (Ritz *et al.*, 1980; Knapp *et al.*, 1982). This antigen, though not present on mature blood lymphocytes, is not leukaemic-cell-specific as it is expressed on a minority of bone marrow precursor cells which are also TdT⁺ and express the Ia antigen. This antigen has also been demonstrated on some clusters of subcapsular thymocytes but the simultaneous expression of T-cell differentiation antigens on these cells has not yet been determined (Hoffmann-Fezer *et al.*, 1983). The presence of elevated levels of TdT measured by immunofluorescence is generally accepted as confirmation of malignant transformation of lymphoid cells even though some myeloid leukaemias are also TdT-positive (Srivastara *et al.*, 1978).

Abbreviations used: ALL, acute lymphoblastic leukaemia; Ig, immunoglobulin.