

**CLONING AND CHARACTERISATION OF**  
**MHC CLASS I GENES IN CATTLE.**

PAMELA BROWN

Ph. D. Thesis

University of Edinburgh

1989



## **Declaration.**

I declare that the work presented in this thesis is my own and that the contribution of others has been clearly indicated.

## **Abstract.**

The bovine MHC class I loci have largely been characterised using immunological techniques. This has identified in excess of thirty alleles all corresponding to one locus, *BoLA-A*. Detailed analysis of bovine MHC class I genes was undertaken using recombinant DNA technology in an effort to expand on this knowledge.

A bovine liver cDNA library was constructed in  $\lambda$ gt10 and screened with a human class I MHC DNA probe. Two bovine class I cDNA clones were isolated. The largest clone, pBoLA-1, size 1263bp, was incomplete lacking a transcription initiation codon and a polyadenylation signal. pBoLA-1 on translation into amino acid sequence could encode a mature protein of 339 amino acids. Comparison of the amino acid sequence of pBoLA-1 to two other bovine cDNA clones revealed that pBoLA-1 may be allelic to one of them. Analysis of these bovine clones facilitated the identification of putative species specific residues. Extensive comparison of pBoLA-1 protein to class I proteins of other species revealed that the protein was more similar to class Ia proteins than to class Ib proteins. The second cDNA clone, pBoLA-2, was approximately 600bp in size. pBoLA-2 mapped to the 3' end of pBoLA-1, extending it to 1.7kb, the approximate size of class I transcripts. The exact nature of the sequence encoded by pBoLA-2 was not established, although Northern blots suggested it may contain a repeated element.

Two bovine genomic clones (which were isolated from a bovine genomic DNA library by Dr Jean-Luc Vilotte using the pBoLA-1 cDNA clone as a probe) were chosen for further characterisation. The first, phage 41, encoded two non-contiguous hybridising pieces of DNA.

Restriction enzyme digests of this phage were compared on a Southern blot to similar digests of genomic DNA, the isolated phage clone appeared to contain a sequence which was <sup>probably</sup> present as 1-1 copy in the genome. The conclusion reached was that phage 41 carries <sup>two class I genes, separated by 1.5kb. This close proximity of class I genes has also been observed</sup> in mouse. The second genomic clone, phage 33, encoded a class I pseudogene. The gene lacked a defined 5' end but shared 52-76.5% nucleotide similarity with pBoLA-1 stretching over exons five to eight.

Lastly, pBoLA-1 was used to dissect the detailed multi-band hybridisation pattern obtained when probing Southern blots. Use of 5' and 3' probes allowed an estimate of at least 2-3 bands recognised in common. Use of a 5' probe at high stringency revealed that pBoLA-1 only detected polymorphic bands, these by inference are class Ia genes with nonpolymorphic bands corresponding to class Ib genes. The polymorphic bands recognised by pBoLA-1 were correlated to the *BoLA-A* serological type of the animal. Bands which segregated with *BoLA-w6.2*, *-w8* and *-w10* were identified.

Knowledge of the class I genes of the bovine MHC was extended from a single serologically defined locus to a large multi-gene family. This multi-gene family consists of class Ia genes which include pBoLA-1 and possibly class Ib genes as defined by genomic clones and Southern blotting analysis of bovine DNA.

## **Acknowledgements.**

Firstly, I would like to acknowledge and thank my supervisor Dr A. John Clark for his help and encouragement. In addition I would like to thank Dr Bob Hill for teaching me the in's and out's of cDNA cloning and Dr Jean-Luc Vilotte for constantly cajoling me and for his contribution to some of the work in this thesis. I also thank Dr R. L. Spooner for acquiring the funding for this project and his help. This project was funded by an A. R. F. C. studentship.

I am grateful to Dr Simak Ali for often lending an ear and to Dr's Maggie McClenaghan and Bruce Whitelaw for constantly reminding me that sometime, someday I would not be writing a Ph. D. thesis. Lastly, I wish to thank Jennifer Anderson, Dr Alan Archibald, John Bowman, Dr Liz Glass, Dr Stephan Harris, Brian Harrower, Dr Lee Innes, Bob Oliver, Morag Robertson, Dr Paul Simons, John Webster and Dr John Williams for their help and support.

Lastly, I wish to thank my mum, brother and dad for their encouragement and patience.

## **Dedication.**

To Norman, with love.

## Abbreviations.

a.a.	amino acid
ATP	adenosine triphosphate
bp	basepair
BAT	HLA-B-associated transcript
B-LCL 721	B-lymphoblastoid cell line 721
BLV	bovine leukaemia virus
$\beta_2m$	$\beta_2$ -microglobulin
CML	cell mediated lympholysis
CIP	calf intestinal phosphatase
cpm	counts per minute
CTL	cytotoxic T-cell
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
ddATP	dideoxyadenosine triphosphate
ddCTP	dideoxycytidine triphosphate
ddGTP	dideoxyguanosine triphosphate
ddTTP	dideoxythymidine triphosphate
DTT	dithiothreitol
DEPC	diethyl pyrocarbonate
dGTP	deoxyguanosine triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dpm	disintegrations per minute
dTTP	thymidine triphosphate
EDTA	ethylene diaminetetra-acetic acid
GTE	glucose, tris.HCl and EDTA

IDDM	insulin dependant diabetes mellitus
IPTG	isopropylthiogalactosidase
Ir	immune response
k	thousand
kb	kilobase pairs
L	liters
Lad	lymphocyte antigen defined
LBM	Luria broth medium
LCMV	lymphocytic choriomeningitis virus
Ld	lymphocyte defined
M	molar
mM	millimolar
$\mu$ g	microgrammes
MHC	Major Histocompatibility Complex
$\mu$ l	microliters
ml	milliliters
MLC	mixed lymphocyte culture
MLR	mixed lymphocyte reaction
nm	nanometer
PBL	peripheral blood lymphocytes
PCR	polymerase chain reaction
pfu	plaque forming units
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
Slp	sex-limited protein



Ss	serum serological
SSC	standard saline citrate
TAE	tris.Acetate and EDTA
TBE	tris.Borate and EDTA
TCA	trichloro acetic acid
TE	tris.HCl and EDTA
TEMED	NNN 'N' -tetramethyl-1, 2-diaminoethane
UV	ultraviolet light
UWGCG	University of Wisconsin Genetics Computer Group
X-GAL	5-Bromo-4-chloro-3-indoyl- $\beta$ -D-galactopyranoside

## Table of contents.

	<u>Page.</u>
<u>Abstract.</u>	i
<u>Acknowledgements.</u>	iii
<u>Dedication.</u>	iv
<u>Abbreviations.</u>	v
<u>Table of contents.</u>	viii
<u>List of figures and tables.</u>	xviii
<u>Chapter 1.</u>	
1.1 Introduction	1
1.2 The major histocompatibility complex in mouse and man.	3
1.2.1. Class I molecules in mouse and man.	3
1.2.1a Immunological characterisation of class I loci.	3
1.2.1a1 Discovery of class I loci in mouse.	3
1.2.1a2 Discovery of class I loci in man.	5
1.2.1b Biochemistry of <i>H-2</i> and <i>HLA</i> class I proteins.	6
1.2.1b1 Structure of the class I molecule.	6
1.2.1b2 Biochemistry of $\beta_2$ -microglobulin.	9
1.2.1b3 Correlation of the X-ray crystallographic structure of the class I molecule with identified areas of variation.	10

1.2.1c Molecular biological characterisation of the MHC in mouse and man.	11
1.2.1c1 Introduction.	11
1.2.1c2 Southern blotting studies of class I genes.	12
1.2.1c3 Molecular cloning of class I genes in mouse.	13
1.2.1c4 Molecular cloning of class I genes in man.	15
1.2.1c5 Class I gene structure.	16
1.2.1c6 Identifying functionally important regions in class I genes.	19
1.2.1c6a Contribution of <i>in vivo</i> generated mutants to mapping functionally important areas of the protein molecule.	20
1.2.1c6b Contribution of exon shuffling experiments in mapping functionally important areas of the molecule.	21
1.2.1c6c Contribution of mutants created by targeted mutagenesis in mapping functionally important areas of the protein molecule.	22
1.2.1c7 What are the underlying mechanisms which may generate the observed polymorphism ?	24
1.2.1c8 A possible function for the class Ib genes ?	27
1.2.1d The cellular biology of MHC class I function in relation to its structure	28a
1.2.2 Class II genes.	29
1.2.2a Immunological characterisation of class II in mouse and man.	29
1.2.2a1 Discovery of class II loci in mouse.	29
1.2.2a2 Discovery of class II loci in man.	30
1.2.2a3 Immune response genes of mouse.	31

1.2.2b Biochemistry of class II molecules in mouse and man.	32
1.2.2b1 Structure and characteristics of class II molecules.	32
1.2.2b2 Polymorphism of the $\alpha$ and $\beta$ chains of class II molecules.	34
1.2.2c Molecular biological characterisation of class II molecules.	35
1.2.2c1 Introduction.	35
1.2.2c2 Southern blotting studies of class II genes.	36
1.2.2c3 Molecular cloning of class II genes.	37
1.2.2c4 Polymorphism of class II genes at the fine structural level.	39
1.2.2c5 Molecular analysis of the association of class II molecules with disease in man and mouse.	40
1.2.2c6 Use of transgenic animals in elucidating the function of the MHC.	41
1.2.3. Class III genes.	42
1.2.3a Immunological characterisation of class III loci.	42
1.2.3a1 Introduction.	42
1.2.3a2 Discovery of class III loci.	42
1.2.3b Molecular biological characterisation of class III genes.	43
1.3 MHC genes of other species.	44
1.3.1 Introduction.	44

1.3.1a Molecular biological characterisation of MHC in other species.	45
1.3.1a1 Introduction.	45
1.3.1a2 Southern blotting studies on class I genes.	45
1.3.1a3 Molecular cloning of class I genes of domestic animals.	46
1.3.1a3a Cloning of class Ia and class Ib genes.	47
1.3.1a4 Southern blotting and cloning of class II genes.	49
1.3.1a5 Molecular cloning of MHC complexes of domestically important animals.	50
1.4 The bovine MHC; <i>BoLA</i> .	52
1.4.1 Introduction.	52
1.4.2 Discovery of class I loci in cattle.	52
1.4.2a Immunological discovery of class I loci.	52
1.4.2b Biochemical description of class I molecules.	53
1.4.3 Class II loci in cattle.	54
1.4.3a Immunological detection of class II loci in cattle.	54
1.4.3a1 Discovery of class II in cattle.	54
1.4.3a2 Immune-responsiveness in cattle.	54
1.4.3b Biochemistry of class II molecules.	55
1.4.3b1 Biochemical characterisation of class II molecules in cattle.	55
1.4.4 Disease resistance/susceptibility studies in cattle.	55
<u>Aims and approaches.</u>	67

<u>Chapter 2.</u>	68
<u>Materials and methods.</u>	
2.1 Sources of deoxyribonucleic and ribonucleic acids.	68
2.2 Recombinant plasmid clones.	68
2.3 Non-recombinant vectors and bacterial host strains used.	68
2.3a Bacterial host strains and genotypes.	69
2.4 Enzymes, antibiotics and specialised chemicals.	69
2.5 Media.	70
2.6 Autoradiography and photography.	71
2.7 Deproteination of nucleic acids using phenol/chloroform extraction.	71
2.8 Ethanol precipitation of nucleic acids.	72
2.9 Spectrophotometric analysis of DNA and RNA samples.	73
2.10 Preparation of DNA.	73
2.10a Genomic DNA preparation.	73
2.10b Minipreparation of plasmid DNA.	73
2.10c Large scale preparation of plasmid DNA.	74
2.11 Preparation of RNA.	76
2.11a Isolation of total RNA	76
2.11.a1 Method 1.	76
2.11.a2 Method 2.	77
2.11b Isolation of polyadenylated RNA.	78
2.12 Restriction enzyme digests of DNA.	79
2.12a Extracted genomic DNA.	79
2.12b Plasmid and phage digests.	79
2.13 Gel electrophoresis.	80
2.13a Non-denaturing agarose gel electrophoresis.	80

2.13b Agarose denaturing gels.	80
2.13c RNA denaturing gels.	80
2.13d Nondenaturing acrylamide gels.	81
2.13e Denaturing acrylamide gels.	82
2.14 Recovery of DNA from gel electrophoresis and purification.	82
2.14a Recovery of DNA from agarose gels.	82
2.14b Recovery of DNA from low melting temperature agarose gels.	83
2.14c Purification of DNA.	83
2.15 Transfer of nucleic acid and immobilisation on a solid support.	84
2.15a DNA transfer.	84
2.15b RNA transfer.	84
2.15c Colony transfer.	85
2.15c1 Nitrocellulose transfer	85
2.15c2 Transfer to Hybond-N membranes.	85
2.16 Radioactive labelling of DNA.	86
2.17 Hybridisation of filters to radioactive probes.	87
2.17a Hybridisation to Hybond-N membranes.	87
2.17b Hybridisation to nitrocellulose membranes.	87
2.17c Washing hybridised filters.	87
2.18 Construction of a bovine liver cDNA library in bacteriophage $\lambda$ gt10.	88
2.18a First strand cDNA synthesis.	89
2.18b Second strand cDNA synthesis.	91
2.18c Nuclease S1 cleavage of double stranded cDNA.	92

2.18d Methylation of <i>EcoRI</i> sites in the cDNA.	93
2.18e T4 polymerase blunt ending of double stranded cDNA.	93
2.18f Kinasing and ligation of linkers to the cDNA.	96
2.18g Ligation of cDNA to bacteriophage $\lambda$ gt10, packaging and plating.	100
2.18h Screening of library.	101
2.18i Storage of library.	103
2.18i1 Equilibrium density centrifugation of $\lambda$ gt10 liver cDNA library.	103
2.18j Phage growth and purification of DNA.	104
2.18j1 Mini-preparation of bacteriophage DNA.	104
2.18j2 Large scale preparation of phage DNA.	105
2.19 Manipulation of DNA fragments.	106
2.19a Phosphatasing of vector DNA.	106
2.19b Ligation of vector DNA to insert DNA.	106
2.20 Transformation of ligations.	107
2.21 Growth and preparation of bacteriophage M13 double stranded vector and single stranded templates.	108
2.21a Growth of single stranded templates.	108
2.21b Large scale production of double stranded M13 bacteriophage.	109
2.22 Sanger dideoxy sequencing.	110
2.23 Processing of sequence information.	111



<u>Chapter three.</u>	112
<u>Characterisation of two bovine cDNA clones and their comparison to other class I genes.</u>	
3.1 Introduction.	112
3.2 Results and discussion.	116
3.2.1 Construction and screening of a bovine liver cDNA library.	116
3.2.2 Characterisation of pBoLA-1.	117
3.2.2a Primary structural features and possible modification sites of pBoLA-1.	117
3.2.2a1 Cysteines.	120
3.2.2a2 Glycosylation sites.	120
3.2.2a3 Other possible modification sites.	120
3.2.2b Alignment of protein domains.	121
3.2.2b1 5' untranslated region and translational start site.	121
3.2.2b2 Signal peptide.	121
3.2.2b3 First, second and third protein domains.	122
3.2.2b4 Transmembrane domain.	122
3.2.2b5 Cytoplasmic domain.	125
3.2.2c Comparison of pBoLA-1 to other bovine class I cDNA's and to class I proteins of other species.	125
3.2.2d The relationship of polymorphism to functional significance.	135
3.2.3 pBoLA-2.	140
3.3 Summary.	147

<u>Chapter 4.</u>	148
<u>Fine structural analysis of two bovine class I genomic clones.</u>	
4.1 Introduction.	148
4.2 Results and discussion.	150
4.2.1 Initial analysis of two genomic clones.	150
4.2.2 Detailed analysis of phage 33.	159
4.2.2a Sequencing of phage 33.	165
4.2.2b Comparison of phage 33 to <i>HLA-A3</i>	171
4.3 Summary.	176
 <u>Chapter 5.</u>	 177
<u>Southern blotting studies on the bovine MHC: dissection of the hybridisation pattern obtained and identification of RFLP's coincident with <i>BoLA-A</i> serological types.</u>	
5.1 Introduction.	177
5.2 Results and discussion.	180
5.2.1 Investigation of the class I region of <i>BoLA</i> by Southern blotting.	180
5.2.1a Human and bovine cDNA probes show that the class I genes are a multi-gene family.	180
5.2.1b Dissection of the class I multi-gene family.	180
5.2.1b1 Evidence for the specific detection of class Ia genes.	180
5.2.1b2 Evidence for class Ib genes in the <i>BoLA</i> class I region.	189

5.2.2 Correlation of RFLP's to serology.	189
5.3 Summary.	202
<u>Conclusion.</u>	203
<u>Future studies.</u>	206
<u>References.</u>	207
<u>Addendum.</u>	250

## List of figures and tables.

	<u>Page.</u>
 <u>Chapter 1.</u>	
Figure 1. Genetic and DNA maps of the <i>H-2</i> , <i>HLA</i> and <i>BoLA</i> major histocompatibility complexes.	57-58
Figure 2. Domain structure of class I and class II proteins.	59-60
Figure 3. The 3D structure of the first two domains of HLA-A2.	61-62
Figure 4. Fine structure of a class I genomic clone.	63-64
Table 1.	65
Table 2.	66
 <u>Chapter 2.</u>	
Figure 1. Alkaline agarose gel loaded with test aliquots of cDNA reactions.	94-95
Figure 2. Nondenaturing acrylamide gel analysis of the end ligation of linkers onto the cDNA.	98-99
 <u>Chapter 3.</u>	
Figure 1. Northern blot analysis.	118-119
Figure 2. Nucleotide sequence and amino acid translation of pBoLA-1 shown with sequencing strategy and restriction map.	123-124
Figure 3. Alignment of pBoLA-1 nucleotide sequence with BL3-6 and BL3-7 nucleotide sequence.	126-127
Figure 4. Alignment of amino acid translated sequence of pBoLA-1, BL3-6 and BL3-7 proteins.	129-130

Figure 5. Percentage homologies of pBoLA-1 and other selected proteins in a domain by domain comparison.	133-134
Figure 6. Alignment of pBoLA-1 protein with other class I proteins.	138-139
Figure 7. Alignment of pBoLA-1 and pBoLA-2.	141-142
Figure 8. Northern blots probed separately with sub-fragments of pBoLA-2 and human cDNA clone pHLA-27.	145-146

#### Chapter 4.

Figure 1. Restriction maps and areas of hybridisation of phages 33 and 41.	152-153
Figure 2. Example of how phage maps depicted in figure 1 were constructed.	154-155
Figure 3. Southern blot of restriction enzyme digests of animal 10160 and phages 33 and 41.	157-158
Figure 4. Orientation of phage clone 33 using 5' and 3' probes derived from pBoLA-1.	161-162
Figure 5. Fine mapping of phage 33: localisation of 5' and 3' end in the <i>EcoRI/Sall</i> subclone.	163-164
Figure 6. Depiction of nucleotide and translated amino acid sequence of phage 33 subclone <i>EcoRI/Sall</i> with areas of similarity to pBoLA-1 indicated.	166-167
Figure 7. Dotplot alignment of region B against nucleotides 1-600 of pBoLA-1.	169-170
Figure 8. Comparison of nucleotide sequence of the gene <i>HLA-A3</i> with regions C, E and F of phage 33.	174-175

## Chapter 5.

Figure 1. Comparison of pHLA-27 and pBoLA-1 when used as probes against bovine genomic DNA.

Figure 2. Comparison of two fragments derived from the 5' and 3' ends of the cDNA clone pBoLA-1 which were used to probe a Southern blot of bovine genomic DNA.

Figure 3. Dissection of the hybridisation pattern of a *TaqI* digest of a Friesian half-sib family using the complete pBoLA-1 cDNA and 5' *EcoRI*/*BglII* specific fragments as probes.

Figure 4. *PvuII* and *EcoRI* digests of a Hereford half-sib family probed with pBoLA-1.

Figure 5. *BamHI* and *PstI* digests of DNA from a Hereford half-sib family probed with pBoLA-1 after Southern transfer to a nylon membrane.

Figure 6. *PvuII* digest of DNA from a Friesian half-sib family probed with pBoLA-1 after Southern transfer.

Figure 7. *PvuII* digest of DNA from two unrelated Friesian half-sib families with the dam as the common parent between the half-sibs, probed with pBoLA-1 after Southern transfer.

Figure 8. Graph of RFLP's corresponding to serological types of the Friesian half-sib family already shown in figures 3 and 6.

# CHAPTER 1.

## 1.1 Introduction.

One of the most fundamental questions asked in eukaryotic biology is how organisms distinguish self from nonself.

Self/nonself discrimination has been observed in sponges (which are primitive metazoa). When two genetically different sponges are apposed, tissue destruction occurs at their boundary (Hildemann *et al.* 1979). Over two hundred genetically distinct sponges were characterised by this rejection phenomenon (Hildemann *et al.* 1979). It became apparent from these studies that the required characteristics for self/nonself discrimination are; a) recognition structures on the cell surface, b) extreme polymorphism in the recognition structures and c) methods of destroying nonself once recognition has been effected. The major histocompatibility loci of mammals fit the above description perfectly; 1), they are cell surface glycoproteins 2), they are extremely polymorphic, mapping to one chromosome arm to form the Major Histocompatibility Complex (MHC) and 3) MHC proteins are recognised by T-cells; if they are foreign a series of reactions are triggered which leads to their elimination.

The MHC is one of the best characterised areas of chromosome in mammals. In mice it is named the *H-2* complex and covers  $2 \times 10^6$  bp on chromosome 17. In man it is named the *HLA* complex covering  $3 \times 10^6$  bp on chromosome 6. The MHC can be sub-divided into three groups of loci; the class I, II and III. The class I loci (Klein 1979) are further sub-divided into class Ia and class Ib (Klein 1986).

The class I loci encode N-glycosylated cell surface proteins which

are composed of a heavy  $\alpha$  chain of 44kd expressed in non-covalent association with a 12kd light chain  $\beta_2$ -microglobulin ( $\beta_2m$ ). In the case of the class Ia, these are expressed on all nucleated cells, whereas the class Ib seem to have a more restricted expression (Flaherty 1980).  $\beta_2m$  is encoded outwith the MHC on chromosome 2 in mice and chromosome 5 in man.

The class II loci (Klein 1977) are also encoded within the MHC and are integral membrane bound cell surface glycosylated proteins. They, by comparison to class I, are restricted in expression to cells of the immune system namely; B-cells, Langerhans cells, macrophages, thymic epithelium and in man on activated T-cells. The class II are expressed as noncovalently associated dimers of heavy  $\alpha$  chain 34kd and light  $\beta$  chain 29kd.

The class III loci are comprised of complement proteins B, C4, C2 and cytochrome P-450 21-hydroxylase. It is argued that these should not be considered as components of the MHC and should lose their class III status (see Klein 1986), but as the MHC is being dealt with as a chromosomal region, it seems plausible to include them in this section. A description of the discovery, biochemistry and molecular elucidation of these loci now follows.



## 1.2 The major histocompatibility complex in mouse and man.

### 1.2.1 Class I molecules in mouse and man.

#### 1.2.1a Immunological characterisation of class I loci.

##### 1.2.1a1 Discovery of class I loci in mouse.

Gorer, Lyman and Snell (1948) were the first to recognise that rejection of tissue transplants between incompatible individual mice was controlled by a group of loci. These loci elicited a specific antigenic response (by generation of antibodies) and all mapped to chromosome 17. Characterisation of these loci with immune sera generated by injection of 1), tumour cells or by 2), inoculation with lymphoid tissue or 3), skin grafting and their subsequent mapping was aided by using inbred (genetically identical) and congenic (genetically identical except for a region of chromosome of interest) strains of mice, see Klein (1975). Development of these congenic mice was due to the work of George Snell (1948,1958), because it was recognised that control of background genetic variation was crucial for definition of the system. The lines were extremely well characterised with distinct combinations of alleles of class I and class II loci. This has enabled individual combinations to be expressed as a haplotype e.g. BALB/c mice are classified as being  $H-2^d$  haplotype. The analysis of intra-H-2 recombinants was first applied by Gorer and Milulska (1959). Subsequent use of these led to the mapping of serological reactions to regions of the chromosome, giving rise to the concept of a two locus model,  $K$  and  $D$ , controlling tissue transplantation in mice (Klein and Shreffler 1972). The existence of a

third locus, the *L* locus, was proposed to explain anomalous antibody typing results in different strains of mice (Snell *et al.* 1974). The *K* and *D* loci are extremely polymorphic with recent figures of 92 alleles at the *K* locus and 63 alleles at the *D* locus (Klein 1986). The *L* locus appears to be less polymorphic with only 2 alleles (Klein 1986). Figure 1 shows the location of these loci on the genetic map of the H-2 complex of the BALB/c mouse.

The express role of these molecules did not become apparent until Zinkernagel and Doherty (1974a) found that virus specific T-cells were dually specific for viral antigens and self surface antigen encoded by class I MHC loci. They injected mice of different *H-2* types intracerebrally with lymphocytic choriomeningitis virus (LCMV). Some were killed seven days later and the cytotoxic activity of their spleen cells was tested on virus-infected L-cells (*H-2<sup>k</sup>*). Only spleen cells from the comparable *H-2<sup>k</sup>* haplotype lysed the cells. This dual specificity was called MHC restriction (Zinkernagel and Doherty 1974b,1979).

The class I loci capable of acting as MHC restriction elements (class Ia) are functionally distinct from another group of class I loci (class Ib). These other loci are much less polymorphic (Michaelson *et al.* 1977; Flaherty 1980), do not seem to elicit a major graft rejection response (Flaherty 1980), do not act as restriction elements for cytotoxic T lymphocytes (CTL's) (Kastner *et al.* 1979b) and were thought to have a limited expression, only appearing at certain times of development and on thymic leukaemias (Old *et al.* 1963; Boyse *et al.* 1963) (it is now apparent that this is not strictly true, see section 1.2.1c8). This second group of loci all map approximately one centimorgan to the right of the *K* and *D* loci on chromosome 17 and

are named the *Qa* and *Tla* (class Ib), (see figure 1 for the location of these loci).

#### 1.2.1a2 Discovery of class I loci in man.

The identification of MHC class I loci in humans lagged behind that of mice and for a while, it was thought that the phenomenon of allograft rejection was restricted to mice (see Klein 1975).

The human leukocyte antigens were first described by J. Dausset, R. Payne and J. Van Rood in 1958 (see Klein [1986] for details), and were subsequently named *HLA*. They are very polymorphic (Dausset 1981) and map to chromosome 6 (Van Someron *et al.* 1974). Their investigation has relied on different methods of generating sera than were used in mice. The main sources of appropriate sera in man have been a), multi-parous women b), people who have received multiple blood transfusions and c) volunteers who have undergone reciprocal skin grafts. Analysis of sera generated by these methods entails large panels of sera being tested against large panels of cells, with similar reacting groups being classified together (see Klein 1982, 1986). The use of family groups has allowed three loci to be characterised, which correspond to the *HLA-A*, *HLA-B* and *HLA-C* class I loci (figure 1)(see Klein 1986). In comparison with the *K* and *D* loci in mice, these are just as polymorphic, with 23 *HLA-A* alleles, 47 *HLA-B* alleles and 8 *HLA-C* alleles (Bodmer and Bodmer 1984).

## 1.2.1b Biochemistry of *H-2* and *HLA* class I molecules.

### 1.2.1b1 Structure of the class I molecule.

As it emerged that class I molecules were pivotal in immune surveillance, it was clearly necessary to determine their structure and relate this to their function. Crucial to this therefore was the isolation of class I molecules "it took a long time and many false starts before biochemists succeeded in extracting the MHC molecules from the membrane" (Klein p.p. 175 1986). Two methods of isolating class I molecules were used; the membrane was either disrupted with a detergent (nonidet-40, *NP-40*) or by cleavage of the external portion of the protein with a proteolytic enzyme (papain). Broadly, the class I polypeptide can be divided into three regions (figure 2); a), an amino terminal extracellular domain of 280 residues b), hydrophobic region of approximately 25 residues and c) carboxy region situated intracellularly. The elucidation of this organisation is described below.

Once solubilised by papain, the extracellular portion of class I proteins can be sub-divided into three domains ( $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ ) comprising amino acid residues 1-90, 91-180 and 181-271 (Ploegh *et al.* 1981). The first domain ( $\alpha_1$ ) consisting of residues 1-90 has an asparagine linked carbohydrate attachment site at amino acid 86 in *H-2* class I molecules and *HLA* class I molecules. The second domain ( $\alpha_2$ ) has a carbohydrate attachment site at amino acid 176 in *H-2* class I molecules but not in *HLA* class I molecules (Nathenson *et al.* 1981). Cysteines capable of forming internal disulphide bridges occur at amino acids 101 and 164 in the second domain and at 203 and 259 in the third domain (Kimball and Coligan 1983) and allowed

delineation of the protein into three separate domains. Only the third external domain ( $\alpha 3$ ) associates non-covalently with  $\beta_2$ -microglobulin on the cell surface (Yokoyama and Nathenson 1983), (figure 2), this stabilises extracellular expression of the molecule. In mouse there is an additional glycosylation site at amino acid 256-Asn, this has only been observed in *H-2K<sup>d</sup>*, *H-2L<sup>d</sup>* and *H-2D<sup>b</sup>*. However, *H-2D<sup>b</sup>* and *H-2L<sup>d</sup>* molecules do not associate efficiently with  $\beta_2$ -microglobulin on the cell surface (Maloy *et al.* 1980; Coligan *et al.* 1980). Perhaps the carbohydrate modification interferes with the interaction of the two molecules. The third domain of class I molecules is similar in size and structure to immunoglobulin constant domains (Orr *et al.* 1979). A similar organisation is also seen for the domains in the class II molecules which are closest to the membrane (see figure 2) (Kaufman *et al.* 1984a).

The existence of a transmembrane region in class I molecules was demonstrated by Walsh and Crumpton (1977). They showed that lactoperoxidase-catalysed iodination of inside-out lymphocyte membrane vesicles only labelled class I molecules and not  $\beta_2$ -microglobulin. By a process of elimination, the membrane binding region was localised to a hydrophobic stretch of residues from 284-307 in *H-2K<sup>b</sup>* (Nathenson *et al.* 1981). Generally, the membrane binding portion is approximately 24 amino acids long in *HLA* and *H-2* class I molecules and contains no charged or polar residues (Ploegh *et al.* 1981). This number of residues is enough to allow it to span the lipid bilayer of a cell if an  $\alpha$ -helical configuration is assumed (Ploegh *et al.* 1981). Immediately outside the lipid bilayer (but inside the

cell) there are basic residues which are postulated to help anchor the molecule in the cell membrane, by interacting with cytoskeletal elements therefore preventing the carboxy terminus from entering the lipid bilayer (Ploegh *et al.* 1981).

The class I protein terminates with a hydrophilic region at its carboxy terminus, which is composed of mostly polar residues (Ploegh *et al.* 1981). The cytoplasmic region is 39 amino acids long for *H-2* (Nathenson *et al.* 1981), 31 amino acids long for *HLA* (Robb 1978). Both have a very high polar amino acid content (up to 50%). Internal to the cytoplasmic region are a number of possible modification sites. Free cysteine residues seen in *HLA-B7* and *H-2K<sup>b</sup>*, although not in similar positions, may allow covalent interaction with cytoplasmic components (Bourguignon and Singer 1977; Koch and Smith 1978). The only post-translational modification sequence present in this region is a phosphorylation consensus sequence ser-asp/glu-x-ser(P)-leu which was identified in *H-2* class I (Rothbard *et al.* 1980) and *HLA* class I (Pober and Strominger 1981). The phosphorylation site is very well conserved across species in class I molecules (Guild and Strominger 1984).

Overall when comparing *H-2* and *HLA* class I molecules in the transmembrane region and the cytoplasmic region of the molecule there is marked divergence. This lack of homology suggests a lack of structural constraint on these regions, they only need to remain hydrophobic and hydrophilic respectively (Ploegh *et al.* 1981; Kimball and Coligan 1983).

### 1.2.1b2 Biochemistry of $\beta_2$ -microglobulin.

The light chain of class I proteins is encoded outwith the MHC region, being found on chromosome 15 in man (Goodfellow *et al.* 1975) and chromosome 2 in mouse (Michaelson 1981; Goding 1981).

$\beta_2$ -microglobulin ( $\beta_2m$ ) was initially discovered in the urine of patients (Berggard and Bearn 1968) and has been characterised essentially as a consequence of its copurification with class I heavy chains. Mammalian  $\beta_2m$  is 99 amino acids long with cysteines at positions 25 and 80 which combine to form a disulphide bond looping out the amino acids between them (see figure 2). The elucidation of the structure of  $\beta_2m$  was done primarily by Peterson *et al.* (1972) and Smithies and Poulik (1972). They both reported the strong similarity that  $\beta_2m$  has to immunoglobulin constant domain structure.

There is also considerable homology between  $\beta_2m$  and the third external domain of heavy class I chains in man and mouse (Nathenson *et al.* 1981). It is now known that this is the domain that  $\beta_2m$  associates with on the cell surface (Yokoyama and Nathenson 1983).

$\beta_2m$  is not polymorphic. Detailed comparison of the protein sequence of  $\beta_2m$  derived from different inbred mouse strains shows a high level of homology exists between them. Two forms of mouse  $\beta_2m$  differed at only one amino acid (Gates *et al.* 1981). This can be contrasted to the large amount of polymorphism displayed by class I polypeptides within a species. The large amount of polymorphism attributable to class I molecules is discussed in further detail below.

1.2.1b3 Correlation of the X-ray crystallographic structure of the class I molecule with identified areas of variation.

Figure 3 presents the structure of the putative antigen binding site of class I molecule HLA-A2. Now that this structure has been elucidated, amino acid positions which were observed to be highly polymorphic can be compared to the structure and determined if they are at sites on the molecule which either interact with antigen or cytotoxic T-cell (CTL) (see Bjorkman *et al.* 1987b for greater detail).

The groundwork for this type of mapping had been laid by numerous previous studies. Detailed sequence comparisons done by Nathenson *et al.* (1981) over the first 100 amino acid residues of *H-2* class I proteins revealed variable clusters at amino acids 61-83 and at 95-99, with minor clusters at 22-24, 30-32, and 41-45. Lopez de Castro *et al.* (1982) compared *HLA-A2* to *HLA-A28* and *HLA-B27* and observed most variation over residues 43-195 which was localised to positions 65-80, 105-116 and 177-194 in the extracellular domain. Krangel *et al.* (1983) compared variants of *HLA-A2* which were serologically identical but varied in CTL recognition, the variation was localised to amino acids 147-157, indicating that this area plays a role in CTL recognition of class I molecules. Kimball and Coligan (1983) compared numerous protein sequences and plotted the regions of diversity, again clustering occurred at defined amino acid positions e.g. 9, 30-32, 62-83, 95-121, 152-157 and 193-198. Some of the areas which these amino acids map to are indicated in figure 3.

Therefore, the three-dimensional structure of HLA-A2 (Bjorkman *et al.* 1987a) confirmed and extended these studies linking the substitution pattern of amino acids in the extra-cellular domains of



class I molecules to functionality. The structure is formed from the first two domains ( $\alpha 1$  and  $\alpha 2$ ) and consists of an intramolecular dimer of  $\beta$ -pleated sheet supporting two  $\alpha$ -helices which combine to form the antigen binding site on the top of the molecule. These areas of structure compare very favourably with the previous mapping of areas of functional importance on the class I molecule (Bjorkman *et al.* 1987b)

### 1.2.1c Molecular Biological characterisation of the MHC in mouse and man.

#### 1.2.1c1 Introduction.

Sections 1.2.1a and 1.2.1b have described the serological and biochemical characterisation of class I molecules. However, there were still major questions to be asked about the class I loci. Only the H-2K<sup>b</sup> molecule had been completely sequenced by biochemical means (Coligan *et al.* 1981). This was a handicap in analysing the possible underlying mechanisms which could contribute to the generation of the large number of class I alleles identified by serology. In addition, the exact relationship of the class Ia to the class Ib loci had not been established. The relatedness of the class Ia molecules (*K*, *D* and *L* in mouse; *A*, *B* and *C* in man) to the class Ib molecules (*Qa* and *Tla*) was not entirely clear when only serological, immunological and biochemical criteria were used. Shared characteristics of the two sets of loci include 1), they both associate noncovalently with  $\beta_2m$  on the cell surface (Flaherty 1980), 2), they have similarly sized polypeptide chains and peptide composition (Soloski *et al.* 1981), but the *Qa* and *Tla* are much less polymorphic (Flaherty 1980), 3),

cell-mediated lymphocytosis (CML) can be directed against *Qa* antigens but is not MHC restricted (Kastner *et al.* 1979a), *Qa* restricted CTL's cannot be generated (Kastner *et al.* 1979b) and 4), the *Qa* and *Tla* molecules are much less polymorphic in wild mice populations (when wild mice were trapped, Klein (1979) counted 17 different *K* alleles, 20 different *D* alleles but only 3 *Qa-1*, 2 different *Qa-2* and 6 different *Tla* alleles).

The answers to these questions could only be provided by analysing the fine structure of the class I genes, this was facilitated by recombinant DNA technology. The first steps in molecular biological characterisation of the MHC involved isolating cDNA clones of class I molecules. Ploegh *et al.* (1980) isolated a cDNA clone corresponding to an *HLA* class I transcript, which they identified by immune precipitation of the *in vitro* translated product. Sood *et al.* (1981) made an oligonucleotide derived from the published protein sequence to prime the synthesis of *HLA-B7* cDNA clones *in vitro*. Once isolated these cDNA clones could be used to build up a picture of class I genes in humans and mice by the dual approach of Southern blotting and genomic cloning.

#### 1.2.1c2 Southern blotting studies of class I genes.

cDNA clones were used as probes in Southern blotting studies of mouse and human genomic deoxyribonucleic acid (DNA). Mouse cDNA clones when used as a probe against mouse DNA detected multiple class-I-hybridising bands (Cami *et al.* 1981) which were not rearranged in different tissues studied (Steinmetz *et al.* 1981a). This suggested that the polymorphism of class I genes was not generated by somatic rearrangement and mutation, unlike immunoglobulin genes

(Tonegawa 1983). When analysing mouse genomic clones Cami *et al.* (1981) reported at least 10-12 genes. This far outweighed the number of known functional class Ia loci (*K*, *D* and *L*). It was possible that most of them were pseudogenes, or they could correspond to the class Ib (*Qa* and *Tla*) genes. A similar large number of class-I-hybridising bands were reported in man (Biro *et al.* 1983). Further hybridisation studies confirmed that most of the hybridising bands mapped to the right or distal end of the MHC (which correspond to the *Qa* and *Tla* genetic regions in mouse) by recombinant inbred strains in mice (Pease *et al.* 1982) and gamma-irradiated deletion mutants in man (Orr and Demars 1983).

#### 1.2.1c3 Molecular cloning of class I genes in mouse.

The results of Southern blotting studies were confirmed by cosmid cloning of the H-2 chromosomal region. Steinmetz *et al.* (1982b) isolated 36 distinct class I hybridising sequences from the BALB/c mouse (H-2<sup>d</sup> haplotype) (now revised to 35 Stephan *et al.* 1986), (see figure 1) and subsequent mapping of these placed 31 in the *Qa* and *Tla* regions (Winoto *et al.* 1983). This was achieved by the availability of recombinant congenic strains of mice. The strategy used was to digest the cosmid clones with infrequently cutting enzymes to construct an overlapping map of cosmids so that the encoded genes could be ordered into clusters. These cosmid clusters were mapped to genetically defined regions of the *H-2* complex by using restriction fragment length polymorphism studies (RFLP). The RFLP mapping strategy relied on the isolation of probes from the cosmids which detect one or a low number of genomic DNA restriction fragments. These probes can then be used against DNA from congenic and

recombinant congenic strains of mice to detect restriction site polymorphisms. Essentially three types of polymorphism can be detected; 1), change in restriction site (mutation). 2), loss of restriction site (deletion) and 3) an increase in number of fragments (duplication). These polymorphisms can then be correlated with serologic polymorphisms therefore allowing mapping of particular fragments to one of the four class I regions characterised *K*, *D*, *Qa* and *Tla*. Using a similar strategy, Weiss *et al.* (1984) isolated 26 class I genes from C57BL/10 (*H-2<sup>b</sup>* haplotype) mouse and also found the majority mapped to the *Qa* and *Tla* regions of chromosome 17.

The techniques of Southern blotting and genomic cloning highlighted several new findings. The first, was that class Ia genes could vary in number between haplotypes in mice. This had been alluded to by Hansen *et al.* (1981) who showed that additional specificities (*M<sup>d</sup>* and *R<sup>d</sup>*) could be immune precipitated from cells (also see review Hansen *et al.* 1984). It is now known that the *D* region of BALB/c mice contains 5 genes (*H-2D<sup>d</sup>*, *D2<sup>d</sup>*, *D3<sup>d</sup>*, *D4<sup>d</sup>* and *H-2L<sup>d</sup>*) (Stephan *et al.* 1986), apart from *H-2D<sup>d</sup>*, only one other *D* gene is expressed (Mann and Forman 1988). All this contrasts with the one gene characterised in the *D* region of C57BL/10 mice (Weiss *et al.* 1984).

The second major finding was that the main source of variation in the numbers of class-I-hybridising bands seen in genomic hybridisation studies was the *Qa* and *Tla* genes. The main characterising features (discussed so far) of the *Qa* and *Tla* genes are their smaller degree of functional polymorphism as detected by serology and biochemistry but their higher degree of polymorphism in numbers as detected at the DNA level. The *Qa* and *Tla* clusters

identified by cloning studies were not completely conserved across inbred lines of mice (Steinmetz *et al.* 1982b; Winoto *et al.* 1983; Weiss *et al.* 1984; Brown <sup>G.D.</sup> *et al.* 1988).

#### 1.2.1c4. Molecular cloning of class I genes in man.

An entirely different approach to large scale mapping and cloning of the class I region has been used in man. Investigators have had to use a cell line which has been  $\gamma$ -irradiated and selected for loss of expression of class I serological specificities (Orr *et al.* 1982) in order to map the order of class I loci.

Using the human B-lymphoblastoid cell line 721 (B-LCL 721), Koller *et al.* (1989) have mapped a region of the HLA complex bounded by the *HLA-B* locus on the proximal end to the *HLA-9.2p* locus at the distal end. The strategy entailed isolating locus-specific DNA markers from genomic clones, which were then hybridised to Southern blots of B-LCL 721 DNA. If they detected a restriction enzyme site polymorphism they were used on further Southern blots of deletion mutant cell line DNA and on an extended HLA-typed pedigree. By these methods they succeeded in characterising 17 class I genes; 11 of these were pseudogenes or fragments, three corresponded to *HLA-A*, *-B* and *-C*, the remaining clones corresponded to *HLA-E* (Koller *et al.* 1988), *HLA-G* (Geraghty *et al.* 1987) and *HLA-F* (Shimizu *et al.* 1988). Figure 1 shows the exact order of these genes. An additional study of the HLA chromosomal region surrounding the *HLA-B* gene has been undertaken by Spies *et al.* (1989).

The mapping of loci in a 435kb region mapping centromeric to the *HLA-B* locus has used DNA probes already available from mice to

isolate genes shared in common e.g. tumour necrosis factor  $\alpha$  and  $\beta$  ( $\text{TNF}\alpha$ ,  $\text{TNF}\beta$ , see section 1.2.3b) but has also used the occurrence of rare restriction enzyme sites. CpG-rich sequences are often associated with the 5' ends of genes and can be identified by cutting with *Bss HII* and *Sac II* restriction enzymes (Bird 1986; Lindsay and Bird 1987). Spies *et al.* (1989) isolated overlapping cosmid clones, those containing *Bss HII* and *Sac II* sites were hybridised to RNA dot blots, a positive signal indicated a transcribed gene. In this way they identified 5 different genes which they named HLA-B-associated transcripts (BAT's) these all map centromeric to the *HLA-B* locus within the MHC complex. Their structure and function is unknown.

The complete mapping of the class I regions in mouse and man reveals that numerous genes are present and that, remarkably, there appears to be some conservation of genes between the species. Of the genes characterised, some correspond to the class Ia, some are class Ib and some are pseudogenes. The BAT genes of man do not appear to be conserved in mouse, they have probably arisen after species divergence.

#### 1.2.1c5 Class I gene structure.

Southern blotting studies in mice highlighted the vast numbers of class-I-hybridising sequences in the genome; the majority of these mapping outside the *K* and *D* regions to the *Qa* and *Tla* regions as described in part 1.2.1c2 of this chapter. One of the first gene sequences to be reported mapped to the *Qa* region in BALB/c mice; *H-2Q7<sup>d</sup>* had eight exons but also had a stop codon which gave a truncated gene product (Steinmetz *et al.* 1981b). The intron/exon

configuration of this clone confirmed the assignment of class I protein structure (which consists of three external domains and a transmembrane and cytoplasmic region) see figure 4. Other genes sequenced included *H-2L<sup>d</sup>* (Moore *et al.* 1982) which was also found to encode eight exons. The fine structure of the sequenced clones showed that the eight exon format was very well conserved between class Ia genes and class Ib genes in mouse (Steinmetz *et al.* 1981b; Evans *et al.* 1982a; Moore *et al.* 1982) and man (Malissen *et al.* 1982; Geraghty *et al.* 1987) and can be dissected as follows (see figure 4):-

5'Untranslated region:- 25 bp in *H-2K<sup>d</sup>* (Lalanne *et al.* 1983), 20 bp in *H-2K<sup>d</sup>* (Kvist *et al.* 1983) normally 5'UT varies from 20-30 bp to allow for area between TATA box and translational start.

Exon 1:- The first exon gives rise to the leader peptide in the immature protein and varies in amino acid length from 21 a.a. in mice (Lalanne *et al.* 1982) to 24 a.a. in man (Srivastava *et al.* 1985) although exceptions include *HLA-A3* which has a leader peptide length of 29a.a. (Strachan *et al.* 1984).

Exons 2, 3 and 4:- Encode the three extracellular domains of the protein which are of length 90, 92 and 92 amino acids respectively. These appear to be very well conserved in length across species.

Exon 5:- This encodes the transmembrane region of the protein of length up to 40 amino acids.

Exon 6:- This exon is normally 11 amino acids long in class Ia genes, class Ib will be discussed below.

Exon 7:- Can vary in length from 13 a.a. (*H-2L<sup>d</sup>*, *H-2K<sup>d</sup>*) to 16 a.a. (*HLA-Cw3*).

Exon 8:- This appears to be the most variable exon in length minimum length is 1 a.a. in *H-2L<sup>d</sup>* and *HLA-A* genes up to 10 a.a. in

*H-2K<sup>d</sup>*.

3' Untranslated region :- can vary in class I genes from approximately 300bp up to 900bp for *Tla<sup>C</sup>*.

Most variation in the length of the molecule is generated by exons 6, 7 and 8. This includes variation in length by alternate splicing of exons (Steinmetz *et al.* 1981a; Archibald *et al.* 1986). Specific examples include those of the *H-2K<sup>q</sup>* genes (Kress *et al.* 1983b) and *H-2K<sup>b</sup>* where *in vitro* mutagenesis of the upstream acceptor site shortened exon 8 from 10 a.a. to 1 a.a. (Handy *et al.* 1988).

Other sources of variation in the length of class I molecules are due to the position of in frame termination codons. This is particularly applicable to *Qa* genes which all seem to encode an 8 exon gene but carry a termination codon in their fifth exon giving rise to truncated proteins (Steinmetz *et al.* 1981b; Mellor *et al.* 1984; Waneck *et al.* 1987; Robinson *et al.* 1988) in mice and *Qa*-like genes in man (Geraghty *et al.* 1987). *HLA-B* genes also have an 8 exon structure but have a stop codon in their seventh exon so are correspondingly three amino acids shorter than *HLA-A* and *HLA-C* gene products (Gussow *et al.* 1987).

Although the gene structure of class Ia genes and *Qa* genes is relatively conserved (see figure 4), this is not the case for the *Tla* genes which have been sequenced so far. In all *Tla* genes sequenced investigators have noted that the first five exons seem to follow the same format as class Ia, but that the sixth exon is much longer. Exon 6 encodes the entire cytoplasmic region of *Tla<sup>b</sup>*, which is two amino acids shorter than that of H-2L<sup>d</sup>, these regions show only 30% similarity (Fisher *et al.* 1985) and the presence of an in frame termination codon in exon 6 means that they are not present in the



protein. Similarly, in *Tla<sup>C</sup>* (*T13<sup>C</sup>*) exon 6 (1040bp) encodes cytoplasmic domain and has 900bp of 3'UT, again sequences similar to exons 7 and 8 in *H-2K<sup>b</sup>* are present but the similarity is very low (Obata *et al.* 1985). In contrast to these results (where the *Tla* genes quoted are more similar to each other than to either the class Ia genes and the *Qa* genes), Brown<sup>G.D.</sup> *et al.* (1988) sequenced and characterised *T13<sup>b</sup>*. *T13<sup>b</sup>* has a termination codon in exon 4, therefore most likely does not express membrane bound protein. The DNA sequence 3' to this is more similar to mouse class Ia genes than to *Tla<sup>b</sup>* and *Tla<sup>C</sup>*. This result suggests that a sub-set of *Tla* genes are more related to class Ia genes and may have arisen by gene conversion or duplication events.

The close examination of the gene structure of the class I genes shows a clear conservation of exon structure between class Ia and class Ib genes, <sup>indicating that</sup> they have probably evolved from the same progenitor gene (Klein 1986). Although it is clear that the class Ib genes seem to encode some unique differences, no function has yet been ascribed to the class Ib genes; this will be discussed in more detail below.

#### 1.2.1c6 Identifying functionally important regions in class I genes.

Once the fine structure of class I genes had been elucidated it was important to determine which were the areas that were needed for correct functioning of the genes in context of restriction, allograft rejection and antibody recognition. Only by combining the results from biochemistry (see section 1.2.1b) and molecular biology would the answers to these questions be obtained. Molecular biological examination consisted of two approaches; 1), analysis of mutants

arising *in vivo* and 2), generation of mutants *in vitro* by a) exon shuffling and b) targeted mutagenesis.

1.2.1c6a Contribution of *in vivo* generated mutants to mapping functionally important areas of the protein molecule.

Studies on *in vivo* generated mutants further substantiated the role of the first and second domains as being the functionally important elements of class Ia genes. The results of many studies on mutants discovered by reciprocal skin grafting and presented by Nathenson *et al.* (1986) are;

1) mutations occurred in the first two external domains [also see Nairn *et al.* (1980); Kimball and Coligan (1983) and Lopez de Castro *et al.* (1983)]

2) identical amino acid substitutions have been detected in altered  $K^b$  molecules of several independently arising mutant mice i.e.  $K^{bm6}$  and  $K^{bm9}$  are identical at nucleotide level.

3) many amino acid substitutions have involved multiple nucleotide alterations per codon.

4) substituted nucleotides and replaced amino acids present in mutant  $K^b$  genes are found at similar positions in other class I genes which suggests they may be donors for the event (an observation originally made by Evans *et al.* 1982a and Pease *et al.* 1983). Thus suggesting that genetic interaction between class I genes may be the driving force behind production of the  $K^b$  mutants and for the generation of diversity of H-2 genes (see section 1.2.1c7).

A similar pattern of amino acid substitution was also seen in human class Ia genes e.g. Vega *et al.* (1985) sequenced *HLA-B27.2* and

found it differed at amino acids 77, 80 and 81 in the first domain. Therefore analysis of mutants arising *in vivo* identified that the important regions for function of class I molecules were the first and second domains.

1.2.1c6b Contribution of exon shuffling experiments, in mapping functionally important areas of the molecule.

Recombinant DNA technology was used to dissect out the relevant functionally important exons of class Ia genes by recombining them in novel and unusual ways but retaining all the necessary information for correct expression. Exon shuffling experiments on *H-2D<sup>d</sup>* and *H-2L<sup>d</sup>* by Evans *et al.* (1982b), Ozato *et al.* (1983), Reiss *et al.* (1983) and Murre *et al.* (1984a), showed that exons 2 and 3 (see figure 4, these correspond to domains 1 and 2) were crucial for correct recognition by alloreactive T cells, while antibody determinants could map independently to any of the three external domains of the protein. Analysis of the interaction between the three external domains to determine if they combine stereospecifically to generate sites, showed that most recognition sites were contained on domains 1 and 2 and the interaction of these with domain 3 had no influence (Allen *et al.* 1984; Bluestone *et al.* 1985). Interaction of the three external domains was also investigated by Arnold *et al.* (1985), they concluded that the first domain of *H-2K<sup>d</sup>* could not function alone as a restriction element but needed a synergistic second domain.

The role of the 3' end of the protein was investigated by Murre *et al.* (1984b) who constructed an *H-2L<sup>d</sup>* gene lacking exons 6, 7 and 8 (cytoplasmic region) replacing these with 3 amino acids and the

polyA+ addition site derived from the *H-2A $\beta^d$*  gene. When the construct was transfected, expression could be detected on the cell surface and it was still recognised by alloreactive CTL's, and virus specific CTL's. All these experiments suggest that antibody targets are directed against individual domains, while alloreactive recognition and killing requires the first two domains of the molecule. These experiments also indicated that the third external domain was not crucial for correct functioning of these molecules.

1.2.1c6c Contribution of mutants created by targeted mutagenesis to mapping functionally important areas of the protein molecule.

Class Ia molecules could also be altered *in vitro* by targeted mutagenesis, therefore allowing an even more detailed examination of the functionally important areas. Simply replacing a.a. 63, 65 and 66 of *L<sup>d</sup>* with *D<sup>d</sup>* showed a gain of nine serological and cellular specificities and that an additional substitution at a.a. 70 conferred even more. The conclusion drawn from this was that a.a. 63-70 is immunodominant in *D<sup>d</sup>* (Koeller *et al.* 1987). Saturation mutagenesis of the first domain uncovered a crucial a.a. at position 27. A Tyr->Asn change completely destroyed all recognition by polyclonal alloreactive CTL (Murray *et al.* 1988). It was postulated that this residue must affect the conformation of the molecule because it is outside the recognised regions of variation described by *in vivo* studies.

Previously, all the evidence from the exon shuffling experiments placed the emphasis for the functionally important area of the

molecule in the first two domains. However, it is now apparent that by more detailed analysis of *in vitro* generated mutants that the third domain also has a crucial role in correct functioning of these molecules. Site directed mutagenesis of a.a. 227 Glu->Lys destroyed CTL recognition of *H-2D<sup>d</sup>* (Potter *et al.* 1987). It appears that the altered residue is recognised by CD8-independent T cells but not CD8-dependent T cells (Potter *et al.* 1989). CD8 is an accessory molecule which is a phenotypic marker of CTL and interacts with class I molecules. A variant of *HLA-A2.1* distinguishable by isoelectric focusing differs in the third external domain at amino acid 236 where a Glu has been substituted for an Ala, this difference gives a different CTL recognition profile (Castano *et al.* 1988).

The conclusions reached from studies on exon shuffling and mutagenesis suggest that the interaction between the first two domains is most important to give rise to conformational sites detected by allogeneic and restricted CTL. In summary; a), small changes in the first or second domains can cause significant alterations in recognition by antibodies or CTL b), the majority of monoclonal antibodies recognise discrete domain specific sites present in the first or second domains. These sites are localised or influenced by polypeptide stretches in the first domain from 70 to 90 and in the second domain from 150 to 180 c), CTL's recognise conformational determinants produced by interaction of a.a. residues located in sites in the first, second and third domains of class I heavy chain polypeptides.

1.2.1c7 What are the underlying mechanisms which may generate the observed polymorphism ?

Having identified that the three external domains were crucial for function, it was of considerable interest to uncover the underlying mechanism which generated polymorphism specifically in these first two domains. Various theories of alternate splicing as is observed in immunoglobulin genes (Tonegawa 1983) and earlier theories suggesting that the polymorphism was generated by controlling the level of expression (Bodmer 1973) had already been discounted. The first clues as to a possible mechanism were gained when detailed investigation of the nucleotide sequence of these genes emerged.

It was observed that the percentage divergence of exons 2 and 3 was 2-3 times greater than that of the introns (Weiss *et al.* 1983a) and that the sequence divergence of two *H-2K* alleles was the same as that between two non-alleles *H-2K<sup>b</sup>* and *H-2L<sup>d</sup>*. Therefore at the DNA level the non-alleles are as similar as alleles (Weiss *et al.* 1983a). This extreme polymorphism of the class Ia meant that at first there was no way that a new allele could be sequenced and automatically classified as to whether it was *K*, *D* or *L* in mice or *A*, *B* or *C* in man. However, it is now emerging that there are some locus specific amino acid residues which will allow assignment of specificity (Szots *et al.* 1986; Gussow *et al.* 1987). These seem to be concentrated to the transmembrane and cytoplasmic domains. Detailed comparison of the nucleotide sequences of class Ia genes reveals that they have a high incidence of replacement substitutions in exons 2 and 3 and a high degree of silent substitutions in exon 4 which was probably indicative of selective pressure to bind  $\beta_2$ -microglobulin

(N'Guyen *et al.* 1985). Therefore as would be expected, the second and third exons seem to be targeted as polymorphic areas of the gene.

Comparison of class Ia and class Ib nucleotide sequence reveals some important differences between them. Mellor *et al.* (1984) compared the *Q10* genes from *H-2Q* and *H-2B* haplotype mice and the comparable *H-2K<sup>q</sup>*, *H-2K<sup>b</sup>* and *H-2K<sup>d</sup>* in all combinations and found that for exons 2, 3 and 4 the class Ia genes displayed 6.8%-9.4% diversity while the *Q10* genes displayed 0.5% variation. Similarly, Lalanne *et al.* (1985) on sequencing the *Q10* molecule from DBA/2 mice and comparing to other strains found that all the differences were in the non-coding region. Therefore the class Ia genes encode much more polymorphism than the class Ib genes.

The underlying mechanism which was generating this diversity, but still obvious similarity, was looked for. The explanation favoured was one of gene conversion, where a unidirectional exchange of information occurs (Bregegere 1983). This was favoured due to the clustering of the nucleotide changes observed (refer back to 1.2.1c6a). Weiss *et al.* (1983b) sequenced the *H-2K<sup>b</sup>* and *H-2K<sup>bm1</sup>* genes and found that the differences between the two genes were over three codons, the differences seen in *H-2K<sup>bm1</sup>* were identical to that of *H-2L<sup>d</sup>* :-

	a.a.	152	155	156
<i>H-2K<sup>b</sup></i>	Glu	Arg	Leu	
	GAA	AGA	CTC	
<i>H-2K<sup>bm1</sup></i>	Ala	Tyr	Tyr	
	GCT	TAT	TAC	
<i>H-2L<sup>d</sup></i>	Ala	Tyr	Tyr	
	GCT	TAT	TAC	

They proposed that *H-2L<sup>d</sup>* was the donor for the gene conversion event. Unfortunately, this gene is absent from the haplotype that the conversion event was proposed to have happened in. However, a putative donor gene was satisfactorily identified by Mellor *et al.* (1983) who suggested the *Q10* gene had gone through unidirectional exchange with *H-2K<sup>b</sup>* in C57BL/10.

However, the role of the class Ib genes as a genetic pool which supplies DNA sequences in a unidirectional exchange with class Ia (Mellor *et al.* 1983; Bregegere 1983) has never been fully accepted. An alternative theory for the marked amount of polymorphism in class Ia genes has newly arisen. It appears that class Ia alleles are ancient and actually predate species divergence (Lawlor *et al.* 1988; Mayer *et al.* 1988). When comparing class Ia cDNA clones of chimpanzee with class I genes of man, 97.8% similarity was noted between *ChLA-A108* and *HLA-A11* (Mayer *et al.* 1988). This very high figure obtained from between species comparison suggests that this class I specificity originally arose in an ancestor of these two species. The small amount of polymorphism between the two species was localised to the first two external domains and in particular to the antigen binding site as proposed by Bjorkman *et al.* (1987a,1987b). However an explanation for the observed large number of alleles at class I loci is still needed. Hughes and Nei (1988) have proposed that the polymorphism is due to over-dominant selection (heterozygote advantage). They calculated the rate of nonsynonymous (amino acid altering) substitution pattern to that of synonymous substitution in the antigen binding site and compared it to that of the remaining amino acids in the first two domains. They found that nonsynonymous substitution was much higher in the antigen binding site, while in the



rest of the domain the reverse was true. Therefore the differences in class I proteins are localised to the functionally important areas of the molecule and probably accounts for the large numbers of alleles in the population via heterozygote advantage.

#### 1.2.1c8 A possible function for the class Ib genes?

The question still remains, do the class Ib genes have any function? As discussed above, the theory of class Ib genes acting as a genetic reservoir of information, to be used to generate polymorphism in class Ia genes, does not seem likely. It remains; that they are nonpolymorphic and probably derived from the same ancestral sequence as the class Ia genes. The documented limited tissue distribution of *Qa* genes (Flaherty 1980) now appears to be due to lack of sensitivity in detection, because RNA transcripts have been found in a wide variety of tissues (Palmer and Frelinger 1987). The limited tissue distribution of the *Tla* genes only appears to be relevant in certain situations. Expression of *Tla* can be induced on T-cells by activation with an alloantigen or a mitogen (Cook and Landolfi 1983). But, swapping 5' regulatory sequences of p20-TL (*Tla* gene) and *H-2K<sup>b</sup>* revealed that after transfection into L-cells the expression levels of the 5' TL+ *H-2K<sup>b</sup>* structural gene is markedly reduced when compared to normal *H-2K<sup>b</sup>* expression. The converse was true for the opposite construct of 5' *H-2K<sup>b</sup>* + TL structural gene (Obata *et al.* 1988). These results imply that *Tla* expression is regulated differently from class Ia expression.

A wider tissue distribution would provide a better base for a new theory which proposes that the class Ib genes act as ligands for a separate class of T-cell receptor. CTL's express receptors on their

cell surface which interact with class Ia molecules, triggering an immune response if the class I molecules are recognised as foreign. These receptors are composed of two chains; the  $\alpha$  and  $\beta$  which are rearranged somatically from a number of gene segments expressed in the T-cell (see review Toyonaga and Mak 1987). Another T-cell receptor has been identified; the  $\gamma\delta$ . The  $\gamma\delta$  receptor is expressed on T-cells which can be subdivided into three different populations (Janeway 1988) and seem to be specialised for immune surveillance of epithelial cells (Bonneville *et al.* 1988; Goodman and Lefrancois 1988). The  $\gamma\delta$  receptor has less combinatorial variation of variable gene segments than the  $\alpha\beta$  (Brenner *et al.* 1988; Davis and Bjorkman 1988) but would seem to have more junctional diversity (Brenner *et al.* 1988; Davis and Bjorkman 1988), suggesting that it may recognise a less polymorphic ligand which could bind antigen (Davis and Bjorkman 1988). Therefore, what is the ligand that this receptor may interact with? Janeway *et al.* (1988) propose that this receptor recognises class Ib molecules. Experimental evidence for this theory was provided by Matis *et al.* (1987). They were able to identify a T-cell line expressing the  $\gamma\delta$  receptor which had MHC linked recognition specificity and cytotoxicity. These functions mapped to the *H-2D*, *Qa*, *Tla* region of the *H-2<sup>b</sup>* haplotype, in addition to this, the cells also lysed *H-2<sup>q</sup>* haplotype cells so had a broad specificity. This result would be entirely compatible with the nonpolymorphic class Ib genes which show appreciable cross-reaction in cell transfection experiments; *Q7<sup>d</sup>* and *Q9<sup>d</sup>* on transfection are both recognised by Qa-2 antibodies (Soloski *et al.* 1988). Although this

possible function for class Ib genes awaits further substantiation, it is an extremely attractive proposition.

#### 1.2.1d The cellular biology of MHC class I function in relation to its structure.

Sections 1.2.1a, 1.2.1b and 1.2.1c have dealt with the immunological, biochemical and fine structural molecular characterisation of the MHC in mouse and man. As already alluded to in section 1.1; the pivotal role of the MHC is to allow an individual to distinguish self from nonself. This is achieved by recognition of class I molecules by T-cells via their T-cell receptor (TCR) triggering an immune response if they are recognised as being foreign. The interaction between T-cells and class I molecules is recognised to be fundamental to the activation and developmental selection of T-cells and will now be discussed.

#### Class I antigen presentation and recognition by T-cells.

Although there are many alleles of MHC class Ia genes in a population, an individual only has two-three functional class Ia genes. The major unanswered question, was how does the small number of MHC class I molecules expressed by an individual interact with the numerous antigens against which immune responses can be made. There are two ways this may be achieved; firstly, reducing the complexity of proteins by proteolytic processing, therefore only presenting short peptides for recognition and secondly, evolving an antigen binding site which can bind a wide variety of peptides with

different sequences and conformations. Townsend *et al.* (1986) were the first to demonstrate, that a prerequisite for MHC-restricted CTL recognition of class I involved antigen processing by the cell and that these antigens were 11-16 amino acids long. While, Bjorkman *et al.* (1987a,b) elucidated the three-dimensional structure of class I molecule HLA-A2 (discussed in section 1.2.1b3) and proposed a putative antigen binding site was formed from the first two domains of the protein. The ability of the antigen binding site to interact with numerous peptides was suggested by the fact that the amino acid polymorphism displayed around the binding site was mainly directed inwards towards the peptide binding groove. Kourilsky and Claverie (1989) noted more than a dozen residues were available to interact with peptide, but only two to four were available for interaction with the T-cell. These observations would fulfil the requirements for antigen processing and binding within a site which could accommodate numerous polymorphic antigens of different conformations.

Having discussed the presentation of antigen, it remains to discuss the recognition of this complex. Although Davis and Bjorkman (1988) have proposed a working model for recognition of the MHC-peptide complex by the TCR, it has been difficult to demonstrate an interaction between peptide and class I molecule, until now. Chen and Parham (1989) demonstrated binding by radiolabelling peptides and detecting low level binding to MHC class I molecules. While, Kane *et al.* (1989) addressed the question of T-cell recognition of the complex by showing that recognition depended on direct binding of influenza peptides to class I molecules. Overall, the interaction between antigen and class I molecule appears to be two-way. Townsend *et al.*

(1989) demonstrated that class I molecules fold up around peptides and by so doing promote their expression and association with  $\beta_2$ -microglobulin on the cell surface. Therefore antigen processing is extremely complex; involving binding of peptide and recognition of the bound complex by the TCR to trigger an immune response.

#### Thymic education of T-cells.

An immune response is only triggered if the complex is recognised as foreign, self antigens do not elicit a response. Therefore, T-cells have to be 'educated' to recognise antigens presented in the context of self MHC molecules, but not to recognise self antigens. This process is thought to occur in the thymus during ontogeny. There are two theories on the thymic education of T-cells, they either advocate positive or negative selection of responding T-cells, the available evidence favours both methods (Sprent *et al.* 1988). The positive selection model states that T-cells which recognise self are eliminated; T-cells which do not recognise anything are also eliminated or remain undetected; but T-cells which have intermediate avidity for self peptides and MHC molecules will be stimulated to proliferate. Recent data obtained from transgenic mice support the positive selection theory. Kisielow *et al.* (1988) introduced an  $\alpha\beta$  T-cell receptor (restricted to class I MHC) which conferred responsiveness to H-Y antigen as a transgene into the germline of mice, the specificity for H-Y antigen was deleted in males but not in females (only males express H-Y antigen). Similarly, Sha *et al.* (1988) introduced a TCR which recognised L<sup>d</sup> as a transgene into mice and found that T-cells in the periphery were deleted for the L<sup>d</sup>

specificity. These experiments only provide an answer for a small part of the puzzle and much remains to be done before T-cell education is completely understood.

### 1.2.2 Class II genes.

#### 1.2.2a Immunological characterisation of class II loci in mouse and man.

##### 1.2.2a1. Discovery of class II loci in mouse.

Concurrent with the serological characterisation of the class I molecules, other investigators were concentrating on culturing lymphocytes. They observed that co-culturing peripheral blood lymphocytes (PBL) from genetically different individuals led to proliferation of the cells. This technique was named mixed lymphocyte culture (MLC) (Bain *et al.* 1963) and the proliferation observed was named the mixed lymphocyte response (MLR). These proliferative effects were mapped in congenic strains of mice to the region between the class I regions *K* and *D* (Dutton 1966). The proliferative response measured between different populations of cells allowed a series of determinants to be characterised which were named lymphocyte antigen determinants (Lad), [also called lymphocyte defined loci (Ld)]. The mapping of these antigens to the region of the class I loci suggested that they could also mediate graft rejection (essentially MLR is an *in vitro* allograft rejection to tissue antigens which differ on the lymphocyte surface). Both class I and class II loci can elicit an MLR response (see review by Shreffler and David 1975; Klein 1986 pp 315-318). However the majority of

stimulation in an MLR is controlled by the *I* region and is due to genetic differences encoded within it (Bach *et al.* 1972; Meo *et al.* 1973). Once it became apparent that further loci mapped to this region, appropriate congenic strains were bred which differed only in these class II determinants. Serological characterisation of class II could only proceed with defined sera (see review by Shreffler and David 1975). Previous attempts at generating sera for class II specificities had been foiled because it was difficult to remove all cross-reacting anti-class I activity.

The class II I-region in mice was split into separate areas which were immunologically defined, they all mapped to the *I* region of mice and consisted of *A*, *B*, *C*, *E* and *J* (Shreffler and David 1975; Klein 1975). Only the *A* and *E* were serologically defined (Klein 1986) (see figure 1 for the location of these loci on the *H-2* map), the rest were identified as cell-cell interactions which seemed to give a differing response.

#### 1.2.2a2 Discovery of class II loci in man.

A similar stimulation between different cultured populations of lymphocytes, as had been observed in mice, was also observed in man (Bach and Amos 1967). The stimulation was due to differences in *HLA* between the cells (Yunis and Amos 1971). Serological definition of Ld determinants in man was more complex due to the lack of inbred lines! This was partly resolved by using B-cell lines from homozygous individuals as reference standards to absorb sera (see review by Winchester and Kunkel 1979) the determinants defined by these sera are named *DR*. Differences determined by MLC were classified as *D* specificities while *DR* specificities were those determined by

serology, [review of *D/DR* region in humans see Winchester and Kunkel (1979)]. Other loci were found to map centromeric to the *DR* locus and also behaved as cellular determinants (Tosi *et al.* 1978; Shaw *et al.* 1980). These are the *HLA-DQ* and *HLA-DP* loci. The *DQ* locus was detected by extra reactions in antisera which did not correspond to the *DR* loci. The *DP* locus has mainly been defined by strong secondary allogeneic proliferative response and cytotoxic response. These loci are indicated on figure 1.

#### 1.2.2a3 Immune Response genes of mouse.

Immune response genes were discovered in studies quite separate from serological and cellular studies. Levine, Ojeda and Benacerraf (1963) demonstrated that the ability to mount an immune response (characterised by detection of antibodies) to a synthetic polypeptide was inherited as a single Mendelian trait. Similarly McDevitt and Sela (1965) found direct genetic control of response to another different synthetic polypeptide in mice. This antibody response to simple synthetic polypeptide antigens was mapped to the region of the *K* and *D* class I loci by McDevitt and Chinitz (1969). At this point the issue then became confused, as to the function of these immune response (*Ir*) genes and to what they actually were. They were shown to be required for the specific activation of T cells (Benacerraf and McDevitt 1972) and to be needed for T and B cell cooperation [though only if autologous antigens were expressed in conjunction with them on the surface of immune cells could a significant antibody response be detected (Katz, Hamaoka and Benacerraf 1973)]. These findings led to the theory that the immune response genes were T-cell recognition structures (Benacerraf and McDevitt 1972) i.e. the T-cell



receptor. This was wrong.

Only once the theory of MHC restriction was proposed for class I genes by Zinkernagel and Doherty (1974b) was the available evidence for the Ir genes reinterpreted by Klein (1976). A unifying concept of MHC restriction emerged (Klein *et al.* 1981). They proposed that class II molecules provided the context for recognition of soluble antigen or natural proteins. For a response under Ir gene control, antigen is only recognised by T-cells in the context of some class II molecules and not in others. This is responsiveness and non-responsiveness.

#### 1.2.2b Biochemistry of class II molecules in mouse and man.

##### 1.2.2b1 Structure and characteristics of class II molecules.

As with class I genes, the mechanisms of how class II genes carry out their function could only be elucidated once they were biochemically isolated.

The class II antigens were purified from cells of the immune system on basis of reactivity with specific antisera once they had been solubilised in detergent (see review by Kaufman *et al.* 1984a). They were found to associate as a cell surface dimer which was integrally membrane bound. The dimer was composed of a heavy  $\alpha$  chain relative mass 33-35kd and a light  $\beta$  chain relative mass 27-29kd (Kaufman *et al.* 1984a) (see figure 2). The separate chains were both mapped to the *I* region of mice by Jones *et al.* (1978) and Cook *et al.* (1979) (unlike class I, where the light chain maps outside the MHC). Of the immunologically defined areas of the mouse *I* region

only two class II molecules were isolated, A and E. In man serology and biochemistry has identified four molecules; 2 DR , 1 DQ and 1 DP. These are equivalent to the mouse *E* and *A* loci respectively with no equivalent of *DP* ( see review by Kaufman *et al.* 1984a).

The domain structure was originally elucidated because the domain boundaries were the only sites accessible to proteolytic digestion (Kaufman and Strominger 1983). Combined amino acid sequencing studies and proteolysis of the dimeric complex showed that they were composed of four domains (Kaufman *et al.* 1984a). The first domain of the  $\beta$  chain ( $\beta$ 1) is 95 amino acids long, shows strong homology to immunoglobulin constant domains with cysteines at positions 15 and 79 which form a disulphide loop. The first domain of the class II  $\alpha$  chain is shorter being 85-88 amino acids and lacks cysteines but has an N-glycosylation site at residue 78 e.g. *DR $\alpha$*  and *E $\alpha$* . In human *DR* molecules the  $\beta$  chain has one glycosylated residue while the  $\alpha$  chain has two (Kaufman *et al.* 1980). The second external domain of both chains ( $\alpha$ 2 and  $\beta$ 2) are 94 amino acids long (Klein 1986). Following on from the external domains towards the carboxy terminus there is a connecting peptide of 13 amino acids in  $\alpha$  and 9 in  $\beta$  (Klein 1986) which leads into a hydrophobic transmembrane region of length 23 amino acids in both chains. This region finishes with a short hydrophilic peptide of 11-13 amino acids thought to associate with structures in the membrane which help anchor it there. The protein concludes with a hydrophilic tail which corresponds to the cytoplasmic region of the protein (Kaufman *et al.* 1984a). Therefore the described structures of class I and class II polypeptides are very

similar (see figure 2). They both comprise an external region, a transmembrane region and a cytoplasmic tail.

### 1.2.2b2 Polymorphism of the $\alpha$ and $\beta$ chains of class II molecules.

Biochemically detectable polymorphism of the class II region can be generated in a variety of ways. One, which is unique to class II molecules is trans-association of the separately encoded chains on the cell surface. Cook *et al.* (1980) showed this for *E* polypeptides where they isolated two different forms;  $E_{\beta}^b E_{\alpha}^d$  and  $E_{\beta}^d E_{\alpha}^b$ .

Similarly, Silver *et al.* (1980) when using an antibody specific to  $A_{\beta}^k$  precipitated a mixture of  $A_{\beta}^k$  and  $A_{\alpha}^k$  and  $A_{\alpha}^b$  class II molecules.

However, the two chains which combine to form class II molecules can differ in the amount of polymorphism they encode. Kaufman *et al.* (1980) noted that on tryptic digestion of class II proteins of man the heavy chains were very similar to each other but the light chains were very dissimilar to each other. While, Cook *et al.* (1980) showed that the A molecules of mice encoded variation in both  $\alpha$  and  $\beta$  chains. To determine if the  $\alpha$  or  $\beta$  chains encoded the functional polymorphism, immune precipitation studies were carried out. Antibodies recognised the  $\beta$  chain of A molecules in mice but not the  $\alpha$  chain. Therefore the alloantigenic specificities on class II A molecules are on the  $\beta$  chain (Silver *et al.* 1980). In man serological typing appears to concur with the polymorphism of the  $DQ_{\beta}$  chain, e.g. the serological type  $DQw2$  types as identical in two cell lines

although the DQ $_{\alpha}$  chain differs (Kappes and Strominger 1988). The exception to this is DQw1 where the  $\alpha$  chain determines serological type (Kappes and Strominger 1988).

Further localisation of the residues in class II molecules which encode the functional polymorphism has been done. The determining of peptide differences between a spontaneous mutant bm12 and the mouse haplotype it arose from, found the differences were localised to the  $\beta$  chain (McKean *et al.* 1981). The observed alteration in Ir-gene phenomena, MLR antigenic determinant and loss of alloantigenicity were therefore all due to a structural mutation on a class II molecule. This demonstrated that *Ir* genes and class II genes were one and the same. Conclusive proof of this was provided by experiments with transgenic mice (see section 1.2.2c6)

However, there are still many unanswered questions about class II molecules. Only two class II polypeptides had been sequenced by protein degradation (DR  $\alpha$  and  $\beta$ ), although immunological studies had shown that there are multiple loci. Most of the detailed information about class II molecules was obtained from applying recombinant DNA techniques to the analysis of the class II region and its products.

### 1.2.2c Molecular biological characterisation of class II molecules.

#### 1.2.2c1. Introduction.

As mentioned above only two class II proteins had been completely sequenced; the DR $_{\alpha}$  and DR $_{\beta}$ , which are both from man. Most information on class II molecules has come from the analysis of class

II cDNA clones (see Klein 1986). One of the first cDNA clones to be isolated corresponded to DR<sub>α</sub> which was selected by the multiple criteria of hybridising to B-cell RNA, translating to give the correct protein sequence in cell free translation experiments and once sequenced, the derived amino acid sequence showed homology to the DR<sub>α</sub> protein sequence (Lee *et al.* 1982). This and other cDNA clones were used in Southern blotting analysis of human genomic DNA to estimate the number of genes in the class II region and were used in large scale genomic cloning studies in both mouse and man.

#### 1.2.2c2 Southern blotting studies of class II genes.

From a variety of studies it was clear that the gene organisation of class II genes in man was more complex than that of mouse. Accumulated evidence suggested that there were multiple bands which hybridised with DR<sub>β</sub> (Wake *et al.* 1982) probably corresponding to three genes (Bohme *et al.* 1983). This is in sharp contrast to the pattern obtained with DR<sub>α</sub> which only hybridised to few bands (Lee *et al.* 1982; Wake *et al.* 1982 and Bohme *et al.* 1983). Analysis of DQ<sub>α</sub> chains showed that there were two α genes (Auffray *et al.* 1983a) and that they mapped to the D region of chromosome six (Auffray *et al.* 1983b). The numerous bands obtained however did not compare with the pattern seen for the class I genes (see section 1.2.1c2) implying that there are fewer class II genes.

### 1.2.2c3 Molecular cloning of class II genes.

Steinmetz *et al.* (1982a) used cDNA clones for human class II  $\alpha$  and  $\beta$  chains to isolate mouse cosmid clones carrying mouse class II genes. Using class II DR $_{\alpha}$ , DR $_{\beta}$ , and DQ $_{\beta}$  cDNA clones enabled them to do a chromosome walk of the *I*-region of BALB/c mice of similar type to that described for class I genes in section 1.2.1c3. They (Steinmetz *et al.* 1982a, 1986; Stephan *et al.* 1986) constructed a map of the region which is shown in figure 1. Other investigators built up this map to include the A $_{\beta 3}$  gene (Widera and Flavell 1984).

The DNA map of the class II region (see figure 1) does not correspond to the genetic map of the region. When using the same strategy which had been successfully employed for the class I gene map, the class II subregions *B*, *C* and *J* all had to be fitted into approximately 3.4kb of DNA (Steinmetz *et al.* 1982a). The existence of these various regions was already in question, as no corresponding protein had ever been isolated. Originally their existence had been invoked to explain immunological phenomena. In the case of the *B* region this has now been disproved (Baxevanis *et al.* 1981). The *C* sub-region, although still with doubts about its' existence may be explained by epistatic effects of the *A* and *E* regions (Murphy 1987). The *J* subregion has some very convincing evidence for its existence but was obviously not present in the *I* region of mice, although *J* effects all mapped to the E $_{\beta}$  gene. It is of note that a hotspot of recombination maps to the second intron of the E $_{\beta 1}$  gene (Steinmetz *et al.* 1984; Kobori *et al.* 1986) but it is not clear how this relates to the *J* subregion. Murphy (1987) proposes that *J* molecules are T-cell

receptors that recognise self-I-region products and/or that recognise receptors for self-I-region products, presumably by positive selection during ontogeny.

The human class II region gene linkage group order has been determined by pulse field gel electrophoresis (Hardy *et al.* 1986; Lawrence *et al.* 1987; Dunham *et al.* 1987 and Chimini *et al.* 1988) this is illustrated in figure 1. It is apparent that there are seven different subregions in the human class II region and accordingly many more genes than in the *I* region of mouse. Of these, the *DR* subregion seems to encode the most variation between haplotypes with different numbers of  $\beta$  genes (Kappes and Strominger 1988).

Of the other subregions, the *DQ* encodes the DQ antigen and another two genes ( $DX_{\alpha}$  and  $DX_{\beta}$ ) which do not appear to be expressed, although no obvious reason for this can be found (Kappes and Strominger 1988). The *DP* subregion encodes the DP antigen and another two pseudogenes. [The mouse homolog of the *DP* genes is  $A_{\beta 3}$  a pseudogene (Widera and Flavell 1984)]. Between these two regions lie the  $DZ_{\alpha}$  and the  $DO_{\beta}$  genes.  $DO_{\beta}$  is not inducible by  $\gamma$ -interferon (Tonnelie *et al.* 1985) like its mouse counterpart  $A_{\beta 2}$  (Wake and Flavell 1985). While  $DZ_{\alpha}$  (which encodes the chain postulated to associate with  $DO_{\beta}$ ) is inducible by  $\gamma$ -interferon (Trowsdale and Kelly 1985). Figure 1 shows the positioning of these genes on the DNA map of the class II regions of mouse and man.

#### 1.2.2c4 Polymorphism of class II genes at the fine structural level.

As in class I genes, clustering of amino acid differences are observed when alleles are compared, but these are all restricted to the first domain. DR<sub>α</sub> is virtually invariant, in DR<sub>β</sub> there are three major variable regions at 9-13, 25-38 and 67-74 (Kappes and Strominger 1988). While in DQ both chains are polymorphic. DQ<sub>α</sub> has a diverse region at 47-53, DQ<sub>β</sub> has diverse regions at 26-37, 52-57 and 70-74 (Kappes and Strominger 1988). Polymorphism of DP is much less than for DR and DQ. Only two different forms of DP<sub>α</sub> exist, DP<sub>β</sub> is more variable than this but the areas often have only two amino acid differences between alleles which does not really qualify as a variable region (Kappes and Strominger 1988). This observation probably explains why DP was mainly defined by secondary MLR typing.

The publication of the x-ray crystallographic structure of HLA-A2 protein and in particular, the elucidation of the putative antigen binding site (Bjorkman *et al.* 1987a, 1987b), has allowed a similar search for features in class II molecules. A hypothetical class II antigen binding site has been proposed by Brown *et al.* (1988), which involves residues 9-17, 28-29, 40-49 and 59-87. These residues correspond to previously identified polymorphic residues (Kappes and Strominger 1988) which are listed above and to functionally important amino acid substitutions (Brown *et al.* 1988).



1.2.2c5 Molecular analysis of the association of class II molecules with disease in man and mouse.

The MHC has been implicated in numerous diseases, either providing protection or rendering an individual more susceptible (for review see Klein 1986). Particularly strong is the association of the MHC with auto-immune disease. Studies of associations have been hampered by a lack of samples and the methods of analysis. One particularly new revelation is the application of the polymerase chain reaction (PCR), which involves using the Klenow fragment of DNA polymerase with synthetic oligonucleotide probes to produce large quantities of the specific target sequence *in vitro* (Saiki *et al.* 1985, 1986). Concentrating on the amino terminal domain of the protein which encodes the functional polymorphism, specific amplification of the DNA sequence using PCR and subsequent sequencing has allowed identification of a number of amino acids in this region which are specifically associated with auto-immune disease. Amino acid 57 of DQ $\beta$  is strongly correlated with resistance/susceptibility to insulin-dependent diabetes mellitus (IDDM), if alanine, valine or serine there is more susceptibility to the disease, if it is an asparagine there is increased resistance to the disease (Todd *et al.* 1987). The mouse model of IDDM also has amino acid 57 implicated in the etiology of the disease (Todd *et al.* 1988). Other associations of class II genes with auto-immune disease include DQw1.9 which is associated with susceptibility to pemphigus vulgaris (Sinha *et al.* 1988 and Todd *et al.* 1988).

1.2.2c6 Use of transgenic animals in elucidating the function of the MHC.

As already discussed throughout this section on the class II molecules, the exact relationship of the class II to the Ir genes had never been fully established to everyone's satisfaction. The correlation of immune response, alloantigenicity and differences in MLR to structural differences in class molecules, had been reported by McKean *et al.* (1981), but, integrally linking one to the other needed a conclusive experiment. This was provided by three different groups; LeMeur *et al.* (1985), Pinkert *et al.* (1985) and Yamamura *et al.* (1985). They each introduced the  $E_{\alpha}$  gene into the germline of mice which did not express this gene, in all of the mice cell surface expression of the  $E_{\alpha}E_{\beta}$  molecule was detected. In addition they showed that they had restored the mouse's ability to respond to a synthetic peptide. These experiments clearly illustrated that nonresponsiveness in these mice had been due to the lack of the  $E_{\alpha}$  gene. These transgenic mice were also used by Flood *et al.* (1986) where it was noted that  $E_{\alpha}$  gene expression influenced expression of the  $J$  locus. In mice lacking the gene no  $J$  expression could be detected, this was cited as evidence that the  $J$  locus may correspond to a T-cell receptor.

### 1.2.3 Class III genes.

#### 1.2.3a Immunological characterisation of class III loci.

##### 1.2.3a1 Introduction.

Apart from the class I and class II loci, there are additional loci which map to the MHC chromosomal region, these additional loci were called class III but this classification is very much in dispute (Klein 1986). The loci comprise components of the complement system and cytochrome P-450 genes. The complement loci which map to the MHC include *B*, *C2* and *C4* (see reviews by Alper 1980; Campbell *et al.* 1988 and Muller-Eberhard 1988). The cytochrome P-450 genes which map to this region encode 21-hydroxylase (*21-OH*) which is involved in steroid biogenesis (Klein 1986).

##### 1.2.3a2. Discovery of class III loci.

These loci were first detected by Shreffler and Owen (1963) who discovered serum serological (*Ss*). This locus segregated as a single gene, associated with the *H-2* complex (Shreffler and Owen 1963) and mapped between the *K* and *D* loci (Shreffler 1964; Shreffler and David 1972) but did not encode a transplantation antigen (Demant *et al.* 1973). *Ss* was hormonally regulated by testosterone, but not to the same degree as another protein, which was also linked to the *H-2* complex (Passmore and Shreffler 1970), but was only expressed in males called sex-limited protein (*Slp*) (Passmore and Shreffler 1971). Both *Ss* and *Slp* were secreted into the blood plasma and are now known to be two separate forms of serum complement *C4*, although *Slp* has no haemolytic activity and no demonstrable function

(Ferreira *et al.* 1978). *C2* (Gorman *et al.* 1980) and *B* (Roos *et al.* 1982) were also mapped to the region between *K* and *D* named the *S* region in mice (see figure 1 for the placement of these genes in the complex). Correct placement of the P-450 genes needed molecular characterisation of this region.

### 1.2.3b Molecular biological characterisation of class III genes.

The *21-OH* gene was mapped to the region 3' of the *C4* gene by Southern blotting (White *et al.* 1984a). White *et al.* (1984b) also provided convincing evidence that the gene was involved in steroid biogenesis thus explaining earlier linkage studies (Dupont *et al.* 1977). Structural mapping of this region of the MHC provided proof that the gene was indeed in the MHC region and that it had been duplicated in man (Carroll *et al.* 1985) and mouse (White *et al.* 1984a; Amor *et al.* 1985). In concert with the duplication of the 21-OH gene the *C4* genes have also been duplicated, both of these are active in man while only *C4A* is active in mouse (Shreffler 1982). cDNA clones have been isolated for *C2*, *C4* and *B* and used to characterise cosmid clones in man (Carroll *et al.* 1984). In mouse the arrangement of the complement genes was established by Chaplin *et al.* (1983). In mouse it appears that the second *C4* gene (*C4b* formally called *S/p*) is inactive and has had hormonal regulation imposed upon it by a retroviral insertion at its 5' end (Stavenhagen and Robins 1988). Other genes which were found to map to this region of the MHC in mouse were the genes for tumour necrosis factor  $\alpha$  and  $\beta$  (Spies *et al.* 1986), the equivalent genes in man were also mapped by Spies *et al.* (1989).

All these loci are indicated on figure 1 and represent the complete map of the MHC of mouse and man.

The argument for the renaming of these loci from class III genes has been strengthened by recent reports of the MHC of chickens (Guilletmot *et al.* 1988), where no complement encoding loci were found. This would substantiate Klein's reasoning that these loci have become entrapped and are not linked to the MHC because of involvement in cell lysis.

### 1.3 MHC genes of other species.

#### 1.3.1 Introduction.

Section 1.2 has discussed the MHC of mouse and man extensively. However, as already mentioned the phenomenon of graft rejection was thought to be peculiar to mice (Klein 1975). Only, once recognition of the importance of the MHC in immune surveillance (see Klein 1986) coupled with the discovery of similar systems in other species, was the true significance of the MHC realised. In strict chronological order, the *B* system of fowl was discovered before that of man (see Klein 1986) and there are now leukocyte antigen systems described in a variety of species. Only those most relevant to study of the bovine MHC (*BoLA*) will be discussed. See Klein (1986) for a fuller description of MHC of other species.

### 1.3.1a Molecular biological characterisation of MHC in other species.

#### 1.3.1a1 Introduction.

The contribution of recombinant DNA technology to characterisation of the MHC of other species has been varied in its application. With the rapid advance of the understanding of the MHC in mouse and man, other species have tended to lag behind. However, with application of the same techniques it is hoped that by comparison to identified features of the MHC in man and mouse that the knowledge of the MHC of other animals will be quickly advanced.

#### 1.3.1a2 Southern blotting studies on class I genes.

With limited serological characterisation of the class I region in most species, it seemed that the greatest amount of information can be gained quickly and easily by the application of Southern blotting. Consequently, hybridising heterologous and homologous probes (where available) to genomic DNA digests of cattle (Vaiman *et al.* 1986; Lindberg and Andersson 1988), sheep (Chardon *et al.* 1985b), miniature pigs (Singer *et al.* 1982), pigs (Chardon *et al.* 1985a; Flanagan *et al.* 1988), horses (Vaiman *et al.* 1986; Alexander *et al.* 1987), rabbits (Rebiere *et al.* 1987), chickens (Warner *et al.* 1986; Goto *et al.* 1988), rats (Palmer *et al.* 1983; Cortese Hassett *et al.* 1986) and cats (Yuhki and O'Brien 1988) revealed that these species have multiple class-I-hybridising sequences. A comparison of the number of class-I-hybridising sequences in different species is presented in table 1 (adapted from Yuhki and O'Brien 1988). From table 1, it seems that the numbers of class-I-hybridising fragments can vary considerably depending on the choice of restriction enzyme and the

species being investigated. Pigs and rabbits have fewer class I genes (Singer *et al.* 1982; Marche *et al.* 1985; Rebiere *et al.* 1987) than man or mouse. While the balkan mole rats and syrian hamster have numerous class-I-hybridising sequences (Nizetic *et al.* 1985, 1988; McGuire *et al.* 1985), both display a distinct lack of polymorphism (see table 1 and references above). The favoured explanation for this in these species, are they have both probably gone through a population bottle neck and have not had sufficient time to diversify.

In some cases the polymorphic class-I-hybridising pattern obtained seems to be in linkage disequilibrium with serological specificities. This has allowed investigators to identify particular hybridising class I bands or haplotypes (defined by sets of bands) as being indicative of serological types, allowing class I typing to proceed by DNA hybridisation alone e.g. cattle (Vaiman *et al.* 1986; Lindberg and Andersson 1988); sheep (Chardon *et al.* 1985b); pigs (Chardon *et al.* 1985a; Flanagan *et al.* 1988); and horses (Alexander *et al.* 1987). The polymorphic class-I-hybridising pattern can also be correlated to actual genetic regions of the chromosome e.g. rats (Cortese Hassett *et al.* 1986) if recombinant inbred lines are used. This is exactly the same strategy as employed for mice by Steinmetz *et al.* (1982b) and Weiss *et al.* (1984).

### 1.3.1a3 Molecular cloning of class I genes of domestic animals.

Although the MHC complexes of man and mouse are the best characterised and understood. Cloning of the MHC of other species will provide additional insights into the evolutionary relationships of these genes and expose any features which are atypical. Domestically

important species are particularly relevant in this type of approach because the results from investigation of smaller animals e.g. mice, rats and rabbits are not always directly applicable to larger animals. This is reflected in studies of MHC influence on disease resistance/susceptibility, because each species has its own particular disease repertoire.

### 1.3.1a3a Cloning of class Ia and class Ib genes.

Several class Ia genes from species other than man and mouse have now been cloned. These include two class I genes from pigs; PD1 (Singer *et al.* 1982) and PD14 (Satz *et al.* 1985). PD1 and PD14 show all the features attributable to class Ia genes and have been expressed in L cells (Singer *et al.* 1982; Satz *et al.* 1985). PD1 has been introduced into mice as a transgene (Frels *et al.* 1985) and shown to function in graft rejection. pR9 is a cDNA which has been cloned from a rabbit cell line RL-5 and conforms to typical class I structure (Tykocinski *et al.* 1984). Likewise the corresponding genomic clone displays the characteristic 8 exon structure of class I genes (Marche *et al.* 1985) previously discussed in section 1.2.1c5. While the class I gene Hm-1.6 cloned from syrian hamster has seven exons, with no discernible remnants of an eighth exon (McGuire *et al.* 1986) but still appears to encode all the features necessary to be functional.

All serological evidence to date supports the suggestion that only one locus is expressed in cattle; *BoLA-A* (see section 1.4.2a). In contrast, cDNA cloning of class I sequences recently presented by Ennis *et al.* (1988) reports cloning of two non-allelic clones from cattle. This is the strongest evidence to date for expression of more than one bovine class I locus (see section 1.4.2a below). Similar



reports on the MHC of rats (RTI) which quote difficulties in raising sera to loci other than *RTI.A* (Misra *et al.* 1987) (although there is also *RTI.E*), can now be explained by peptide mapping.  $A^U$  and  $A^N$  are very different while  $A^U$  and  $E^U$  are very similar (Misra *et al.* 1987). It appears that rat class I molecules are very similar within haplotypes but very different between them, explaining the difficulty in raising sera to loci other than *RTI.A*. Only one *RTI* class Ia cDNA clone has been published to date, but the clone is incomplete lacking a 5' end (Kastern 1985).

All the above sequences are most similar to class Ia genes of mouse and man. Now, putative class Ib genes have also been cloned from other species. A rat cDNA clone RTI.2 encodes a stop codon in exon 5 (Kastern 1985). This gives it the same structure as the *Qa* genes of mice and the one *Qa*-like sequence from man (Geraghty *et al.* 1987; Koller *et al.* 1989). The region of the rat MHC which the clone derives from is not known but a region (*RTI-C*) similar to the *Qa* and *Tla* regions of mouse exists in rat and has been shown not to restrict CTL (Gunther and Wurst 1984).

On a similar theme to that of rat, two unusual rabbit cDNA clones have been isolated from the same cell line as class Ia pR9. pR26 has a stop codon in frame within the leader peptide portion of the sequence it is therefore not clear if it is translated (Tykocinski *et al.* 1984). pR27 is unusually spliced and expressed in transformed T-cell lines and thymus as two large transcripts of 2.8 and 3.9kb (Rebiere *et al.* 1987). Horses have soluble class I antigens in their serum (Lew *et al.* 1986b) at levels comparable to that of mouse (Mellor *et al.* 1984; Devlin *et al.* 1985), but no class Ib sequence has been isolated to date. Rats also have soluble class I molecules in their serum, they arise

exclusively from the liver and kidney and are of relative mass 40kd (Spencer and Fabre 1987). They do not appear to be class Ib molecules because they all map to the *RT1.A* locus (Spencer and Fabre 1987)

#### 1.3.1a4 Southern blotting and cloning of class II genes.

As in studies of class I, the class II genes of other species have been extensively investigated by Southern blotting. A resume of the current numbers and their possible type are given in table 2. The underlying theme in most animals is that they are more similar to man than mouse in their arrangement and number of class II genes. This is particularly true of cattle, where the arrangement seems to mimic that of man (Andersson *et al.* 1988), with a recombination hot spot occurring within the class II region [M, A, C4, DRa DRb, DQa, DQb]\*[DYa, DYb, DOb ] (Andersson *et al.* 1988). The DY genes reported do not seem to correspond to any defined class II genes of man but cross react with DQ and DP probes in Southern blotting studies (Andersson *et al.* 1988) (see figure 1 for comparison of the arrangement of these genes with the MHC's of mouse and man).

Only five class II gene sequences have been reported; *RT1-A $\beta$*  from rats (Eccles and McMaster 1985), an *RLA-DQ  $\alpha$*  gene from rabbits (LeGuern *et al.* 1987), a class II  $\beta$  chain from chickens (Bourlet *et al.* 1988) and two from cattle; a DR $\beta$ -like pseudogene (Muggli-Cockett and Stone 1988) and a transcribed class II DR $\beta$ -like gene (Muggli-Cockett and Stone 1989).

### 1.3.1a5 Molecular cloning of MHC complexes of domestically important animals.

Large scale genomic cloning equivalent to experiments performed in mouse (Steinmetz *et al.* 1982a, 1982b; Weiss *et al.* 1984) has only been carried out in rabbits (Sittisombut and Knight 1986) and chickens (Guilletmot *et al.* 1988).

In rabbits the clones covered 300kb of DNA and in chickens the area cloned was 320kb of DNA. Neither of these are complete overlapping maps, but they both have yielded a great deal of information. In rabbits the hybridising genes separate into sub-regions as defined by differential hybridisation to class II probes from man, also the general arrangement of class II genes seems to be similar to that for man (Sittisombut and Knight 1986; Rebiere *et al.* 1987). The numbers of genes are given in table 2. The general conservation suggests that the array of sub-regions and similarity predate the divergence of man, mouse and rabbit. It appears to be a completely different story for chickens.

The chicken *B* complex comprises a class I region called B-F where the class I genes cloned by Guilletmot *et al.* (1988) mapped, there is also the B-L region which encodes the class II molecules and a B-G which encodes erythrocyte specific class IV molecules peculiar to the chicken (Pink *et al.* 1977). Guilletmot *et al.* (1988) identified six class I genes, five class II genes and seven which did not fall into any category of MHC gene. They found that the gene size was much smaller than that of man and mouse and that the class I and II genes were scattered amongst each other and did not fall into separate regions as is seen in man and mouse. The chicken class I protein B-F12 showed 43% amino acid similarity to HLA-B27 (Guilletmot *et al.* 1988) while

the chicken class II  $\beta$  gene sequenced showed 63, 66 and 62% similarity to the second, third and fourth exons of  $DQ\beta$  (Bourlet *et al.* 1988). Guilletmot *et al.* (1988) were not able to identify any class II  $\alpha$  genes by transfection of cosmids into cells or any class III genes by hybridisation. The failure to discover class II  $\alpha$  genes suggests that the *B* complex is arranged differently, the  $\alpha$  and  $\beta$  genes are not present as tandemly duplicated pairs. Likewise, the failure to discover any class III genes suggests the current classification of these as MHC class III genes may well be wrong (as argued by Klein 1986). It seems they may have been inserted there at a later date, after the divergence of birds from mammals which was estimated to have occurred 250 million years ago.

The question remains as to the identity of the 7 unassigned genes that Guilletmot *et al.* (1988) characterised. They could be candidates for genes which manifest physiological effects that have been mapped to the MHC. The *Ped* gene which maps to the *Qa-2* subregion of *Qa* and controls pre-implantation embryo development in mice has two alleles, a slow and a fast which when expressed on pre-implantation mouse embryos seem to influence the rate of cleavage (Warner *et al.* 1986). Previous observations that embryos slow to implant are lost would suggest that this could be of commercial significance in domestically important species. A report of MHC influence on the growth of eggs in chickens and embryo survival in pigs would indicate the presence of a similar gene in these animals (Wilmot *et al.* 1985). These genes could also influence disease resistance. The *B* region of chickens seems to encode differential resistance to a number of viral diseases. Genetic resistance to Marek's disease in chickens is



associated with the possession of  $B^{21}$  (Padzerka *et al.* 1975; Longenecker *et al.* 1977; Briles *et al.* 1977) while genetic resistance to fowl cholera is associated with  $B^1$  (Lamont *et al.* 1987).

## 1.4. The bovine MHC: *BoLA*.

### 1.4.1 Introduction.

Having discussed the human and mouse MHC extensively and having indicated the current molecular biological knowledge of the MHC's of other species, it now seems appropriate to introduce the elucidation of the bovine MHC. The serological and immunological knowledge of *BoLA* is scant when compared to that of man and mouse, but a general theme has emerged; *BoLA* is also extremely polymorphic and displays most of the features associated with these other species. How this conclusion was reached will be discussed below.

### 1.4.2 Discovery of class I loci in cattle.

#### 1.4.2a Immunological discovery of class I loci.

The class I region of *BoLA* named *BoLA-A* was identified by Amorena and Stone (1978) and Spooner *et al.* (1978) by use of serology. Sera obtained from multiparous cattle, by alloimmunisation and by separation from colostrum were used to identify lymphocyte determinants. There are now 34 serologic specificities identified which all map to the one locus *BoLA-A* (Bull *et al.* 1989). Therefore the class I region seems to encode one locus which is extremely polymorphic (as would be predicted for a MHC specificity). Some of

these serological specificities can be sub-divided by utilisation of different detection techniques. The *W6* specificity can be split serologically into two sub-groups *W6.2* and *W6.4* (Spooner and Morgan 1981), likewise previously separate specificities are found to form a cross reactive group e.g. *W4*, *W7* and *W10* (Spooner 1986). Population studies on *BoLA-A* reveal differences in the frequency of *BoLA-A* specificities in different breeds of European cattle (Oliver *et al.* 1981) and in different breeds throughout the world (Spooner *et al.* 1987b; Kemp *et al.* 1988). As yet serology has provided no conclusive evidence for the existence of any other class I loci (this may be because the only locus detected by sera is the most antigenic one). Other methods of detecting *BoLA-A* have included the use of alloreactive T cells which seem to confirm the serological groupings (Teale *et al.* 1986a; Spooner *et al.* 1987a).

#### 1.4.2b Biochemical description of class I molecules.

The molecular size of class I molecules in cattle was determined by precipitation with immune sera and analysis on SDS-polyacrylamide gels (Hoang-Xuang *et al.* 1982b). The size of 44kd appears to be similar to that described for other species. Biochemical evidence for the existence of more than one class I locus has been gained by sequential precipitation with sera. Analysis of the precipitate on gels seems to indicate that two distinct species of chain can be identified and that up to five different molecules are expressed on the lymphocyte surface (Bensaid *et al.* 1988) [compare this to the cloning of two class I cDNA clones by Ennis *et al.* (1988) discussed previously].

### 1.4.3 Class II loci in cattle.

#### 1.4.3a Immunological detection of class II loci in cattle.

##### 1.4.3a1 Discovery of class II in cattle.

Class II region of *BoLA* named *BoLA-D* was described by Usinger *et al.* (1977) as comprising of at least two polymorphic loci which could cross stimulate in MLR reactions. The linkage of the class I and class II regions was proven by studies on full sib families (Spooner *et al.* 1978; Usinger *et al.* 1981). Further characterisation of the class II region has proceeded with the use of monoclonal antibodies (mAb). mAb's raised in other species do cross react with cattle (Spooner *et al.* 1983; Lewin *et al.* 1985), and specific mAb's to cattle antigens have been raised (Letesson *et al.* 1983; Lalor *et al.* 1986). However, more successful characterisation of class II has proceeded with alloreactive bovine T-cell clones (Teale *et al.* 1986b). In addition, their subsequent use in detecting class II polymorphism (Teale *et al.* 1987) will probably help to determine the functionally important epitopes on class II molecules.

##### 1.4.3a2. Immune-responsiveness in cattle.

Characterisation of immune-responsiveness in cattle is at an early stage. To date, differences have been noted in generating an immune-response to horse serum albumin (HSA). The class I specificity *W16* was correlated with a high response to HSA while *W2* seemed to correlate with low responsiveness (Lie *et al.* 1986). *In vitro* T-cell proliferative responses have been measured in cattle (Glass and Spooner 1989). They show a bimodal response in unrelated cattle, which in half sib groups segregates with MHC haplotype (E. J. Glass personal communication).

#### 1.4.3b Biochemistry of class II molecules.

##### 1.4.3b1. Biochemical characterisation of class II molecules in cattle.

Huang-Xuang *et al.* (1982a) successfully isolated class II chains by using human mAb which cross reacted with cattle class II molecules separated by 2-D gel analysis. The sizes obtained of 27kd and 34kd correspond to those reported for man and mouse.

#### 1.4.4 Disease resistance/susceptibility studies in cattle.

Previously serologically defined bovine class I products have been shown to be the principle recognition structures for alloreactive bovine T-cell lines and clones (Teale *et al.* 1986a; Teale *et al.* 1987; Spooner *et al.* 1987). Distinct evidence for class I restriction of CTL was obtained by investigation of the immune response to *Theileria parva*. This is a parasite which invades and transforms lymphocytes and elicits expression of infection specific antigens on the cell surface, these are detected as foreign and generate a strong CTL response (Eugi and Emery 1981). This can be inhibited by blocking with w6/32 (Goddeeris *et al.* 1986) which reacts with a conserved determinant on class I molecules (Chardon *et al.* 1983). It has been noted that this strong response in cattle can be biased to certain class I specificities (Morrison *et al.* 1986).

A more convincing association of *BoLA* and disease resistance was obtained by investigation of bovine leukaemia virus (BLV). The first clinical signs of infection by BLV are the development of persistent lymphocytosis (a polyclonal expansion of BLV- infected peripheral B cells). It was noted that *W8.1* was negatively associated with the



presence of antibodies to BLV and that *W12.1* was associated with increased susceptibility to polyclonal expansion of B cells (Lewin *et al.* 1988).

Figure 1. Genetic and DNA maps of the *H-2*, *HLA* and *BoLA* major histocompatibility complexes.

This figure portrays the genetic and DNA maps of the *H-2*, *HLA* and *BoLA* major histocompatibility complexes. The *BoLA* complex has been included purely for comparison, none of the genes indicated have been physically mapped, their existence has been implied from Southern blotting studies (see section 1.3.1a4).

The *H-2* complex is shown with the genetic map above the DNA map. Breaks in the DNA map indicate that these regions have not been linked by overlapping genomic clones. The map was compiled using information from Stephan *et al.* (1986) and Muller *et al.* (1987).

The *HLA* complex is also shown with the genetic regions above the DNA map. All the distances are approximately correct, the only exception being the loci which map to the right of *HLA-B*. All loci which are indicated to the right of *HLA-B* have been shown to be linked and to occur in this order. However, their actual physical relationship to each other and the exact distance that separates them is not known at this time. The overall layout of the *HLA* complex was determined by pulse field gel electrophoresis (Hardy *et al.* 1986; Dunham *et al.* 1987; Lawrence *et al.* 1987 and Chimini *et al.* 1988). The map was compiled using information from Kappes and Strominger (1988) (class II), White *et al.* (1984) and Dunham *et al.* (1987) (class III) and Lawrence *et al.* (1987) (RS5 clone), Spies *et al.* (1989) and Koller *et al.* (1989) (class I).



Figure 2. Domain structure of class I and class II proteins.

Figure 2, redrawn from Kappes and Strominger (1988), depicts the domain structure of class I and II proteins. The separate domains are labelled according to nomenclature. S-S indicates a disulphide bond formed between two cysteines. •- indicates where carbohydrate side chains are added onto the amino acid backbone, those which are striped are not conserved across species and to date have only been observed in mouse (see section 1.2.1b1).

Figure 2

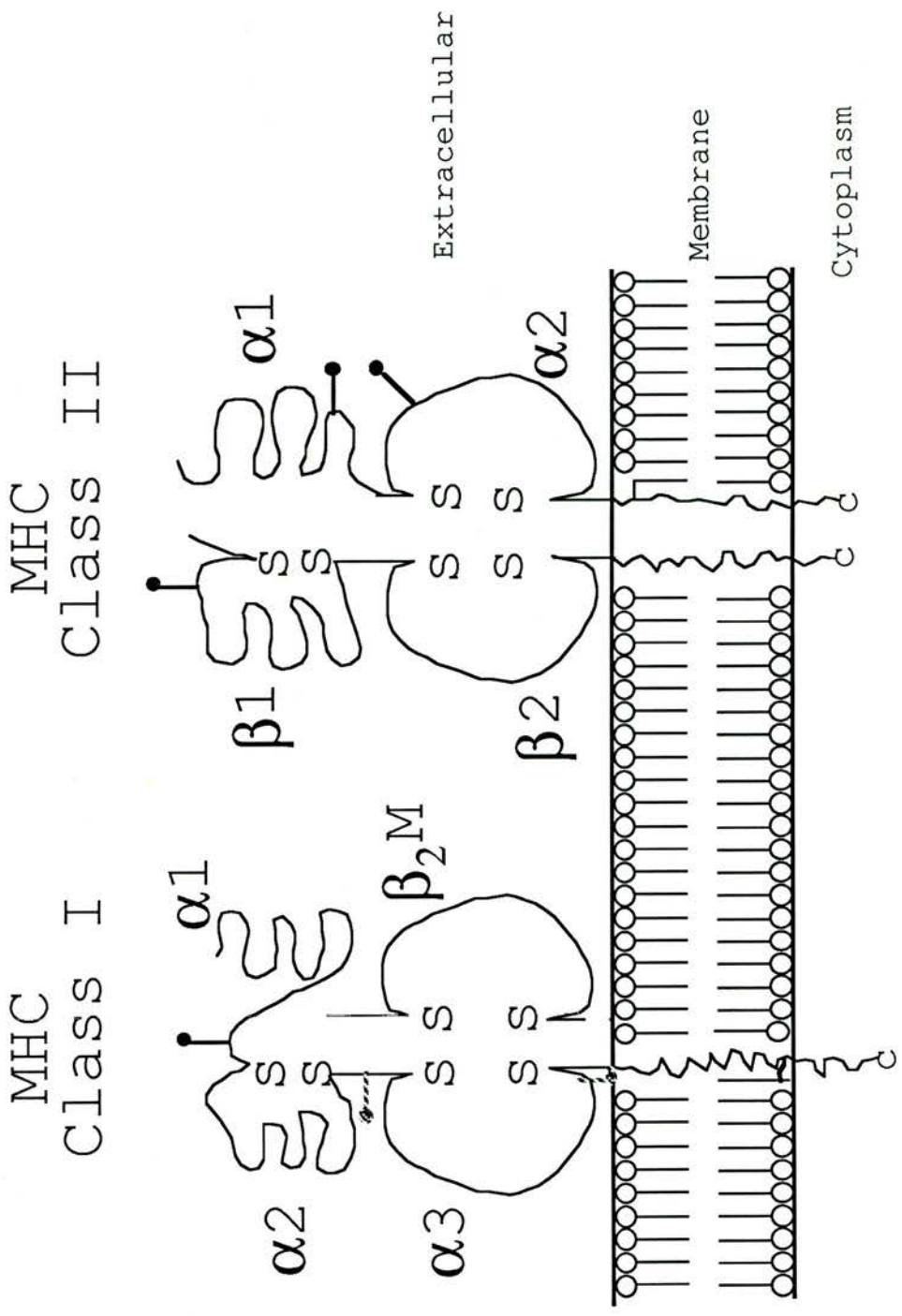


Figure 3. The 3D structure of the first two domains of HLA-A2.

Figure 3, redrawn from Bjorkman *et al.* (1987b), depicts the folding pattern of the first two domains of the class I molecule. Relevant amino acids are numbered. Figure 3a depicts the  $\alpha$ -helical regions determined by Bjorkman *et al.* (1987a, b) which occur at residues 50-55, 58-84, 138-148, 151-161, 163-173, and 175-180.

Figure 3b shows the same structure with the regions identified by Nathenson *et al.* (1981) (22-24, 30-32, 41-45, 61-83 and 95-99); Lopez de Castro *et al.* (1982), (65-80, 105-116 and 177-194); Krangel *et al.* (1983), (147-157) and Kimball and Coligan (1983), (9, 30-32, 62-83, 95-121 and 152-157) depicted as shaded areas.

Figure 3a

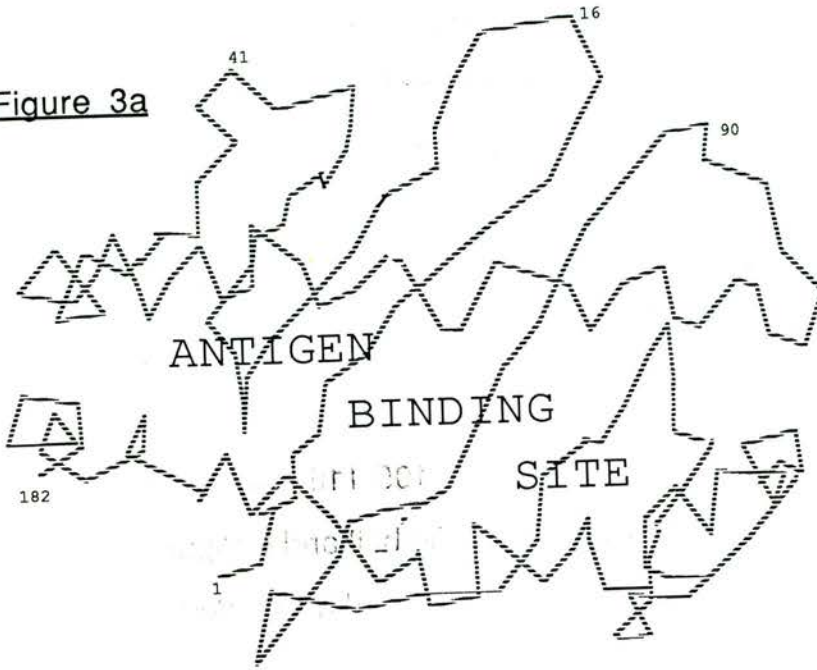


Figure 3b

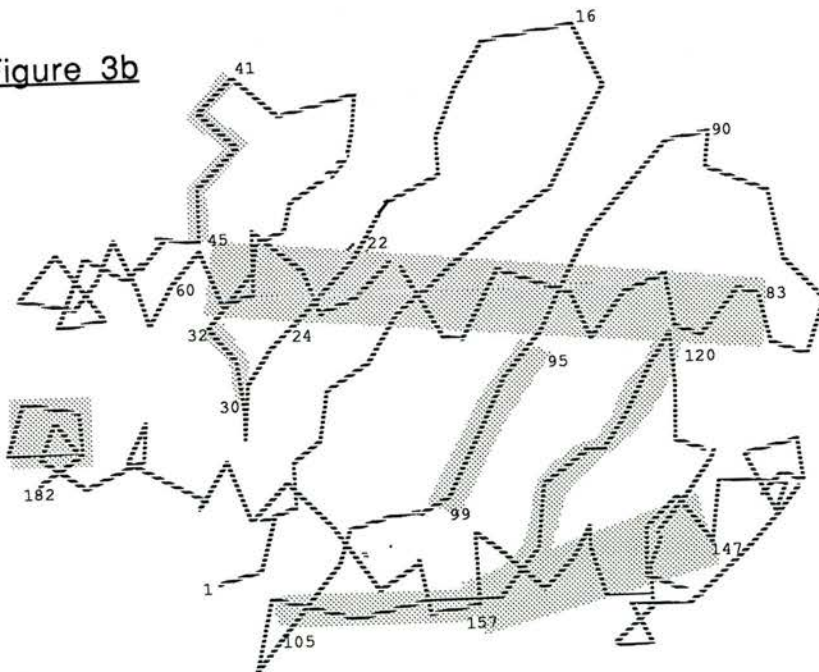


Figure 4. Fine structure of a class I genomic clone.

The intron/exon configuration of *H-2Q7<sup>d</sup>* redrawn from Steinmetz *et al.* (1981b) is indicated. The exons are solid black boxes with the 3' untranslated region shown as cross hatched. How these exons correspond to the amino acid sequence of H-2K<sup>b</sup> is also indicated. Amino acids 340-346 do not seem to have an equivalent region in the genomic clone.



Figure 4

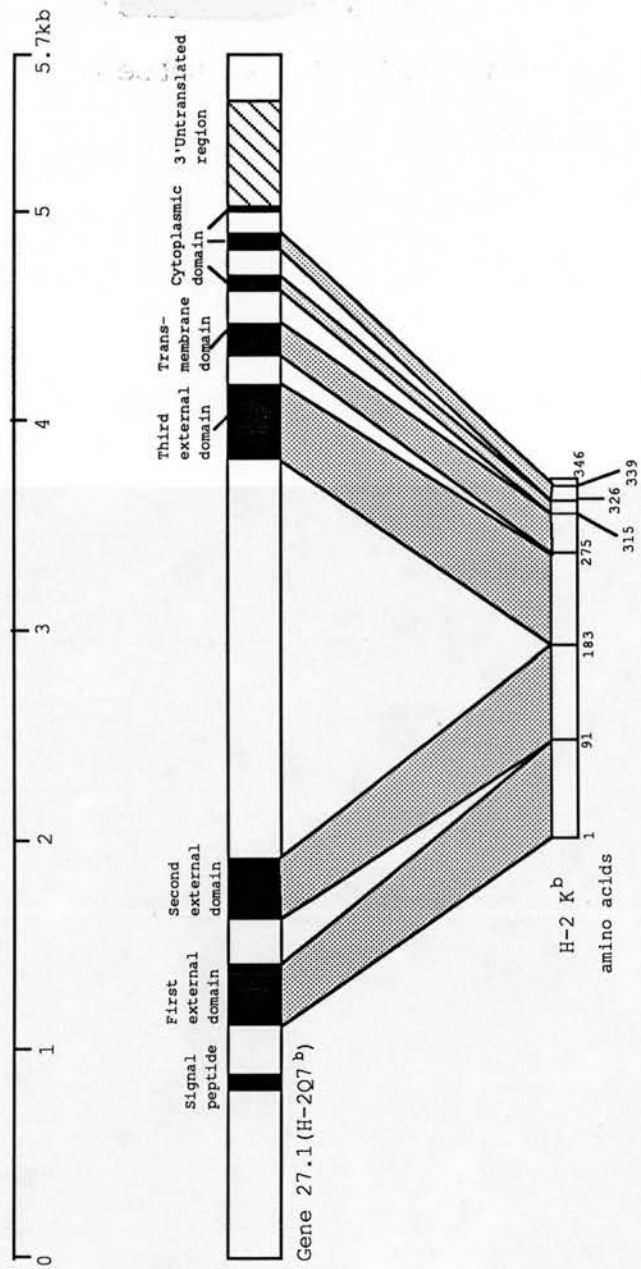


Table 1. A comparison of the numbers of hybridising bands detected when probing genomic Southernns with class I probes.

The table compares the numbers of fragments obtained when probing genomic DNA with class I MHC probes. The DNA was cut with the restriction enzyme indicated, fractionated on an agarose gel and transferred to a solid support before probing with a radioactively labelled class I probe. Each column is headed by the species which the experiments have been done in. The figures represent the number of bands seen per animal, with the figures in brackets detailing how many of these bands were polymorphic i.e. not shared by every animal. The references the results were taken from are indicated below the table. ND= not determined.

Table 1

Class	Restriction enzyme	Species											
		Cattle	Sheep	Pig	Horse	Cat	Man	Mouse	Rat	Mole rat (Israeli)	Mole rat (Balkan)	Syrian Hamster	Rabbit
I	<i>Fst I</i>	ND	ND	ND	ND	14-17 (7)	ND	16-17 (18)	ND	ND	16-17 (1)	ND	ND
	<i>Bam HI</i>	ND	ND	4-7 (7)	ND	10-14 (5)	ND	17-18 (14)	13-20	10-14 (1)	15-16 (1)	13	ND
	<i>Eco RI</i>	13-14 (16)	16 (8)	7-10 (22)	20-25	7-9 (4)	11 (1)	11-15 (11)	10-20	12-17 (16)	16-17 (1)	18	ND
	<i>Eco RV</i>	ND	ND	ND	ND	7-8 (2)	12-13 (6)	11-13 (5)	ND	ND	ND	ND	ND
	<i>Hind III</i>	ND	ND	6-11 (11)	24-33 (16)	9	20-22 (11)	ND	ND	ND	ND	ND	ND
	<i>Pvu II</i>	10-20 (15)	ND	7-14	23-30	ND	ND	ND	ND	ND	ND	ND	8-13 (7)

- ① Valman et al. (1986)
- ② Lindberg and Andersson (1988)
- ③ Chardon et al. (1985b)
- ④ Chardon et al. (1985a)
- ⑤ Alexander et al. (1987)
- ⑥ Yuhki and O'Brien (1988)
- ⑦ Cohen et al. (1983)
- ⑧ Data from AKR and BALB/c mice (see Yuhki and O'Brien 1988)
- ⑨ Palmer et al. (1983)
- ⑩ Nizetic et al. (1985)
- ⑪ Nizetic et al. (1988)
- ⑫ McGuire et al. (1985)
- ⑬ Reblere et al. (1987)

Table 2. Comparison of the number of class II genes in different species as detected by direct cloning or deduced by Southern analysis of genomic DNA.

This table shows the numbers of class II genes in different species as deduced by either Southern blotting analysis or molecular cloning. The class II nomenclature is that adopted for humans enabling the different species to be compared to mouse and to each other. The appropriate references are indicated in the right of the table.

Table 2

	DQ		DR		DP		DO $\beta$	DZ $\alpha$	
	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$			
CATTLE	1-2	1-2	1	>3	?	?	1	1	Andersson and Rask (1988); Andersson <u>et al.</u> (1988)
SHEEP	>3	>2	1	>2	?	?	1	1	Scott <u>et al.</u> (1987)
PIGS	1	>2	1	3	1	?	1-2	1	Sachs <u>et al.</u> (1988)
RABBITS	1	1	1	>3	2	1	?	1	LeGuern <u>et al.</u> (1985) Sittisombut and Knight (1986)
MAN	2	2	1	1-4	2	2	1	1	Rask <u>et al.</u> (1985); Trowsdale <u>et al.</u> (1985); Hardy <u>et al.</u> (1986)
MOUSE	1	1	1	2	?	1	1	?	Kobori <u>et al.</u> (1986); Steinmetz <u>et al.</u> (1986)
CAT	?	?	1	1	1	?	1	?	Yuhki and O'Brian (1988)
DOG	?	2	?	2	?	2	1	?	Sarmiento and Storb (1988)

### Aims and approaches.

On commencement of this project, there had been very little molecular genetic analysis of the bovine MHC. Immunological analysis had identified one class I locus and had indicated the presence of class II loci, but with no characterisation.

The main thrust of this project was to apply recombinant DNA techniques to analysis of the class I loci in cattle. In addition, the intention was to investigate the use of DNA probes for *BoLA* typing. It was hoped that isolating class I cDNA clones would identify functional genes, these would then be used to isolate genomic clones. The analysis of these would indicate how the bovine MHC relates to other species. The genes would then be used as probes to dissect the complicated multi-band pattern obtained when hybridising to Southern blots. Lastly, the aim was to establish DNA typing of *BoLA*, rather than serological typing. This was to be achieved by looking for RFLP's, after hybridisation of Southern blots with class I probes, which correlated with the animals *BoLA* type.

## Chapter Two.

### Materials and Methods.

#### 2.1 Sources of deoxyribonucleic acids and ribonucleic acids.

Deoxyribonucleic acid (DNA) was prepared from bovine peripheral blood lymphocytes (PBL) separated from whole blood. Blood was collected in either 7ml or 20ml evacuated tubes containing heparin sodium diluted in 0.9% NaCl to a final concentration of 2-3 units/ml blood. The average yield of DNA was 50 $\mu$ g/ml blood.

Ribonucleic acid (RNA) was either prepared from bovine PBL or liver. Whole blood was collected in 1 litre evacuated bottles containing 0.2 volumes of acid citrate dextrose. Liver was frozen as small individual pieces in liquid nitrogen. PBL yielded 200 $\mu$ g total RNA from  $1 \times 10^9$  cells, while tissue realised 5mg/g. Total RNA was also extracted from C603 a *Theileria parva* transformed bovine lymphoblastoid cell line (kindly supplied by Dr A. J. Teale, ILRAD, Kenya) the yield obtained was 600 $\mu$ g from  $6.7 \times 10^7$  cells.

#### 2.2 Recombinant plasmid clones.

pHLA-27; a human class I MHC cDNA of length 1200bp, supplied by Dr Per Peterson, Uppsala, Sweden.

#### 2.3 Non-recombinant vectors and bacterial host strains used.

pPoly1; 2068bp plasmid, carrying ampicillin resistance (Lathe *et al.* 1987).

pUC18; 2686bp plasmid, carrying ampicillin resistance

(Yanisch-Perron *et al.* 1985).

lambda gt10 (Huynh, Young and Davis 1986) was the lambda phage vector used to construct the cDNA library.

EMBL301 was the phage vector used by Dr Jean-Luc Vilotte to construct the bovine genomic library (Lathe *et al.* 1987).

M13tg130 and M13tg131 were the phage vectors used for generating sequencing templates (Kieny *et al.* 1983).

### 2.3a Bacterial host strains and genotypes.

**JM83** (*ara*,  $\Delta$ (*lac-pro*), *strA*, *thi* I, 80d*lac* I<sup>q</sup> Z  $\Delta$ M15)

**DH1** [*F*<sup>-</sup>, *rel* A1, *end* A1, *gyr* A96, *thi* -1, *hsd* R17(*r*<sup>-</sup><sub>k</sub>, *m*<sup>+</sup><sub>k</sub>), *sup*E44, *rec*A1]

**JM101** ( $\Delta$  (*lac-pro*), *sup* E44, *thi* 1, *F'* *tra* D36, *pro* AB, *lac* I<sup>q</sup>, Z $\Delta$ M15)

**BNN102** (*hsd* (R<sup>-</sup>, M<sup>+</sup>), *sup* E, *thr*, *leu*, *thi*, *lac* Y1, *ton* A21, *hfl*A150 [*chr*::Tn10])

### 2.4 Enzymes, antibiotics and specialised chemicals.

Amersham:- <sup>35</sup>SdATP 3000ci/mmol,  $\alpha$ <sup>32</sup>PdCTP 410ci/mmol and 3000ci/mmol,  $\gamma$ <sup>32</sup>PdATP 3000ci/mmol, T4 DNA ligase, ØX 174, Hybond-N.

Becton Dickinson:- BBL trypticase.

Biolabs:- *E.coli* DNA ligase, 12 mer phosphatased *EcoRI* linkers

Biorad:- ultra pure DNA grade agarose, low gelation temperature agarose

Boehringer Mannheim:- dNTP's, ddNTP's, DNA polymerase I,  $\beta$ -NAD, RNasin, T4 polymerase, T4 polynucleotide kinase, RNase A, wt lambda, proteinase K, X-GAL, IPTG, calf intestinal phosphatase (CIP) and



caesium chloride (CsCl).

Difco:- bactotryptone, yeast extract.

Fluka:- Guanidinium thiocyanate, guanidine hydrochloride.

Life Sciences:- X. L. reverse transcriptase.

Pharmacia:- Oligod(T) cellulose type 77F, oligo d(t)<sub>12-18</sub>, universal primer M13, Klenow fragment DNA polymerase I, random digest of calf thymus DNA into oligonucleotides pd(N)<sub>6</sub>

Schleicher and Schuell:- Nitrocellulose membranes, Elutip-d.

Serva:- acrylamide, N,N', methylene-bisacrylamide

Sigma:- diethylpyrocarbonate (DEPC), type 1 agarose, polyvinylpyrrolidone, ficoll 400, TEMED, sodium pyrophosphate, sodium dodecyl sulphate (SDS), ampicillin, tetracycline, chloramphenicol, antifoam A and thiamine.

Stratagene:-  $\lambda$ gt10 phosphatased arms, Gigapack *in vitro* phage packaging kit.

Worthington Biochemicals:- Deoxyribonuclease I.

All chemicals were analytical grade and supplied by British Drug Houses (BDH) unless otherwise stated.

### 2.5 Media.

LBM plates were poured using 10g bactotryptone, 5g yeast extract, 5g NaCl, 15g agar per litre, with 10 mls of 1M MgSO<sub>4</sub>7H<sub>2</sub>O being added just before pouring. Top agar overlays were poured from 10g bactotryptone, 5g yeast extract, 8g agar or 7g agarose, 5g NaCl per litre and 10 mls 1M MgSO<sub>4</sub>7H<sub>2</sub>O added just before pouring.

BBL plates consisted of 10g BBL trypticase, 5g NaCl, 10g agar per

litre. BBL top was made from 6.5g agar, 5g NaCl and 10g trypticase per litre.

The bacterial strain JM101 was maintained on minimal medium plates, consisting of 0.2 volumes 5xM9 salts (10mM NaCl, 60mM Na<sub>2</sub>HPO<sub>4</sub>, 27.5mM KH<sub>2</sub>PO<sub>4</sub>, 25mM NH<sub>4</sub>Cl), 0.001 volumes 1M MgSO<sub>4</sub>, 0.001 volumes 0.1M CaCl<sub>2</sub>, 0.001 volumes 10mg/ml thiamine (vitamin B), 0.01 volumes 20% glucose were added to 400mls of 1.5% melted agar.

All *E. coli* strains were grown in either LBM medium (10g tryptone, 5g yeast, 5g NaCl, 2g MgCl<sub>2</sub>.6H<sub>2</sub>O per litre) or in 2 x TY (10g yeast extract, 10g NaCl, 16g bactotryptone per litre) liquid medium.

### 2.6 Autoradiography and Photography.

Agarose gels stained with 0.5µg/ml of ethidium bromide were photographed using Polaroid type 667 film in a land camera, illuminating from below using a transilluminator of wavelength 302nm.

Autoradiographs were done using Agfa-gevaert film RP1, films were preflashed using a hand held flash gun, placing the flashed side in contact with the film. Films were exposed at either room temperature or at -70<sup>o</sup> C using Dupont intensifying screens.

### 2.7 Deproteination of nucleic acid using phenol/chloroform extraction.

DNA samples were deproteinised using phenol/chloroform extraction. The number and form of the deproteination steps could vary depending on the quality of the sample. A phenol extraction

entailed adding an equal volume of water saturated phenol to the DNA solution, mixing thoroughly and spinning in a centrifuge to separate the organic phase from the aqueous phase. The aqueous (DNA containing phase) was then transferred to a clean tube. A phenol/chloroform extraction entailed adding 0.25 volumes of phenol to the sample, this was thoroughly mixed before 0.25 volumes of chloroform/isoamyl alcohol were added (24:1). Again the phases were mixed and separated by centrifugation and the aqueous phase removed to a clean tube. Chloroform extraction would then be performed to remove residual traces of phenol. An equal volume of chloroform/isoamyl alcohol (24:1) was added, the phases thoroughly mixed and then separated by brief centrifugation. The aqueous phase was removed to a clean tube and was ready for subsequent manipulation.

### 2.8 Ethanol precipitation of nucleic acids.

This was used to concentrate DNA and RNA.

Ethanol precipitation of nucleic acids involved adding 2.5-3 volumes of 100% ethanol with 0.3 volumes of 3M NaAcetate pH 5, mixing and incubating at either -20°C or -70°C for minimum periods of 3 hours or 30 minutes respectively. The DNA or RNA could then be recovered by centrifugation for an allotted time, followed by decanting off the ethanol supernatant. To facilitate dissolving of the DNA or RNA pellet, excess ethanol was evaporated off either at room temperature or in a vacuum drier. The pellet was then dissolved in the appropriate buffer and stored at either 4°C (genomic DNA samples only) , -20°C (DNA samples) or -70°C (RNA samples only).

## 2.9 Spectrophotometric analysis of DNA and RNA samples.

The concentration of DNA and RNA samples was determined by analysing an aliquot in a spectrophotometer with a deuterium light source at 260 and 280 nm. An optical density reading of 1 at 260 nm equals 50µg DNA/ml; 40µg RNA/ml and 20µg oligonucleotides/ml. The 260/280 nm ratio gives the base content, this is 1.8 for DNA and 2.0 for RNA.

## 2.10 Preparation of DNA.

### 2.10a Genomic DNA preparation.

DNA was prepared from PBL separated from whole blood by spinning at 3k rpm (revolutions per minute) and removing the buffy coat to a fresh tube. Contaminating red cells in the buffy coat were lysed by adding 1.8% saline, vortexing briefly, then adjusting back to isotonic conditions using sterile water. The buffy coat was collected as a pellet by spinning at 1.5k rpm for 10 minutes, the supernatant decanted off and the cells lysed by addition of 0.5 mls of 10% SDS. The volume was adjusted to 5mls by addition of 4.5 mls digestion buffer (0.15M NaCl, 10mM tris.HCl, 10mM EDTA/Na pH8) with 200µg/ml proteinase K. The crude preparation was incubated at 37°C for at least 3 hours before phenol/chloroform/isoamyl alcohol (25:24:1) extracting twice, chloroform/isoamyl (24:1) extracting once and precipitating by addition of two volumes of absolute ethanol. The fluffy DNA precipitate was hooked out and left to air dry before resuspending in 1ml of TE (10mM tris pH7.4, 1mM EDTA pH 8) at 4°C.

### 2.10b Minipreparation of plasmid DNA.

This was prepared using a modification of the method of Birnboim

and Doly (1979).

3 ml culture of *E. coli* containing plasmid <sup>mb5</sup> grown up in LBM at 37°C overnight with shaking and the appropriate antibiotic. 1.5 mls of this was harvested in the morning into eppendorf centrifuge tubes and spun for 1 minute. The supernatant was removed using suction and the pellet resuspended in 100µl of GTE (50mM glucose, 25mM tris.HCl pH 7.9, 10mM EDTA/Na pH 8, with 2mg/ml lysozyme freshly added). The tubes were incubated on ice for 30 minutes before the addition of 200µl of 9:1 0.22M NaOH/10% SDS (freshly made up), the tubes were inverted to mix and left on ice for a further 5 minutes. Once the cells were disrupted, high molecular weight DNA and proteins were precipitated by the addition of 150µl 3M NaAcetate pH5 with incubation on ice for a further 60 minutes. The debris was pelleted by centrifugation for 10 minutes. Approximately 380µl of supernatant was removed to a fresh tube and the DNA precipitated by addition of 2 volumes of 20:1 absolute ethanol, 3M NaAcetate pH5 with incubation at -70°C for 15 minutes. The plasmid DNA was collected by centrifugation for 5 minutes, the pellet was either ethanol precipitated again or subjected to a phenol/chloroform/isoamyl alcohol extraction (25:24:1) (this is advisable if the *E. coli* strain used was rec<sup>-</sup>). After further ethanol precipitation the pellet was dried and resuspended in 50-100µl of TE (10 mM tris.HCl, 1mM EDTA/Na pH8). RNase A at a concentration of 1 µg/ml was added to remove contaminating RNA. The plasmid DNA could now be examined on an agarose gel.

#### 2.10c Large scale preparation of plasmid DNA.

This follows the previous method except that the volumes were

scaled for larger quantities. The plasmid was purified by density equilibrium centrifugation in CsCl.

An overnight culture of *E. coli* containing plasmid was set up in 10 mls of LBM with antibiotic. This was used to inoculate a large 2 litre flask of LBM and grown to stationary phase overnight. The cells were harvested by centrifugation at 6k rpm for 10 minutes in a Sorvall GSA rotor. The supernatant was discarded and the pellet resuspended in 25 mls GTE (see section b) with the addition of 5 mls 10mg/ml lysozyme and left at room temperature for 10 minutes. The cells were lysed by addition of 60 mls 9:1 0.22 M NaOH/10% SDS, and incubated on ice for 5 minutes before adding 30 mls 5M KAcetate pH5, a further incubation on ice of 15 minutes ensures precipitation of protein and high molecular weight DNA. The solution was cleared of debris by centrifuging for 10 minutes and filtering the supernatant through gauze. The plasmid DNA was then precipitated by addition of 0.6 volumes of isopropanol and centrifugation for 10 minutes at 8k rpm in a Sorvall GS3 rotor. The pellet was resuspended in 8 mls TE pH8 and phenol/chloroform then chloroform extracted twice before ethanol precipitation. The pellet was redissolved in 5.6mls TE with 4.75g CsCl and 250 $\mu$ l 10mg/ml ethidium bromide, this was transferred to quick seal centrifuge tubes and spun at a temperature of 20 $^{\circ}$ C at 50k in a Beckman Vti 65 rotor overnight. On centrifugation the buoyant density of the plasmid DNA is different from that of RNA and chromosomal DNA and it appears as a thick band approximately 1/3-1/2 way down the tube. The plasmid band was withdrawn with a needle and syringe and the ethidium bromide removed by dissolving in butan-1-ol, mixing and removing the top organic layer. The CsCl is removed by ethanol precipitation from 70% ethanol (0.66 volumes ethanol, 0.33 volumes

TE pH 8) and the resultant plasmid pellet either further purified by phenol extraction or used directly. The concentration and purity was determined by reading the optical density of an aliquot at 260nm and 280nm.

### 2.11 Preparation of RNA.

All solutions, glassware and plastics were soaked in a 0.1% solution of DEPC, then autoclaved.

#### 2.11a Isolation of total RNA.

Two methods were used, the first was generally used to process small quantities of tissue, the second was used for larger amounts.

##### 2.11.a1 Method 1.

The first method was as outlined by Chirgwin *et al.* (1979). No more than 1g of tissue was homogenised in 4mls of 4M guanidinium thiocyanate (4M guanidinium thiocyanate, 0.5% sodium *N*-lauroylsarcosine, 25 mM NaCitrate pH 7, 0.1 M 2-mercaptoethanol, 0.1% 30% Antifoam A, filtered and pH adjusted to 7), this was layered onto 1.2 mls of 5.7M CsCl (5.7M CsCl, 25 mM NaAcetate pH 5) and centrifuged in a swing out Sorvall AH650 rotor for 12 hours at 20°C at a speed of 36k rpm. The resultant RNA pellet was resuspended in 1ml (maximum) of 7.5M guanidine hydrochloride (7.5M guanidine hydrochloride pH 7, 0.025 volumes 1M NaCitrate, 5 mM dithiothreitol) and precipitated by addition of 0.025 volumes 1 N acetic acid and 0.5 volumes of ethanol. After incubation at -70°C for at least 1 hour the RNA was pelleted in a Sorvall HB4 rotor for 10 minutes at 10k rpm at -10°C and resuspended in 1-2mls water. Further ethanol precipitation steps were performed by addition of 0.3 volumes 3M NaAcetate pH5 and 3 volumes ethanol. This was repeated once more. The RNA was

resuspended in H<sub>2</sub>O and the optical density of an aliquot measured at 260nm to determine the concentration. The integrity of the RNA was roughly estimated by electrophoresis in formaldehyde gels, staining with 10µg/ml ethidium bromide, then destaining for 3-4 hours in water. The presence of 28s and 18s ribosomal RNA ( rRNA) bands was used as a rough indication that the total RNA prepared is of reasonable quality.

#### 2.11.a2 Method 2.

The second method was an adaptation of Cox (1968). 8M guanidine hydrochloride (10mM tris.HCl, 10mM EDTA/Na, 8M guanidine hydrochloride pH 7.5) was used as a chaotropic agent to denature RNases. Liquid nitrogen frozen tissue was slightly thawed on ice before homogenising at high speed for 1 minute, approximately 1g of tissue was homogenised per 20 mls 8M guanidine hydrochloride. The *undesirable residue* was pelleted from the homogenate by spinning at 9k rpm for 5 minutes in Sorvall GSA rotor. The supernatant was decanted to a fresh tube containing 0.5 volumes of ethanol. This was incubated at -20°C for 45 minutes before pelleting at 9k rpm for 10 minutes at 4°C. The pellet was resuspended in half its initial volume in 6M guanidine hydrochloride (6M guanidine hydrochloride, 10 mM tris.HCl, 10 mM EDTA/Na pH 7.5) with the aid of a dounce homogeniser, 0.5 volumes of ethanol were added to this, mixed in and incubated at -20°C for 45 minutes. This step was repeated one more time before resuspending the pellet in DEPC treated water. The RNA was precipitated from the water by addition of 0.5 volumes of 7.5M ammonium acetate and 2.5 volumes of ethanol. The RNA was pelleted by centrifuging for 10 minutes at 10k rpm in a Sorvall HB4 rotor. The pellet was resuspended



in water and reprecipitated two to three times more with 0.1 volumes of 2M sodium acetate pH 5 and 2.5 volumes of ethanol, incubating at  $-70^{\circ}\text{C}$  for at least one hour before centrifuging. The resultant RNA pellet was finally resuspended in DEPC treated water and an aliquot analysed spectrophotometrically.

#### 2.11b Isolation of polyadenylated RNA.

Using the protocol of Aviv and Leder (1972) with some of the modifications of Chirgwin *et al.* (1979) polyadenylated RNA was selected.

The oligo(dT)-cellulose was first rinsed in 0.1M sodium hydroxide and allowed to settle, the wash was removed and the oligo(dT) thoroughly rinsed with DEPC treated water. The oligo(dT) was allowed to settle as a 1 cm bed volume in a Biorad 10cm column and washed thoroughly with DEPC treated water using a peristaltic pump. The buffers were made as follows; Binding buffer1 (BB1) = 50mM tris.HCl pH 7.4, 0.4M LiCl, 0.5% SDS , Binding buffer 2 (BB2)= 50mM tris.HCl pH 7.4, 0.4M LiCl, 0.1% SDS and Elution buffer (EB) = 20mM tris.HCl pH 7.4, 0.1% SDS.

RNA was first dissolved in BB1 at a concentration of 2.5mg/ml, heated to  $68^{\circ}\text{C}$  for 10 minutes, cooled on ice and then adjusted to 0.4 M LiCl. The RNA and BB1 are allowed to run through slowly, the run through was reapplied therefore ensuring that a high proportion of the polyadenylated RNA was bound to the column. The column was washed thoroughly with BB1 monitoring the optical density (OD) of the eluate until it was close to zero. The column was then washed with BB2. The polyadenylated RNA was eluted by addition of 10 mls of EB collecting the eluate in fractions and reading them at 260nm. The fractions were then pooled and ethanol precipitated.

## 2.12 Restriction enzyme digests of DNA.

### 2.12a Extracted genomic DNA.

10 $\mu$ g of DNA was digested in a volume of 200 $\mu$ l with 2-3 units of enzyme/ $\mu$ g of DNA according to the manufacturers directions. For each new DNA sample processed 10 $\mu$ l of the digest was removed and run on an agarose gel to check for complete digestion, if incomplete more enzyme was added and the incubation time prolonged. On complete digestion, the DNA was phenol/chloroform extracted and ethanol precipitated with addition of 10 $\mu$ g of 10mg/ml Dextran sulphate at -70 $^{\circ}$ C for 30 minutes. The DNA was collected after centrifugation for 5 minutes in an eppendorf microfuge. The pellet was washed in 70% ethanol before drying under vacuum. The dry pellet was resuspended in 40 $\mu$ l water before adding 10 $\mu$ l DNA loading dye (30% ficoll 400, 0.1% SDS, 40 mM EDTA/Na pH 8, 1.2 mg/ml bromophenol blue) and heated to 65 $^{\circ}$ C for 10 minutes before loading on an agarose gel.

### 2.12b Plasmid and phage digests.

Restriction digests of plasmid or phage DNA were carried out at a concentration of <1  $\mu$ g/10 $\mu$ l with the concentration of enzyme not exceeding 10%. The incubation time varied from 1 to 2 hours. Small aliquots of these digests were checked on agarose minigels for completeness of digestion.

Large scale preparative digests were also done this way by simply scaling up the amounts used and digesting phage or plasmid at concentrations >100  $\mu$ g/ml.

## 2.13 Gel electrophoresis.

### 2.13a Non-denaturing agarose gel electrophoresis.

Appropriate percentage gels were cast in gel forming plates. Minigels generally used 1-1.5% agarose and were electrophoresed at 10-12 volts/cm. Large 250ml-350ml gels (20cm x 20cm) were cast as 0.8% concentration and run either slowly overnight at 1-2 volts/cm or during the day at 3-6 volts/cm. Preparative gels used either ultrapure agarose or low melting temperature agarose (Biorad) at the appropriate concentration and were electrophoresed overnight at 1-1.5 volts/cm.

In all cases the gel buffer used consisted of 40mM tris acetate, 2.5mM EDTA/Na pH7.7 and all gels were run with 0.5µg/ml ethidium bromide in the gel and gel buffer. All samples were loaded onto the gels containing 0.2 volumes of tracking dye (30% ficoll, 0.1% SDS, 40 mM EDTA/Na, 1.2 mg/ml bromophenol blue).

### 2.13b Agarose denaturing gels.

These were run exactly according to the specifications set out in Maniatis *et al.* (1982).

### 2.13c RNA denaturing gels.

300ml RNA denaturing gels were cast in 20 x 20 cm plates, with 1.5% agarose, 2.2 M formaldehyde, 10 mM NaPO<sub>4</sub> pH 7 (stock= 1M Na<sub>2</sub>HPO<sub>4</sub> and 1M NaH<sub>2</sub>PO<sub>4</sub> mixed to give pH7). The gel was run in 10mM NaP buffer at 1-2 volts/cm. RNA samples were treated at 65°C for 15 minutes in RNA denaturant (50% formamide, 2.2 M formaldehyde, 0.5 mM EDTA/Na pH8, 10 mM sodium phosphate) were quenched on ice before adding 0.2 volumes of RNA tracking dye (0.5% SDS, 0.025% bromophenol blue, 25% glycerol, 25 mM EDTA/Na pH 8) added before

loading.

### 2.13d Nondenaturing acrylamide gels.

All acrylamide gels (acrylamide /N,N' , methylene-bisacrylamide, 29:1) were cast as 0.4 mM thick gels using a Raven vertical sequencing gel apparatus. The gel mix consisted of; (8% gel) 10 mls 40% acrylamide, 2.5 mls 20 x TBE pH 8.8, 37.5 mls H<sub>2</sub>O. TBE buffer consisted of 125mM Tris.Borate pH 8.8, 89mM boric acid, 1 mM EDTA/Na. The appropriate percentage of acrylamide was selected using the guidelines in Maniatis *et al.* (1982) depending on the sizes of fragments to be separated. Before pouring the gel the glass plates were steeped overnight in 1% solution of detergent then thoroughly cleaned and one plate siliconised. These plates were then taped together using 10cm scotch tape and 0.4mM spacers. The acrylamide mix was poured between these plates ensuring that no air bubbles formed and was polymerised using 0.3mls 10% ammonium persulphate and 20-30µls of TEMED. Once poured the comb was inserted and the gels were left at a slight angle covered with Saran wrap with a weight placed onto the plates to ensure good contact. The gels were left to polymerise for an hour. Just before use the wells were flushed with 1 x TBE and clamped to the gel tank. These gels were run at approximately 30 watts. Once the electrophoresis run was complete the gels were fixed in 10% acetic acid and placed onto sheets of 3 mm chromatography paper and covered with Saran wrap. The gel was dried onto the backing sheets in a Biorad slab gel drier at 80°C for at least one hour. Once dry the Saran wrap was removed and the gel exposed to film in a cassette at either room temperature or at -70°C with intensifying screens.

### 2.13e Denaturing acrylamide gels.

Denaturing acrylamide gels were poured using the same gel apparatus, except that the gel mix contained 8M urea. Once poured the gels had a sharktooth comb inserted if they were to be used to resolve sequencing reaction mixes, or an ordinary separate well comb. Once polymerised the wells were flushed with 1 x TBE and clamped to the gel apparatus. These gels were electrophoresed at 40-45 watts, until the samples were thought to be resolved adequately. The gel was then removed from the plates and fixed for 10 minutes in 10% acetic acid, before placing on 3 mm chromatography paper with a Saran wrap cover. As described in d) above the gels were dried and exposed to film. If the gel had been used to resolve  $^{35}\text{S}$ dATP labelled sequencing reactions then it was exposed overnight at room temperature without intensifying screens.

### 2.14 Recovery of DNA from gel electrophoresis and purification.

#### 2.14a Recovery of DNA from agarose gels.

Once DNA had been adequately resolved in an agarose gel the desired fragment was visualised using a hand held UV lamp of long wavelength (366nm). The DNA band was then electrophoresed onto a piece of single thickness dialysis membrane in an adaptation of the method outlined by Maniatis *et al.* (1982). Dialysis tubing was boiled with 1-2 mM EDTA/Na pH 8 for 5 minutes before opening down one side. A slot was cut in the gel in front of the DNA fragment and the dialysis membrane inserted such that it enveloped the gel. The gel was then electrophoresed at 9 volts/cm until all the DNA had left the gel and become trapped by the membrane, this was then quickly removed after reducing the voltage to 0.75 volts/cm. The DNA was

washed off the membrane into a sterile tube using 5 mls of TE+0.06M NaCl or low salt (0.2M NaCl, 20 mM tris. HCl pH 7.5, 1mM EDTA/pH8) buffer if the DNA was to be recovered using an Elutip-d column.

#### 2.14b Recovery of DNA from low melting temperature agarose gels.

The desired DNA fragment was visualised using a hand held UV lamp and excised from the gel, being careful to trim off excess gel from the slice. The gel slice containing the DNA was melted at 65°C with two volumes of TE (10 mM tris.HCl pH 7.4, 1 mM EDTA/Na pH 8) and 0.06M NaCl if it was to be column purified. Once melted the DNA was phenol/chloroform extracted, with back extractions of the phenol phase and ethanol precipitated. If the DNA was to be purified over a DE52 column then the ethanol precipitation step was omitted.

#### 2.14c Purification of DNA.

DNA was either purified using Elutip-d (Schleicher and Schuell) according to manufacturers instructions or purified by passing over a DE52 column. The DE52 column was constructed in a 1 ml plastic syringe with 100µl of DE52 with a siliconised glass wool plug [as described by Maniatis *et al.* (1982)]. The DNA to be purified was adjusted to a salt concentration of 0.06 M NaCl while suspended in TE, this was then dripped slowly over the column, before washing with 0.3 M NaCl+TE and eluting with 0.6 M NaCl+TE. At the elution stage 75µl was dripped over first and discarded before adding a further 200µl. The eluate was diluted to a NaCl concentration of 0.3M before adding three volumes of ethanol and 1-10µg of 10mg/ml dextran sulphate. The DNA was then collected by centrifugation and its

concentration determined by running an aliquot on an agarose gel with concentration markers.

## 2.15 Transfer of nucleic acid and immobilisation on a solid support.

### 2.15a DNA transfer.

DNA fractionated on agarose gels was either transferred to nitrocellulose membranes (Schleicher and Schuell) or to Hybond-N (Amersham) using the strategy devised by Southern (1975). The gels were pretreated before transfer according to Wahl *et al.* (1979).

The gels were immersed in 0.22 M HCl (this was omitted if the fragments to be transferred were less than 5kb in size) and gently agitated for 2 x 15 minutes. This solution was aspirated off and replaced by denaturant (freshly made up each time) 0.5M NaOH, 1.5M NaCl. The gel was bathed in this solution for 2 x 15 minutes. The neutralisation solution consisted of 1.5M NaCl, 0.5M tris.HCl pH 7.4, the gels were neutralised for one hour before transfer. The transfer solutions used 20 x SSC (1xSSC= 150mM NaCl, 15mM Na<sub>3</sub>Citrate pH 7-7.5) in the bottom reservoir with 3 sheets of 3 mm chromatography paper cut to the same size as the gel presoaked in 2xSSC laid on top of the gel membrane sandwich. Layers of towels were placed on top of this to draw the solution up through the gel.

Once transfer was complete the nitrocellulose membranes were baked at 80°C under vacuum for two hours, Hybond-N membranes were wrapped in Saran wrap and fixed by UV-irradiating on a transilluminator wavelength 302nm for 45 seconds.

### 2.15b RNA transfer.

After fractionation on formaldehyde denaturing gels the RNA was

transferred to a solid support by first soaking the gel in 3xSSC (1 x SSC = 15 mM NaCl, 15mM Na<sub>3</sub>Citrate pH7-7.5) then transferring using a setup as described above and previously described by Southern (1975)<sup>with</sup> 20xSSC in the bottom reservoir and 3xSSC soaked 3 mm paper on top of the membrane. Once transfer was complete the membranes were treated as for DNA transfer.

#### 2.15c Colony transfer.

Phage plaques were transferred either to nitrocellulose as described by Benton and Davis (1977) or to Hybond-N (Amersham) as per manufacturers instructions.

##### 2.15c1 Nitrocellulose transfer.

Briefly, duplicate filters lifts were taken, these were in contact with the plate for a period of one minute and three minutes. The lifts were orientated to each other and to the plate by marking through the filter into the agar below with indelible ink on a 18 gauge needle, the lifts were then peeled off the plates and rinsed (contact side up) in 0.5M NaOH, 1.5M NaCl for one minute followed by one minute in 0.2M tris.HCl pH7.4 and one minute in 2xSSCP (150mM NaCl, 15mM NaCitrate, 13mM KH<sub>2</sub>PO<sub>4</sub>, 1mM EDTA/Na pH7.2). Lastly the filters were air dried on sheets of 3 mm chromatography paper. Once dry the filters were sandwiched between sheets of 3 mm before baking in a vacuum oven for two hours at 80°C.

##### 2.15c2 Transfer to Hybond-N membranes.

The membrane was left in contact with the plaques on the plate for varying lengths of time depending on how many replicas of one plate were needed. The first lift was left on normally for one minute and orientated by piercing through the membrane with an 18 gauge needle



coated with indelible ink. The membrane was then peeled off the plate and placed plaque side up onto pads of 3 mm paper soaked in 1.5M NaCl, 0.5M NaOH for 7 minutes. The membrane was neutralised by placing on a pad soaked in 1.5M NaCl, 0.5M tris.HCl pH 7.2, 1mM EDTA/Na for three minutes. This step was repeated before transferring the membrane to a 2xSSC wash. The membranes were wrapped in Saran wrap and fixed on a UV transilluminator for 45 seconds.

### 2.16 Radioactive labelling of DNA.

Labelling of DNA followed the Random priming method devised by Feinberg and Vogelstein (1983).

10-20ng of DNA was boiled for three minutes with 125ng of p(dN)<sub>6</sub> (Pharmacia). This was quenched on ice before addition of 0.2 volumes of buffer ( 330mM tris.HCl pH 7.5, 30mM MgCl<sub>2</sub>), 0.2 volumes of d(NTP) mix (0.25 volumes of 0.6mM dATP, 0.6mM dTTP, 0.6mM dGTP, 50mM DTT) , 10-50μCi α<sup>32</sup>PdCTP and 1 unit of Klenow fragment DNA polymerase I. The reaction was incubated at room temperature.

Incorporation was assayed by sampling 1μl of the reaction after three hours into 1ml of 10% sodium pyrophosphate. The labelled DNA was precipitated by adding 1μl 10mg/ml Dextran sulphate and 300μl 50% trichloroacetic acid. The mix was incubated on ice before collecting by suction onto a 2.5cm glass fibre filter and Cherenkov counted. Incorporation was routinely in excess of  $1 \times 10^9$  <sup>dpm</sup>/μg DNA.

## 2.17 Hybridisation of filters to radioactive probes.

Two hybridisation solutions could be used depending on the type of experiment being undertaken.

### 2.17a Hybridisation to Hybond-N nylon membranes.

Once the DNA was UV-fixed to the membrane it was prehybridised in 0.5 M sodium phosphate buffer pH 7.2, 7% SDS, 1mM EDTA/Na pH8 (Church and Gilbert 1984). The labelled probes were denatured by adding 0.25 volumes of 2M NaOH (freshly made up) on ice for 5 minutes before adding to hybridisation buffer (recipe as before). All hybridisations were carried out in heat sealed plastic bags and were incubated at 65°C overnight with agitation.

### 2.17b Hybridisation to nitrocellulose membranes.

Nitrocellulose membranes were incubated for three hours in 10xDenhardtts (1xDenhardtts = 0.02% ficoll 400, 0.02% polyvinylpyrrolidone, 0.02% Bovine serum albumin fraction V), 4xSSC (1xSSC = 150mM NaCl, 15mM Na<sub>3</sub>Citrate) and 0.1% SDS at 65°C. They were then ready for prehybridisation in; 4xSSC (as before), 10xDenhardtts (as before), 0.1% sodium pyrophosphate, 0.1% SDS, 0.5 µg/ml poly A (10mg/ml concentration) and 25µg/ml denatured salmon sperm DNA (prepared as described in Maniatis *et al.* 1982). The membranes were prehybridised for at least one hour before being set up for hybridisation with the same solutions as before, except denatured probe was added at  $<1 \times 10^6$  /ml.

### 2.17c Washing hybridised filters.

Washing conditions for filters which had been hybridised varied depended on the amount of similarity between the probe and the sequence of interest. The filters were removed from the bag

containing the probe and agitated quickly and vigorously in washing solution, this was decanted off and fresh solution added. Washes consisted of one five minute wash, and two to three twenty minute washes all done at 65°C. In general, low stringency washes consisted of 2 x SSC, 0.1% SDS. High stringency washes were 0.1-0.2 x SSC, 0.1% SDS. Once hybridised and washed filters were wrapped in Saran wrap before autoradiography.

### 2.18 Construction of a bovine liver cDNA library in bacteriophage $\lambda$ gt10.

There are numerous problems associated with the construction of cDNA libraries, the two main ones are; obtaining full length cDNA's and generating representative libraries. Okayama and Berg (1982) recognised these problems and devised a method for cloning of full length cDNA's. Their method entailed cloning directly into a eukaryote expression vector, by ligating the poly(A)<sup>+</sup> tail of the mRNA to a poly(dT) tail in the vector. This poly(dT) tail was then used to prime synthesis of the first strand. The first strand was then oligo(dC) tailed with dCTP before ligating to dGTP tailed plasmid. The remaining RNA strand (which had been used as the template for the first strand) was then replaced using the enzymes DNA polymerase I, *E. coli* RNase H and *E. coli* DNA ligase. The hallmarks of this method were that full length copies of mRNA's were obtained, because addition of homopolymer tails is facilitated by base paired double stranded molecules (full length transcripts). Also *E. coli* RNase H nicks the remaining template RNA strand providing suitable primers for the synthesis of the second cDNA strand. Lastly, the use of *E. coli*

DNA ligase ensures that adjacent RNA and DNA strands are not ligated together unlike T4 DNA ligase. The drawbacks to this method were that transformation efficiencies of plasmids were low when compared to the packaging efficiency of bacteriophage (although this is not now the case).

The methodology which I used to construct the cDNA library was originally devised by Gubler and Hoffman (1983). This was outlined as a protocol by Huynh, Young and Davis (1986) and Watson and Jackson (1986). The main feature of the method was the use of *E. coli* DNA ligase, *E. coli* RNase H and DNA polymerase I to synthesise the second cDNA strand, as devised by Okayama and Berg (1982). The Gubler and Hoffman (1983) method did not use the elaborate vector-primer method devised by Okayama and Berg (1982), it also does not use nuclease S1 to cleave the hairpin loop which forms after first strand synthesis (this will be discussed more later on). The second feature of the chosen method, was that the double stranded cDNA was ligated into the *EcoRI* site of bacteriophage  $\lambda$ gt10 and selected in bacterial strain BNN102. This is particularly useful, because the *EcoRI* site which the cDNA's are ligated into is within the *cl* gene, obligating the phage to undergo lytic growth. The bacterial strain BNN102 carries the *hfl*<sup>-</sup> mutation which prevents establishment of lysogens in its' genome. Therefore this provides a selection system for recombinant bacteriophage. These bacteriophage must undergo lytic growth, any phage lacking an insert is not permitted to lysogenise by the bacterial host strain so these are selectively lost.

#### 2.18a First strand cDNA synthesis.

Polyadenylated RNA was prepared as described in section 2.9b. The

reaction mix for synthesis of the first cDNA strand consisted of the following;

2mM dithiothreitol  
250 $\mu$ m d(NTP's)  
3 mM KCl  
100 $\mu$ g/ml oligo d(T)<sub>12-15</sub>  
100 units/ml RNasin  
50 $\mu$ g/ml poly(A)<sup>+</sup> RNA  
100 $\mu$ Ci/ml  $\alpha$ <sup>32</sup>PdCTP (410 $\mu$ Ci/mmol)

this was incubated at 43<sup>o</sup>C for two minutes before addition of 2000 units/ml X.L. reverse transcriptase.

The various components of the reaction mix were added in the order listed and incubated at 43<sup>o</sup>C for one hour. The yield of first strand cDNA was calculated by removing an aliquot of mix before addition of reverse transcriptase and on completion of the reaction. These were analysed for incorporation by trichloroacetic acid precipitation as described in section 16. The yield of first strand varied from 5-25%.

The first strand cDNA mix was phenol extracted twice and ether extracted once (this is the same principle as chloroform extraction, except that ether floats on top of the solution facilitating easier removal), excess ether was removed by heating the tube for a short period of time at 65<sup>o</sup>C. The first strand was precipitated by addition of 0.25 volumes of 7.5M NH<sub>4</sub>Acetate and 2.5 volumes of alcohol and incubating at -70<sup>o</sup>C for thirty minutes. The first strand cDNA was collected by centrifugation for five minutes in a microfuge, the ethanol decanted and the wet pellet resuspended in 200 $\mu$ l TE (10mM

tris.HCl pH 7.4, 1mM EDTA/Na pH8). The ethanol precipitation was repeated. After the final precipitation an aliquot of approximately 40,000 cpm was removed for analysis on an alkaline agarose gel (figure 1).

#### 2.18b Second strand cDNA synthesis.

The first cDNA strand with template mRNA strand was then used to prime synthesis of a double stranded cDNA molecule. The reaction volume for the second strand was always 100 $\mu$ l and the reagents were added in the following order;

20mM tris.HCl pH 7.5  
5mM MgCl  
100mM KCl  
10mM NaAcetate  
0.01% gelatin  
50 $\mu$ m dNTP's  
7.5 $\mu$ m  $\beta$ NAD  
9 units/ml *E. coli* RNase H  
250 units/ml DNA polymerase I (holoenzyme)  
4 units/ml *E. coli* DNA ligase  
100 $\mu$ Ci/ml  $\alpha$ <sup>32</sup>PdCTP (410 $\mu$ Ci/mmol).

The tube was incubated at 12<sup>o</sup>C for one hour then moved to 22<sup>o</sup>C for an additional hour. As before, samples were removed at the start and finish of the incubation period and assayed for radioactive incorporation. The yield of second strand was 100% (within the limitations of the procedure used to assay incorporation).

The second strand was deproteinised by phenol extraction as described for the first strand, phenol was removed by dissolving in

ether. Before deproteination the reaction was stopped by addition of 5 $\mu$ l 0.5M EDTA/Na pH 8. The double stranded cDNA was precipitated by addition of NH<sub>4</sub>Acetate to 2M and ethanol (as described for first strand). For the second ethanol precipitation 0.1 volumes of 3M sodium acetate pH5.1 was added with 3 volumes of ethanol. An aliquot of cDNA corresponding to approximately 40,000 cpm's was removed for analysis (see figure 1).

#### 2.18c Nuclease SI cleavage of double stranded cDNA.

Cleavage of the cDNA with nuclease SI was included in this protocol. This was because despite claims to the contrary a transient hairpin loop does form and can be used to prime the synthesis of the second strand, for an example of this see Gubler and Hoffman (1983). The presence of a hairpin will interfere with subsequent ligation of linkers on to the cDNA, cleavage of the structure by nuclease SI will remove this. In addition nuclease SI digestion can be used as an indicator of the quality of the cDNA, if it is completely double stranded then it should be resistant to further cleavage once the hairpin loop has been removed (Maniatis et al.1982).

The correct nuclease SI concentration was first determined by titrating dilutions against aliquots of the cDNA. This ensured that overdigestion could not occur. Nuclease SI concentrations used included 2, 4, 6 and 8 units, with the optimum being chosen for large scale digestion.

The reaction volume was usually 50 $\mu$ l with 10 $\mu$ l nuclease SI buffer (1mM ZnCl<sub>2</sub>, 250mM NaCl, 30mM NaAcetate pH 5.5) and 3 units of SI nuclease. This was incubated at 37<sup>o</sup>C for 20 minutes.

The reaction was stopped by addition of 1 $\mu$ l 0.5M EDTA/Na pH8, SDS to 1% with 0.5 volumes phenol followed by heating at 65 $^{\circ}$ C for two minutes. The tube was subsequently centrifuged and the upper phase phenol/chloroform extracted, chloroform extracted and precipitated with 3 volumes of ethanol, 0.1 volume 3M NaAcetate pH5.1 and 1 $\mu$ l 10mg/ml dextran sulphate. As before 40,000 cpm were removed for analysis on an alkaline agarose gel.

Figure 1 indicates that the cDNA was indeed resistant to S1 degradation, so was suitable for further manipulation.

#### 2.18d Methylation of *EcoRI* sites in the cDNA.

The double stranded cDNA was dissolved in 35 $\mu$ l water to which 10 $\mu$ l of 5xbuffer (1x = 10 mM tris.HCl pH 7.9, 10mM EDTA/Na pH 8), 5 $\mu$ l 100 $\mu$ M S-adenylosyl methionine and 5 units of *Eco RI* methylase were also added. The reaction was incubated at 37 $^{\circ}$ C for thirty minutes, 70 $^{\circ}$ C for five minutes, then phenol/ether extracted and ethanol precipitated using 1 $\mu$ l of 10mg/ml dextran sulphate as a carrier.

#### 2.18e T4 polymerase blunt ending of double stranded cDNA.

The blunt ending reaction was carried out in a volume of 30 $\mu$ l, to this was added 3 $\mu$ l 10xT4 polymerase buffer (1x = 66mM KAc, 33mM tris.Acetate pH 7.9, 10mM MgAcetate), 1.5 $\mu$ l 20xdNTP's (mix consists of 0.2 volumes of 10mM dTTP, 10mM dATP, 10mM dGTP, 10mM dCTP and 0.1 volumes 100mM DTT) and 10 units of T4 DNA polymerase. The reaction was incubated at 37 $^{\circ}$ C for thirty minutes.

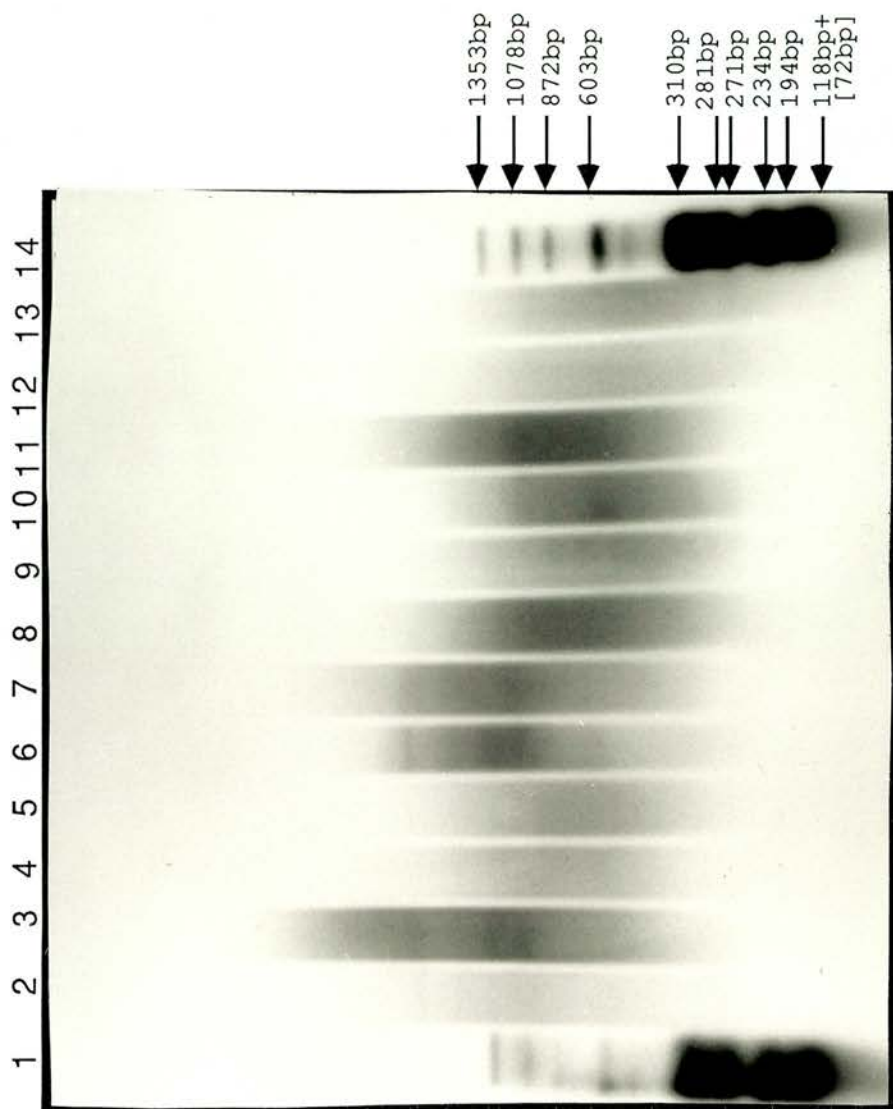
The reaction was stopped by adding 1 $\mu$ l 0.5 M EDTA/Na pH 8 and then phenol/chloroform extracted and ethanol precipitated as before



Figure 1. Alkaline agarose gel loaded with test aliquots of cDNA reactions.

Approximately 40,000cpm were removed from each reaction and loaded on an alkaline agarose gel, the gel was exposed to x-ray film at room temperature. After analysis of this gel, only 3 units of nuclease S1 were used to digest the cDNA, all the reactions were taken through to the next steps in the cloning procedure. The samples were loaded from left to right as follows:- end labelled *HaeIII* digest of  $\phi$ x174; first reaction 1st strand, 2nd strand, S1 nuclease treated (3 units), S1 nuclease treated (6 units); Second reaction 1st strand, 2nd strand, S1 nuclease treated (3 units), S1 nuclease treated (6 units); Third reaction 1st strand, 2nd strand, S1 nuclease treated (3 units), S1 nuclease treated (6 units); end labelled *HaeIII* digest of  $\phi$ x174.

Figure 1



with addition of carrier. The blunt ended cDNA was separated from contaminants by dripping over a DEAE 52 column as described in section 12c. The DNA eluted from the column was ethanol precipitated with carrier before proceeding to the next stage in it's manipulation.

#### 2.18f Kinasing and ligation of linkers to the cDNA.

Before the linkers could be ligated to the cDNA they firstly had to be end-labelled, this ensures that they will ligate and their progress in subsequent ligation reactions can be easily monitored. 5µg of phosphatased *Eco RI* linkers were kinased in a reaction volume of 50µl to which the following were also added, 5µl 10xpolynucleotide kinase buffer (1x = 70mM tris.HCl pH7.6, 10mM MgCl, 5mM DTT), 8% PEG<sub>6000</sub>, 6 units T4 polynucleotide kinase and 50µCi  $\gamma^{32}$ PATP. The reaction was incubated at 37<sup>o</sup> C for 30 minutes, before a further 6 units of enzyme was added and 5µl of 10mM rATP, the reaction was incubated for a further 30 minutes. The preparation was stored frozen at -20<sup>o</sup> C. Addition of polyethylene glycol (PEG<sub>6000</sub>) appears to stabilise the T4 polynucleotide kinase (Harrison and Zimmerman 1986).

Ligation of the linkers to themselves was checked before attempting to ligate them to the cDNA. A typical ligation reaction consisted of;

0.25µg linkers end labelled

2µl 10mM rATP

2µl 10xligase buffer (1x = 10mM tris.HCl pH7.4,

10mM MgCl<sub>2</sub>, 50mM NaCl)

2µl 50mM DTT

8 $\mu$ l 40% PEG<sub>6000</sub>

1 unit T4 DNA ligase (Amersham),

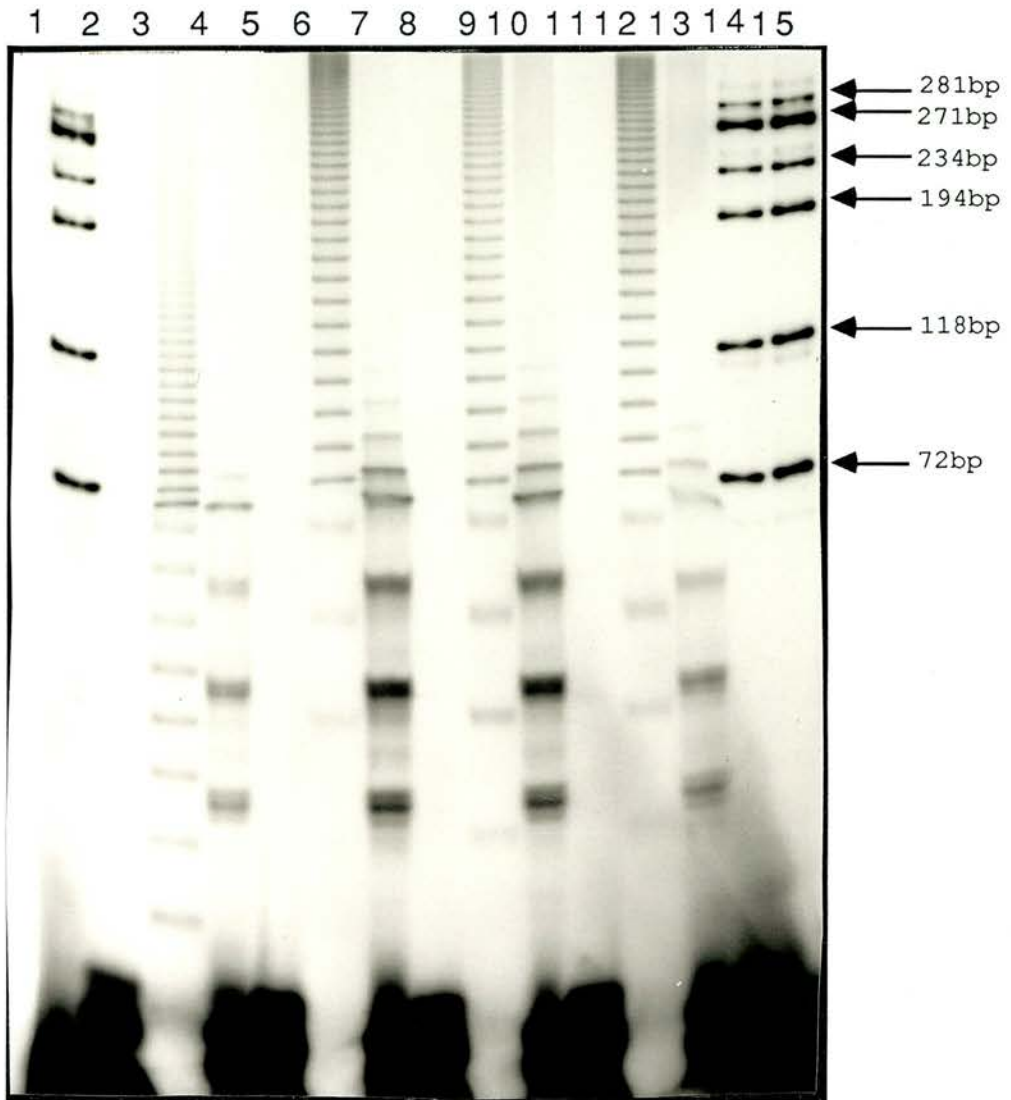
this was made up to a reaction volume of 20 $\mu$ l and incubated overnight at 14°C. In the morning ligase is inactivated by heating to 70°C for 5 minutes before adding *Eco RI* restriction enzyme (usually 100 units) and incubating at 37°C for 1-2 hours. Samples of the reaction were removed before addition of T4 ligase, after the ligation reaction and after digestion by *Eco RI* and loaded onto a non-denaturing 10% acrylamide gel. If the ligation and digestion were judged successful then the linkers were ligated to the cDNA.

0.5 $\mu$ g of end labelled linkers were ligated to the cDNA using the same reaction conditions as previously described. After ligation, excess linkers were digested away by addition of 120 units of *Eco RI* enzyme for 45 minutes before adding another 120 units of enzyme. At each stage samples were taken to be analysed on an 8% acrylamide gel (figure 2). Excess linkers were removed by ethanol precipitation (two times) and by fractionating the cDNA on a 1.5% low melting temperature agarose gel. After electrophoresis overnight at 2 volts/cm the cDNA track was size cut at approximately 400bp, the agarose plug was melted at 68°C and phenol/chloroform extracted (as described before). The cDNA was purified over a DEAE 52 column as previously described and ethanol precipitated using dextran sulphate as a carrier in two steps to ensure that all the cDNA had pelleted to the bottom of the tube.

Figure 2. Nondenaturing acrylamide gel analysis of the end ligation of linkers onto the cDNA.

Aliquots of the various ligation reactions were removed for analysis (see text) and loaded onto an 7.5% acrylamide gel. The size markers are indicated on the right of the figure. The gel was loaded from left to right as follows:- Track1 end labelled *HaeIII* digest of  $\phi$ x174; Track2 unligated linkers; Track3 ligated linkers; Track4 ligated linkers digested with *EcoRI* ; Track5 unligated linkers and third reaction cDNA; Track6 linkers ligated onto third reaction cDNA; Track7 cleavage of excess linkers from third reaction cDNA with *EcoRI* ; Track8 unligated linkers and first reaction cDNA; Track9 linkers ligated onto first reaction cDNA; Track10 excess linkers cleaved from first reaction cDNA using *EcoRI* ; Track11 unligated linkers and second reaction cDNA; Track12 linkers ligated onto second reaction cDNA; Track13 excess linkers cleaved from cDNA using *EcoRI* ; Track14 and 15 end labelled *HaeIII* digest of  $\phi$ x174.

Figure 2



2.18g Ligation of cDNA to bacteriophage  $\lambda$ gt10,  
packaging and plating.

To ensure that all the cDNA was ligated as single inserts into bacteriophage  $\lambda$ gt10 a rough estimate of the yield was needed. This was judged from the initial yield of the synthesis reaction allowing for losses through the subsequent steps such that a 1:1 molar ratio of  $\lambda$ gt10 arms to cDNA inserts was obtained. The ligation reaction was set up as previously described with 0.5 $\mu$ g of phosphatased  $\lambda$ gt10 arms /10 $\mu$ l reaction. In total three of these reactions were set up, a 2 $\mu$ l aliquot was removed and packaged. This yielded 80,000 plaques. Therefore subsequent packaging reactions of the library used 3 and 4  $\mu$ l aliquots. The packaging reactions were plated on 20 cm x 20 cm plates. The packaging mixes were obtained from Stratagene, 4 $\mu$ l of ligation mix was added to the packaging reaction and left to incubate at 22 $^{\circ}$ C for two hours, this was then added to BNN102 cells. The plating cells were prepared by inoculating a colony into LBM containing 15 $\mu$ g/ml tetracycline, 10mM MgSO<sub>4</sub>, 0.2% maltose and grown until their optical density at 600nm was 0.5. Upon reaching 0.5 the cells were harvested by spinning at 3k rpm for 10 minutes in a Sorvall HB4 rotor and resuspended in 0.4 volumes of 10mM MgSO<sub>4</sub>. Approximately 1ml of BNN102 cells were mixed with a single packaging reaction and absorbed for 20 minutes at 37 $^{\circ}$ C before adding 35mls of top agarose and pouring onto a level prewarmed plate. Once the top agarose was set the library was incubated upside down at 37 $^{\circ}$ C for 6-7 hours, at such time the plates were almost confluent with the plaques just touching each other. The total library size was 2.5x10<sup>5</sup> plaque forming units (pfu's). The plates were cooled to 4 $^{\circ}$ C

before plaque lifts were taken.

#### 2.18h Screening of library.

The library was transferred onto nitrocellulose by the method of Benton and Davis (1977). Briefly, duplicate filters lifts were taken these were in contact with the plate for a period of one minute and three minutes before being rinsed in 0.5M NaOH, 1.5M NaCl for one minute followed by one minute in 0.2M tris.HCl pH 7.4 and one minute in 2xSSCP (150mM NaCl, 15mM NaCitrate, 13mM  $\text{KH}_2\text{PO}_4$ , 1mM EDTA/Na pH7.2) lastly the filters were air dried on sheets of 3 mm paper. Once dry the filters were sandwiched between sheets of 3 mm before baking in a vacuum oven for two hours at 80°C. The filters were then hybridised as outlined in section 17b using the human pHLA-27 cDNA class I clone at a concentration of  $10^6$ cpm/ml and washed in 3xSSC, 0.1% SDS, 0.1% sodium pyrophosphate at 65°C for 30 minutes followed by two 30 minute washes in 2xSSC, 0.1% SDS, 65°C. The filters were attached to sheets of 3 mm paper with radioactive orientation marks before being covered with Saran wrap and exposed to film with intensifying screens.

A ten day exposure of the library revealed three positives which were picked into 1ml of phage buffer (10mM tris.HCl pH 7.5, 10mM  $\text{MgSO}_4$ , 0.01% gelatin) using the wide end of a Sarstedt 1ml blue plastic tip. The phage were allowed to diffuse into the buffer overnight at 4°C after which an aliquot was diluted and titred. The phage were replated onto 90mm plates at a density of 500 plaques /plate. Nitrocellulose lifts were taken of these plates, one for each probe. The lifts were probed as outlined in section 17b with pHLA-27



and pBR322 plasmid. This identified two of the positives as pBR322. The positive plaques from this screen were subjected to a further round of screening. The positive plaques from this third screen were picked only if they were well separated from the rest.

In an effort to identify any additional positives in the library which may have been missed because a heterologous probe had been used to screen the library. A screen of the library was undertaken using the positively hybridising phage clone isolated in the first screen (this positive insert had been subcloned into pPoly1 before the insert was excised and purified and used as a clone). This identified one additional positive which was plaque purified as described above. Both of these positives were grown up as bulk cultures by the method devised by Blattner *et al.* (1977) which is described in section i below.

At this point it must be stated that from a library size of  $2.5 \times 10^5$  pfu, more than two clones should have been obtained. Sood *et al.* (1981) estimate an individual class I mRNA at an abundance of 0.01-0.05% of poly(A)<sup>+</sup> RNA, this would have given approximately 25-125 class I positives per expressed transcript. Screening the library with a housekeeping gene ( $\beta$ -actin) revealed no positives. A possible explanation for the small number of class I positives was that the library had become contaminated with *EcoRI* linkers. This did not happen because of incomplete digestion of ligated linkers (see figure 2), but may have happened at the gel purification stage, due to linkers being trapped in the gel slot (result not shown).

### 2.18i Storage of library.

Once screening of the library had been completed, the plates were overlaid with 50mls of phage buffer and left agitating overnight at 4°C. The buffer was decanted off with the phage into centrifuge bottles and centrifuged at 4k rpm for 10 minutes in a Sorvall HB4 rotor. The supernatant was decanted off and stored over 0.1% chloroform. Half of this was subjected to CsCl equilibrium density centrifugation as outlined in Maniatis *et al.* (1982) and described in detail below.

#### 2.18i1 Equilibrium density centrifugation of $\lambda$ gt10 liver cDNA library.

8mls of library was underlaid by CsCl of three different densities (2.5mls 1.3g/ml CsCl, 1.5mls 1.5g/ml CsCl and 1.5mls 1.7g/ml CsCl). The gradient was spun in a SW28 rotor for two hours fifteen minutes at 23k rpm at a temperature of 4 C. The bluish phage band was removed by puncturing the tube with a wide gauge needle and sucking off. The phage were then suspended in 40mls of phage buffer with added CsCl to give a density of 1.5g/ml. This was sealed in polycarbonate tubes and spun in a Ti70 rotor in a Beckman ultracentrifuge for 24 hours at 38k rpm, 4°C. The centrifuge tube containing the library was punctured through the bottom and 1ml aliquots were dripped slowly into eppendorf tubes, being careful not to upset the banding of the phage. The 1ml aliquots were diluted thousand fold in phage buffer and 0.001 of this was spotted onto a gridded agar plate which had been freshly overlaid with top agarose containing BNN102. These plates were incubated at 37°C overnight and the fraction containing the banded library identified by presence

of plaques on the sectored plate, the titre equalled  $5 \times 10^8$  p.f.u./ml. Once the correct fractions had been identified they were stored as a CsCl mix at 4°C.

#### 2.18j Phage growth and purification of DNA.

90mm plate lysates of the purified plaques were made by absorbing in excess of  $1 \times 10^3$  phage to 0.1ml of plating BNN102 and incubating overnight. This was cooled and overlaid with 5mls of phage buffer and left agitating overnight at 4°C. In the morning the overlay was scraped off and cleared of debris by centrifugation. This plate lysate stock was stored over 0.1% chloroform. Before use the plate lysate was titred and found to contain  $1.5-2.5 \times 10^{11}$  phage/ml.

#### 2.18j1 Mini-preparation of bacteriophage DNA.

Often it was necessary to ascertain if the phage contained inserts and to determine how many were different from each other. 400µl of the plate lysate was removed to an eppendorf tube, to this 2µl of DEPC, 10µl 10% SDS and 50µl 2M tris, 0.2M EDTA/NA pH 8.5 was added. The mix was then incubated for 5 minutes at 70°C. This was cooled before addition of 50µl 5M KAc pH5 and left on ice for 40 minutes. The preparation was then centrifuged for 15 minutes in a microfuge and 470µl removed to a fresh tube. The supernatant was phenol/chloroform extracted once, chloroform extracted once before ethanol precipitation. Once ethanol was added the tube was mixed and immediately spun in an eppendorf centrifuge for 6 minutes. The ethanol was decanted and the phage pellet resuspended in 50µl TE and 1mg/ml of RNase A added. This was heated to 70°C for 10 minutes. The phage DNA was then ready for manipulation.

### 2.18j2 Large scale preparation of phage DNA.

To obtain enough phage DNA in order to excise the cloned cDNA insert a large scale bulk culture was started (Blattner *et al.* 1977).  $2 \times 10^5$  phage were absorbed to 32mls of BNN102 cells which had been concentrated into 0.2 volumes of 10mM  $MgSO_4$ . The absorption was left at 37°C for 20 minutes before being added to 500mls of prewarmed LBM. This was shaken at 37°C until lysis occurred, after which 5mls of chloroform was added, shaking continued for 30 minutes before 3mg of RNase and 3mg of DNase were added per flask. The flask was left to sit at room temperature for 40 minutes before being cleared of debris by centrifuging at 10k rpm for 10 minutes in a Sorvall GSA rotor. 10% PEG<sub>6000</sub> and 1M NaCl were added to the phage containing supernatant, dissolved and left at 4°C for at least 2 hours. The precipitated phage were pelleted by centrifugation at 10k rpm for 15 minutes at 4°C in a Sorvall GSA rotor. The phage pellet was dissolved in 8mls of phage buffer, residual debris was removed by mixing vigorously with an equal volume of chloroform and clearing by brief centrifugation at 3k rpm for 2 minutes. The top phase was removed and underlaid by three different concentrations of CsCl as described in section 18i. After centrifugation the bluish phage band was removed as before and dialysed against two changes of 50mM tris.HCl pH 7.9, 10mM  $MgCl_2$ , 10mM NaCl. The dialysed phage were then phenol/chloroform extracted twice and chloroform extracted twice. After ethanol precipitation the phage were pelleted at 8k rpm for 10 minutes in a Sorvall HB4 rotor and redissolved in TE pH 7.4. The quality of the preparation was checked on an agarose gel and contaminating RNA removed by RNase A treating ( heating phage DNA

solution at 68°C for 10 minutes with 10 µg/ml RNase A). This protocol was also followed when preparing DNA from EMBL301 genomic clones, except that the initial inoculum of phage for absorption to the cells was  $3 \times 10^7$  -  $5 \times 10^8$ . Upwards of 1 mg of phage DNA was obtained per litre of culture using this method.

## 2.19 Manipulation of DNA fragments.

### 2.19a Phosphatasing of vector DNA.

This was exactly as described by Maniatis *et al.* (1982) for both 5' and 3' overhangs, except that two concentrations of phosphatase were used; 1x and 10x the amount suggested in Maniatis *et al.* (1982). A typical reaction consisted of;

pPoly1 *Eco RI* cut (0.8 µg) 10 µl

5 µl of 10xCIP (1x= calf intestinal phosphatase buffer 50mM

tris.HCl pH9, 1mM MgCl<sub>2</sub>, 0.1mM ZnCl<sub>2</sub>, 1mM spermidine)

CIP (calf intestinal phosphatase @ 20units/µl) used 0.02 and 0.2 units

H<sub>2</sub>O 35 µl

to give a total volume of 50 µl. This would be incubated at 37°C for 30 minutes before another aliquot of phosphatase was added and the incubation continued for another 30 minutes. This was only applicable for removing the terminal phosphate from 5' overhangs. The CIP was inactivated and removed from the reaction as detailed in Maniatis *et al.* (1982).

### 2.19b Ligation of vector DNA to insert DNA.

If the vector had self-ligatable ends then it was phosphatased, this was unnecessary if incompatible ends were present. All ligations of

plasmid DNA used 50ng of vector all M13 bacteriophage ligations used 100ng of vector. Ligations were set up as follows;

vector 50 or 100ng

insert (3:1 molar excess over vector)

1 x ligation buffer (1x= 10mM MgCl<sub>2</sub>, 50mM NaCl, 10mM tris.HCl pH 7.4)

1 x DTT (1x = 5mM DTT)

1 x rATP (1x = 1mM rATP)

1 unit T4 DNA ligase

*adjusted to a volume of 20µl*  
 this was mixed and incubated at 14°C overnight for blunt ended ligations or three hours for 5' or 3' overhangs.

In addition to the above ligation a number of controls were always set up. These include a transformation control of unrestricted vector (usually 20ng with no ligase), a background control for percentage of vector actually restricted (50ng restricted vector with no T4 ligase), a control for self-ligation (50ng restricted vector with T4 ligase) and a control for the success of the phosphatasing reaction (50ng phosphatased vector with T4 ligase).

### 2.20 Transformation of ligations.

A culture of the bacterial strain to be used for transforming into was inoculated into a sterile conical flask with LBM. The cells were grown until their optical density was 0.5 at 540nm. The cells were harvested by centrifuging at 8k rpm for 8 minutes at 4°C in a Sorvall HB4 rotor. The cell pellet was either gently resuspended into 0.5 volumes of 50mM icecold CaCl<sub>2</sub> or 0.2 volumes CM1 [10mM NaAcetate pH 5.6, 5mM NaCl, 50mM MnCl<sub>2</sub> freshly made up]] and left on ice for

twenty minutes before repelleting the cells at 8k rpm for 8 minutes. The final cell pellet was resuspended in either 0.1 volumes of 50mM CaCl<sub>2</sub> or 0.02 volumes of CM2 (10mM NaAcetate pH 5.6, 5% glycerol, 70mM CaCl<sub>2</sub>, 5mM MnCl<sub>2</sub>). The advantage of using CM1 and CM2 meant that the competent cells could be stored frozen at -70<sup>o</sup> C for periods up to one month before use.

Normally 0.1 volumes of ligation mix (never exceeding 5µl from a 20 µl mix) would be added to either 200µl of CaCl<sub>2</sub> treated competent cells or 100µl of CM1/CM2 treated cells. The cells were incubated on ice for 30 minutes, then transferred to 42<sup>o</sup>C for two minutes; at this point the treatment of different vectors can vary as to which antibiotic resistance gene they carry. Ampicillin resistant vectors are incubated at 37<sup>o</sup>C for one hour before plating out aliquots on selective plates, tetracycline resistant plasmids need only be incubated for 30 minutes both of these have 1ml of LBM added to them and bacteriophage M13 needs no further incubation and LBM is not added . M13 must be plated immediately with 3.7mls BBL top agarose containing 200µl JM101 (100µl if used CM2 method) untreated cells at 0.5 OD, 20µl IPTG (100mM solution) and 50µl 20% X-GAL (dissolved in dimethylformamide).

## 2.21 Growth and preparation of bacteriophage M13 double stranded vector and single stranded templates.

### 2.21a Growth of single stranded templates.

Using a colourimetric assay for recombinants (X-GAL) only white plaques were toothpicked into 1.5 mls of LBM containing 0.01 volumes of freshly grown JM101. The tubes were grown at 37<sup>o</sup>C for 5 hours

with vigorous shaking, after which the broth was harvested into eppendorf tubes. The broth was cleared of bacteria by spinning at full speed in an eppendorf microfuge for 5 minutes. The supernatant was transferred to a fresh tube to which 150µl of PEG<sub>6000</sub> mix was added (2.5M NaCl, 20% PEG<sub>6000</sub> ). The tubes were incubated on ice for 10 minutes before pelleting the phage precipitate by spinning for 10 minutes. The PEG solution was drawn off using a suction pump, the tubes ~~were~~ re-spun for 30 seconds and residual PEG solution sucked off. The phage pellet was resuspended in 200µl of TE and phenol/chloroform extracted once followed by a chloroform extraction. The M13 recombinant DNA was ethanol precipitated and resuspended in 20-30µl of seqTE (10mM tris.HCl pH7.4, 0.1mM EDTA/Na pH 8). 3µl of this was used in subsequent sequencing reactions.

#### 2.21b Large scale production of double stranded M13 bacteriophage.

A single colony of JM101 was toothpicked and inoculated into 2 x TY medium and grown overnight. A single well separated M13 plaque was toothpicked into 1.5mls of a 0.01 dilution of an overnight culture of JM101 and grown at 37°C for 2-3 hours. If recombinants were to be grown then a white plaque was picked, if not then a blue plaque was picked. This starter culture was then transferred into 1 mls of 2 x TY containing 0.01 dilution of uninfected JM101 (this was set up at the same time as the starter culture) and shaken for a further 4-5 hours. After this time the cells are pelleted and the supernatant reserved as an inoculum for large scale growth.

Two large 2 litre flasks containing 500mls of 2 x TY were



inoculated with JM101 and incubated with shaking until the OD of the cells equalled 0.5 at 600nm. At this point 5mls of harvested supernatant phage preparation were added and the incubation continued for 4 hours, for the last 30-40 minutes of incubation chloramphenicol was added to 25 $\mu$ g/ml. The addition of chloramphenicol drives all the replication intermediates into double stranded form. After 4 hours the cells were harvested and treated exactly the same as for a large scale plasmid preparation see section 10c.

The yield of double stranded M13 bacteriophage DNA was in excess of 500 $\mu$ g/litre of culture.

### 2.22 Sanger dideoxy sequencing.

All DNA sequencing reactions were performed according to the method devised by Sanger *et al.* (1977). <sup>35</sup>SdATP was used as the radioactive label as devised by Biggin *et al.* (1983). The exact method followed was one elucidated by Dr S. Harris.

3 $\mu$ l of single stranded template prepared by method detailed above in section 21a was annealed to 1 $\mu$ l universal primer (0.03 OD units/ml Pharmacia), 2 $\mu$ l annealing mix (100mM tris.HCl pH8, 100mM MgCl<sub>2</sub>) and 2 $\mu$ l of H<sub>2</sub>O by combining in an eppendorf tube and cooling from 70°C for 90 minutes in a water bath.

2 $\mu$ l of the annealed template was subsequently added to separate tubes containing 2 $\mu$ l each of G, A, T and C mix;

	G mix	A mix	T mix	C mix
0.5mM dGTP	2.5 $\mu$ l	25 $\mu$ l	25 $\mu$ l	25 $\mu$ l
0.5mM dTTP	25 $\mu$ l	25 $\mu$ l	2.5 $\mu$ l	25 $\mu$ l
0.5mM dCTP	25 $\mu$ l	25 $\mu$ l	25 $\mu$ l	2.5 $\mu$ l
10mM ddGTP	3.5 $\mu$ l	-	-	-
1 mM ddATP	-	0.25 $\mu$ l	-	-
10mM ddTTP	-	-	4 $\mu$ l	-
10mM ddCTP	-	-	-	1.5 $\mu$ l
Seq TE buffer	4 $\mu$ l	24.75 $\mu$ l	66 $\mu$ l	46 $\mu$ l
Total volume	100 $\mu$ l	100 $\mu$ l	100 $\mu$ l	100 $\mu$ l

in addition to the above mixes 2 $\mu$ l of Klenow mix (1 $\mu$ l  $^{35}$ SdATP, 7.2 $\mu$ l seqTE, 4 units Klenow fragment DNA polymerase I) was added. The mixes were combined by spinning very briefly in an eppendorf microfuge and reaction started by incubating at 37 $^{\circ}$ C for 20 minutes. The sequencing reaction was chased by adding 2 $\mu$ l of cold dNTP's (0.25mM each nucleotide) and incubating at 37 $^{\circ}$ C for a further 35 minutes. The reaction was stopped by adding 4 $\mu$ l of tracking dye (96% deionised formamide, 20mM EDTA/Na pH 8, 0.003% bromophenol blue, 0.003% xylene cyanol FF). The sequencing reactions were loaded onto a denaturing acrylamide gel after heating to 100 $^{\circ}$ C for 3 minutes.

### 2.23 Processing of sequence information.

All sequence information was processed on a Vax computer network using University of Wisconsin genetics computer group (UWGCG) programmes.

## Chapter Three.

### Characterisation of two bovine cDNA clones and their comparison to other class I genes.

#### 3.1 Introduction.

The number of functional transplantation antigens varies in different inbred strains of mice with three, *K*, *D* and *L* in haplotype d (BALB/c), but only two, *K* and *D*, are present in haplotypes b (C57BL/10) and k, (Hansen *et al.* 1981; Demant and Ivanyi 1981). The number of functional transplantation antigens in man does not appear to vary between individuals; three (*A*, *B* and *C*) are detected (Ploegh *et al.* 1981).

Analysis of class I at the molecular level shows that the number of hybridisable class I like sequences in mouse is in considerable excess of functional transplantation antigens (Cami *et al.* 1981, see also chapter 1 section 1.2.1c3). 15-20 bands can be detected per individual mouse haplotype in inbred mouse strains (Cami *et al.* 1981; Steinmetz *et al.* 1981b); whereas up to 30 different hybridising bands are detected when probing Southern blots of human DNA (Biro *et al.* 1983). These results therefore indicate, that in man and mouse class I genes are part of a multigene family (Orr and Demars 1983; Pease *et al.* 1982). Similar complexity in numbers of class I like hybridising bands has been reported for a variety of species (see chapter 1, section 1.3.1a2 for full listing) e.g. rats (Palmer *et al.* 1983), horses (Alexander *et al.* 1987), sheep (Chardon *et al.* 1985b), pigs (Chardon *et*

*al.* 1985a) and in cattle (Vaiman *et al.* 1986; Lindberg and Andersson 1988).

In an attempt to resolve this apparent contradiction between the numbers of functional class I genes and number of detectable class I-like sequences, investigators have cloned and mapped the entire *H-2* complex containing the class I regions *K*, *D*, *Qa* and *Tla* in the mouse. Steinmetz *et al.* (1982b) cloned the *H-2* complex of the BALB/c mouse and identified thirty six class I genes [this was later revised to thirty five (Stephan *et al.* 1986)]. However, of these only five mapped to the *K* and *D* regions, while the rest mapped to the *Tla* and *Qa* regions (Winoto *et al.* 1983). Of the five clones which mapped to the *K* and *D* regions, three separately encoded *K*, *D* and *L* genes, the remaining two were nonfunctional in the cell transfection assay used (Goodenow *et al.* 1982b). Correspondingly, Weiss *et al.* (1984) isolated overlapping clones of the *H-2* complex of inbred mouse strain C57BL/10 and identified twenty six different class I genes.

Early attempts at locating cosmid clones which carried functional genes, entailed transfecting the cloned cosmid DNA into L-cells. This approach identified ten, unfortunately their identity could not be determined by serological means (nor could the gene encoding the Qa-2 antigen be identified) (Goodenow *et al.* 1982b). Analysis of one particular gene cloned from the BALB/c mouse cosmid library showed that it encoded a defective transmembrane domain and mapped to the *Qa* region (Steinmetz *et al.* 1981b, see also chapter 1 figure 3). It has been demonstrated by DNA transfection that this gene,  $Q7^d$  (see figure 1, chapter 1 for location on DNA map), encodes the nonpolymorphic Qa-2 antigen as does its homolog in the *H-2<sup>b</sup>* haplotype ( $Q7^b$ ) (Waneck *et al.* 1987). Upon transfection,  $Q7^d$  can

generate two forms of Qa-2 antigen, either as a secreted form in L-cells, or as a membrane bound form in liver cells (this accounts for the previous failure to detect it) (Stroynowski *et al.* 1987). However, it now appears that in addition to  $Q7^b$ ,  $Q9^b$  also expresses a product biochemically similar to Qa-2 antigen after transfection (Soloski *et al.* 1988). Therefore numerous *Qa* genes appear to express products which are serologically indistinguishable from each other. This would be in agreement with their observed lack of polymorphism (Flaherty 1980).

A comparable amount of complexity was found when identifying HLA genes. Two of the first clones isolated were found to be genomic pseudogenes; HLA 12.4 (Malissen *et al.* 1982) and pLNIIA (Biro *et al.* 1983). The exon/intron structure of pLNIIA (Biro *et al.* 1983) was atypical of human class I genes (Strachan *et al.* 1984). Possible human counterparts of the murine *Qa* genes have now been isolated (Geraghty *et al.* 1987; Koller *et al.* 1989). Other isolated genes include clone RS5, which appears to be functional but is non-ubiquitously expressed (it was not expressed in the cell line it was cloned from) and has a nonconserved amino acid structure (Srivastava *et al.* 1987). Clone RS5 appears to be very alike another class I sequence, which is transcribed (Mizuno *et al.* 1988) and which has been characterised at the genomic level (Koller *et al.* 1988). This new gene (HLA-6.0) maps to a completely separate locus within the *HLA* complex in man named *HLA-E*. The complete class I region of man has now been cloned and shown to contain 17 class I genes (Koller *et al.* 1989, see chapter 1, section 1.2.1c4). The human equivalent of the murine *Tla* region may be the gene complex *CD1* (Calabi *et al.* 1986). Although *CD1* does not map to the same chromosome as *HLA*, its' encoded genes do associate

with  $\beta_2$ -microglobulin and share twenty one per cent homology at the amino acid level with the fourth exon in MHC class I genes (Martin *et al.* 1986) which bind  $\beta_2$ -microglobulin non-covalently (Yokoyama and Nathenson 1983). However it is now known that there is a separate CD1 complex in mice (Bradbury *et al.* 1988). Therefore from the examples given above it appears that the *H-2* and *HLA* complexes may be more similar than at first realised, both having class Ia and putative class Ib genes. This is contrary to Rogers (1985b) conclusions that the *Qa* and *Tla* loci were not conserved across species. Although *Tla* loci have not been found in man, possible *Qa* loci have. It is therefore highly likely that other species may also have similar clusters of MHC class I genes.

In contrast to results with genomic cloning, cDNA cloning is more likely to lead to the isolation of DNA sequences encoding transplantation antigens. This is because they seem to be the most abundant class I transcripts in a cell (Kress *et al.* 1983c; Mizuno *et al.* 1988; Koller *et al.* 1988). The first isolation of HLA cDNA clones and their verification was facilitated by amino acid data on class I transplantation antigens. Sood *et al.* (1981) used specific oligonucleotides to prime synthesis of DNA complementary to *HLA-B7* mRNA; while Ploegh *et al.* (1980) used cell-free translation and immunoprecipitation to verify their cDNA clones. Other approaches include using cross species homology to isolate murine cDNA clones with a human class I cDNA probe (Steinmetz *et al.* 1981a), this is possible because considerable homology exists at the nucleotide level between murine and human class I genes, 82.3% in exon two rising to 85.2% in exon four between *H-2L<sup>d</sup>* and *HLA* class I sequences

(Srivastava *et al.* 1985).

It was decided to isolate a bovine cDNA. This would be feasible if it was assumed there was appreciable cross species similarity between cattle and human class I sequences. Appreciable cross species hybridisation of class I sequences had been demonstrated by two groups; Vaiman *et al.* (1986) and Lindberg and Andersson (1988) and was further verified by using a human HLA cDNA as a probe against Southern blotted bovine genomic DNA (see figure 1, chapter <sup>Five</sup>). This chapter describes the isolation and characterisation of two bovine cDNA clones isolated from a bovine liver cDNA library and their comparison to class I genes from other species.

## 3.2 Results and Discussion.

### 3.2.1 Construction and screening of bovine liver cDNA library.

A bovine liver cDNA library was constructed using the technique outlined by Huynh, Young and Davis (1986) and Watson and Jackson (1986); in the lambda immunity insertion vector gt10 [Stratagene limited; Huynh, Young and Davis (1986)], see chapter 2 section 2.18 for the detailed methodology used.

Nitrocellulose lifts were taken of the unamplified library by the method of Benton and Davis (1977) and probed with a HLA-27 cDNA clone (provided by Dr Per Peterson). One positive clone was detected. This was plaque purified and grown up in bulk culture (Blattner *et al.* 1977). The insert was excised and used to reprobe the filters. One additional positive was detected.

Restriction with *EcoRI* cut out the complete inserts from gt10, giving sizes of approximately 1.2 kb and 0.6 kb. The 1.2 kb cDNA was hybridised to one half of a Northern blot of bovine RNA (figure 1). The HLA-27 cDNA clone was hybridised to the other half of the Northern. Each cDNA only hybridised to one band of size approximately 1.7 kb, this is the expected size of class I mRNA ( Ploegh *et al.* 1980; Morrello *et al.* 1982 ). The largest bovine cDNA insert (pBoLA-1) was chosen for further study partly due to its size and because subcloning of the smaller cDNA insert (pBoLA-2) was not achieved.

### 3.2.2 Characterisation of pBoLA-1.

The pBoLA-1 cDNA clone was subcloned into the *EcoRI* site of vector pPoly1 (Lathe *et al.* 1987). This vector contains the same polylinker as M13 vectors tg130 and tg131 (Kieny *et al.* 1983) and therefore facilitates subcloning for sequencing. The restriction map and the sequencing strategy are shown in figure 2, along with the complete sequence.

The cDNA translates (using the longest open reading frame) to nucleotide 1086, after this, there is 177bp of 3' untranslated (3'UT) sequence. The sequence is not full length. There is no translational start site and no polyadenylation site. Therefore the cDNA is truncated at both the 5' and 3' ends.

#### 3.2.2a Primary structural features and possible modification sites of pBoLA-1.

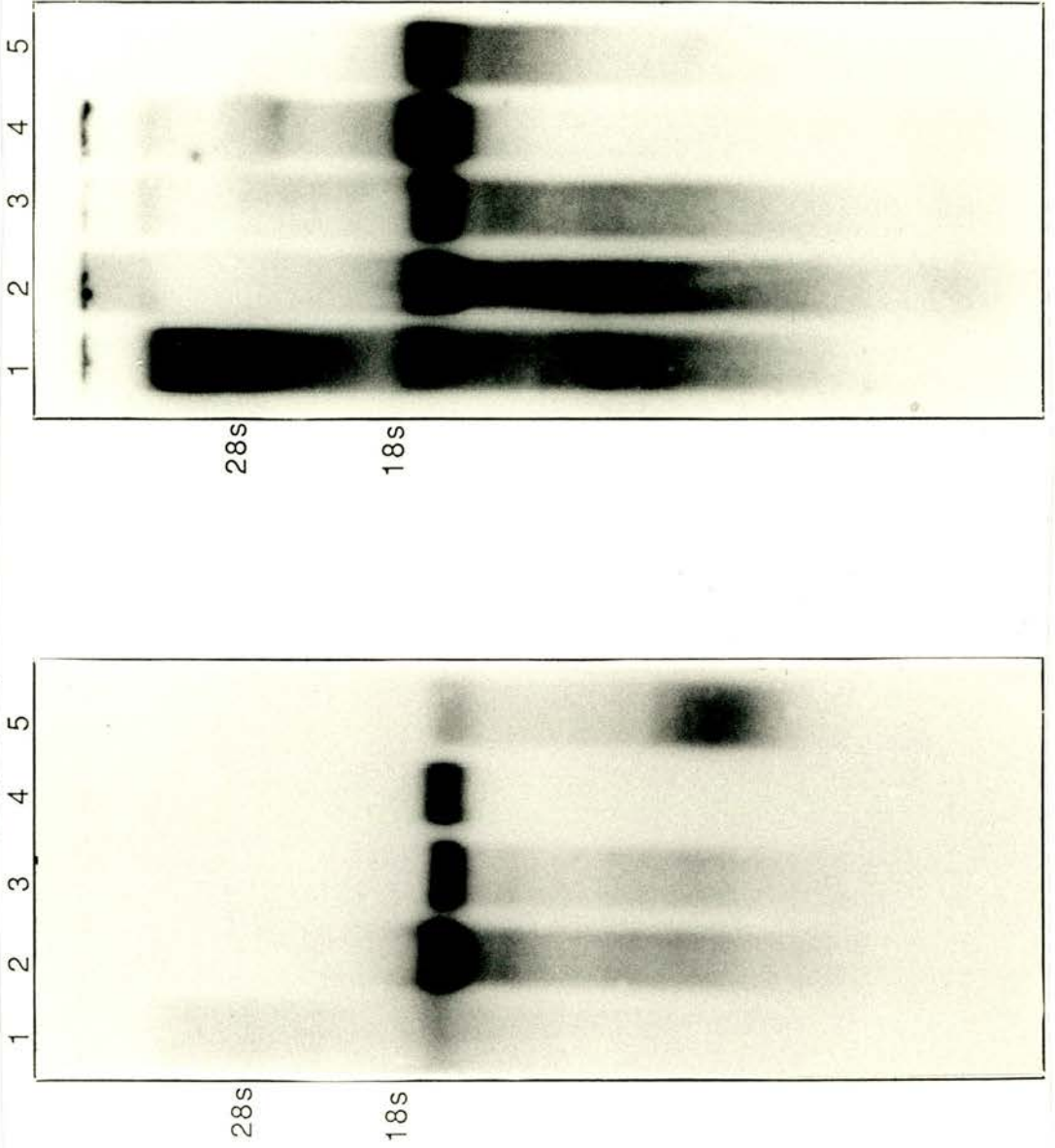
Structural features which are apparent when comparing other class I sequences also seem to be present in pBoLA-1, (figure 2).



Figure 1. Northern blot analysis.

This figure shows a Northern blot probed as two separate halves with pBoLA-1 shown on the left and pHLA-27 shown on the right. Probe amounts and specific activities were not equivalent, pHLA-27 was of a higher specific activity than pBoLA-1. The northern was washed in 2xSSC, 0.1% SDS at 65°C. The ribosomal 28s (roughly equivalent to 5kb in size) and 18s (roughly equivalent to 2kb in size) bands are indicated as reference size markers. Lane 1, 40µg mouse total liver RNA; lane 2 and 3, 40µg bovine total liver RNA; lane 4, 20ug total RNA extracted from C603 *Theileria parva* transformed lymphoblastoid cell line; lane 5, 2µg bovine poly(A)<sup>+</sup> liver mRNA (degraded in pBoLA-1 blot).

Figure 1.



### 3.2.2a1 Cysteines.

Firstly, cysteines which form internal disulphide bridges (Kimball and Coligan 1983), are present at amino acid's 101 and 164 in the second protein domain, and at amino acids 203 and 259 in the third protein domain in pBoLA-1. Amino acid 101 has previously been shown to be crucial for forming the secondary structure of the class I protein in man. Changing it from cysteine to serine by *in vitro* mutagenesis, resulted in a protein structure that was not recognised by the majority of monoclonal antibodies tested (Shiroshi *et al.* 1984). Other evidence for the importance of the cysteines, includes that of pseudogene HLA-12.4 which has phenylalanine at amino acid 259 instead of cysteine. This is postulated as the mutation which rendered the gene non-functional (Malissen *et al.* 1982), again due to loss of structural rigidity.

### 3.2.2a2 Glycosylation sites.

One possible glycosylation site, recognised as asn-X-ser/thr, is apparent in the cDNA clone at amino acid 86, indicated in figure 2. This site appears to be the only one completely conserved across species (Tykocinski *et al.* 1984; Lew *et al.* 1986a).

### 3.2.2a3 Other possible modification sites.

Other amino acids which may be targets for possible modification are at position 318 (tyrosine), this is conserved and is phosphorylated *in vitro* (Guild *et al.* 1983). Additionally, two conserved serines are present at amino acid 330 and 333. These are phosphorylated in *HLA-B7* and are identified by the consensus ser-asp/glu-x-ser(P)-leu, these sites are conserved across species but have not been shown to be of functional significance (Guild and Strominger 1984). Only one other possible post-translational

modification site exists, a cysteine at amino acid position 306 in pBoLA-1. This is present in the transmembrane domain and has been shown to incorporate a fatty acid via a thioester bond into HLA proteins (Kaufman *et al.* 1984b).

### 3.2.2b Alignment of protein domains.

Translation of pBoLA-1 into its corresponding amino acid sequence (figure 2) reveals several conserved features of class I sequences and allows identification of possible protein domains by comparison to other class I antigens.

#### 3.2.2b1 5' untranslated region and translational start site.

The 5' end of pBoLA-1 is missing, so consequently there are no 5' untranslated sequences (5'UT). Mouse class I cDNA's have a 5' untranslated region of twenty five basepairs (Lalanne *et al.* 1983). There is also no distinguishable start site for translation, signalled in eukaryotes as a methionine amino acid (ATG).

#### 3.2.2b2 Signal peptide.

Following the 5'UT region there is a signal peptide in class I proteins, the presence of this directs the attached protein to the rough endoplasmic reticulum (Walter *et al.* 1984). HLA proteins mostly have a signal peptide of length twenty four amino acids, *HLA-A3* differs with twenty nine amino acids (Strachan *et al.* 1984). Murine signal peptides appear to be shorter than this mostly being of length twenty one amino acids. The signal peptide in pBoLA-1 is of length twenty three amino acids (see figure 2) and has the features necessary to be functional; a hydrophobic core and an alanine residue -1 from the cleavage site (Von Heijne 1983). Comparing the length of

the signal peptide in pBoLA-1 to other class I proteins would suggest that this region either has one or several amino acids missing.

### 3.2.2b3 First, second and third protein domains.

H-2 and HLA class I proteins were found to subdivide into three external regions when solubilised by papain (see chapter 1, section 1.2.1b1), these were 1-90; 91-182; and 183-274 (Orr *et al.* 1979; Lopez de Castro *et al.* 1979; Nathenson *et al.* 1981). The amino acid sequence of pBoLA-1 shown in figure 2, when aligned with other MHC proteins (figure 6) shows distinct conservation at these protein domain boundaries and allows the division of pBoLA-1 into three external domains.

### 3.2.2b4 Transmembrane domain.

HLA and H-2K<sup>b</sup> proteins have hydrophobic membrane binding regions of twenty four amino acids (Ploegh *et al.* 1981), which serve to anchor the protein in the cell membrane (Ploegh *et al.* 1981). The protein encoded by *H-2Q7<sup>d</sup>* has an insertion of a hydrophilic amino acid in this transmembrane region (Steinmetz *et al.* 1981b), which renders it non-functional for membrane insertion. This protein uses a phospholipid tail as alternative means of anchoring in the cell membrane (Stroynowski *et al.* 1987).

There is an unbroken stretch of twenty six hydrophobic amino acids in pBoLA-1, which would facilitate membrane binding. In pBoLA-1 the total transmembrane region is thirty seven amino acids. This region can be separated from the cytoplasmic region in pBoLA-1 by comparing it to the intron/exon organisation of a class I genomic clone (see chapter 4). Therefore, on this basis, pBoLA-1 encodes a protein which anchors in the cell membrane via an intact transmembrane region.

Figure 2. Nucleotide sequence and amino acid translation of pBoLA-1 shown with sequencing strategy and restriction map.

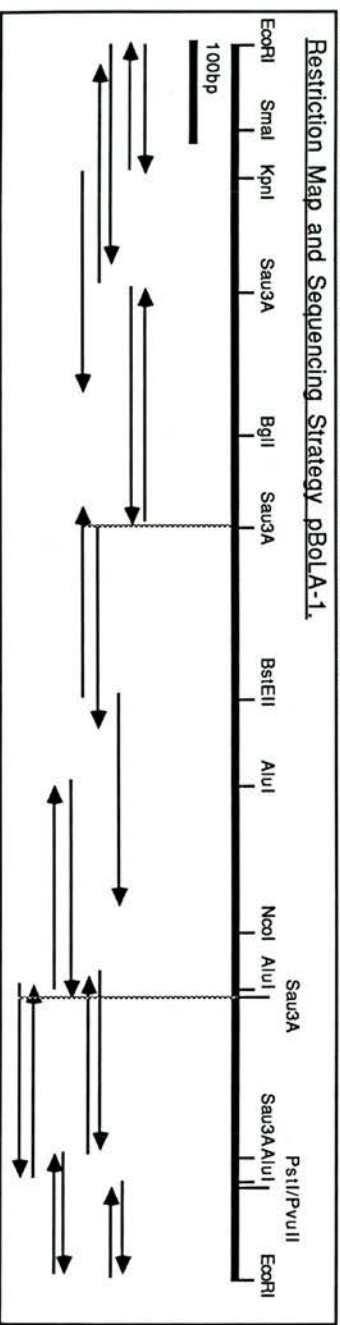
The complete nucleotide sequence with amino acid translation of pBoLA-1 are shown in this figure. 92% of the cDNA clone has been sequenced on both strands.

The putative leader peptide corresponding to the first 70bp of sequence is underlined. The probable domain structure of pBoLA-1 is also indicated above the nucleotide sequence. This comprises of three external domains ( $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$ ), transmembrane (Tm) and cytoplasmic (Cyt) regions. Cysteine residues which form disulphide bridges in class I proteins are underlined in domains two and three. The possible post-translational modification sites are also indicated by underlining. These include; an Asp N-linked glycosylation site in the first domain, a cysteine in the transmembrane domain, a tyrosine in the cytoplasmic region and a possible phosphorylation site in the cytoplasmic region.

The sequencing strategy with the relevant restriction sites is shown below the sequence.

Figure 2.

10 30 50 70 1st Domain 90 110  
 GCCTCTGAGTCTGGAGAACACAGACTCTGCTCTGGGGTCTGCTCCAGACCCGGGCTCCACTCCGATGAGGATTTTCAACCCGGCTGTCCCGGGCCGCT  
 ProLeuSerLeuGlyAtrCAsnArVAlLeuLeuLeuSerGlyValLeuAlLeuThrGluThrArAlaGlySerHisSerMetArGlyRheSerThrAlaValSerArGpRogLyLe  
 130 150 170 190 210 230  
 CCGGGAGCCCGGTACTGGAAAGTGGGTACTGTCAGCACACAGCATTCGTCGGGGTTCGACAGCCGCCCCAATCCGAGATGGAGCCCGGGGGGGTGGAGCAGAGGGCC  
 uGlyGuproArGlyrLeuGluValGlyTyValAspAspThrGlnLeuArGpHeAspSerArAlaProAsnProArGmetGuproArGlyArGTrValGlyGlnGlyPr  
 250 270 290 310 330 2nd Dom350  
 CGAGTATGGATCAGAGACCCGAAAGCCAGCCAGCCAGACACTTCCTCCGGGAACTGMACTGACCTGCTACTACACAGCAGCCGCGGTCTACACACCCCTTCACGTG  
 oGlyTyTrpAspGlnGluThrArGlyAlaIalysGlyThrAlaGlnThrPheArGlyLaAsnLeuAsnIleAlaLeuGlyTyTrpAsnGlnSerGluIalysSerHisSThrPheGlnTr  
 370 390 410 430 450 470  
 GATGTACGGCTCGACGTGGGGCCGAGGGGGCTCCGGGGGTTTCAGTACGGCTACAGCGGAGAGATTACATCCGCTGAAACAGACCTGCGCTCTGACCCGGGGCGGA  
 pHeTyrgLyCyAspValGlyProAspGlyArGleuArGlyrGlyPheMetGlnTyrgLyTrpAspGlyArGAspTyrlleAlaleuAsnGluAspLeuArGserTrpThrAlaAlaAs  
 490 510 530 550 570 590  
 CACGGGGCTCAGATCACCMAAGCCGAAAGTGGAGCCGACAGTGGAGCCGAGAGACAGAGAACTACTGAGGACAGTGGCTGGAGTGGCTCCGCAATACCTGGAGACCCGGGAAGA  
 pThrAlaIalaglnIleThrIyArGlyArGlyAlaIalaglyGluAlaGluArGlnArGAspTrpLeuGlyThrCyValGluTrpLeuArGlyrGlyrleuGlnuThrGlyLyAs  
 610 630 650 670 690 710  
 CACGCTGCTGCCCGACACCTCCAAAGCCACATGTGACCATCTCCATCTGCTGTGTCATGAGTCCACTCCGAGGCTGGGGCTGAGCTTACCCCTGAGACATCTGACTGACCTG  
 pThrLeuLeuArGlylaAspProProLyAlaHisValThrHisSerIleSerGlyHisGlyValIThrLeuArGlyTrpAlaleuGlyPheTyroGluAspIleSerLeuThrTr  
 730 750 770 790 810 830  
 GCAGCCAAATGGGAGACACAGACCCAGACATGAGCTTGTGGAGCACAGCCCTTCAGGGGAGCAGAAACTCCAGAAAGTGGGGCCGCTGGTGTGCTTCGAGAGAGCAGAAATA  
 pGlnArGAsnGlyGluAspGlnThrGluAspMetGluLeuValGluThrArGpRoserGlyAspGlyAsnPheGlnIlyTrpAlaAlaLeuValAlaProserGlyGluGlnGlnIlySTy  
 850 870 890 Tm region 910 930 950  
 CACATGCCGAGTGCACAGAGGGGCTTCAGAGGCCCTCAACCTGMAATGGAACCTCCACAGCCCTCCCTCCAGATGGGCATTCATGTTGGCCCTGCTCCCTCGTGGTCACTGG  
 rThrCyAspValGlnHisGlnGlyLeuGlnLeuProLeuThrIleuIlySTpGluProProGlnProSerPheLeuThrMetGlyIleIleValGlyLeuValLeuLeuValValThrGI  
 970 990 Cyt region 1010 1030 1050 1070  
 AGCTGTGGTGGCTGAGTGTGATCTGCATGAGAAAGCCCTCAGTGGTGAATAACAGAGGACTTATATCCAGGCTTCAAGCAGTGAAGTCCGACAGTGGCTCCAGCGTTCC  
 yAlaValAlaIalaglyValIleCySerMetIySerArGserGlyGluIyArGlyThrIyIleGlnAlaSerSerSerArAspValSerLeuThrValPr  
 1090 1110 1130 1150 1170 1190  
 TAAAGTGTGAGACACCCGGCTTGGGGGACGTGATGATCTTCATCCCGGTATGTGACATCAGATCCGACACCCCTCTTCTGCAAGCTGATGTAAATGTCTGTGTCCCTATTTGGC  
 oLyValEnd  
 1210 1230 1250  
 ATAACTAGAGACTTGGAGACTGGCCACCCATCCACTGTCTGCCCTCCACCCCTGAC



### 3.2.2b5 Cytoplasmic domain.

This region can be variable in length (Kimball and Coligan 1983) and is characteristically hydrophilic in nature (Ploegh *et al.* 1981). There are cDNA clones which have undergone alternate splicing in this region see Kress *et al.* (1983b) and Cosman *et al.* (1982) and this probably accounts for the reported heterogeneity of the 3' end of class I proteins. The cytoplasmic domain in pBoLA-1 is comprised of hydrophilic amino acids and is twenty eight amino acids in length, being of an equivalent length to HLA-A and HLA-C proteins.

This bovine cDNA clone encodes a full length mature class I protein of 339 amino acids. The clone encodes all the features associated with functional class Ia proteins.

### 3.2.2c Comparison of pBoLA-1 to other bovine class I cDNA's and to class I proteins of other species.

Since isolation of pBoLA-1 the sequence of two other class I cDNA clones has been published by Ennis *et al.* (1988). The clones BL3-6 and BL3-7 were isolated from a bovine B lymphocyte cell line of unknown serological type.

Upon detailed comparison of nucleotide and translated amino acid sequence of these three bovine cDNA clones some interesting features emerge. Figure 3 shows the comparison of the nucleotide sequence of these three clones. BL3-6 and BL3-7 share 92.8% and 89.2% nucleotide similarity with pBoLA-1. This can be compared to 83.6% similarity of HLA-A2 to pBoLA-1 and 80.5% to HLA-B7. Comparison of only the 3'UT regions of these three clones reveals homology of 89% and 88.8% respectively to pBoLA-1 and 86.2% to each other. The 3'UT region can be locus specific in man and mouse (Koller *et al.* 1984; Lew



Figure 3. Alignment of pBoLA-1 nucleotide sequence with BL3-6 and BL3-7 nucleotide sequence.

The nucleotide alignment of the three bovine cDNA clones is shown. When nucleotides are not identical the differences are indicated, black dots signify that the sequence does not extend to these regions, or that for maximum alignment a gap has been created. Putative start codons for BL3-6 and BL3-7 are indicated by shaded boxes. The common translational stop codon is indicated by a box over all three sequences.

Maximum alignment of pBoLA-1 and BL3-6 and BL3-7 (Ennis *et al.* 1988) was achieved using UWGCG computer programmes on a Vax network.



*et al.* 1986a). The high level of homology displayed by the bovine clones is expected, but the 3' UT region figures are probably too low to allow assignment of any of the sequences as alleles, where values of 93.7% to 98.3% are required in man (Ennis *et al.* 1988).

Comparison of the amino acid sequence of these clones is shown in figure 4. The cDNA's do not all have the same translational start site, BL3-7 has an ATG codon giving a leader peptide of length 24 amino acids, while BL3-6 has a leader peptide of 21 amino acids which is more reminiscent of mouse class Ia proteins. The heterogeneity of translational start sites exhibited by these clones requires some explanation. Visual comparison of the leader peptides would suggest that pBoLA-1 has a 5' end which is significantly different from the other two (see below). It may be that some opportune ligation of a spurious DNA fragment has occurred during the construction of the cDNA library, or that pBoLA-1 may have a leader peptide of 29 amino acids.

There are unique amino acids in the three clones (not seen in any other class I proteins to date) which may be considered as species specific residues. These are Phe113, Gly117 and Asn158 which are shared by all three cDNA's while Arg110 and Gly196 are shared exclusively by pBoLA-1 and BL3-6. The position of these amino acids can be compared to the regions of high amino acid variability in class Ia proteins, they all seem to occur within, or very close to the boundaries of these regions (this will be discussed in detail for pBoLA-1 in detail later). Overall amino acid homologies are 92.6% for BL3-6 and 89% for BL3-7 when compared to pBoLA-1. All the clones encode a 339 amino acid mature protein.

Figure 4. Alignment of amino acid translated sequence of pBoLA-1, BL3-6 and BL3-7 proteins.

The amino acid translation of sequences pBoLA-1, BL3-6 and BL3-7 (Ennis *et al.* 1988) are shown aligned in this figure. The sequence has been divided into domains and these are depicted as separate blocks. Regions of known high amino acid variability in other class Ia proteins are indicated at amino acids 9, 22-24, 31-32, 37-45, 65-80 in the first domain, at amino acids 95-99, 106-116, 148-157, 163-173 and 177-182 in the second external domain and lastly from 183-198 in the third domain as shaded regions. These regions are as defined by Kimball and Coligan (1983) and Nathenson *et al.* (1986).

Figure 4.

**Leader Peptide**

pBoLA-1 .PLSLGRNRVLLLSGVLVLTETRA  
 BL3-6 ..MGPRALLLLLSGVLILTETRA  
 BL3-7 MRVMPRTLLLLLSGVLVLTETLA

**First Domain**

pBoLA-1 GSHSMRYFSTAVSRPGLGEPRLYEVGYVDDTQFVRFDSAPNPRMGPRAPWVEQEGPEYWDQETRKAKGTAQTFRANINIALGYYNQSEA 90  
 BL3-6 GSHSLRYFSTAVSRPGLGEPRLYEVGYVDDTQFQFDSAPNPRMEPRARWVEQEGPEYWDNRNTRNAKNAQSFVWINTLRGYYNQSEA  
 BL3-7 GSHSLRYFYTGVSRRPGLGEPRLYAVGYVDDTQFVRFDSAPNPREEPVPMWMEQEGPEYWDNRNTRLYKDTAQIFRVDLNTLRGYYNQSET

**Second Domain**

pBoLA-1 GSHTFQWYGCDDVGPDGRLLRGMQYGYDGRDYIALNEDLRSWTAADTAAQITKRWEEAAGEAERQRYNILEGTCVEWIRRYLETGKDTLLRA 170  
 BL3-6 GSHTLQWMSGCDVGPDCALRRGMQYGYDGRDYIALNEDLRSWTAGETEAAQITKRWEEAAGYAEVQRNYLEGECEVENLRRYLIENKDKTLLRA  
 BL3-7 GSHNIQAMYGCDDVGPDGRLLRGMQYGYDGRDYIALNEELRSWTAADTAAQITKRWEEAAGAAETWRNYLEGECEVEWLRRYLIENKDKTLLRA

**Third Domain**

pBoLA-1 DPPKAHVYHHSISGHEVTLRCWALGFYPEDISLTWQRNGEDQTDMELVETRP SGDGNFQKWAALVVP SGHEQLYTCRVQHEGLQEP LTIKW 270  
 BL3-6 DPPKAHVTHHPISGREVTLRCWALGFYP EISLTWQHDGEDQTDMELVETRP SGDGTFFQKWAALVVP SGDEQRYTCRVQHEGLQEP LTLRW  
 BL3-7 DPPKAHVTHHSISDREVTLRCWALGFYP EISLTWQREGEDQTDMELVETRP SGDGTFFQKWAALVVP SGHEQRYTCRVQHEGLQEP LTLRW

**Transmembrane and Cytoplasmic regions**

pBoLA-1 EPPQPSFLTMGIIIVGLVLLVVTGAVVAGVVICMKKRSGEKRGTYIQASSSDSAQGS DVS LITVPKV 330  
 BL3-6 EPPQPSFLTMGIIIVGLVLLVVTGAVVAGVVICMKKRSGEKGGNYIQASSSDSAQGS DVS LITVPKV  
 BL3-7 EPPQTSFLIMGIIIVGLVLLV . .ALVAGAVIWRKRSGEKGRITYTQAASSSDSAQGS DVS LITVPKV

Detailed separate protein domain comparisons of the three clones to each other and to other species is presented in figure 5. Comparisons to selected HLA, selected H-2 and to the available sequence of class I proteins from SLA, RLA, HMA and RTI are also included as reference. Figure 5 illustrates that the percentage similarity of the leader peptide of pBoLA-1 when compared to bovine class I proteins is higher than to other species, but is still lower than the comparison of BL3-6 to BL3-7 (85.7%). The figures when comparing to other species are consistently low and may reflect that the leader peptide needs only to be hydrophobic. The figures however are still not inconsistent with the fact that there may be an anomaly at the 5' end of the putative leader peptide of pBoLA-1. Surveillance of the percentage similarity of the first protein domain of the selected proteins with pBoLA-1 reveals that BL3-6 shows most homology. Other proteins which show high homology for this region include PD14, RLA-1, H-2D<sup>d</sup> and BL3-7. The fact that the first domain of pBoLA-1 is more homologous to class I proteins from other species than to a comparable bovine cDNA, further illustrates the extreme variability displayed by these proteins.

Higher homologies are displayed when comparing the second and third domain of the bovine clones to other class I proteins. The second domain of pBoLA-1 shows most homology to BL3-6, closely followed by *HLA-CW1*. Ennis *et al.* (1988), when comparing the two bovine clones to other class I proteins, found that homologies throughout the bovine clones were not consistently high with any one particular class I protein. The same phenomenon seems to be occurring here (although the comparison list is not so extensive), pBoLA-1 appears to show high homology only to BL3-6 over the first two domains. In the

third domain, the bovine cDNA clones are 10% more homologous to each other than to any other class I protein compared. The probable explanation for this is that the third domain is under strong conservative influence to maintain sequences which will enable association with bovine  $\beta_2$ -microglobulin.

The transmembrane domains of pBoLA-1 and BL3-6 are the only ones to show 100% homology. This is one of the regions which can contain locus specific residues which are apparent when comparing alleles (Szots *et al.* 1986; Gussow *et al.* 1987). The 100% homology is strong evidence that pBoLA-1 and BL3-6 are alleles (the nucleotide homology for this region is 99.1%). Other protein sequences which also show reasonably high figures when compared to pBoLA-1 are PD1 and PD14, all the other proteins show consistently low figures for this domain when compared. As already mentioned (see chapter 1 section 1.2.2a and this chapter) the transmembrane region of class I proteins needs only to be hydrophobic (Ploegh *et al.* 1981). The requirement for hydrophobicity suggests that there are no structural constraints on this domain, this is usually reflected in protein sequence comparisons of the region which vary widely between species (figure 6 also illustrates this, the regions vary more than the external domains when different species are compared).

Lastly, comparison of the cytoplasmic region again shows that BL3-6 is most homologous to pBoLA-1. The cytoplasmic region of class I proteins is the converse of the transmembrane region being predominantly hydrophilic. The homology figures reflect that between species there is less conservation than within a species. Experiments have proven that this region is not essential for correct functioning

Figure 5. Percentage homologies of pBoLA-1 and other selected proteins in a domain by domain comparison.

The separate domain structures of pBoLA-1 protein indicated in figures 2 and 3 have been compared to the separate domains of other class I proteins. The protein sequence of pBoLA-1 was compared to; bovine BL3-6 and BL3-7 (Ennis *et al.* 1988), *HLA-A2* (Koller and Orr 1984), *HLA-A3* (Strachan *et al.* 1984), *HLA-B7* (Sood *et al.* 1985), *HLA-Cw1* (Gussow *et al.* 1987) from man, *H-2L<sup>d</sup>* (Evans *et al.* 1982a), *H-2K<sup>b</sup>* (Weiss *et al.* 1983), *H-2D<sup>d</sup>* (Sher *et al.* 1985) from mouse, PD1 and PD14 (Satz *et al.* 1985) from pigs, RLA-1 (Tykocinski *et al.* 1984) from rabbits, HM-1.6 (McGuire *et al.* 1986) from syrian hamster and RT1.1 (Kastern 1985) from rats.

The comparisons were done using UWGCG programmes on a Vax computer network.



Figure 5

pBoLA-1	BL3-6	BL 3-7	H1A-A2	H1A-A3	H1A-B7	H1A-CW1	H-2Ld	H-2Kb	H-2Dd	PD1	PD14	RLA-1	HM-1.6	RT1.1
Leader peptide	76.1%	69.6%	69.6%	69.6%	65.2%	65.2%	45.0%	57.0%	52.2%	57.0%	60.9%	47.8%	58.8%	-
1st Domain	90.1%	84.4%	82.2%	82.2%	80.0%	78.9%	78.9%	83.3%	85.6%	78.9%	86.9%	86.9%	82.2%	-
2nd Domain	92.3%	90.2%	82.6%	85.9%	85.9%	89.1%	80.4%	81.5%	82.6%	84.8%	85.9%	86.9%	78.3%	-
3rd Domain	97.8%	98.9%	85.9%	86.9%	84.8%	86.9%	78.2%	79.3%	82.6%	88.0%	89.1%	89.1%	82.6%	78.3%
Tm region	100%	85.7%	64.7%	64.7%	45.9%	48.6%	58.3%	43.2%	45.7%	75.7%	75.7%	56.8%	43.2%	48.6%
Cyt region	92.9%	82.1%	75.0%	75.0%	80.0%	67.8%	64.0%	64.0%	68.0%	67.9%	71.4%	56.0%	68.0%	60.7%

of class I molecules (Murre *et al.* 1984b), therefore it is likely that it is not conserved to the same degree as the external domains.

It would seem reasonable from the homology figures displayed in figure 5 to describe pBoLA-1 as an allele of BL3-6.

#### 3.2.2d The relationship of polymorphism to functional significance.

Sequence comparisons indicate that the cDNA clone shares most homology with all class Ia molecules sequenced to date, rather than the class Ib molecules, both are shown aligned in figure 6.

In figure 6 the first alignment shows the comparison of pBoLA-1 protein to all other class Ia proteins. Differences only are marked on this figure and they occur at regions where polymorphism is well documented in MHC class Ia proteins. Overall a.a 1-180 seem to be variable (Orr *et al.* 1979). This large variation in the two external domains can be further sub-divided to give clusters of variation as defined by Nathenson *et al.* (1986), (see chapter 1 section 1.2.1b3 and 1.2.1c6). These regions were previously predicted to form  $\alpha$ -helical secondary structure (Vega *et al.* 1984; Gussow *et al.* 1987) and are localised between amino acids 63-85 in the first domain and between amino acids 146-160 in the second domain. Amino acids 60-80 are postulated to be of extreme functional importance (Nathenson *et al.* 1986), while residues 146-160 are implicated in allograft rejection and cytotoxic T cell recognition (Nairn *et al.* 1980; Krangel *et al.* 1983; Lopez de Castro *et al.* 1983). The other shaded regions have all been shown to be highly polymorphic in sequence comparisons (Kimball and Coligan 1983, N'Guyen *et al.* 1985). These regions are noticeable when compared to the amino acid sequence of pBoLA-1 and

suggest that it has variation in areas of functional significance.

Recently the tertiary structure of HLA-A2 was reported (Bjorkman *et al.* 1987a). The previously predicted areas of  $\alpha$ -helix mentioned above, were found to map to the 'top' of the HLA-A2 molecule and are postulated to form the antigen binding site; capable of binding peptides of 8-20 a.a. long ( Bjorkman *et al.* 1987a, b). The similar spread of amino acid polymorphism seen when comparing pBoLA-1 protein to other class Ia proteins is a further indication that pBoLA-1 may encode a functional class Ia antigen. All amino acids conserved in the antigen binding site at amino acids 57, 58, 61, 64, 68, 72, 73, 75, 78, and 84 in the first external domain; and at positions 146, 150, 154, 155, 162, 165, and 169 in the second external domain are also conserved in pBoLA-1. It is postulated that these conserved amino acids make up one face of an  $\alpha$ -helix and are conserved to preserve the  $\alpha$ -helical configuration of the antigen binding site (Gussow *et al.* 1987). Class I proteins which have not conserved these sites include those encoded by *Tla<sup>b</sup>*, *Tla<sup>c</sup>*, *Q7<sup>b</sup>*, *Q7<sup>d</sup>*, *Q10<sup>b</sup>*, *Q10<sup>q</sup>*, clone RS5 and *HLA-328*, all of these are either pseudogenes or class Ib genes . The only difference apparent in pBoLA-1 is the substitution of an asparagine at amino acid 158 (codon=AAC) in pBoLA-1 compared to the normal alanine (codon=GCG, or GCT, or GCC). The alanine is conserved in all other sequences except *Tla<sup>b</sup>* and *Tla<sup>c</sup>* which encode threonine, *HLA-12.4* with valine and *RLA-2* with serine, all of these have not been shown to function as restriction elements. However with the recent publication of another two class I cDNA clones from cattle which also show this substitution (figure 4) it now appears that this is probably a species specific residue (Klein and Figueroa

1986).

Further alignments of pBoLA-1 to *Qa* encoded proteins of mouse and putative *Qa* encoded proteins of man and rat are presented in figure 6. The alignment shows only differences from pBoLA-1 protein and illustrates that these *Qa* proteins have similar amino acid substitution patterns, different from pBoLA-1, even when the comparison is across species. This may formally exclude pBoLA-1 from their classification eg see residues at positions 105, 110, 113 and 190. An alignment of pBoLA-1 with *Tla* encoded proteins is also shown on figure 6. It is more difficult to formally exclude pBoLA-1 from this group because the only representatives sequenced all derive from mouse, therefore there will be strong sequence conservation between the two mouse clones. However, in mouse it is known that they are more similar to each other than to any other class I proteins (Fisher *et al.* 1985; Obata *et al.* 1985). pBoLA-1 does not appear to display any particularly strong homology to these clones.

Figure 6. Alignment of pBoLA-1 protein with other class I proteins.

The alignment of pBoLA-1 with other class I proteins is shown. The alignment of pBoLA-1 is shown as three separate parts. These are split into alignment with class Ia proteins of all species and alignment with class Ib proteins. The alignment of class Ib proteins is further sub-divided into the derived amino acid sequences of the *Qa* genes of mice and *Qa*-like genes from other species and the *Tla* genes of mice.

When aligning pBoLA-1 with class Ia proteins of other species only differences are shown. The sequence has been split into the separate domain structure of class I proteins. Where no consensus was achieved an • is used instead. This is only applicable to the first three domains, thereafter this signifies a position where the sequence has been adjusted so to achieve maximum alignment. The proteins are an HLA consensus sequence (for details see below), a rabbit class Ia clone RLA-1 (Tykocinski *et al.* 1984), two pig class Ia sequences (Satz *et al.* 1985), a H-2 consensus sequence (see below for details), a rat class Ia sequence RT1.1 (Kastern 1985) and a syrian hamster HM-1.6 (McGuire *et al.* 1986) protein sequence.

Capital letters signify residues held in common in each of the derived consensus sequences. The HLA protein consensus was derived from the following sequences HLA-A2 (Koller and Orr 1985), HLA-A3 (Strachan *et al.* 1984), HLA-AW24 (N'Guyen *et al.* 1985), HLA-A28 (Lopez de Castro *et al.* 1982), HLA-B7 (Sood *et al.* 1985), HLA-B27 (Seeman *et al.* 1986, Szots *et al.* 1986), HLA-B40 (Lopez de Castro *et al.* 1983), HLA-B44 (Kottman *et al.* 1986), HLA-BW58 (Ways *et al.* 1985), HLA-CW1 and HLA-CW2 (Gussow *et al.* 1987), HLA-CW3 (Sodoyer *et al.* 1984) and HLA-328 (Srivastava *et al.* 1985).

The consensus for H-2 proteins was derived from the following sequences; H-2K<sup>b</sup> (Weiss *et al.* 1983a), H-2K<sup>d</sup> (Kvist *et al.* 1983; Lalanne *et al.* 1983), H-2L<sup>d</sup> (Evans *et al.* 1982a; Moore *et al.* 1982), H-2D<sup>b</sup> (Maloy *et al.* 1982; Reyes *et al.* 1982), H-2D<sup>d</sup> (Sher *et al.* 1985), H-2K<sup>k</sup> (Arnold *et al.* 1984) and H-2K<sup>w28</sup> (Morita *et al.* 1985).

In the alignment of pBoLA-1 with class Ib sequences, again only differences are shown. The sequences used were; H-2Q4<sup>b</sup> (Robinson *et al.* 1988), H-2Q7<sup>b</sup> (Waneck *et al.* 1987), H-2Q7<sup>d</sup> (Steinmetz *et al.* 1981b), H-2Q10<sup>b</sup> (Mellor *et al.* 1984), HLA-6.0 (HLA-G) (Geraghty *et al.* 1987) and RT1.2 (Kastern 1985). The alignment of pBoLA-1 with Tla protein sequences includes Tla<sup>b</sup> (Obata *et al.* 1985) and Tla<sup>c</sup> (Fisher *et al.* 1985).

The alignment was done using UWGCG programmes on a Vax computer network.

**First Domain.** 10 30 50 70 90  
pBoLA-1 GSHSMRYEFAVSRPGLGEPRLGVGYVDTGTVRFSDAPNPRGPRAPWVEQEGPEYWDQETKAKGTAQIFRANLNALGYNQSEA  
HLA.CON g mr l tsvsr gr EFA E f s aaSp re ap l qe pe re kkykaqaqtd vn rrlrg ea  
RLA-1 S S EFA E f s aaSp re ap l qe pe re kkykaqaqtd vn rrlrg ea  
PD1 P L S DR DS EFA E f s aaSp re ap l qe pe re NV E YGVC ELR  
PD14 P L DR EFA E f s aaSp re ap l qe pe re PM N IY G RT  
H-2.CON p l a p lge yia hte rf e YE ar m eg Ere gi gnE .f Vd Rtlr ag  
RT1.1 .....  
HMI.6 L VII W D V M ..... LE ..... M ER G DWE I VD RLP G

**Second Domain.** 110 130 150 170  
pBoLA-1 GSHTQWVGCDDVGDGRLRGFMQSYDGRDYIALNEDLRSWTAADTAQITKRKAAGEAERGNYLEGTCVEWLRRLYLGKDTLLRA  
HLA.CON thqr y dv b grll yd ya K ia n r adt a itqr sea rv egl a lecte ewlrryrlgn etlgrt  
RLA-1 T H E W A FFH YR A A A N Q HA FE e wlrlyrlgn etlgrt  
PD1 L S YL L LH YR DA A A M Q D R S Q L S M Q  
PD14 I I L L YS ET A A D R S Q L S OK M Q  
H-2.CON s l rmy dvg g L YE A d c KT a maal rr Qa a ry A e ry h Na l T  
RT1.1 .....  
HMI.6 I I S V H E L YS DA D T T R DL V HE A IE Q NR EQ T

**Third Domain.** 190 210 230 250 270  
pBoLA-1 DPPKAHVYHHSISGHEVTLCRWALGFYPEDISLTWRNGEDQTDMLVETRPSGDGNFQKWAALVVP SGEELYTCRVOHEGLQEP LTLK  
HLA.CON dp thvT pi D At a g ya E T w D dq T A gtfqk a v e r chw g pk ltlrw  
RLA-1 T PA DR A AE T T G T V R R P T  
PD1 L T TR PS DLO RE E Q S T P S H R  
PD14 E T TR PS DLO KE E Q S T P S H R  
H-2.CON s Th .rpe. A T L elt d a t sv LK . hvy . Pe R  
RT1.1 .....  
HMI.6 G T NPEGT A T K DE EL M A T VE I R HE S R

**Trn and Cyt Regions.** 290 310 330 350  
pBoLA-1 ..EPPQPSFLTMGIIVGLVLLVVTGAVVAGVVICMKKRSRGEKRGTYIQASSSDSAQGSVSLTVPKV.....  
HLA.CON epsqqtptivgavglavlvavvavvaavmcrkssggkgsysqaassdsaggsdvalta...  
RLA-1 EF A TA IV VA VLGVLILGA VA RRR HS DG G R TP AGHRD D MP...  
PD1 D A PVPVIV VL A M WR T G S T AG D KDRPV  
PD14 D PVPVIV VL A WR G S T AG KDRPV  
H-2.CON eppstvtntviiavlvlglaaivtga vaf mkrtrnt Glgqd AL PG qtsld LpDckvmvhdphsla  
RT1.1 EPS STQ NMETTV YVILGA AII A II AVVAVMKRRNTGGKGVGYAPARSDSSSDVSLSDCKA  
HMI.6 K STVPS AV AV GAVTIIG V AVVR R NTGGKRGNYAQAPDRDSAQ SD SLET..

**First Domain.** 10 30 50 70 90  
pBoLA-1 GSHSMRYEFAVSRPGLGEPRLGVGYVDTGTVRFSDAPNPRGPRAPWVEQEGPEYWDQETKAKGTAQIFRANLNALGYNQSEA  
H-2Q4b P LLS M S EFA E f s aaSp re ap l qe pe re ER D NE I V RTL S M AG  
H-2Q7b Q LQ M S W F I S N H EFA E f s aaSp re ap l qe pe re ER DJ HES GS RT QS KG  
H-2Q7d H-2Q10b Q LQ M S W F I S N H EFA E f s aaSp re ap l qe pe re ER DM HES GS RT QS KG  
HLA-6.0 P S A R F I I M EFA E f s aaSp re ap l qe pe re ER DR NE S HVS RTL H S  
RTB ..... EE NT AH D M QTLR

**Second Domain.** 110 130 150 170  
pBoLA-1 GSHTQWVGCDDVGDGRLRGFMQSYDGRDYIALNEDLRSWTAADTAQITKRKAAGEAERGNYLEGTCVEWLRRLYLGKDTLLRA  
H-2Q4b VIS E M L YL FA E KT M R R Q AT KSA A QS Q S L T  
H-2Q7b M L YL FA E KT V M R R Q I KDA A MOS Q S L T  
H-2Q7d K L YL FA E KT V M R R Q I KDA A MOS Q S L T  
H-2Q10b S I L L YE A K L KT V I R S Q A Y A A L MOS L Q S L T  
HLA-6.0 S I L L YE A K L KT V I R S Q A Y A A L NV QR A H V M Q P  
RTB ..... E P

**Third Domain.** 190 210 230 250 270  
pBoLA-1 DPPKAHVYHHSISGHEVTLCRWALGFYPEDISLTWRNGEDQTDMLVETRPSGDGNFQKWAALVVP SGEELYTCRVOHEGLQEP LTLK  
H-2Q4b TC HR DGI AN I L EL T A T SV L K N H H P R  
H-2Q7b T PR YGA A T L EL T A T SV L K N H H P R  
H-2Q7d T PR YGA A T L EL T A T SV L K N H H P R  
H-2Q10b T T PG EGD P A T L EL T Q A T SV L K N H Y P R  
HLA-6.0 T T PVFDY A AE I D V A T V R H P M R  
RTB ..... TI PRPEGT A T L L L A T E V L K N L E P R

**Transmembrane and Cytoplasmic Regions.** 290 310 330  
pBoLA-1 EPPQPSFLTMGIIVGLVLLVVTGAVVAGVVICMKKRSRGEKRGTYIQASSSDSAQGSVSLTVPKV.....  
H-2Q4b PSTVSN ANVAV V GAWPSLQLWLL.....  
H-2Q7b GRWE PPY VSNMATIAVV DL IIGAVVAVFMNRR.....  
H-2Q7d GRWE PPY VSNMATIAVV DL AIIIGAVVAVFMNRR.....  
H-2Q10b PSTDSI SH AD LWPSLKLWYL.....  
HLA-6.0 K. S LP IP MGIVAG LA T AAVAV WRKSSD.....  
RTB ..... SKEGAFYRP.....

**First Domain.** 10 30 50 70 90  
pBoLA-1 GSHSMRYEFAVSRPGLGEPRLGVGYVDTGTVRFSDAPNPRGPRAPWVEQEGPEYWDQETKAKGTAQIFRANLNALGYNQSEA  
Tlab L L AIS W I A L EFA E f s aaSp re ap l qe pe re AR EIVTSN F E Q T M D L GN  
Tlac L L AIS W I A L A A E T G T Y KLS AR EIVTSN F E Q T M D L GN

**Second Domain.** 110 130 150 170  
pBoLA-1 GSHTQWVGCDDVGDGRLRGFMQSYDGRDYIALNEDLRSWTAADTAQITKRKAAGEAERGNYLEGTCVEWLRRLYLGKDTLLRA  
Tlab V E E F S F AYE H Q KT E RS Q Y T LR T RDS L NR Q ECT  
Tlac V E E F S F AYE H Q KT M E RS Q Y T LR T RDS L NR Q ECT

**Third Domain.** 190 210 230 250 270  
pBoLA-1 DPPKAHVYHHSISGHEVTLCRWALGFYPEDISLTWRNGEDQTDMLVETRPSGDGNFQKWAALVVP SGEELYTCRVOHEGLQEP LTLK  
Tlab T T PRPEGY R A T L ELI T A T V L K K H Y P R  
Tlac T T ARPEGD AH T L ELI T A T V L K K H Y P R

**Transmembrane and Cytoplasmic Regions.** 290 310 330  
pBoLA-1 EPPQPSFLTMGIIVGLVLLVVTGAVVAGVVICMKKRSRGEKRGTYIQASSSDSAQGSVSLTVPKV.....  
Tlab S MPNRTTVRA LGAMILLGFS S .. MWMRKNNNG NGDDNTAAAYQNEREHL SG..  
Tlac S MPNRTTVRA LGAMILLGFS S .. MWMRKNNNG NGDDNTAAAYQNEREHL SG..

### 3.2.3 pBoLA-2.

The second smaller cDNA clone, isolated in  $\lambda$ gt10, would not subclone into a more manipulable vector. First attempts met with no result in vectors pUC18 (Yanisch-Perron *et al.* 1985), pPoly1 (Lathe *et al.* 1987), M13 vectors tg130 and tg131 (Kieny *et al.* 1983), when cloning as a complete 0.6 kb cDNA clone. Cloning with small flanking regions of  $\lambda$ gt10 still attached to the cDNA also produced no result in pUC18 and pPoly1. Restriction mapping of the cDNA gave internal *PstI*, *PvuII* and *SmaI* sites. Directional cloning using these sites into M13 tg130 and tg131 realised only one type of clone flanked by *EcoRI* and *PstI* restriction sites, which was approximately 0.08 kb in length.

When sequenced by Sanger dideoxy sequencing (Sanger *et al.* 1977), pBoLA-2 was found to correspond to sequences at the 3' end of the large cDNA clone pBoLA-1, see figure 7. There is one nucleotide difference between the two clones at position 1151, where a transversion of A->G has occurred. Although only a small region of pBoLA-2 has been characterised, one nucleotide difference from pBoLA-1 would suggest that it corresponds to another mRNA transcript or may possibly be an error of reverse transcription. Szots *et al.* (1986) described one cDNA clone corresponding to *HLA-B27*, which had a stop codon at amino acid 274, this would give rise to a greatly truncated protein. Other clones which overlapped this region were found to code for an amino acid at this site and so it was concluded that it had been an error in reverse transcription. The complete clone pBoLA-2 maps to the 3' end of pBoLA-1 (figure 7) and could extend the pBoLA-1 cDNA clone by 400bp, this would give a transcript size of 1.7kb the approximate size of class Ia transcripts.

Figure 7. Alignment of pBoLA-1 and pBoLA-2.

The nucleotide sequences of pBoLA-1 and pBoLA-2 are shown aligned. The sequenced area of pBoLA-2 is indicated. Vertical lines connect the two sequences indicating where a match occurs. The mapped restriction sites and the orientation of the sequenced area of pBoLA-2 are also shown.

The alignment and arrangement were carried out using UWGCG programmes on a Vax computer network.

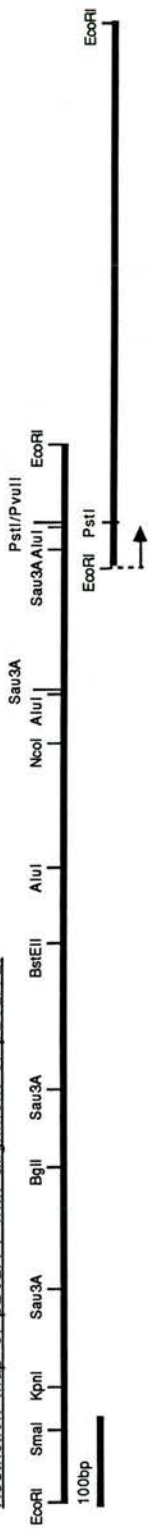


Figure 7

pBoLA-1  
 10 GCCTCTGAGTCTGGGAAGAACAGAGTCTTCTCGGGGGTCTGTGGTCTGACCGAGACCCGGGCTGGCTCCACTCGATGAGGTATTTTCAGCACCGCCGCTGCCGGGAGCCCGCGGTACCTGGAAAGTCCGGCTA 150  
 30 70 110 130 150  
 170 CGTGACGACACCGCAGTTTCGTGCGGTTTCGACAGCAGCCGCGGAAATCCGAGGATGAGCCCGGGCCGGGTGGGTGGACAGGAGGGCCCGAGTATTGGGATCAGAGAGACGCGAAAGGCCAACGCGCACAGACTTTCGGGGGAA 230  
 190 210 230 250 270 290  
 310 CCTGAAACATCGCACTCGGCTACTACMACCAGAGCGAGGGCGGGTCTCACACCTTCCAGTGTACGGCTGCGAGCTGGGGCCGGACGGGCGCTTCCGGCCGGGGTTTCATGCAGTACGGCTACGACGGGACAGATTACATCGGCCCTGAA 430  
 330 350 370 390 410 430 450  
 470 CGAGGACCTGGCTCTGGACCGCGGGGACACGGCGGCTCAGATCACCAAGCGCAAGTGGGAGGGCGGAGGTAGGGCGGAGAGACAGAGGAACTACTCTGGAGGGCACGTTGGAGTGGCTCCGGCAGATACCTGGAGACCGGGGAGGA 530  
 490 510 530 550 570 590  
 610 CACGCTGCTGCGCGCAGACCCCTCCAAAGGCAATGTGACCCATCACTCCATCTCTGTCATGAGGTACCCCTGAGGTGAGGTGGGCGCTGGGCTTCTACCCCTGAGGACATCTCACCTGGCAGCGCAATGGGGAGGACACGACCCAGGA 690  
 710 730 750  
 770 CATGGAGCTTGTGGAGACCCAGGCTTCAGGGGACGGAACTTCAGAAAGTGGGGGCGCTGGTGGTCCCTCTGGAGAGGCAAAATACACATGCCGAGTGCAGCACGAGGGGCTTCAGGAGGCCCTCACCCCTGAAATGGGAACCTCC 890  
 910 930 950 970 990 1010 1030 1050  
 TCAGCCCTCCTCCCTCACCATGGGCATCAATTGTGGCTGGTTCTCTCGTGTCACTGGAGCTGTGGCTGGAGTTGTGATCTGCATGAAAGAGCGGCTCAGGTGAAAGAGGGGACTTATATCCAGGCTTCAAGCACTGACAGTGC  
 1070 1080 1100 1110 1130 1150 1170 1190  
 CCAGGGCTGTGATGTCTCTCACGGTTCCTAAAGTGGTACACCGGGCTTGGGGGACTGAGTGTGATGCTTCAATCCCGGCTATFGACATCAGATCCCGACCCCTCTTTTGGGCTTGGCATGTGAAATGTGTCTGTGCCCTTATTCGC  
 1210 1230 1250  
 ATAAAGTGGAGTTAGGAGACTGGCCACCCACCTGCCCCACTGCTGCCCTCCCCACCCCTGACC

pBoLA-2  
 1210 1230 1250  
 CCGGGCTTGGGGGACTGAGTGTGATGCTTCAATCCCGGCTATGTGACATCAGATCCCGGACCCC

Restriction Map of pBoLA-1 with alignment of pBoLA-2.



100bp

Unfortunately since only a small portion has been sequenced the exact nature of sequences 3' can only be surmised. Other class I cDNA clones have been shown to encode long 3' UT regions with insertions of B2 SINE elements in mice (Lalanne *et al.* 1982; Kress *et al.* 1984). B2 SINE elements are dispersed repetitive elements equivalent to type 2 Alu-like repeats found in man, they are flanked by short repeats but are bordered on 3' side by an A-rich region (see Jelineck and Schimid 1982, for fuller description). Genomic clones of human class I sequences have also been shown to contain SINE elements in their 5' and 3' regions (Geraghty *et al.* 1987), introns (Koller *et al.* 1988) and in the 3' UT region only (Srivastava *et al.* 1987). In an attempt to establish if any class I sequences were present in the remaining larger *PstI/EcoRI* fragment, probes were made from agarose gel purified insert restricted with *PstI*. These consist of the 0.08 kb fragment characterised, and the larger 0.52 kb fragment. The probes were used separately against bovine RNA in a Northern blot (see figure 8). Both fragments were found to recognise the same size band of approximately 1.7kb perhaps arguing against insertion of a SINE element. But, the larger fragment did seem to hybridise as a smear, suggesting that it was detecting a heterogeneous population of transcripts. If these are transcripts, then they appear to be concentrated in the smaller size range of of the poly(A)<sup>+</sup> track. This would not be inconsistent with Kramerov *et al.* (1982) who reported that B2 repeats are present as transcripts of size 200-400 nucleotides long in poly(A)<sup>+</sup> RNA in mice and other rodents. They also noted that these transcripts are more numerous in tumour cell lines. Unfortunately, a simpler explanation for the result may be that the RNA is degraded. This could be resolved by reprobing the Northern

with a different probe, although this has not yet been carried out. Therefore conclusions about the cDNA clone cannot be drawn, except that this 3' region still detected RNA at the correct size for class I transcripts.

Figure 8. Northern blots probed separately with sub-fragments of pBoLA-2 and human cDNA clone pHLA-27.

A 1.5% formaldehyde gel was ran with triplicate sets of samples. This was Northern blotted and probed as three separate parts. The first panel was probed with the human cDNA pHLA-27, the second with a 0.08kb *EcoRI/PstI* gel purified fragment of pBoLA-2 and the third panel with the large remaining *PstI/EcoRI* fragment of pBoLA-2. The probe specific activities were not the same. The samples are; track 1, 30µg total bovine liver RNA, track 2, 2µg poly (A)<sup>+</sup> bovine liver RNA and track 3, 20µg of C603 total RNA isolated from a *Theileria parva* transformed lymphoblastoid cell line.

pBoLA-2  
0.52kb



pBoLA-2  
0.08kb

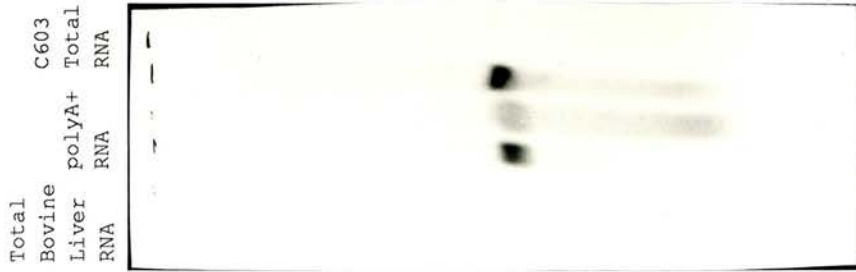
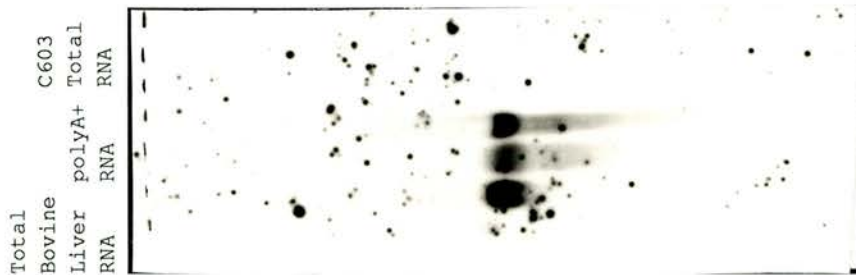


Figure 8

HLA-27



### 3.3 Summary.

This chapter describes two bovine class I cDNA clones which were isolated from a bovine liver cDNA library.

The cDNA clone, pBoLA-1, has all the features of a class Ia protein. Firstly, comparison to other class I proteins reveals that pBoLA-1 protein appears to be more similar to class Ia proteins and not to class Ib. This was deduced from comparison of its encoded amino acid sequence to other class I proteins. It appears to encode a full length mature class I transcript of 339 amino acids. pBoLA-1 may be an allele of another published bovine cDNA clone BL3-6 due to the 100% amino acid homology that the two clones share in the transmembrane domain. These two clones share a number of amino acid residues which seem to be species specific, the most notable of these is Asn158 which occurs on the conserved face of one of the  $\alpha$ -helices which make up the putative antigen binding site of class Ia molecules (Bjorkman *et al.* 1987a,b).

The second cDNA clone isolated, pBoLA-2, maps to the 3' end of pBoLA-1 and may extend this clone to a length of 1.7kb. pBoLA-2 was not further characterised by sequencing, only by hybridisation to a Northern blot. The pattern of hybridisation obtained suggests that it may encode a repetitive element.

## Chapter 4.

### Fine structural analysis of two bovine class I genomic clones.

#### 4.1 Introduction.

Probing Southern blots of both mouse and human DNA with class I cDNA probes revealed numerous cross-hybridising sequences (Steinmetz *et al.* 1981a; Biro *et al.* 1983). This led to the conclusion that class I genes in mouse and man were part of a multi-gene family (Pease *et al.* 1982; Orr and Demars 1983). Detailed characterisation of the class I region in BALB/c mice revealed 35 class I genes (Stephan *et al.* 1986) and 26 in C57BL/10 mice (Weiss *et al.* 1984). Detailed characterisation of one gene from the BALB/c library by Steinmetz *et al.* (1981b) revealed a sequence which was most probably a pseudogene. Likewise, Malissen *et al.* (1982) and Biro *et al.* (1983) also characterised human class I genomic clones which were pseudogenes.

In an effort to circumvent these problems, investigators started to use a functional test to identify class I genes. Goodenow *et al.* (1981) identified  $L^d$  by transfection of cosmid clones into mouse L-cells and screening for cell surface expression. Therefore the functional  $H-2L^d$  gene was correctly identified before being fully characterised by Moore *et al.* (1982). In follow up experiments, Goodenow *et al.* (1982a) used cell transfection to identify a further two functional class Ia genes in the cosmid library made from BALB/c mouse DNA. Investigators of the HLA complex also used transfection to identify clones of interest. Strachan *et al.* (1984),

Sodoyer *et al.* (1984) and N'Guyen *et al.* (1985) used cell transfection to identify the *HLA-A3*, *HLACw3* and *HLA-AW24* genes. As discussed in chapter 1 section 1.3, investigators of other species utilised the information provided by studies of mouse and man and applied it to their systems. <sup>For example</sup> Satz *et al.* (1985) isolated strongly hybridising positives from a pig cosmid library which they transfected into cells and identified PD14 (a class Ia pig gene). The unifying theme throughout these experiments was the use of cell transfection to identify functional genes of interest.

Unfortunately, due to a shortage of time a cell transfection assay for cattle class Ia genes was not set up. Therefore alternative methods of identifying possible functional class Ia genes in a bovine genomic library had to be found. One way was to look for conservation of restriction enzyme sites between cDNA's and their corresponding genomic clone. This has already been used effectively by Marche *et al.* (1985). They isolated the rabbit genomic clone equivalent to the rabbit cDNA clone pR9 (Tykocinski *et al.* 1984), by analysing genomic clones for a diagnostic restriction fragment internal to the pR9 cDNA. Analysing clones for conservation of restriction sites was also used by Szots *et al.* (1986) to classify class Ia cDNA clones into groups of *HLA-A*, *-B* and *-C* in man.

This chapter describes the use of pBoLA-1 as a probe against a bovine genomic library constructed by Dr Jean-Luc Vilotte and the subsequent characterisation of two of the genomic clones selected. Previous experiments by Vaiman *et al.* (1986), Lindberg and Andersson (1988) and data presented in chapter 5 of this thesis showed that numerous class-I-hybridising bands were detected using a heterologous probe, this was also true of a homologous probe (chapter



5 only). It therefore seems reasonable to assume that class I genes in cattle are part of a multi-gene family and that the number of cross-hybridising class I sequences in a genomic library will be large. It was hoped that the advantage of using a homologous probe would help classify these cross-hybridising sequences. The isolation strategy relied on using sequence homology between the class Ia cDNA clone pBoLA-1 (see chapter 3) and its genomic counterpart as a distinguishing feature. Only strongly hybridising positives were to be selected from the genomic library. The strongly hybridising positives were then to be further classified by restriction mapping them and looking for restriction sites which were shared with pBoLA-1.

## 4.2 Results and Discussion.

### 4.2.1 Initial analysis of two genomic clones.

A genomic DNA library had been constructed using DNA obtained from fresh peripheral blood lymphocytes from animal 10160 and the phage vector EMBL 301 (Lathe *et al.* 1987), by Dr Jean-Luc Vilotte (see Lathe *et al.* 1987). The library of size  $2.6 \times 10^5$  p.f.u., representing 65% of the bovine genome, was amplified before being screened by Dr Jean-Luc Vilotte with the complete cDNA clone pBoLA-1. Twelve of the strongest hybridising positives were picked by Dr Vilotte and rescreened until purified as single plaques. Mini-plate lysates were made of these twelve clones and a small amount of phage DNA was extracted (chapter 2, section 2.18j1). Mapping of the clones with the restriction enzyme *EcoRI* revealed that only two types of clones were

present. Isolation of identical phage clones was probably due to the fact that the library had been amplified before screening, since amplification can lead to over-representation of some clones in a library.

One of each type of clone was picked and grown up as a bulk culture (Blattner *et al.* 1977). These two clones, phage 33 and phage 41, were mapped by single and double restriction digests to yield the maps shown in figure 1. Phage 33 has an insert of approximately 16kb, phage 41 has an insert of approximately 20kb. The regions hybridising to pBoLA-1 were mapped by Southern blotting the gels used to construct the restriction maps (see figure 2 for example of how this was done). Hybridising fragments are indicated on figure 1 under the restriction maps. On digestion with *EcoRI* and *Sall*, phage 33 yielded an internal fragment of 5.4kb which hybridised to the cDNA. In contrast phage 41 yielded two non-contiguous fragments which hybridised to the cDNA. The two fragments could be roughly located within the phage clone by *KpnI* digestion and were finely mapped by *PvuII* and *EcoRI* digestion (figure 1). The presence of two non-contiguous class I-hybridising regions in phage 4 suggests that the clone contains two distinct class I genes, separated by approximately 15kb. This close proximity of class I genes has also been observed in mouse where *K* genes and *Qa* region genes are each separated by 15kb (Weiss *et al.* 1984) and in man, the *HLA-AW24* is flanked by other class I genes (N'Guyen *et al.* 1985).

The continuity of two phages was checked by comparing hybridisation patterns of restriction digests to hybridisation patterns obtained from equivalent restriction digests of DNA from animal 10160 (the DNA used to make the library) with the same restriction

Figure 1. Restriction maps and areas of hybridisation of phages 33 and 41.

This figure shows the restriction sites mapped in phages 33 and 41. The map for phage 41 shows sites which are indicated <sup>above and</sup> below the phage insert. <sup>have been</sup> The restriction sites indicated above the insert <sup>mapped at the 5' ends.</sup> are <sup>mapped</sup> throughout the phage. <sup>mapped</sup> While those below have only been <sup>mapped</sup>. The areas which hybridised to pBoLA-1 are indicated below the phage insert as cross-hatched bars.

The map for phage 33 shows sites which were mapped in this phage in the large genomic clone. The sub-clone *EcoRI/Sall* derived from phage 33 has a more detailed restriction map (not all of these enzymes were used to map the large clone). The regions labelled A, B, C, D, E and F in the *EcoRI/Sall* subclone are have been partially sequenced. The restriction enzymes sites were mapped using procedures outlined in Maniatis *et al.* (1982).

Figure 1

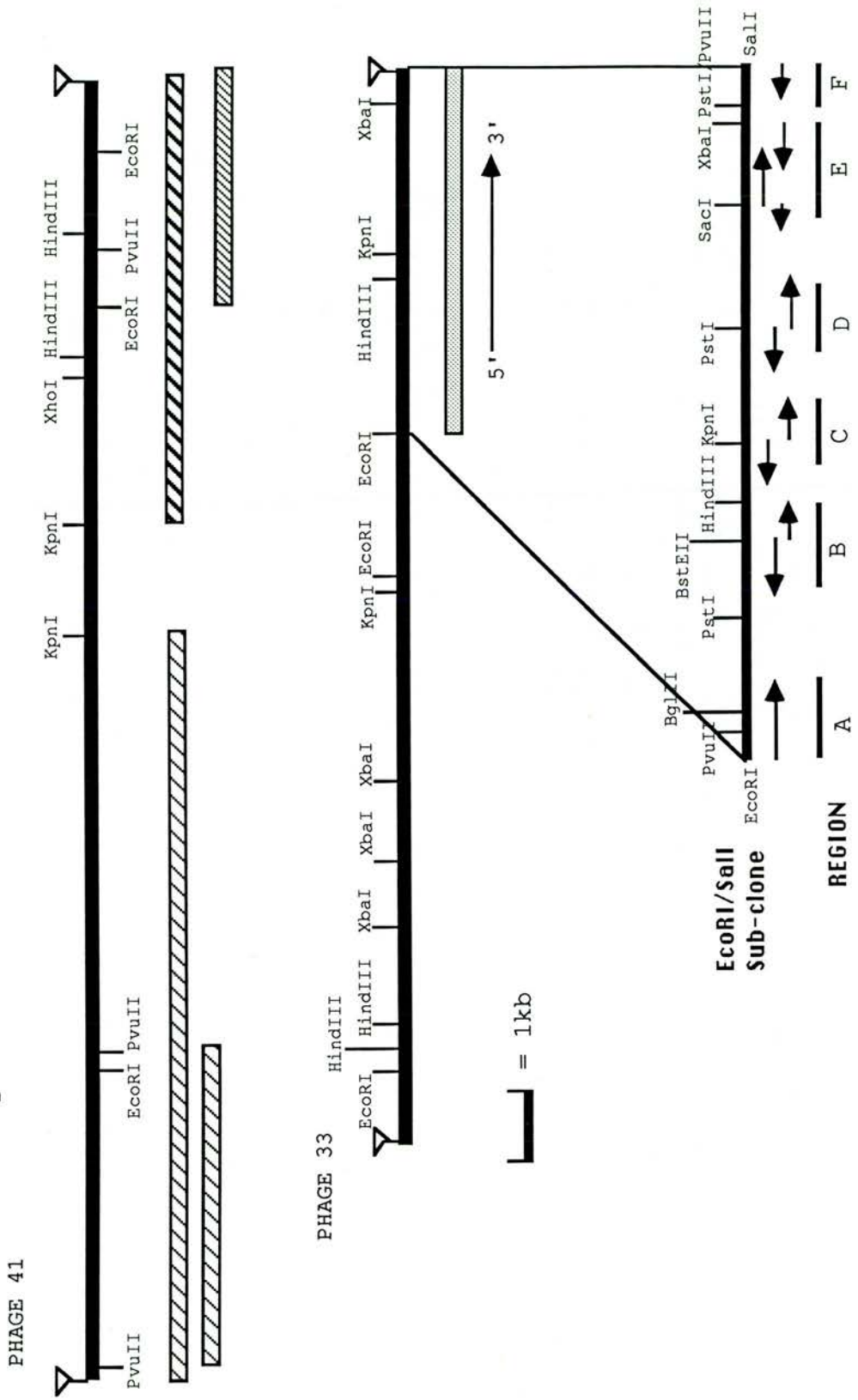
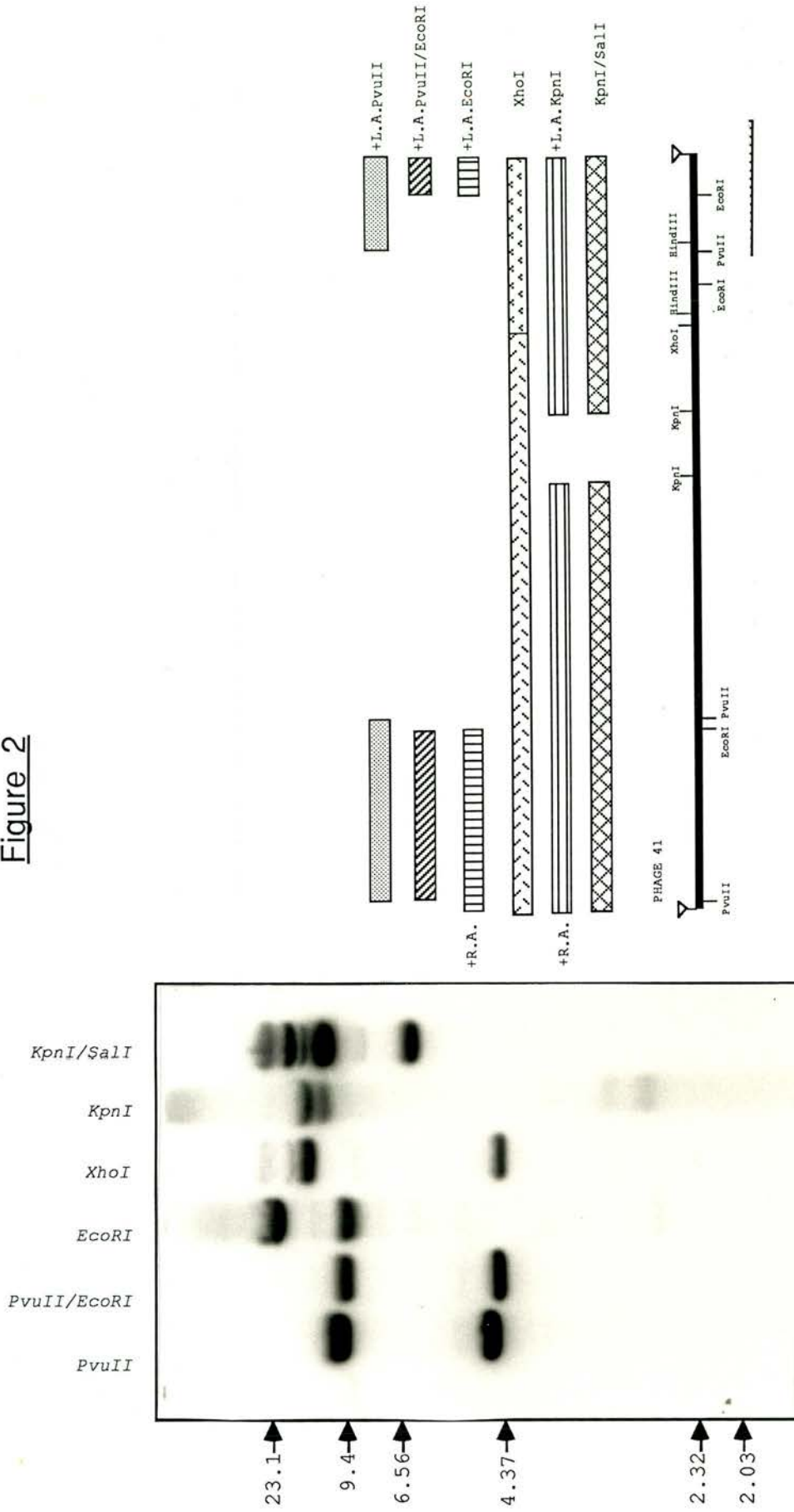


Figure 2. Example of how phage maps depicted in figure 1 were constructed.

The figure depicts how the restriction sites were mapped for phage 41. The left hand panel shows a photograph of a Southern blotted gel which has been probed with pBoLA-1. Therefore, only hybridising bands are visible. The right hand panel shows how these hybridising bands relate to the phage insert. The restriction enzymes used are listed from left to right on the top of the Southern, these relate to the bars listed above the restriction map from top to bottom. A few faint bands were detected on the Southern blot which are probably spurious, these occur in the *EcoRI* lane and the *KpnI* lane. In addition, the *KpnI/SalI* digest is a partial, with larger incomplete digests being detected by the probe.

Figure 2



enzymes. Figure 3 shows the results of this comparison. Tentative conclusions can be drawn from this exercise. Internal hybridising bands are generated by *PvuII* digestion of phage 33 and *EcoRI/XbaI* digestion realises a band of 4.8kb (see figure 3) which appears to correspond to an equivalent band in the genome of animal 10160. The intensity of the hybridisation signal would suggest that this band is present as a single copy in the genome and, therefore, that phage 33 corresponds to a single gene.

The result for phage 41 was not so clear. Hybridisation of an *XbaI/EcoRI* restriction digest to pBoLA-1 revealed three internal fragments of sizes 2.4kb, 1.3kb and 0.1kb. The 2.4kb fragment can be accounted for by internal *EcoRI* sites in the phage. Unfortunately the *XbaI* sites have not been mapped in this phage, but there are no sites in the vector (Lathe *et al.* 1987). The intensity of the hybridisation signal obtained from the comparable fragments in 10160 suggest that they are <sup>probably</sup> present as 1 copy in the genome.

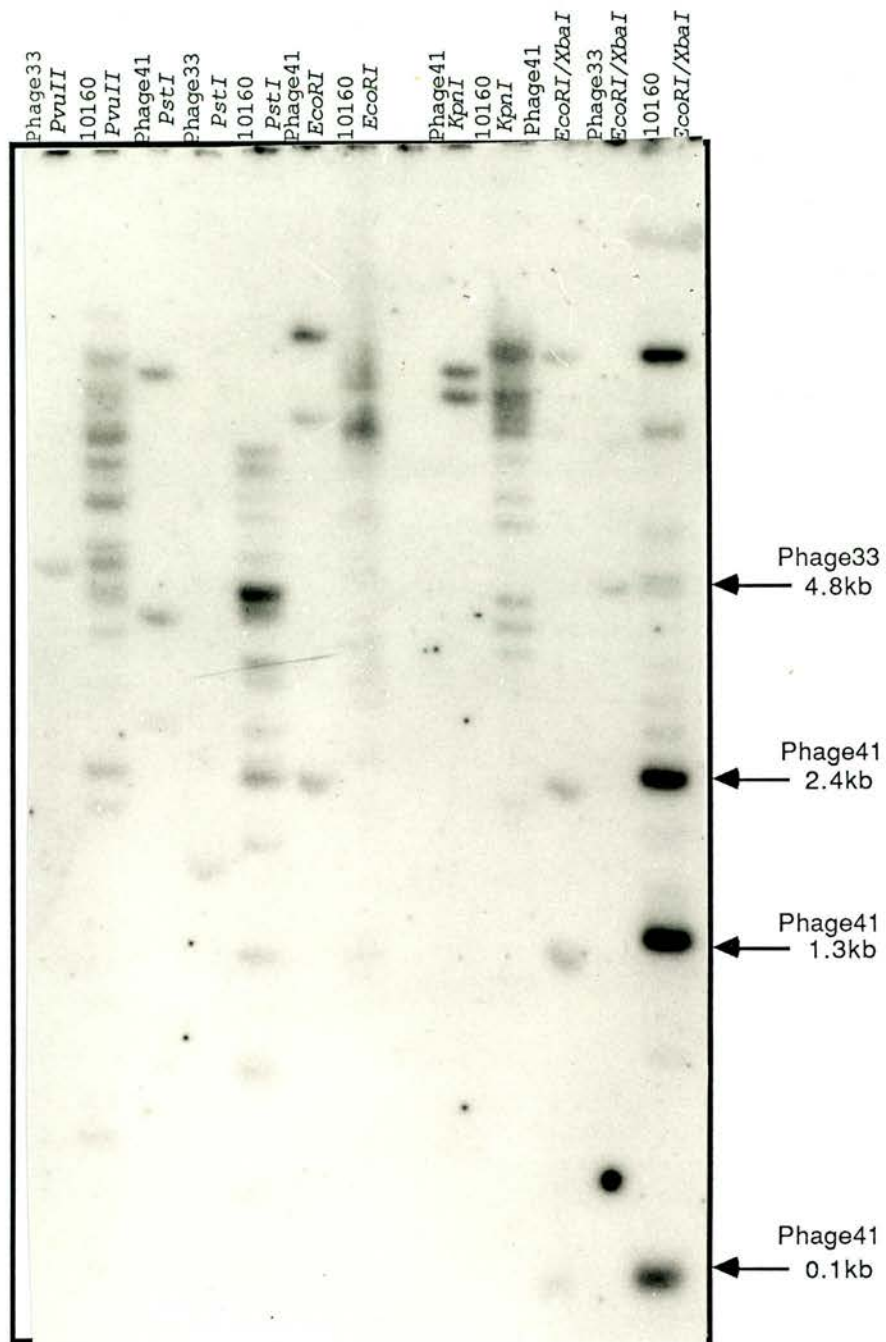
Figure 3. Southern blot of restriction digests of animal 10160 and phage 33 and 41.

This figure shows the results of hybridising pBoLA-1 to a Southern blot of DNA extracted from animal 10160 with phage DNA.

10µg aliquots of DNA from animal 10160 were digested with *PvuII*, *PstI*, *EcoRI*, *KpnI* and *XbaI/EcoRI*. Phage DNA of clones 33 and 41 were also digested with the same restriction enzymes. The restriction digests were then grouped according to the enzyme used and run on an 0.8% agarose gel and Southern blotted. The amount of phage DNA loaded onto the gel in each lane was equivalent to one copy of a gene 5kb in size, in addition to this 10µg of salmon sperm DNA was added as a carrier. As a result of using carrier DNA, the phage bands appear to migrate marginally faster than those of the genomic DNA. The blot was probed with pBoLA-1 and washed to high stringency in 0.1xSSC 0.1%SDS at 65°C. The autoradiograph was exposed with intensifying screens at -70°C for 7 days.



Figure 3



#### 4.2.2 Detailed analysis of phage 33.

As described above, Southern blotting experiments suggest that phage 33 corresponds to a single copy gene. Further experiments were undertaken to determine if the clone comprised a complete gene.

Restriction enzyme digests of phage 33 were probed with separate 5' and 3' specific probes prepared from pBoLA-1 (as illustrated in figure 4). Both the 5' and 3' probes hybridised to a 5.5kb *EcoRI / Sall* fragment. After *XbaI* digestion of the phage, both 5' and 3' probes hybridised to a 10kb doublet. The restriction map shows that two 10kb fragments can be generated by *XbaI* digestion and can be separated by a double digest of *EcoRI / XbaI*. This revealed that the 5' probe hybridised to a 4.3kb fragment, while the 3' probe still hybridised strongly to the 10kb fragment and weakly to the 4.3kb fragment. This enabled orientation of the phage clone with the 5' region roughly mapping between *KpnI* and *SacI* and the 3' region mapping 3' of the *XbaI* site. The *EcoRI/Sall* fragment of phage 33 was sub-cloned into the vector pPoly1 (Lathe *et al.* 1987) for easier manipulation.

The procedure described above was repeated for restriction digests of the *EcoRI/Sall* sub-clone of phage 33. This further localised the 3' end to a *PstI* fragment at the right end of the insert (see figure 5). The exact location of the 5' end was not better defined, other than it mapped 3' of the *KpnI* site (this could be inferred because the 0.75kb *BstEII/KpnI* fragment was not detected by hybridisation). Unfortunately, apart from hybridising to fragments specific to the sub-clone, the 5' probe seems to show appreciable cross reaction with pPoly1 vector sequences and it also appears to mimic the hybridisation pattern obtained with the 3' probe. This cross-reaction

probably results from some entrapment of 3' sequences and pPoly1 DNA chains by the 5' cDNA fragment during the gel isolation procedure. With hindsight, the conclusion which should have been reached from this experiment is that this clone lacks a 5' end. But, the hybridisation results presented in figure 4 contradict this because the hybridisation patterns of the two probes differ suggesting the presence of a 5' end. Overall, when comparing the two probes the strength of hybridisation was much stronger at the 3' end. This would be expected because the 3' region used as a probe corresponds to 3'UT which is very conserved in class I genes (Strachan *et al.* 1984). The weakly hybridising 5' end and strongly hybridising 3' end could only be further characterised by sequencing the sub-clone.

Figure 4. Orientation of phage clone 33 using 5' and 3' probes derived from pBoLA-1.

DNA from phage 33 was digested with the named restriction enzymes and fractionated on an agarose gel. Duplicate sets of samples were loaded on the same gel. This gel was Southern blotted then halved. One half was probed with a 5' *EcoRI/BglI* fragment of pBoLA-1 (corresponding to the leader peptide and first two domains of pBoLA-1 protein) and washed to low stringency in 2xSSC, 0.1% SDS at 65°C, this is panel A. The second halve was probed with a *EcoRI/PstI* fragment of pBoLA-1 (corresponding to the 3'UT region) and washed to high stringency in 0.2xSSC, 0.1% SDS at 65°C this is panel C. Panel B shows a higher stringency wash (0.5xSSC, 0.1%SDS at 65°C) of a duplicate gel probed with the 5' fragment of pBoLA-1. In each case the size marker was a *HindIII* digest of  $\lambda$  phage whose sizes are shown on the left of each figure.

Figure 4

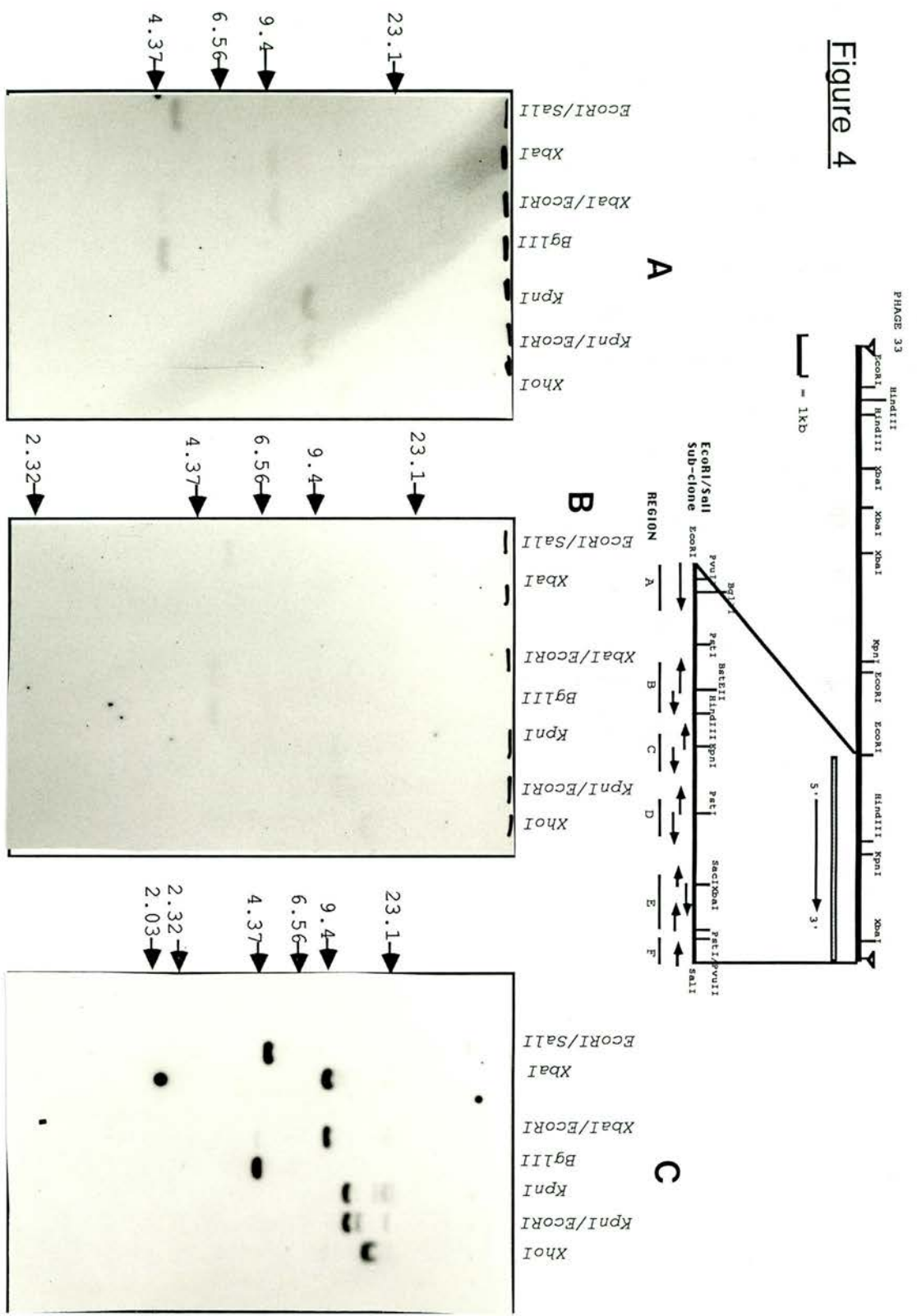


Figure 5. Fine mapping of phage 33: localisation of 5' and 3' end in the *EcoRI/SalI* subclone.

Once the *EcoRI/SalI* hybridising fragment of phage 33 was subcloned, the 5' and 3' ends were localised to more defined regions. This figure shows the results obtained from the analysis of the subclone.

Restriction digests of the subclone were loaded in duplicate, run overnight and Southern blotted. The membrane was halved, one half being probed with the 5' probe, the other probed with the 3' probe. Panel A shows the results obtained when a 3' probe (*EcoRI/PstI* fragment) obtained from pBoLA-1 was used as a probe. The Southern was washed to high stringency in 0.2xSSC, 0.1% SDS at 65°C. Panel B shows the results when a 5' probe was hybridised and washed at low stringency in 2xSSC, 0.1%SDS at 65°C. Unfortunately each of the probes were prepared from pPoly1 (Lathe *et al.* 1987) and there appears to be some contamination with plasmid because vector sequences are hybridised to in each case.

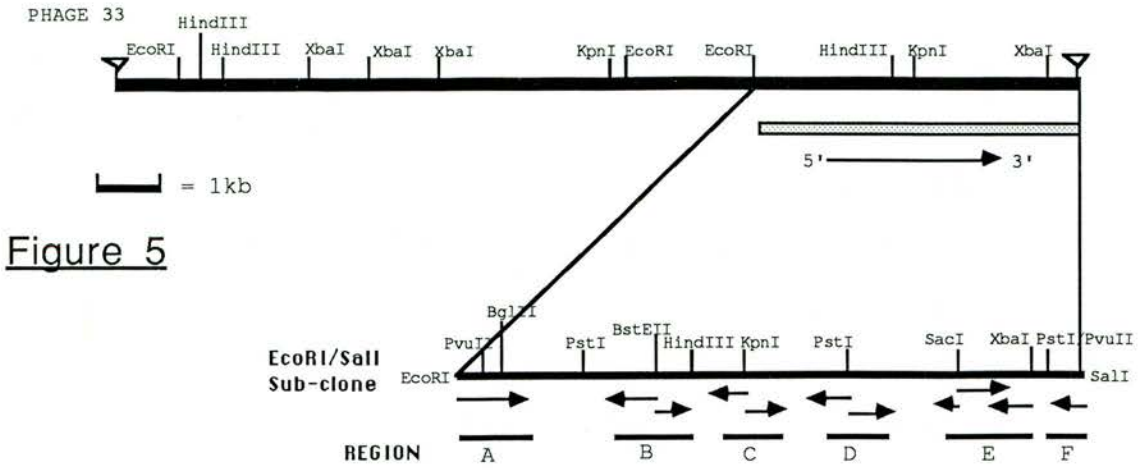
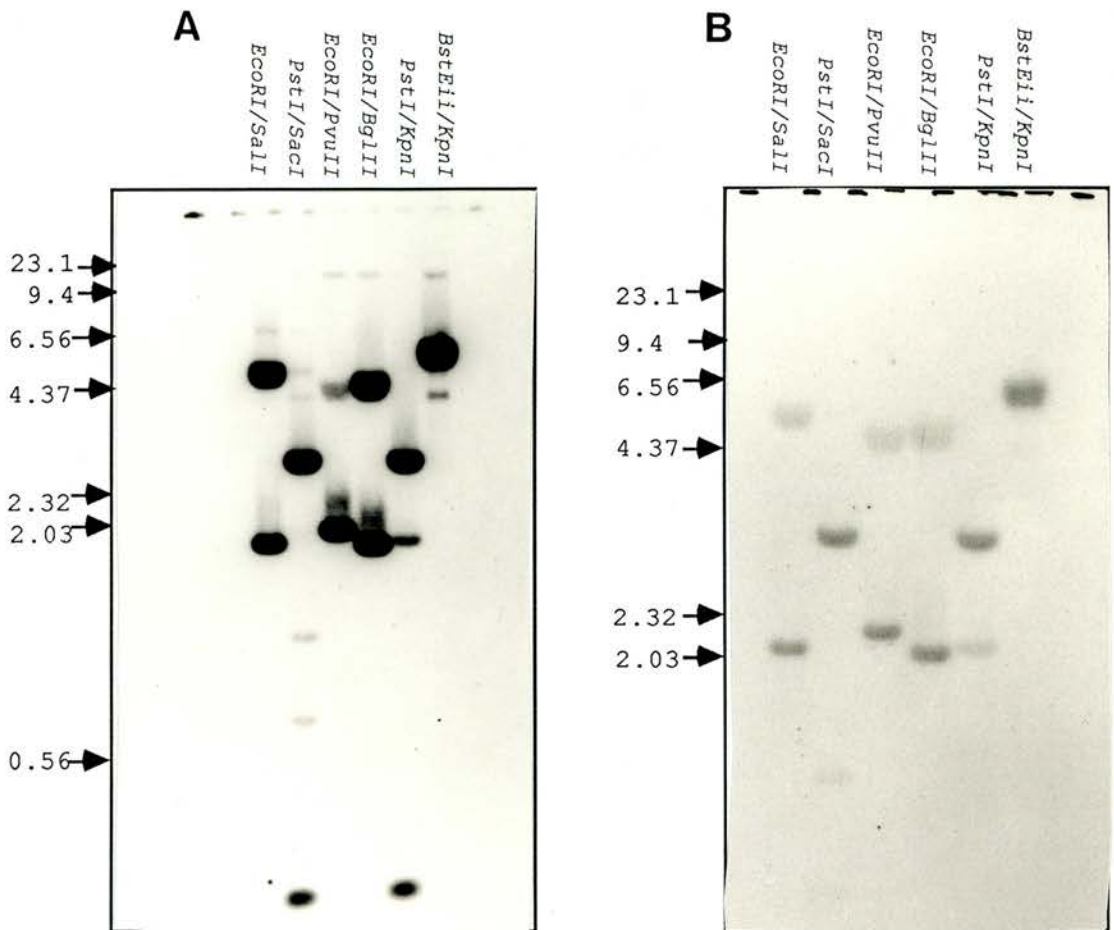


Figure 5



#### 4.2.2a Sequencing of phage 33.

Figure 1 indicates A, B, C, D, E and F; the regions of the *EcoRI* / *Sall* sub-clone of phage 33 which have been sequenced. The sequence is incomplete, with only two thirds having been sequenced. However when comparing the sequence of the six regions to pBoLA-1, homology was detected. The homology was strongest at the 3' end, as the previous hybridisation experiments had predicted. Figure 6 shows the detailed comparison of the sequenced regions of phage 33 with the cDNA clone pBoLA-1. Sequenced regions E and F contain exons 5, 6, 7 and 8 and allow placement of these exons in the cDNA. These areas respectively show 52%, 75%, 75% and 76.5% homology to pBoLA-1. Exon 7 in phage 33 seems to be four nucleotides larger than the homologous sequence in pBoLA-1, but there is still clear conservation of intron/exon splice signals (Breathnach and Chambon 1981) for this exon and the others identified, these are indicated on figure 6. Translation into amino acids of exonic sequences identified in regions E and F revealed no premature stop codons and conservation of the last 11 amino acids and stop codon. The divergence of amino acid sequence 5' to this point corresponds directly to the insertion point of the four nucleotides in the 5' region of exon 7. Although unlikely, these nucleotides could correspond to a sequencing error as only one strand has been sequenced. The observed conservation of amino acids at the 3' end encoded by regions E and F would suggest that the 3' end of this gene is most like class Ia sequences.

Identification of the 5' end of this gene by sequencing has proved more difficult. Comparing the restriction map for the sub-clone to the restriction map of pBoLA-1 suggested that the area surrounding the *KpnI* site in the sub-clone may be the 5' end of the clone. If phage



Figure 6. Depiction of nucleotide and translated amino acid sequence of phage 33 sub-clone *EcoRI/Sall* with areas of similarity to pBoLA-1 indicated.

This figure shows comparisons of regions E and F to pBoLA-1. Firstly the restriction map of phage 33 is given to enable direct visualisation of where regions E and F map on the phage clone. The complete nucleotide sequence of regions E and F are shown with the intron/exon splice sites as shaded areas, these denote the boundaries of the exons. The regions are not shown as one contiguous sequence because it is not clear how much DNA lies between them which has not been sequenced.

Having identified probable exons, these are then shown aligned with the nucleotide sequence of pBoLA-1. In this alignment the divisions between the exons are indicated by black bars, the stop codon is indicated by an asterisk. Lastly the amino acid translation of pBoLA-1 is shown aligned to the amino acid translation of exons 5, 6, 7 and 8 from regions E and F.

All these sequences were aligned using UWGCG programmes on a Vax computer network.



33 contained the genomic equivalent of the cDNA clone the restriction maps should be very similar. All the hybridisation evidence presented above suggested that the *KpnI* site marked the 5' border of the clone. In an effort to identify regions 5' of exon 5, hybridisation of M13 sub-clones cloned from regions 5' of this were identified in region B (clones covered the *BstEII/KpnI* area) and region C (clones covered the *KpnI/PstI* area). Sequencing of hybridising M13 clones from these two regions localised a small region of homology which is indicated in figure 7. Figure 7 is a dotplot comparison of region B and the 5' end of pBoLA-1. The areas of homology are shown below, with the corresponding pBoLA-1 sequence.

```

Region of homology detected in the leader peptide (first exon) of pBoLA-1:-
pBoLA-1 bp 20  TCCTGCTGCTCTCGGGGTCCTGGTCCTGACCGAGACCCCGGGCTGGCTCCCACT
                |   |   |   |   |   |   |   |   |   |   |   |   |   |
region B bp 500      TGATTACCTACCTTGCCCTGAACGAGACCTGTGCTCCTGACCGCAG
Region of homology detected in the second domain (third exon) of pBoLA-1:-
pBoLA-1 bp350->590 (2nd domain) against region B bp420->603, area in fig7 is underlined:-
ACCTTCCAGTGGATGTACGGCTGCGACGTGGGGCCGGACGGGCGTCTCCCGCGGGGTTTCATGCAGTACGGCTACGACGGC.AGAGAT
                |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
...AGTCTTCTCTGAAAGAAGTATAATTCATATTTGAAAAAGTTATTGCCTCAGGTTCTGGTAGATTGCTTATGATGGCCAATGAT
TACAT...CGCCCTGAACGAGGACCTGCGCTCCTGGACCGGGCGGACACGGCGGCTCAGATTCACCAAGCGCAAGTGGGAGGGCGCA
||| |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
TACCTACCTTGCCCTGAACGA.GACCTGTGCTCCTG.ACCGCAGT.GACA.GGTG.CTCAGACTCAGCTCA.CAGGCATGTACATACA
GGTGAGGGCGGAGAGAC.AGAGAACTACCTGGAGGGCACGTGCGTGGAGTGGCTCCGCAGATACCTGGAGACCGG pBoLA-1
                |   |   |   |   |   |
TACAGTCACAGTACAGTAGAC...TACCT..... region B

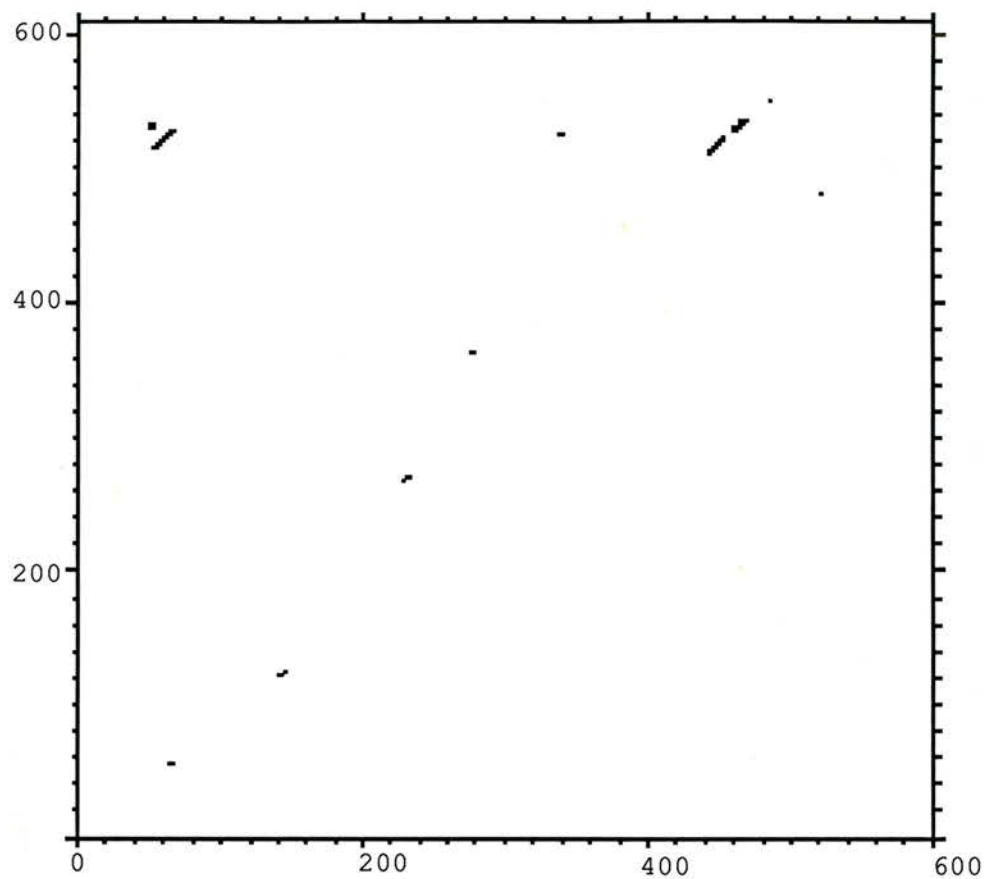
```

Figure 7 indicates where these areas of homology are in pBoLA-1. The area detected in the second domain is longer with more matches between the sequences. However, the homology is confined to within the sequence coding for the second domain (the sequence above corresponds to the whole of the third exon of pBoLA-1) with the underlined area showing 75% similarity. The alignment may be spurious, this could only be answered after more sequencing. The fact that this region does show some homology may explain some of the anomolous results obtained when using a 5' probe to map the phage insert (see figures 4 and 5).

Figure 7. Dotplot alignment of Region B against nucleotides 1-600 of pBoLA-1.

The first six hundred nucleotides of pBoLA-1 were compared to the complete sequence of region B (see figure 1 for area that region B maps too in phage clone 33). The results are shown as a dotplot with pBoLA-1 forming the abscissa and region B forming the ordinate axis. The putative exonic structure of pBoLA-1 is drawn below the abscissa. Two areas detected, one which maps to the leader peptide of pBoLA-1 and one which maps to an area in the second domain of pBoLA-1. The area found in the second domain of pBoLA-1 was found to correspond to a conserved area of class I proteins. If only this conserved area is used then class I genes are primarily detected in a computer search (see text).

The dotplot programme and search criteria were done using UWGCG programmes on a Vax computer network.

Figure 7R  
E  
G  
I  
O  
N  
B

pBoLA-1 (1-600bp)

LP	1st Domain	2nd Domain
----	------------	------------

The gene may be a pseudogene, although the evidence from the 3' region suggests that this area has been largely conserved except for the 5' area of exon 7. The sequence could correspond to a truncated class I gene. 5' and 3' gene fragments have been identified in large scale cloning of the MHC regions of BALB/c mice (Steinmetz *et al.* 1982b) and C57BL/10 mice (Weiss *et al.* 1984), this possibility will be discussed more below.

#### 4.2.2b Comparison of phage 33 to *HLA-A3*.

Comparisons of phage 33 to the pBoLA-1 cDNA clone rely on exons being present in the phage. The homology of regions E and F to pBoLA-1 has allowed the last 4 exons to be correlated with the sequence of the cDNA. Since the sequence of phage 33 is incomplete, correct placement of exons 1, 2 and 3 may be missed. Therefore the six regions which have been sequenced were compared to a genomic clone. A large comparison against the human gene *HLA-A3* was undertaken.

Figure 8 shows the alignments obtained when the regions A,B,C,D,E and F were compared to the gene *HLA-A3*; only C, E and F showed significant homology in the alignment. Region C on comparison maps 3' to the third exon area of *HLA-A3*. Region C has 55% identity with *HLA-A3*, the aligned similarity does not stretch over the whole region, but breaks down at the 5' boundary of the fourth exon in *HLA-A3*. This marked break suggests that either this corresponds to an insertion or deletion site in the region or that the area sequenced in the bovine clone is probably still intronic. This intron can vary in length from 550bp in *HLA-A2* (Koller and Orr 1985) and 577bp in *HLA-A3* (Strachan *et al.* 1984) up to 1669bp in *H-2D<sup>d</sup>* (Sher *et al.*

1985). The presence of a large intron at this point in class I genes may explain the occurrence of 5' and 3' gene fragments observed in mouse (Steinmetz *et al.* 1982b; Weiss *et al.* 1984) the two parts of the gene are probably separated by a recombination event. It is possible with the apparent lack of homology to the third exon of *HLA-A3* that this gene encodes a long intron and may be a 3' gene fragment.

The next alignment of sequences covers the clones which stretch from just 5' of the *SacI* site in the sub-clone to the 3' boundary of the phage clone. These are regions E and F. These two regions show appreciable homology to the human genomic clone upwards of 68% and 64% respectively and seem to show good conservation of intron/exon splice signals (Breathnach and Chambon 1981). This confirms the conclusion reached when comparing these regions to pBoLA-1; that exons 5-8 are present in this gene. When comparing region E encoded exon 7 to exon 7 in the *HLA* genomic clone there is again a four base pair mismatch, this is the same number of non-aligned nucleotides as was obtained in comparison to pBoLA-1. However, the comparison shows that the non-aligned nucleotides occur internal to the exon and not at the 5' boundary as before, obviously exon 7 is atypical in this bovine gene. All the exons show reasonable levels of homology to *HLA-A3*; 45%, 84%, 82% and 65% these figures can be compared to the figures quoted previously for pBoLA-1; 52%, 75%, 75% and 76.5%. The percentage similarity for exons 6 and 7 is higher when the sequence is compared to *HLA-A3* than when it is compared to pBoLA-1. In contrast, exon 8 which includes the 3'UT region displays 10% more homology to pBoLA-1 than to *HLA-A3*. The percentage similarity for the last exon which includes the 3'UT region suggests that this

sequence is not an allele of pBoLA-1, but is still more related to cattle sequences than to *HLA-A3* in this region. 99.5% identity in this region has been reported for *HLA -A* class I alleles (Koller and Orr 1985).

Region B did not give any significant alignment when compared on the UWGCG gap programme to *HLA-A3*, the small area of homology previously noted when compared to pBoLA-1 was detected when visualised as a dotplot comparison (data not shown). Direct comparison of this 39bp stretch of sequence to the comparable area in *HLA-A3* shows a percentage similarity of 70.3% this compares to 71% between this area and pBoLA-1, but is still significantly less than 97.8% similarity between *HLA-A3* and pBoLA-1.

Region D consistently showed no homology to any class I sequence with which it was compared, although it corresponds to an area around the *Pst*I site which maps between regions C and E/F which do show homology. Reasons for this lack of homology could be that either there has been an insertion/deletion event in this gene or this intron is too far diverged to be detected in comparison programmes. Insertion events of Alu repeats are well documented in class I genes (Kress *et al.* 1984; Geraghty *et al.* 1987 etc, see chapter 3 section 3.23), therefore this sequence was compared to the available sequences of repeats contained within the data base. No homology was detected. The results discussed above certainly suggest the existence of a diverged 5' end and a conserved 3' end in the phage clone. Perhaps the unlinking of the two maps to region D, but this is not clear when there is conservation of region C which maps 5' to this area.



Figure 8. Comparison of nucleotide sequence of the gene *HLA-A3* with regions C, E and F of phage 33.

All the regions sequenced in phage 33 were compared to *HLA-A3* but only C, E and F showed significant similarity. The alignment of these regions with the human gene are shown. Exons one to eight are shown in *HLA-A3* as shaded boxes. Regions C, E and F overlap with some of these exonic sequences and are indicated. The area of region C which overlaps with exon 4 of *HLA-A3* has not been included within the shaded box as being exonic sequence. Where a match occurs between the two sequences it is indicated as a vertical bar. Dots have been introduced to maximise the alignment of the two sequences.

The comparisons were done using UWGCG computer programmes on a Vax network.



### 4.3 Summary.

This chapter describes the isolation of two phage clones which do not overlap. Phage 41 contains two non-contiguous hybridising-areas within it. This clone appears to correspond to an amplified structural block of class I genes suggesting it may contain the bovine equivalents of the mouse *Qa* and *Tla* genes.

Phage 33 appears to contain a gene which is present as a single copy. Further examination of phage 33 using detailed hybridisation analysis with 5' and 3' probes suggested the existence of a 3' end and possibly also a 5' end. Sequence analysis of phage 33 confirmed that it may have a remnant of a 5' end and definitely established that it has a 3' end. The 3' end showed upwards of 75% similarity with pBoLA-1 and in excess of 65% similarity with a human gene *HLA-A3* for exons 6, 7 and 8. The spread of regions of homology throughout the clone is quite confusing and at first it would appear that the clone may only be a 3' gene fragment, but areas of similarity corresponding to sequences present in the 5' end of class I genes are present. The conclusion reached was that phage 33 contains a piece of DNA which corresponds to a class I pseudogene.

## **Chapter 5.**

### **Southern blotting studies on the bovine MHC: dissection of the hybridisation pattern obtained and identification of RFLP's coincident with *BoLA-A* serological types.**

#### **5.1 Introduction.**

Serological analysis of the MHC complexes of man and mouse has revealed extreme polymorphism in numbers of serologic alleles determined. Klein (1986) reports 92 alleles at the *K* and 63 alleles at the *D* loci in populations of mice. This contrasts with the limited amount of serologic polymorphism determined for other loci which map to the *H-2* region namely the *Qa* and *Tla* (Flaherty 1980) (see chapter 1, section 1.2.1). The *Qa* and *Tla* loci although associated with  $\beta_2$ -microglobulin, have never been shown to function as viral restriction elements (Kastner *et al.* 1979a, 1979b; Flaherty 1980); unlike the *K*, *D* and *L* loci (Zinkernagel and Doherty 1979).

Southern blotting studies of mouse genomic DNA revealed a reasonable degree of complexity with 10-12 bands per individual (Steinmetz *et al.* 1981a). This was confirmed by Cami *et al.* (1981), who found complementary copies of class I-like sequences when they screened a mouse genomic library. The identity of these cross-hybridising sequences was resolved by complete cloning of the class I region of BALB/c mice (Steinmetz *et al.* 1982b), where thirty five genes were cloned (Stephan *et al.* 1986) but thirty one mapped to the *Qa* and *Tla* regions (Winoto *et al.* 1983). Therefore although the *K*, *D* and *L* class Ia loci are functionally distinct from the *Qa* and *Tla*

loci, they cross-hybridise appreciably at the DNA level.

Several findings emerged which enabled some distinction of the class Ia and the class Ib genes at the DNA level. In general, polymorphic restriction fragments mapped to the *K* and *D* regions (Steinmetz *et al.* 1982b; Weiss *et al.* 1984) and non-polymorphic bands mapped to the *Qa* and *Tla* regions (Steinmetz *et al.* 1982b; Winoto *et al.* 1983; Weiss *et al.* 1984; Fisher *et al.* 1985). Using specific probes in Southern blotting studies on overlapping cosmid DNA clones meant investigators were able to correlate restriction fragment length polymorphism's (RFLP's) of bands with individual genes in recombinant inbred congenic lines of mice. So a complete map of the *H-2* class I region was constructed. This type of experimental analysis cannot be applied to cattle due to the absence of recombinant inbred strains. Therefore results of Southern blotting studies and gene mapping in man may be more applicable.

Until recently (Koller *et al.* 1989; Spies *et al.* 1989) no large scale cloning and characterisation project had been completed on the *HLA* class I regions, so the exact numbers of class I genes were not known. A variety of approaches were used in efforts to overcome this; namely Southern blotting in conjunction with deletion mutant analysis and RFLP analysis.

Biro *et al.* (1983) reported at least 30 class I genes in man, this was probably an overestimate but did confirm that the class I genes in man were also a multi-gene family. Mapping of these genes used deletion mutant analysis (Orr *et al.* 1982) where loss of a band was correlated with loss of expression. As in mouse the majority of the bands were found to map to the telomeric side of the class Ia genes (Orr and De Mars 1983). Cann *et al.* (1983) were the first to correlate

a restriction enzyme polymorphism to a serological type, an 8.6kb fragment generated by *EcoRV* digestion of genomic DNA correlated with *HLA-B8*. Extension of these restriction fragment length polymorphism studies led investigators to see if they could sero-type an individual on the pattern obtained in RFLP studies. It was found that multiple bands could be assigned to one serological specificity leading to the formation of an alloantigen (Cohen *et al.* 1985). However this type of assignment has been complicated by the observation that more RFLP types seem to correlate to the *A* locus specificity which is less serologically polymorphic than the *B* locus (Cohen *et al.* 1985). An explanation for this could be that the *A* locus is flanked by more polymorphic markers than the *B* locus. N'Guyen *et al.* (1985) found differences in the number of *A* region genes in different individuals as did Chimini *et al.* (1988). Therefore the existence of more polymorphic flanking regions of DNA seems to give better serological correlation, but it means that a direct correlation of restriction fragment with serological type cannot be made.

For the bovine system where serology is perhaps less advanced than in man, at least for additional loci. The aim was to use a combination of approaches to better characterise the class I genes and ultimately to use this information to type for class I specificities.

## 5.2 Results and Discussion.

### 5.2.1 Investigation of the class I region of *BoLA* by Southern blotting.

#### 5.2.1a Human and bovine cDNA probes show that the class I genes are a multi-gene family.

Fractionating total bovine lymphocyte DNA on an agarose gel and probing the Southern with pHLA-27 (a class I cDNA clone from man) gave the result shown in figure 1. The major noticeable feature was the multi-band pattern of hybridisation obtained, implying that the *BoLA* class I region contains multiple genes or lots of similarly-hybridising fragments. However, the pattern of hybridisation is complicated when using this probe because of cross-hybridisation to a repeated element in the cattle genome. Background cross-hybridisation disappears if a homologous bovine class I probe (pBoLA-1) is used (see chapter 3 for the full description of this probe). Figure 1 presents a comparison of the hybridisation pattern between a human probe and a bovine probe. The bovine probe gives less background hybridisation and is ultimately of more use in Southern blotting studies.

#### 5.2.1b Dissection of the class I multi-gene family.

##### 5.2.1b1 Evidence for the specific detection of class Ia genes.

The multi-band hybridisation pattern obtained when using a bovine class I cDNA as a probe, can be further dissected by using specific 5' and 3' fragments derived from the cDNA clone. This strategy of using probes derived from specific regions of class I genes has been

Figure 1. Comparison of pHLA-27 and pBoLA-1 when used as probes against bovine genomic DNA.

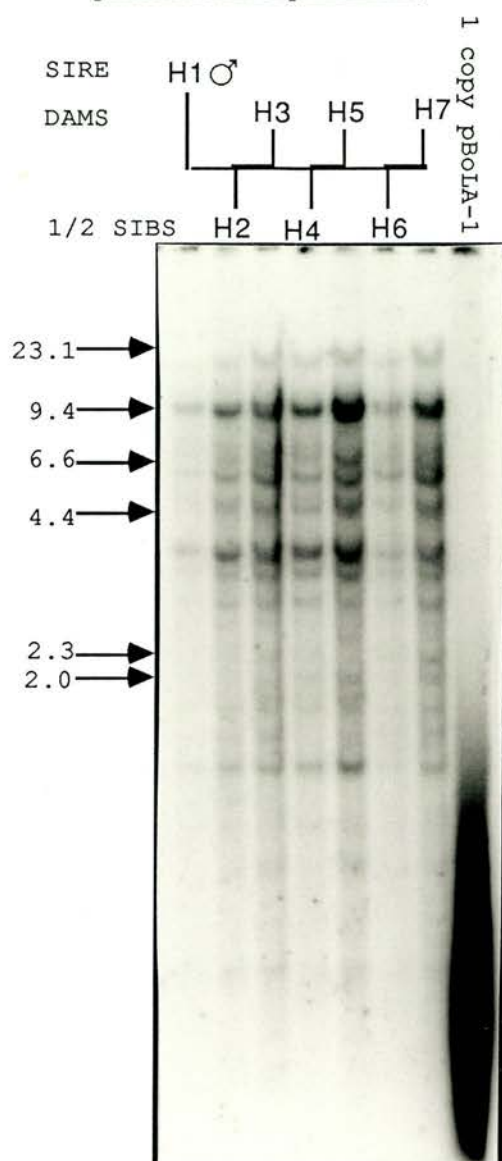
Bovine lymphocyte DNA was digested with *PvuII* before being loaded onto a 0.8% agarose gel. Duplicate samples were loaded. After electrophoresis the gel was transferred, the membrane was halved so that a direct comparison of two different probes could be made. The specific activities of the two probes were not equal, both hybridisations were washed to 2xSSC, 0.1% SDS at 65°C.

The gel was loaded from left to right with 10µg of DNA per well. The DNA was from a Hereford half-sib family; H1(9447 Sire), H2 (9956 sib), H3 (9127 Dam), H4 (9876 sib), H5 (9216 Dam), H6 (9802 sib) and H7 (9204 Dam). The last track on each gel contains the equivalent of one copy of the cDNA clone pBoLA-1 with 10µg of salmon sperm DNA as a carrier.

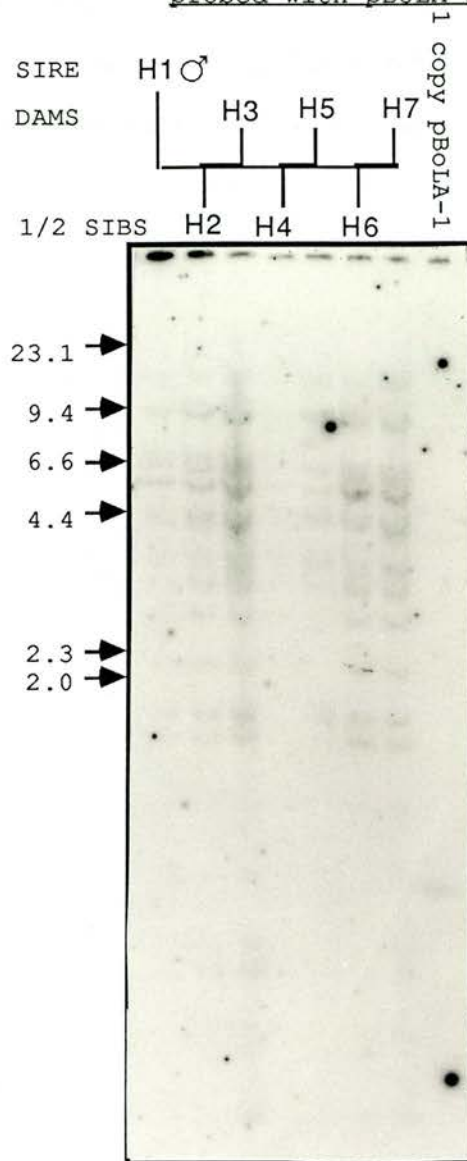


Figure 1

*PvuII* digest of Hereford family  
probed with pHLA-27.



*PvuII* digest of Hereford family  
probed with pBoLA-1.



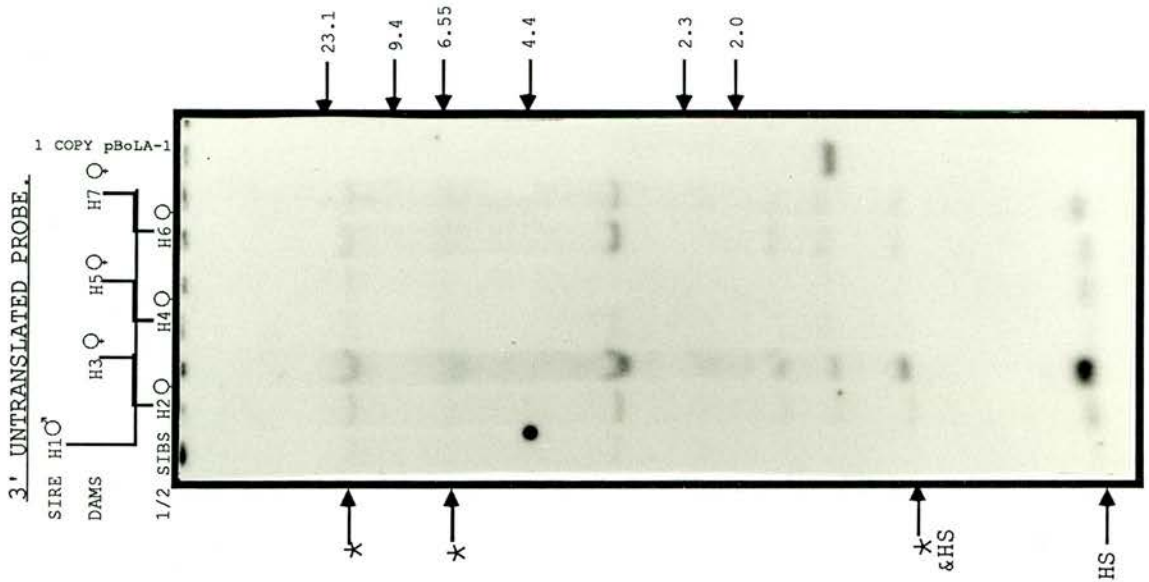
successfully applied by other investigators (Steinmetz *et al.* 1982b; Koller *et al.* 1984; Rogers 1985a, 1985b), who were able to dissect the otherwise complicated hybridisation patterns obtained. Figure 2 presents the results of this investigation. Firstly, it is apparent that the 3' untranslated (3'UT) probe obtained by *PstI/EcoRI* digestion of pBoLA-1 is more specific in hybridisation (particularly after a high stringency wash), only recognising a few bands. This agrees with the results of Koller *et al.* (1984) who reported that the 3'UT region of *HLA-A* and *HLA-B* were locus specific in the *HLA* class I region. Since then, it has been noted that these do cross-hybridise at high stringency, but only to *HLA-C* genes (Gussow *et al.* 1987; Chimini *et al.* 1988). The relatively simple pattern of hybridisation obtained by hybridising with the 3' UT probe is in marked contrast to the pattern obtained using a 5' probe (5' probe derived by *EcoRI/SacI* digest of pBoLA-1 which realises a 214bp fragment which corresponds to the leader peptide and first domain of the protein), see figure 2. This confirms that the 3'UT probe is specific because it recognises a subset of bands detected by the 5' probe. Only two bands are recognised at high stringency, suggesting at least two class Ia genes in cattle. Tallying the number of bands recognised by the 5' and 3'UT probes in common reveals 2-3 bands, suggesting that these are candidates for complete class Ia genes. This figure is probably an underestimate of class Ia genes because it ignores intra-genic restriction enzyme sites.

Figure 3 presents a Southern blot of a *TaqI* digest of bovine DNA probed with the complete cDNA clone pBoLA-1, then stripped and reprobbed with the 5' end. The 5' end of the cDNA seemed to hybridise to a subset of fragments, and interestingly, these corresponded to

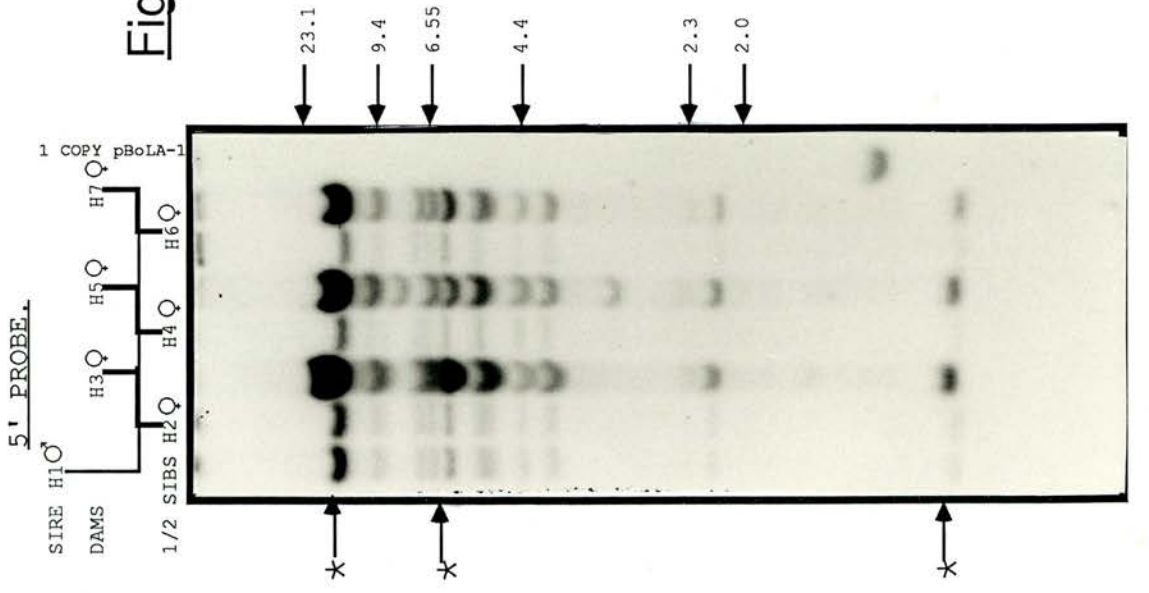
Figure 2. Comparison of two fragments derived from the 5' and 3' ends of the cDNA clone pBoLA-1 which were used to probe a Southern blot of bovine genomic DNA.

This figure compares two probes derived from pBoLA-1, a 5' *EcoRI* /*SacI* probe comprising the leader peptide and first domain of the pBoLA-1 protein and a 3' UT probe obtained by *PstI* /*EcoRI* digest of the cDNA (see chapter 2, figure 2 for exact placement of these restriction enzyme sites on the cDNA clone). Bovine genomic DNA was digested with *PvuII* before being fractionated on a 0.8% agarose gel and transferred to a nylon membrane. The samples were loaded in duplicate so the membrane could be probed as two separate halves. The probes were not of equivalent specific activity, but each blot was first washed to low stringency in 2xSSC, 0.1% SDS at 65°C and exposed for 5 days (the results of which are shown here) before being washed to high stringency in 0.1xSSC, 0.1% SDS. The marker sizes of a *HindIII* digest of lambda DNA are indicated. The fragments recognised in common at low stringency by the two probes are indicated by arrows with an asterisk. The fragments specifically recognised by the 3' UT probe at high stringency are indicated by the arrows with HS.

The animals are from a Hereford half-sib family. In each case the relationship between sib and dam is indicated by a connecting bold line. The samples from left to right are the same as those given in figure 1; 9447 sire (H1), 9956 sib (H2), 9127 dam (H3), 9876 sib (H4), 9216 dam (H5), 9802 sib (H6), 9204 dam (H7) and 1 copy equivalent of the cDNA clone pBoLA-1.



**Figure 2**



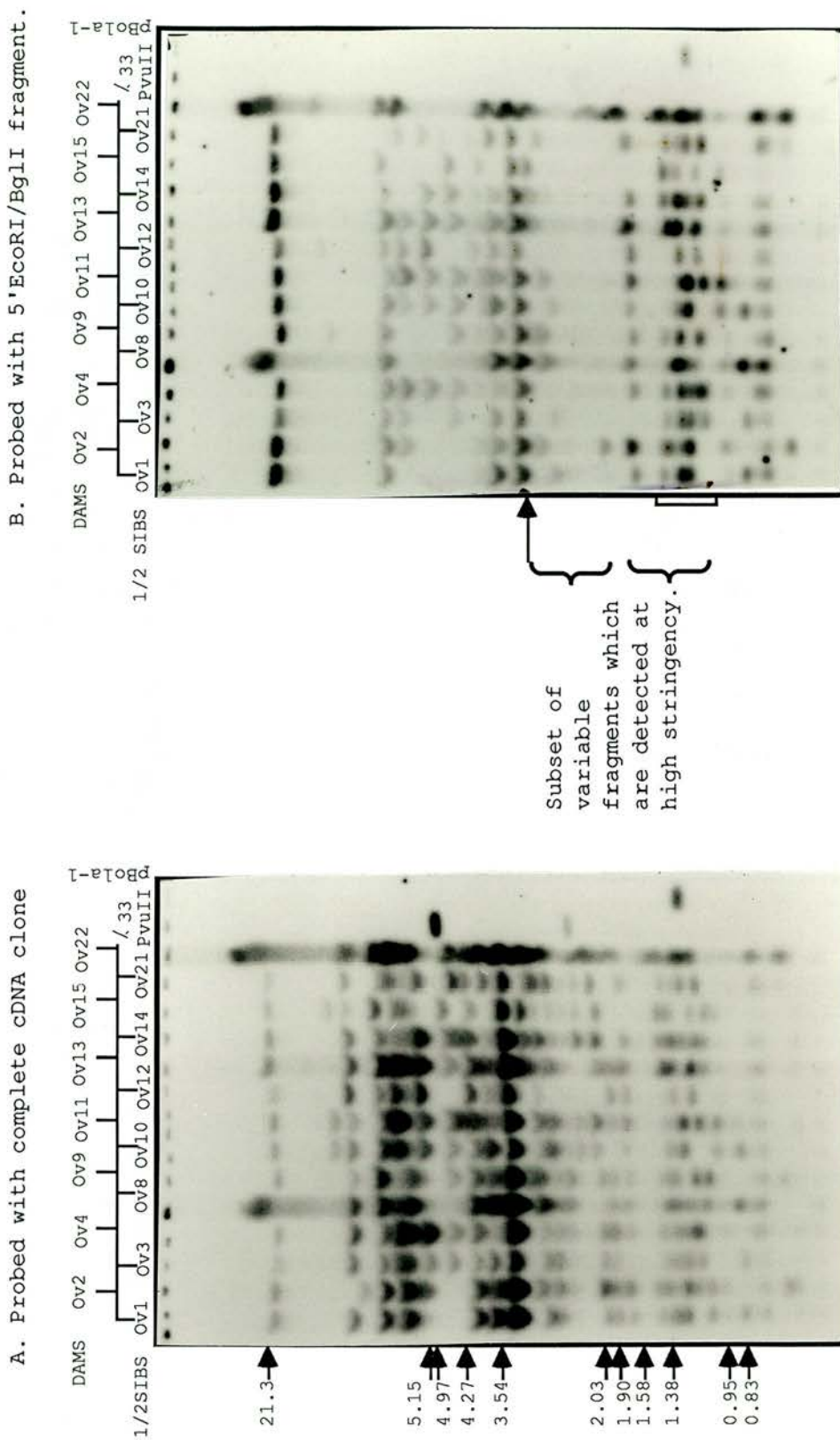
some of the variable fragments on the Southern. Two points of information can be gained from this; 1), it illustrates that pBoLA-1 specifically recognises variable fragments and 2), these are of a small size range 1-1.5kb. The small size of hybridising fragments can be explained by the use of *TaqI* to digest the DNA. *TaqI* recognises a four base pair sequence, TCGA, which should occur every 250bp on average. We would expect a skewed distribution of these sites towards the 5' end of the gene because class I genes are very G+C rich in their 5' regions (Lindsay and Bird 1986). But on the other hand CpG, which is a part of the recognition sequence for this enzyme, is highly mutable (Barker *et al.* 1984) if methylated undergoing C-> T transitions. There is direct evidence for methylation of the *HLA* class I region as shown by restriction enzyme digestion using methylation sensitive enzymes. Only a proportion of the region was restricted and therefore not methylated and this was correlated to the small proportion of class I genes which are actually expressed (Chimini *et al.* 1988). This accumulated evidence could suggest that class Ia genes may have more *TaqI* sites in their 5' regions. These are undermethylated so C->T transitions are less likely than compared to class Ib genes, ultimately giving rise to more *TaqI* sites in the 5' regions of class Ia genes. Even more variation to the restriction fragment length polymorphism patterns of class Ia genes will be introduced, because this is where the functional variation between class Ia proteins map. So consequently, more polymorphism is detected in restriction enzyme analysis of the DNA of class Ia genes.

Figure 3. Dissection of the hybridisation pattern of a *TaqI* digest of a Friesian half-sib family using the complete pBoLA-1 cDNA and 5' *EcoRI* / *BglI* specific fragment as probes.

DNA was digested to completion with *TaqI* before being fractionated on a 0.8% agarose gel, this was then Southern transferred to a nylon membrane. Photograph A shows the hybridisation pattern obtained when the full length pBoLA-1 cDNA was used as a probe. This membrane was washed to low stringency in 2xSSC, 0.1% SDS at 65°C and exposed for 7 days. Photograph B shows the same Southern which has been stripped and reprobbed with a specific 5' fragment of pBoLA-1 obtained by *EcoRI* / *BglI* digestion (see chapter 2, figure 2 for location of restriction enzymes sites on pBoLA-1) which corresponds to the leader peptide and the first two domains of pBoLA-1 protein. This time the Southern was washed to low stringency (2xSSC, 0.1% SDS) at 65°C, exposed for 7 days (the result being illustrated) and then washed more stringently in 0.2xSSC, 0.1% SDS at 65°C before being exposed for 12 days. The fragments still recognised at high stringency are indicated on the figure.

The family used comprised of Friesian half-sibs, in each case the offspring occurs to the left of the dam e.g. Ov1 is the offspring of dam Ov2. The samples were loaded left to right: 5604 sib, *BoLA* type -/- (Ov1); 7123 dam, *BoLA* type -/- (Ov2); 5612 sib, *BoLA* type 20/- (Ov3); 7437 dam, *BoLA* type 11/20 (Ov4); 5622 sib, *BoLA* type 6.4/- (Ov8); 7315 dam, *BoLA* type 6.4/8 (Ov9); 5623 sib, *BoLA* type -/- (Ov10); 7298 dam, *BoLA* type 5/Eu12 (Ov11); 7480 sib, *BoLA* type 11/- (Ov12); 7393 dam, *BoLA* type 8.2/11 (Ov13); 5605 sib, *BoLA* type 8/- (Ov14); 7361 dam, *BoLA* type 8/- (Ov15); 7492 sib, *BoLA* type 10/- (Ov21); 7320 dam, *BoLA* type 10/- (Ov22); 1 copy equivalent of the subcloned *EcoRI* / *Sall* phage 33 (see chapter 4) digested with *PvuII* and 1 copy equivalent of cDNA pBoLA-1.

**Figure 3**



### 5.2.1b2 Evidence for class Ib genes in the *BoLA* class I region.

Figure 3 also illustrates the presence of two separate groups of hybridising fragments; a polymorphic group and a non-polymorphic group. Steinmetz *et al.* (1982b), Winoto *et al.* (1983) and Weiss *et al.* (1984) all noted that the class Ib genes tended to have less restriction fragment length polymorphism and reports by Rogers (1985a, 1985b) confirm that the class Ia and class Ib genes can be separated into differently hybridising groups. The results presented in figure 3 indicate that the *BoLA* class I region also contains class Ib genes. To date, class Ib genes have also been identified in rats (Kastern 1985) and in man (Geraghty *et al.* 1987; Koller *et al.* 1989).

Although this result does not substantiate that the class Ib genes are conserved across species, it would contradict the conclusion reached by Rogers (1985b) that they are not! Further evidence for possible bovine class Ib genes is presented in chapter 4 figure 3.

### 5.2.2 Correlation of RFLP's to serology.

Finally, the correlation of class I-hybridising bands to the serological types of animals was examined. Previously, specific hybridising bands have been correlated to serological types in man (Cann *et al.* 1983), sheep (Chardon *et al.* 1985b), pigs (Chardon *et al.* 1985a) and horses (Alexander *et al.* 1987) (see chapter 1, section 1.3.1a2 for a more extensive list).

In an effort to correlate restriction fragment pattern with serological specificity various restriction digests on a model half-sib Hereford family were undertaken. Of all the enzymes tested; *PvuII*, *PstI*, *EcoRI* and *BamHI* (see figures 4 and 5), only *PvuII* yielded variation between individual animals which correlated with serologic



specificities (figure 4 ). A 3.1kb *PvuII* band correlated with *BoLA-w6.2*. However, the other enzymes did detect some notable features. The *EcoRI* digest of the DNA generated a completely different band pattern from that obtained with the other enzymes (figure 4). It appears that most of the class I genes occur as either 8, 11 or 15kb transcription units. This has also been documented in man where 7 and 8kb transcription units are detected (Srivastava *et al.* 1987).

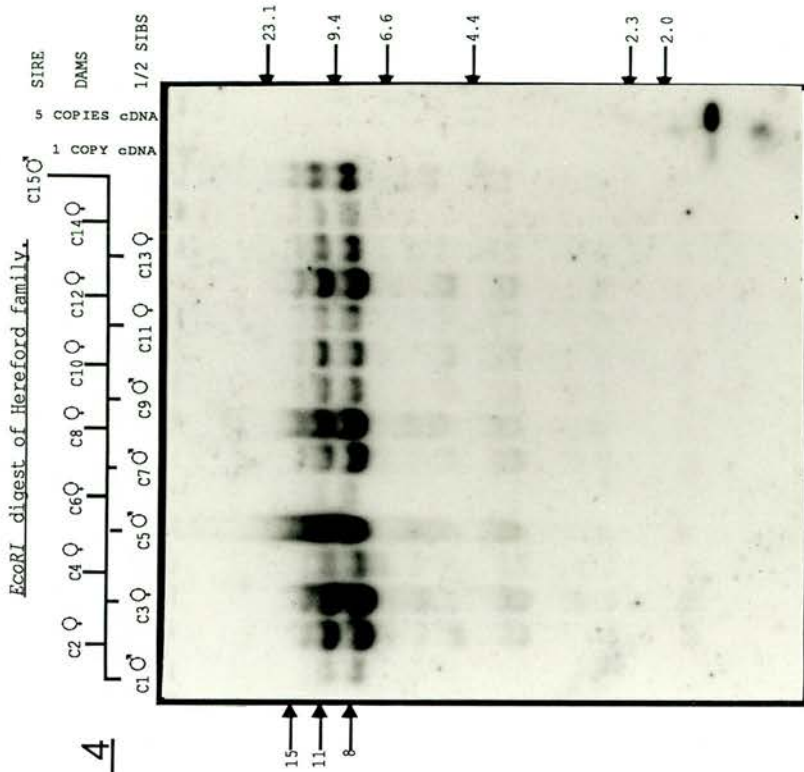
Further studies in half-sib family groups have allowed more correlation of restriction fragment band and serologic specificity. Figure 4 showed the correlation of a 3.1kb *PvuII* fragment with *BoLA-w6.2* in Herefords. Figure 6 depicts a *PvuII* digest of a Friesian half-sib family. On this occasion correlation of a 2.9kb fragment with *BoLA-w10* and 10.5kb *PvuII* fragment with *BoLA-w8* was seen. Finally figure 7 shows a dam half-sib family. This family differs from the previous families because the common parent is the dam rather than the sire. The dam has the *BoLA* type *w6.2/w10*, this is coincident with the presence of two *PvuII* fragments of 2.9kb (*w10*) and 3.0kb (*w6.2*). Each one of the half-sibs has inherited one allele or the other, and correspondingly one band or the other. Track four on figure 7 shows that this half-sib has inherited one allele from the sire, so displays the same two bands as the dam. The previous correlation of an 11kb *PvuII* band (10.5kb on figure 6, assume this is due to error in accurately measuring large fragment sizes) with *BoLA-w8* is also apparent in this family and in the second dam family shown in figure 7.

Figure 4. *PvuII* and *EcoRI* digests of a Hereford half-sib family probed with pBoLA-1.

The digested DNA was fractionated overnight on an 0.8% agarose gel before being transferred to a nylon membrane. The membrane was probed with pBoLA-1 and washed in 0.5xSSC, 0.1% SDS at 65°C before being exposed for 7 days. The size markers are a *HindIII* digest of lambda phage DNA. The 3.1kb fragment which correlates with *BoLA-w6.2* is indicated by an arrow.

The second panel shows the results obtained when the same DNA is digested with *EcoRI* and then probed with pBoLA-1. This membrane was washed to 0.5xSSC, 0.1% SDS at 65°C and exposed for 7 days. The size markers are a *HindIII* digest of lambda phage DNA. The transcription units of 8, 11 and 15kb are indicated by arrows.

Each gel was loaded left to right in the following order: 9898 sib, *BoLA* type 20/- (C1); 9241 dam, *BoLA* type 1/13/20 (C2); 9897 sib, *BoLA* type 9/20 (C3), 9480 dam, *BoLA* type 6/6.2/20 (C4); 9869 sib, *BoLA* type 9/20 (C5); 9315 dam, *BoLA* type 7/20 (C6); 9865 sib, *BoLA* type 7/20 (C7); 9066 dam, *BoLA* type 9/20 (C8); 9847 sib, *BoLA* type 1/13/20 (C9), 9027 dam, *BoLA* type 1/13/- (C10); 9844 sib, *BoLA* type 9/20 (C11); 9072 dam, *BoLA* type 9/20 (C12); 9797 sib, *BoLA* type 9/20 (C13); 9169 dam, *BoLA* type 1/13/20 (C14); 9547 sire, *BoLA* type 9/20 (C15); 1 copy equivalent of pBoLA-1 and 5 copies of pBoLA-1.



**Figure 4**

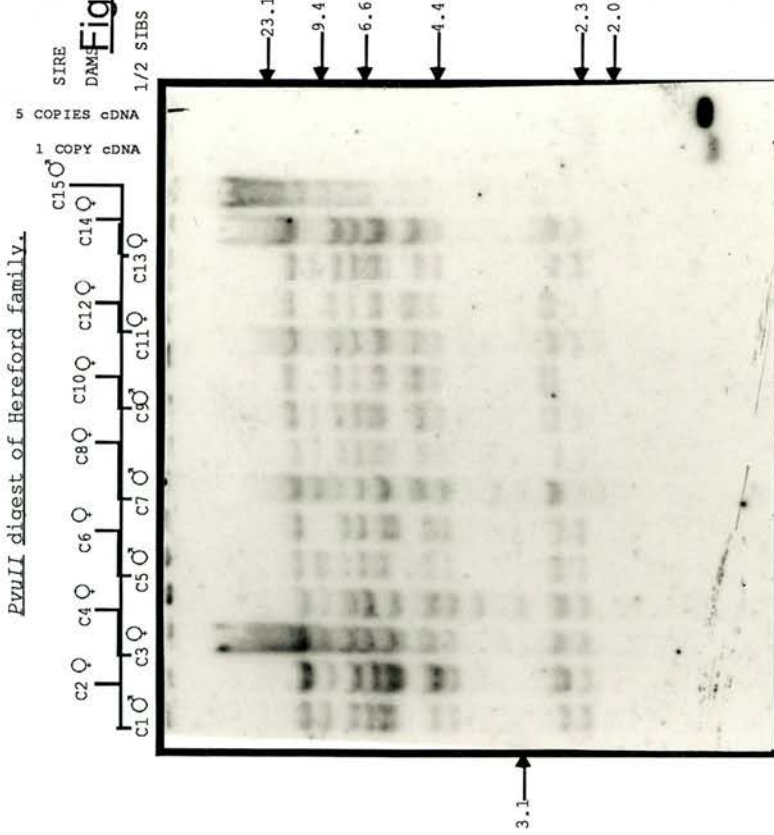


Figure 5. *Bam*HI and *Pst*I digests of DNA from a Hereford half-sib family probed with pBoLA-1 after Southern transfer to a nylon membrane.

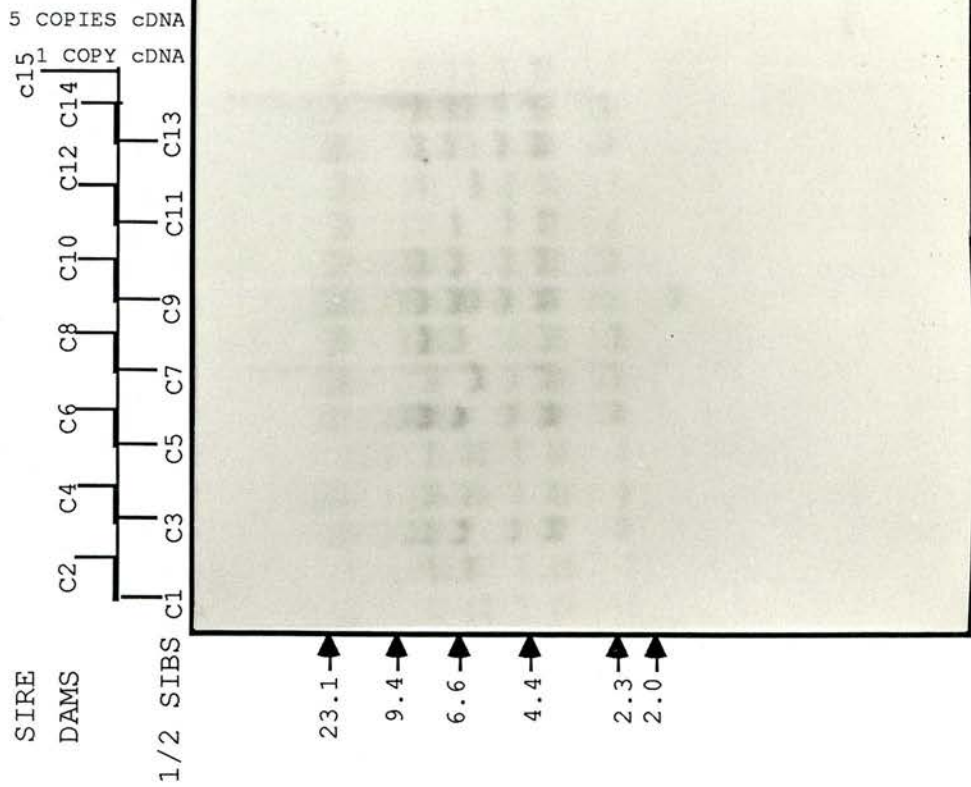
The first panel shows *Bam*HI digested DNA fractionated on an agarose gel, transferred and probed with pBoLA-1. The blot was washed to 0.5xSSC, 0.1% SDS at 65°C and exposed for 10 days at -70°C. The size markers are a *Hind*III digest of lambda phage DNA and are indicated on the figure.

The second panel shows another gel of the same DNA samples which were digested with *Pst*I. Although there is significant variation in the band pattern on each of these gels, no correlation to the serological type of the animal could be found.

Each gel was loaded left to right: 9898 sib, *BoLA* type 20/- (C1); 9241 dam, *BoLA* type 1/13/20 (C2); 9897 sib, *BoLA* type 9/20 (C3), 9480 dam, *BoLA* type 6/6.2/20 (C4); 9869 sib, *BoLA* type 9/20 (C5); 9315 dam, *BoLA* type 7/20 (C6); 9865 sib, *BoLA* type 7/20 (C7); 9066 dam, *BoLA* type 9/20 (C8); 9847 sib, *BoLA* type 1/13/20 (C9), 9027 dam, *BoLA* type 1/13/- (C10); 9844 sib, *BoLA* type 9/20 (C11); 9072 dam, *BoLA* type 9/20 (C12); 9797 sib, *BoLA* type 9/20 (C13); 9169 dam, *BoLA* type 1/13/20 (C14); 9547 sire, *BoLA* type 9/20 (C15); 1 copy equivalent of pBoLA-1 and 5 copies of pBoLA-1.

Figure 5

BamHI digest of Hereford family C.



PstI digest of Hereford family C.

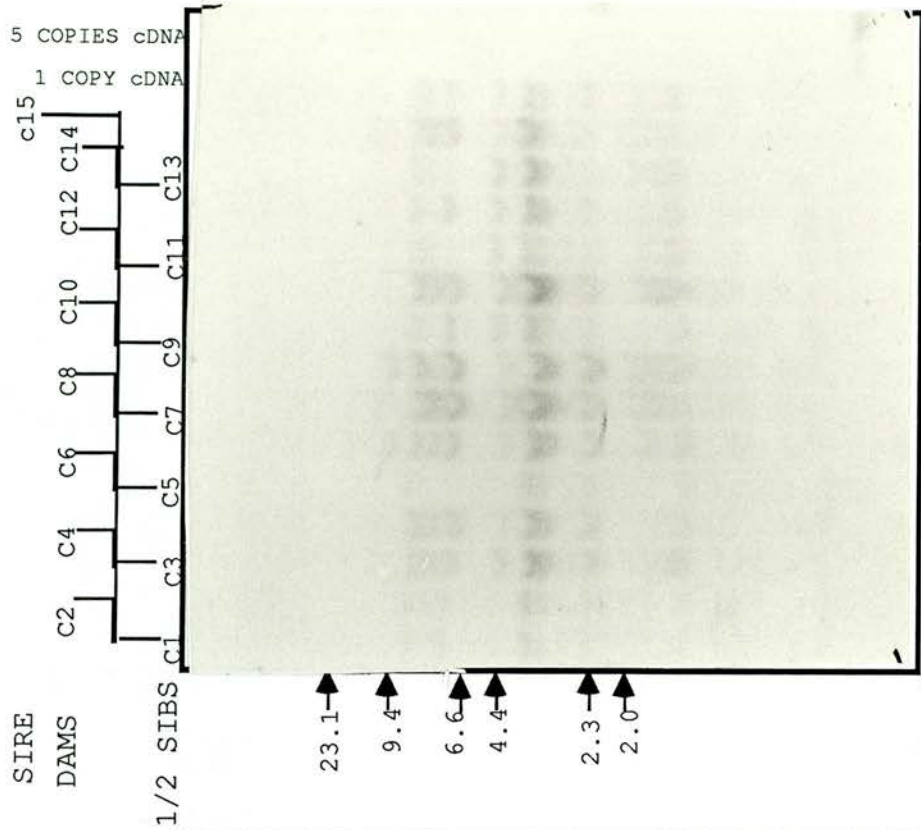


Figure 6. *PvuII* digest of DNA from a Friesian half-sib family probed with pBoLA-1 after Southern transfer.

DNA was extracted from peripheral blood lymphocytes of these animals before being digested with *PvuII*. The digests were fractionated on an agarose gel, transferred to nylon membrane and probed with pBoLA-1. The membrane was washed to low stringency in 2xSSC, 0.1% SDS at 65°C. The size marker was a *HindIII* digest of lambda phage DNA and the sizes are indicated on the left of the figure. The restriction fragments which correlate with serological types of the animals are indicated on the left of the figure. These are; 10.5kb with *BoLA-w8* and 2.9kb with *BoLA-w10*.

The gel was loaded with 10µg of DNA from left to right: 5604 sib, *BoLA* type -/- (Ov1); 7123 dam, *BoLA* type -/- (Ov2); 5612 sib, *BoLA* type 20/- (Ov3); 7437 dam, *BoLA* type 11/20 (Ov4); 5622 sib, *BoLA* type 6.4/- (Ov8); 7315 dam, *BoLA* type 6.4/8 (Ov9); 5623 sib, *BoLA* type -/- (Ov10); 7298 dam, *BoLA* type 5/Eu12 (Ov11); 7480 sib, *BoLA* type 11/- (Ov12); 7393 dam, *BoLA* type 8.2/11 (Ov13); 5605 sib, *BoLA* type 8/- (Ov14); 7361 dam, *BoLA* type 8/- (Ov15); 7492 sib, *BoLA* type 10/- (Ov21); 7320 dam, *BoLA* type 10/- (Ov22); 5625 sib, *BoLA* type 8/- (Ov5); 5608 sib, *BoLA* type 10/20 (Ov19) and 5613 sib, *BoLA* type 20/- (Ov20).

Freisian 1/2 sib family digested with PvuII  
and probed with pBoLA-1.

**Figure 6**

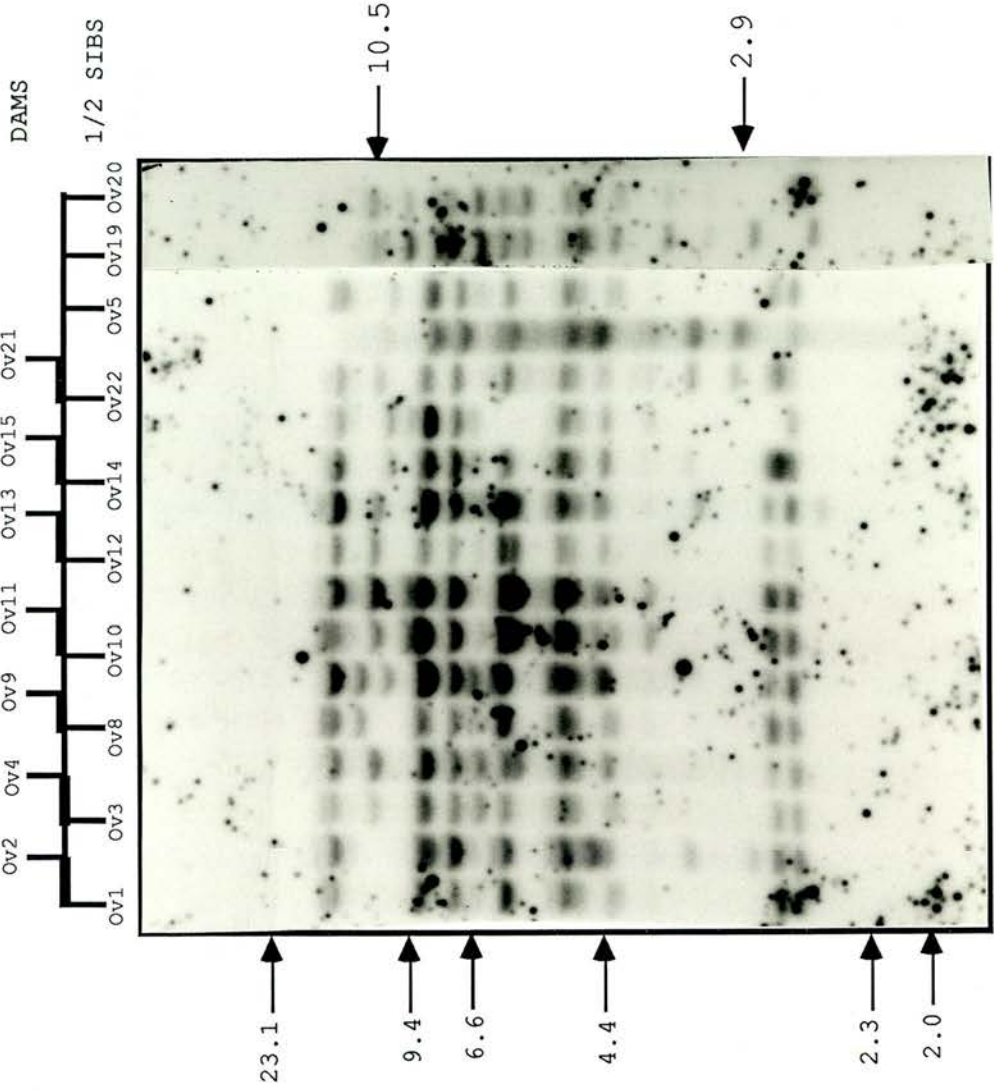


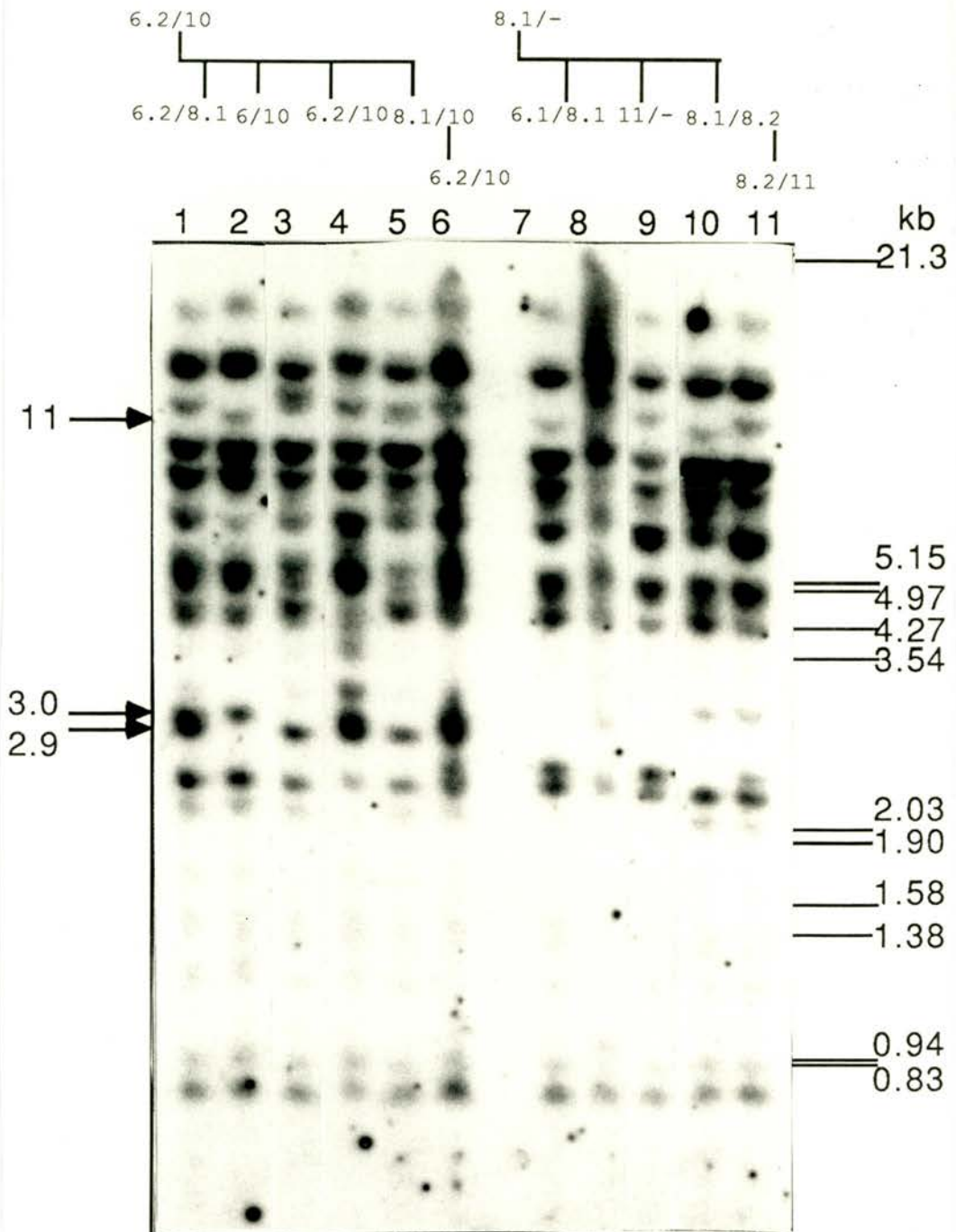
Figure 7. *PvuII* digest of DNA from two unrelated Friesian half-sib family with the dam as the common parent between the half-sibs, probed with pBoLA-1 after Southern transfer.

DNA was obtained from these animals and digested with *PvuII*. The digests were run on an agarose gel and transferred to nylon membrane. The blot was probed overnight, washed in 2xSSC, 0.1% SDS at 65° C and exposed for seven days. These extended pedigrees are different from the previous families because the common parent between the half-sibs is the dam, in addition to this second generation animals were available. In each case, track 5 is the mother of track 6 as is track 10 of track 11. The size marker is a *HindIII / EcoRI* digest of lambda phage DNA and the sizes are indicated on the right of the figure. The fragments in these animals which correlate with serological specificities are indicated on the left of the figure; 11kb fragment correlates with *BoLA-w8.1*, 3.0kb with *BoLA-w6.2* and 2.9kb with *BoLA-w10*.

The gel was loaded with 10µg of DNA per track, from left to right; 3420 dam, *BoLA* type 6.2/10 (track1); 10668 sib, *BoLA* type 6.2/8.1 (track2); 10442 sib, *BoLA* type 6/10 (track3); 10202 sib, *BoLA* type 6.2/10 (track4); 6707 sib (and dam of track 6), *BoLA* type 8.1/10 (track5); 10578 sib, *BoLA* type 6.2/10 (track6); 3893 dam, *BoLA* type 8.1/- (track 7); 10673 sib, *BoLA* type 6.1/8.1 (track 8); 10435 sib, *BoLA* type 11/- (track 9); 6637 sib (dam of track 11), *BoLA* type 8.1/8.2 (track 10) and 10555 sib, *BoLA* type 8.2/11 (track 11).



Figure 7



An in depth analysis of restriction fragment patterns has been carried out for the family in figures 3 and 6 and is presented in figure 8. Figure 8 illustrates the grouping of similar band lengths plotted against similar typed animals in an effort to determine if any other correlation can be made i.e. can serological types be defined by a variety of bands, giving an allo genotype. The correlations previously discussed can be seen on this type of graph as is the correlation of an extra 6.4kb band with *BoLA-w20*, but it is not apparent in every case. Unfortunately, the sample size for this type of study was probably too small to allow any meaningful conclusion to be reached, but it does enable easier examination of the results. The grouping of RFLP patterns in this way has previously been done in man (Cohen *et al.* 1985). Correlation of serological specificity to a restriction fragment detected within a family is useful, but is limited when a population is being examined. Oliver *et al.* (1989) (bound at back of thesis) present the results of a study in unrelated animals. They see more varied correlation of bands with various restriction enzymes but conclude that the technique is only valid if linked to isoelectric focusing studies of expressed antigens and to serology (if it is available).

Do the serological types show linkage to the RFLP bands ?

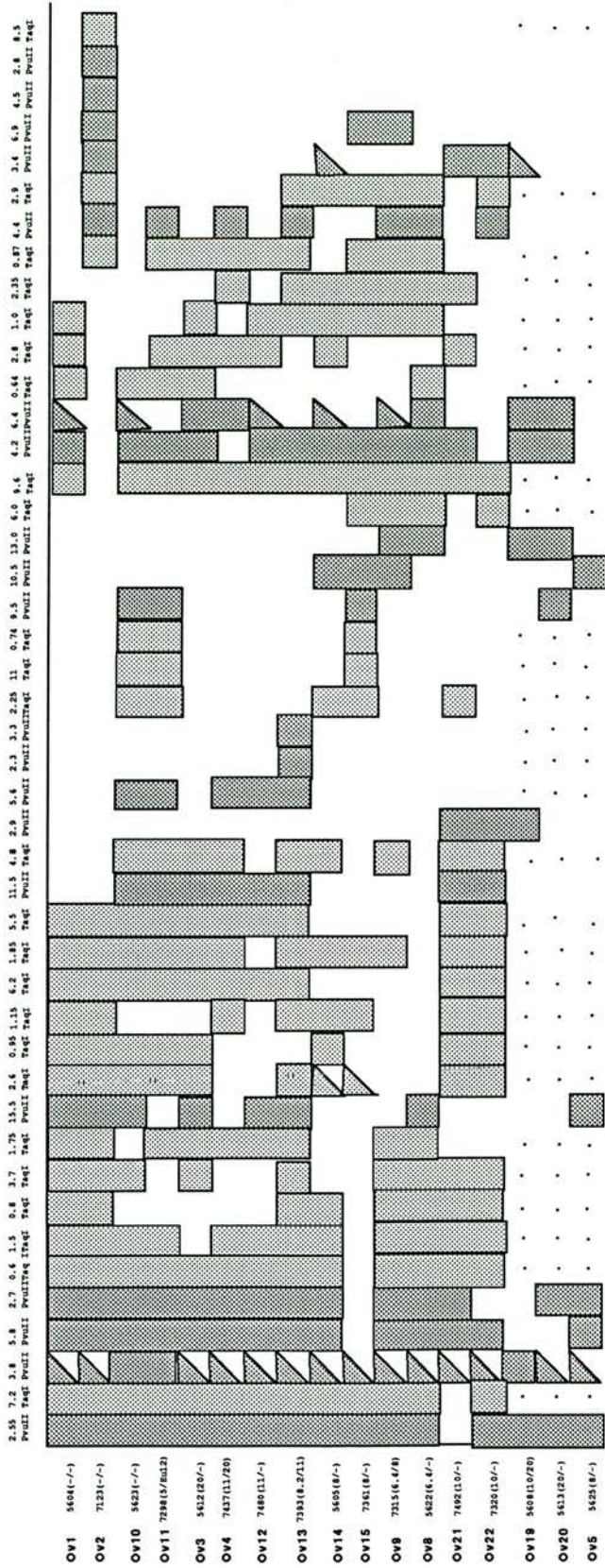
To demonstrate linkage between these serological types and RFLP fragments entails determining values for recombination fraction ( $\theta$ ); the hypothesis to be tested is one of free recombination  $\theta=0.5$  versus that of linkage  $\theta<0.5$ . Likelihood is defined as the probability with which given observations occur; the value of interest being the ratio of linked  $L(\theta<0.5)$ /unlinked  $L(\theta=0.5) = \lambda$ . The ratio  $\lambda$  is called the odds ratio with the  $\log$ (to the base 10) of this named the lod score.

Unfortunately, the animals depicted in figure 7 do not show any recombination between the correlated specificities *w6.2/w10* and the *Pvull* fragments of 3.0kb/2.9kb. This would imply very close linkage, except that the small sample size precludes any deduction of this sort. Assuming that the size of the MHC is roughly 1-5 centimorgans this is 1-5% recombination throughout its length. Therefore before a lod score could be determined, a minimum sample size of twenty animals would be required before you would expect to detect 1 recombinant.

Figure 8. Graph of RFLP's corresponding to serological types of the Friesian half-sib family already shown in figure 3 and 6.

This figure attempts to graphically depict the information already shown in figures 3 and 6. The family used consists of dams and half-sibs whose DNA has been digested with *TaqI* and *PvuII*, the results of these were shown in figures 3 and 6. The ordinate axis shows the sizes of the RFLP's detected when these Southern blots were probed with a bovine class I probe. The names of the animals and their serological types are marked on the figure. Similar serological types have been grouped together as have similar RFLP patterns. When a band is scored as being one copy it is represented as a box, a half copy is a triangle. Not all the animals have been analysed for both the restriction enzymes, where data is missing a dot has been inserted.

**Figure 8**



### 5.3 Summary.

This chapter presents the results of Southern blotting analysis of the *BoLA* class I genes with the bovine cDNA clone pBoLA-1. A heterologous probe cross-hybridised with a repeated element in the bovine genome. Use of the homologous probe (pBoLA-1) allowed a tally of the numbers of genes (5-7 with the 3' UT probe) and dissection of the MHC class I genes into polymorphic class Ia and nonpolymorphic class Ib genes. The 5' end of the cDNA probe appears to preferentially recognise class Ia genes at high stringency, suggesting that this cDNA is indeed a transcript of a class Ia gene.

Southern blotting studies on family groups has allowed correlation of RFLP's to serological types. Specific bands for *BoLA-w6.2*, *-w8* and *w10* have been identified.

## Conclusion.

Previously, it had become clear that cattle as an outbred species, with a long generation time, were not particularly well suited to immunological studies, unlike mice. Cattle therefore are more akin to man, both essentially being an outbred species. But, unlike man, cattle have not been as intensively investigated. Because of these problems the knowledge of the bovine MHC was scant in comparison to mouse and man. The most appropriate way to bridge the gap in knowledge was to use recombinant DNA technology; capitalising on the information available for mouse and man and applying it to cattle.

Heterologous class I cDNA probes were not suitable for use on Southern blots of bovine DNA (see chapter 5, figure 1) and were limited in their application. Therefore the first aim was to obtain a functional gene. A bovine liver cDNA library was constructed and screened with a heterologous probe. Two positives were obtained. The larger clone was further characterised and named pBoLA-1. The clone was not full length but did encode a mature class I protein. This was classified as a class Ia protein by a number of criteria, these included; 1), the clone showed greatest similarity to class Ia genes and proteins of other species (see chapter 3, figure 6) 2), it shared 100% amino acid homology with the transmembrane region of another bovine class I protein, encoded by the cDNA clone BL3-6 (see chapter 3, figures 4 and 5) 3), in Southern blotting studies at high stringency washing a probe derived from the 3' UT region detected specific bands (see chapter 5, figure 2), mimicking the results already seen for other class Ia genes in man (Koller *et al.* 1984) and 4) a 5' specific probe

corresponding to the leader peptide and first two domains of pBoLA-1 protein detected a polymorphic subset of class I-hybridising bands on a Southern blot, suggesting that it specifically detected class Ia genes (chapter 5, figure 3).

Once establishing that this clone was most probably a copy of class Ia mRNA it was used as a probe to screen a genomic library. The intention was to identify the genomic counterpart of pBoLA-1, unfortunately this was not the case. One of the most likely reasons for the failure, was that the genomic library used only corresponded to 65% of the bovine genome and had been amplified before being screened. Of the class I-hybridising phage which were detected two were selected for further study. The first, phage 41, contained an insert of approximately 20kb which contained two non-contiguous hybridising regions within it. Comparison of the hybridisation pattern obtained from *EcoRI/XbaI* restriction digests of the clone and DNA used to construct the library revealed that the clone appeared to have an insert which was <sup>probably</sup> present as 1 copy in the bovine genome (chapter 4, figure 3).

The theory that the bovine genome contains class Ib genes, is gained support from hybridisation studies of Southern blots. On digestion of cattle DNA with *TaqI* and probing with pBoLA-1 two different sets of bands emerged, a polymorphic and nonpolymorphic set. Having already established that pBoLA-1 preferentially identifies class Ia genes at high stringency, then by inference, the nonpolymorphic bands are most probably class Ib genes. The presence of nonpolymorphic bands which mostly correspond to the nonpolymorphic *Qa* and *Tla* genes has also been seen in mice



(Steinmetz *et al.* 1982b; Weiss *et al.* 1984). The second phage chosen for analysis was two thirds sequenced and found to have the remnants of a 5' end and a 3' end composing of exons 5-8 with an anomaly in its seventh exon (chapter 4, figures 6 and 8). This phage clone appears to contain a pseudogene and is very like the MHC gene fragments seen in mouse (Steinmetz *et al.* 1982b; Weiss *et al.* 1984) and man (Koller *et al.* 1989).

Having identified class Ia and class Ib genes on Southern blots, it remained to be determined if the polymorphic fragment pattern obtained could be used to type the animals. RFLP analysis of the bovine MHC is still at a very preliminary stage, but the results obtained did prove encouraging. The best restriction enzyme seemed to be *PvuII*. None of the other enzymes tried gave bands which could be used to type the animals, although they were polymorphic. The best correlation obtained was between a *PvuII* 3kb fragment and *BoLA-w6.2*. This correlation held across two different breeds of cattle and was inherited in a Mendelian fashion. The inheritance pattern suggested that either this fragment corresponds to the *-w6.2* specificity itself, or more likely it is in strong linkage disequilibrium with it. RFLP's which correlated with *BoLA-w8* and *-w10* were also identified. These RFLP studies have now been extended to include different *BoLA* types and a wider range of restriction enzymes [(Oliver *et al.* 1989), bound at back of thesis]. The conclusion reached in these further studies was that complete DNA typing of the bovine MHC was probably unrealistic, but used in conjunction with other techniques it is a useful tool.

Therefore the class I genes of the bovine MHC have been expanded from one serologically defined locus, *BoLA-A*, to a multi-gene family

consisting of class Ia and class Ib genes. It has emerged that *BoLA* is not atypical in its complement of class I genes and actually appears to be very similar to other species.

#### Future studies.

These results are really only the beginning of the molecular characterisation of the bovine MHC. It would have been better if the genomic counterpart of pBoLA-1 had been isolated and this would probably be one of the first objectives to be pursued. Ideally complete characterisation of the bovine MHC by isolation of overlapping genomic clones is probably the only way to identify how similar the bovine MHC is to that of mouse and man. During this work it also became clear that an assay for functional genes was needed, this could be provided by cell transfection of class I genes and screening for expression. But, as discussed throughout this thesis, this technique is also fraught with difficulty, especially when in cattle there is good serology for only one class I locus. It seems there are no easy options in identifying genes of interest in a multi-gene complex, as investigators of the MHC's of mouse and man have found. If real progress is to be made in understanding the bovine MHC, complete molecular characterisation will have to be undertaken. Only after this will informed immunological analysis of the bovine MHC be possible.

## References.

- Alexander, A. J., Bailey, E. and Woodward, J. G.: Analysis of the equine lymphocyte antigen system by Southern blot hybridisation. *Immunogenetics* 25 : 47-54, 1987
- Allen, H., Wraith, D., Pala, P., Askonas, B. and Flavell, R. A.: Domain interaction of H-2 class I antigens alter cytotoxic T-cell recognition sites. *Nature* 309 : 279-281, 1984
- Alper, C. A.: Complement and the MHC. In : *The role of the major histocompatibility complex in immunobiology* , M. E. Dorf (ed). Garland, New York, pp 173-220,1980
- Amor, M., Tosi, M., Duponchel, C., Steinmetz, M. and Meo, T.: Liver mRNA probes disclose two cytochrome P-450 genes duplicated in tandem with the complement *C4* loci of the mouse *H-2S* region. *Proc. Natl. Acad. Sci. U. S. A.* 82 : 4453-4457, 1985
- Amorena, B. and Stone, W. H.: Serologically defined (SD) locus in cattle. *Science* 201 : 159-160, 1978
- Andersson, L., Lunden, A., Sigurdardottir, S., Davies, C. J. and Rask, L.: Linkage relationships in the bovine MHC region. High recombination frequency between class II subregions. *Immunogenetics* 27 : 273-280, 1988
- Andersson, L. and Rask, L.: Characterisation of the MHC class II region in cattle. The number of *DQ* genes varies between haplotypes. *Immunogenetics* 27 : 110-120, 1988
- Archibald, A. L., Thompson, N. A. and Kvist, S.: A single nucleotide difference at the 3' end of an intron causes differential splicing of two histocompatibility genes. *EMBO J.* 5 : 957-965, 1986

- Arnold, B., Burgert, H. G., Archibald, A. L. and Kvist, S.: Complete nucleotide sequence of the murine *H-2K<sup>k</sup>* gene. Comparison of three H-2K locus alleles. *Nucleic Acids Res.* 12 : 9473-9487, 1984
- Arnold, B., Horstmann, U., Kuon, W., Burgert, H. G., Hammerling, G. J. and Kvist, S.: Alloreactive cytolytic T-cell clones preferentially recognise conformational determinants on histocompatibility antigens: analysis with genetically engineered hybrid antigens. *Proc. Natl. Acad. Sci. USA* 82 : 7030-7034, 1985
- Auffray, C., Ben-Nun, A., Roux-Dosseto, M., Germain, R. N., Seidman, J. G. and Strominger, J. L.: Polymorphism and complexity of the human DC and murine I-A $\beta$  chain genes. *EMBO J.* 2 : 121-124, 1983a
- Auffray, C., Kuo, J., DeMars, R. and Strominger, J. L.: A minimum of four human class II  $\alpha$ -chain genes are encoded in the HLA region of chromosome 6. *Nature* 304 : 174-177, 1983b
- Aviv, H. and Leder, P.: Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. U. S. A.* 69 : 1408-1412, 1972
- Bach, F. H. and Amos, D. B.: Hu-1: Major histocompatibility locus in man. *Science* 156 : 1506-1508, 1967
- Bach, F. H., Widmer, M. B., Bach, M. L. and Klein, J.: Serologically defined and lymphocyte-defined components of the major histocompatibility complex in the mouse. *J. Exp. Med.* 136 : 1430-1444, 1972
- Bain, B., Vas, M. and Lowenstein, L.: A reaction between leukocytes in mixed peripheral blood cultures. *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 22 : 428 (abstr), 1963
- Barker, D., Schafer, M. and White, R.: Restriction sites containing CpG

- show a higher frequency of polymorphism in human DNA. *Cell* 36 : 131-138, 1984
- Baxevanis, C. N., Nagy, Z. A. and Klein, J.: A novel type of T-T cell interaction removes the requirement for *I-B* region in the *H-2* complex. *Proc. Natl. Acad. Sci. U. S. A.* 78 : 3809-3813, 1981
- Benacerraf, B. and McDevitt, H. O.: Histocompatibility-linked immune response genes. *Science* 175 : 273-279, 1972
- Benoist, C. O., Mathis, D. J., Kanter, M. R., Williams II, V. E. and McDevitt, H. O.: Regions of allelic hypervariability in the murine  $A_\alpha$  immune response gene. *Cell* 34 : 169-177, 1983
- Bensaid, A., Naessens, J., Kemp, S. J., Black, S. J., Shapiro, S. Z. and Teale, A. J.: An immunochemical analysis of class I (BoLA) molecules on the surface of bovine cells. *Immunogenetics* 27 : 139-144, 1988
- Benton, W. D. and Davis, R. W.: Screening  $\lambda$ gt recombinant clones by hybridisation to single plaques in situ. *Science* 196 : 180-182, 1977
- Berggard, I. and Bearn, A. G.: Isolation and properties of a low molecular weight  $\beta_2$ -globulin occurring in human biological fluids. *J. Biol. Chem.* 243 : 4095-4103, 1968
- Biggin, M. D., Gibson, T. J. and Hong, G. F.: Buffer gradient gels and  $^{35}\text{S}$  label as an aid to rapid sequence determination. *Proc. Natl. Acad. Sci. USA* 80 : 3963-3965, 1983
- Bird, A. P.: CpG-rich islands and the function of DNA methylation. *Nature* 321 : 209-213, 1986
- Birnboim, H. C. and Doly, J.: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7 :

1513-1518, 1979

- Biro, P. A., Pan, J., Sood, A. K., Kole, R., Reddy, V. B. and Weissman, S. M.: The major histocompatibility complex. *Cold Spring Harbour Sym. Quant. Biol.* 47 : 1079-1086, 1983
- Bjorkman, P. J., Saper, M. A., Samaroui, B., Bennett, W. S., Strominger, J. L. and Wiley, D. C.: Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329 : 506-512, 1987a
- Bjorkman, P. J., Saper, M. A., Samaroui, B., Bennett, W. S., Strominger, J. L. and Wiley, D. C.: The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329 : 512-518, 1987b
- Blattner, F. R., Williams, B. G., Blechl, A. E., Denniston-Thompson, K., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Sheldon, E. L. and Smithies, O.: Charon phages: safer derivatives of bacteriophage lambda for DNA cloning. *Science* 196 : 161-163, 1977
- Bluestone, J. A., Foo, M., Allen, H., Segal, D. and Flavell, R. A.: Allospecific cytolytic T lymphocytes recognise conformational determinants on hybrid mouse transplantation antigens. *J. Exp. Med.* 162 : 268-281, 1985
- Bodmer, W. F.: A new genetic model for allelism at histocompatibility and other complex loci: polymorphism for control of gene expression. *Transplantation Proc.* 5 : 1471-1475, 1973
- Bodmer, J. and Bodmer, W.: Histocompatibility 1984. *Immunol. Today* 5 : 251-254, 1984
- Bohme, J., Owerbach, D., Denaro, M., Lernmark, A., Peterson, P. A. and Rask, L.: Human class II major histocompatibility antigen  $\beta$ -chains

- are derived from at least three loci. *Nature* 301 : 82-84, 1983
- Bonneville, M., Janeway, Jr, C. A., Ito, K., Haser, W., Ishida, I., Nakanishi, N. and Tonegawa, S.: Intestinal intraepithelial lymphocytes are a distinct set of  $\gamma\delta$  T cells. *Nature* 336 : 479-481, 1988
- Bourguignon, L. Y. W. and Singer, S. J.: Transmembrane interactions and the mechanism of capping of surface receptors by their specific ligands. *Proc. Natl. Acad. Sci. U. S. A.* 74 : 5031-5035, 1977
- Bourlet, Y., Behar, G., Guillemot, F., Frechin, N., Billaut, A., Chausse, A. M., Zoorob, R. and Auffray, C.: Isolation of chicken major histocompatibility complex class II (B-L)  $\beta$  chain sequences: comparison with mammalian  $\beta$  chains and expression in lymphoid organs. *EMBO J.* 7 : 1031-1039, 1988
- Boyse, E. A., Old, L. J. and Luell, S.: Antigenic properties of experimental leukemias. II. immunological studies *in vivo* with C57BL/6 radiation-induced leukemias. *J. Natl. Cancer Instit.* 31 : 987-995, 1963
- Bradbury, A., Belt, K. T., Neri, T. M., Milstein, C. and Calabi, F.: Mouse CD1 is distinct from and co-exists with TL in the same thymus. *EMBO J.* 7 : 3081-3086, 1988
- Breathnach, R. and Chambon, P.: Organisation and expression of eucaryotic split genes coding for proteins. *Ann. Rev. Biochem.* 50 : 349-383, 1981
- Bregegere, F.: A directional process of gene conversion is expected to yield dynamic polymorphism associated with stability of alternative alleles in class I histocompatibility antigens gene family. *Biochimie* 65 : 229-237, 1983

- Brenner, M. B., Strominger, J. L. and Krangel, M. S.: The  $\gamma\delta$  T cell receptor. *Advan. Immunol.* 43 : 133-192, 1988
- Briles, W. E., Stone, H. A. and Cole, R. K.: Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science* 195 : 193-195, 1977
- Brown, G. D., Choi, Y., Egan, G. and Meruelo, D.: Extension of the H-2 Tl<sup>b</sup> molecular map. *Immunogenetics* 27 : 239-251, 1988
- Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J. and Wiley, D.C.: A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature* 332 : 845-850, 1988
- Bull, R. W., Lewin, H. A., Wu, M. C., Peterbaugh, K., Antczak, D., Bernoco, D., Cwik, S., Dam, L., Dacies, C., Dawkins, R. L., Dufty, J. H., Gerlach, J., Hines, H. C., Lazary, S., Leibold, W., Leveziel, H., Lie, ø., Lindberg, P. G., Meggiolaro, D., Meyer, E., Oliver, R., Ross, M., Simon, M., Spooner, R. L., Stear, M. J., Teale, A. J. and Templeton, J. W.: Joint report of the third international bovine lymphocyte antigen (BoLA) workshop, Helsinki, Finland, 27 July 1986. *Animal Genetics* 20 : 109-132, 1989
- Calabi, F. and Milstein, C.: A novel family of human major histocompatibility complex-related genes not mapping to chromosome 6. *Nature* 323 : 540-543, 1986
- Cami, B., Bregegere, F., Abastado, J. P. and Kourilsky, P.: Multiple sequences related to classical histocompatibility antigens in the mouse genome. *Nature* 291 : 673-675, 1981
- Campbell, R. D., Law, S. K. A., Reid, K. B. M. and Sim, R. B.: Structure, organisation and regulation of the complement genes. *Ann. Rev.*



*Immunol. 6* : 161-195, 1988

- Cann, H. M., Ascanio, L., Paul, P., Marcadet, A., Dausset, J. and Cohen, D.: Polymorphic restriction endonuclease fragment segregates and correlates with the gene for HLA-B8. *Proc. Natl. Acad. Sci. USA 80* : 1665-1668, 1983
- Carroll, M. C., Campbell, R. D., Bentley, D. R. and Porter, R. R.: A molecular map of the human major histocompatibility complex class III region linking complement genes C4, C2 and factor B. *Nature 307* : 237-241, 1984
- Carroll, M. C., Campbell, R. D. and Porter, R. R. : Mapping of steroid 21-hydroxylase genes adjacent to complement component C4 genes in HLA , the major histocompatibility complex in man. *Proc. Natl. Acad. Sci. USA 82* : 521-525, 1985
- Castano, R., Ezquerro, A., Domenech, N. and Lopez de Castro, J.: An HLA-A2 population variant with structural polymorphism in the  $\alpha 3$  region. *Immunogenetics 27* : 345-355, 1988
- Chaplin, D. D., Woods, D. E., Whitehead, A. S., Goldberger, G., Colten, H. R. and Seidman, J. G.: Molecular map of the murine S region. *Proc. Natl. Acad. Sci. USA 80* : 6947-6951, 1983
- Chardon, P., Kali, J., Leveziel, H., Colombani, J. and Vaiman, M.: Monoclonal antibodies to HLA recognise monomorphic and polymorphic epitopes on BoLA. *Tissue Antigens 22* : 62-71, 1983
- Chardon, P., Vaiman, M., Kirszenbaum, M., Geffrotin, C., Renard, C. and Cohen, D.: Restriction fragment length polymorphism of the major histocompatibility complex of the pig. *Immunogenetics 21* : 161-171, 1985a
- Chardon, P., Kirszenbaum, M., Cullen, P. R., Geffrotin, C., Auffray, C.,

- Strominger, J. L., Cohen, D. and Vaiman, M.: Analysis of the sheep MHC using HLA class I, II and C4 cDNA probes. *Immunogenetics* 22 : 349-358, 1985b
- Chimini, G., Pontarotti, P., Nguyen, C., Toubert, A., Boretto, J. and Jordan, B. R.: The chromosome region containing the highly polymorphic HLA class I genes displays limited large scale variability in the human population. *EMBO J.* 7 : 395-400, 1988
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J.: Isolation of biologically active ribonucleic acid from sources enriched in ribonucleases. *Biochemistry* 18 : 5294-5299, 1979
- Church, G. M. and Gilbert, W.: Genomic sequencing. *Proc. Natl. Acad. Sci. USA* 81 : 1991-1995, 1984
- Cohen, D., Paul, P., Font, M. P., Cohen, O., Sayagh, B., Marcadet, A., Busson, M., Mahouy, G., Cann, H. M. and Dausset, J.: Analysis of HLA class I genes with restriction endonuclease fragments: implications for polymorphism of the human major histocompatibility complex. *Proc. Natl. Acad. Sci. USA* 80 : 6289-6292, 1983
- Cohen, D., Paul, P., Le Gall, I., Marcadet, A., Font, M., Cohen-Haguenaer, O., Sayagh, B., Cann, H., Lalouel, J. and Dausset, J.: DNA polymorphism of HLA class I and class II regions. *Immunol. Rev.* 85 : 87-105, 1985
- Coligan, J. E., Kindt, T. J., Nairn, R., Nathenson, S. G., Sachs, D. H. and Hansen, T. H.: Primary structural studies of an H-2L molecule confirm that it is a unique gene product with homology to H-2K and H-2D antigens. *Proc. Natl. Acad. Sci. USA* 77 : 1134-1138, 1980
- Coligan, J. E., Kindt, T. J., Uehara, H., Martinko, J. and Nathenson, S. G.: Primary structure of a murine transplantation antigen. *Nature* 291 :

35-39, 1981

- Cook, R. G., Capra, J. D., Bednarczyk, J. L., Uhr, J. W. and Vitetta, E. S.: Structural studies on the murine Ia alloantigens VI. Evidence that both subunits of the I-A alloantigen are encoded by the I-A subregion. *J. Immunol.* 123 : 2799-2803, 1979
- Cook, R. G., Vitetta, E. S., Uhr, J. W. and Capra, J. D.: Structural studies on the murine Ia alloantigens VII. Further evidence for two-gene control of the E/C alloantigens. *J. Immunol.* 124 : 1594-1597, 1980
- Cook, R. G. and Landolfi, N. F.: Expression of the thymus leukemia antigen by activated peripheral T lymphocytes. *J. Exp. Med.* 158 : 1012-1017, 1983
- Cortese Hassett, A. L., Stranick, K. S., Locker, J., Kunz, H. W. and Gill III, T. J.: Molecular analysis of the rat MHC. 1. Delineation of the major regions in the MHC and in the *grc*. *J. Immunol.* 137 : 373-378, 1986
- Cosman, D., Khoury, G. and Jay, G.: Three classes of mouse H-2 messenger RNA distinguished by analysis of cDNA clones. *Nature* 295 : 73-76, 1982
- Cox, R. A.: The use of guanidinium chloride in the isolation of nucleic acids. *Methods Enzymol.* 12B : 120-129, 1968
- Dausset, J.: The major histocompatibility complex in man. *Science* 213 : 1469-1474, 1981
- Davis, M. M. and Bjorkman, P. J.: T-cell antigen receptor genes and T-cell recognition. *Nature* 334 : 395-402, 1988
- Demant, P., Capkova, J., Hinzova, E. and Voracova, B.: The role of the histocompatibility-2-linked *Ss-Slp* region in the control of mouse complement. *Proc. Natl. Acda. Sci. U. S. A.* 70 : 863-864, 1973

- Demant, P. and Ivanyi, D.: Further molecular complexities of H-2 *K*- and *D*- region antigens. *Nature* 290 : 146-149, 1981
- Devlin, J. J., Weiss, E. H., Paulson, M. and Flavell, R. A.: Duplicated gene pairs and alleles of class I genes in the Qa2 region of the murine major histocompatibility complex: a comparison. *EMBO J.* 4 : 3203-3207, 1985
- Dunham, I., Sargent, C. A., Trowsdale, J. and Cambell, R. D.: Molecular mapping of the human major histocompatibility complex by pulsed-field gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 84 : 7237-7241, 1987
- Dupont, B., Oberfield, S. E., Smithwick, E. M., Lee, T. D. and Levine, L. S.: Close genetic linkage between HLA and congenital adrenal hyperplasia (21-hydroxylase deficiency). *Lancet* 2 : 1309-1312, 1977
- Dutton, R. W.: Spleen cell proliferation in response to homologous antigens in congenic resistant strains of mice. *J. Exp. Med.* 123 : 665-671, 1966
- Eccles, S. J. and McMaster, W. R.: DNA sequence analysis of a rat RT1 class II *A $\beta$*  gene. *Immunogenetics* 22 : 653-663, 1985
- Ennis, P. D., Jackson, A. P. and Parham, P.: Molecular cloning of bovine class I MHC cDNA. *J. Immunol* 141 : 642-651, 1988
- Eugui, E. M. and Emery, D. L.: Genetically restricted cell-mediated cytotoxicity in cattle immune to *Theileria parva* . *Nature* 290 : 251-254, 1981
- Evans, G. A., Margulies, D. H., Camerini-Otero, R. D., Ozato, K. and Seidman, J. G.: Structure and expression of a mouse major histocompatibility antigen gene, *H-2L<sup>d</sup>*. *Proc. Natl. Acad. Sci. USA*

79 : 1994-1998, 1982a

- Evans, G. A., Margulies, D. H., Shykind, B., Seidman, J. G. and Ozato, K.: Exon shuffling: mapping polymorphic determinants on hybrid mouse transplantation antigens. *Nature* 300 : 755-757, 1982b
- Feinberg, A. P. and Vogelstein, B.: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* 132 : 6-13, 1983
- Ferreria, A., Nussenzweig, V. and Gigli, I.: Structural and functional differences between the H-2 controlled Ss and Slp proteins. *J. Exp. Med.* 148 : 1186-1197, 1978
- Fisher, D. A., Hunt III, S. T. and Hood, L.: Structure of a gene encoding a murine thymus leukemia antigen and organisation of *Tla* genes in the BALB/c mouse. *J. Exp. Med.* 162 : 528-545, 1985
- Flaherty, L.: The *Tla* region antigens. In: *The Role of the Major Histocompatibility Complex in Immunology*, M. E. Dorf (ed.), STPM Press, Garland, New York : 33-57, 1980
- Flanagan, M. P., Jung, Y. C., Rothschild, M. F and Warner, C. M.: RFLP analysis of *SLA* class I genotypes in Duroc swine. *Immunogenetics* 27 : 465-469, 1988
- Flood, P. M., Benoist, C., Mathis, D. and Murphy, D. B.: Altered I-J phenotype in  $E_{\alpha}$  transgenic mice. *Proc. Natl. Acad. Sci. USA* 83 : 8308-8312, 1986
- Frels, W. I., Bluestone, J. A., Hodes, R. J., Capecchi, M. R. and Singer, D. S.: Expression of a microinjected porcine class I major histocompatibility complex gene in transgenic mice. *Science* 228 : 577-580, 1985
- Gates III, F. T., Coligan, J. E. and Kindt, T. J.: Complete amino acid

- sequence of murine  $\beta_2$ -microglobulin: structural evidence for strain-related polymorphism. *Proc. Natl. Acad. Sci. USA* 78 : 554-558, 1981
- Geraghty, D. E., Koller, B. H. and Orr, H. T.: A human major histocompatibility complex class I gene that encodes a protein with a shortened cytoplasmic segment. *Proc. Natl. Acad. Sci. USA* 84 : 9145-9149, 1987
- Glass, E. J. and Spooner, R. L.: Requirement for MHC class II positive accessory cells in an antigen specific bovine T cell response. *Res. Vet. Science* 46 : 196-201, 1989
- Goddeeris, B. M., Morrison, W. I., Teale, A. J., Bensaid, A. and Baldwin, C. L.: Bovine cytotoxic T-cell clones specific for cells infected with the protozoan parasite *Theileria parva* : parasite strain specificity and class I major histocompatibility complex restriction. *Proc. Natl. Acad. Sci. USA* 83 : 5238-5242, 1986
- Goding, J. W.: Evidence for linkage of murine  $\beta_2$ -microglobulin to H-3 and Ly-4. *J. Immunol.* 126 : 1644-1646, 1981
- Goodenow, R. S., McMillan, M., Orn, A., Nicolson, M., Davidson, N., Frelinger, J. A. and Hood, L.: Identification of a BALB/c H-2L<sup>d</sup> gene by DNA-mediated gene transfer. *Science* 215 : 677-679, 1982a
- Goodenow, R. S., McMillan, M., Nicolson, M., Sher, B. T., Eakle, K., Davidson, N. and Hood, L.: Identification of the class I genes of the mouse major histocompatibility complex by DNA-mediated gene transfer. *Nature* 300 : 231-237, 1982b
- Goodfellow, P. N., Jones, E. A., Heyningen, V. V., Solomon, E., Bobrow, M., Miggiano, V. and Bodmer, W. F.: The  $\beta_2$ -microglobulin gene is on chromosome 15 and not in the HL-A region. *Nature* 254 : 267-269,

1975

- Goodman, T. and Lefrancois, L.: Expression of the  $\gamma\delta$  T-cell receptor on intestinal CD8<sup>+</sup> lymphocytes. *Nature* 333 : 855-857, 1988
- Gorer, P. A., Lyman, S. and Snell, G. D.: Studies on the genetic and antigenic basis of tumour transplantation. Linkage between a histocompatibility gene and 'fused' in mice. *Proc. Roy. Soc. (Biol)*. 135 : 499-505, 1948
- Gorer, P. A. and Milulska, Z. B.: Some further data on the *H-2* system of antigens. *Proc. Royal Soc. (Series B)* 151 : 57-69, 1959
- Gorman, J. C., Jackson, R. J., DeSantola, J. R., Shreffler, D. C. and Atkinson, J. P.: Development of the hemolytic assay for mouse C2 and determination of its genetic control. *J. Immunol.* 125 : 344-351, 1980
- Goto, R., Miyada, G., Young, S., Wallace, R. B., Abplanalp, H., Bloom, S. E., Briles, W. E. and Miller, M. M.: Isolation of a cDNA clone from the *B-G* subregion of the chicken histocompatibility (*B*) complex. *Immunogenetics* 27 : 102-109, 1988
- Gubler, U. and Hoffman, B. J.: A simple and very efficient method for generating cDNA libraries. *Gene* 25 : 263-269, 1983
- Guild, B. C., Erikson, R. L. and Strominger, J. L.: HLA-A2 and HLA-B7 antigens are phosphorylated *in-vitro* by Rous sarcoma virus (pp60<sup>V-SRC</sup>) at a tyrosine residue encoded in a highly conserved exon of the intracellular domain. *Proc. Natl. Acad. Sci. USA* 80 : 2894-2898, 1983
- Guild, B. C. and Strominger, J. L.: Human and murine class I MHC antigens share conserved serine 335, the site of HLA phosphorylation *in vivo* . *J. Biol. Chem.* 259 : 9235-9240, 1984

- Guilletmot, F., Billaut, A., Pourquie, O., Behar, G., Chausse, A. M., Zoorob, R., Kreibich, G. and Auffray, C.: A molecular map of the chicken major histocompatibility complex: the class II  $\beta$  genes are closely linked to the class I genes and the nucleolar organiser. *EMBO J. 7* : 2775-2785, 1988
- Gunther, E. and Wurst, W.: Cytotoxic T lymphocytes of the rat are predominantly restricted by RT1.A and not RT1.C-determined major histocompatibility antigens. *Immunogenetics 20* : 1-12, 1984
- Gussow, D., Rein, R. S., Meijer, I., Hoog, W., Seemann, G. H. A., Hochstenbach, F. M. and Ploegh, H. L.: Isolation, expression and the primary structure of *HLA-Cw1* and *HLA-Cw2* genes: evolutionary aspects. *Immunogenetics 25* : 313-322, 1987
- Handy, D. E., McCluskey, J., Lew, A. M., Coligan, J. E. and Margulies, D. H.: Signals controlling alternative splicing of major histocompatibility complex H-2 class I pre-mRNA. *Immunogenetics 28* : 81-90, 1988
- Hansen, T. H., Ozato, K., Melino, M. R., Coligan, J. E., Kindt, T. J., Jandinski, J. J. and Sachs, D. H.: Immunochemical evidence in two haplotypes for at least three *D* region-encoded molecules, D, L and R. *J. Immunol. 126* : 1713-1716, 1981
- Hansen, T. H., Spinella, D. G., Lee, D. R. and Shreffler, D. C.: The immunogenetics of the mouse major histocompatibility gene complex. *Ann. Rev. Genet. 18* : 99-129, 1984
- Hardy, D. A., Bell, J. I., Long, E. O., Lindsten, T. and McDevitt, H. O.: Mapping of the class II region of the human major histocompatibility complex by pulsed field gel electrophoresis. *Nature 323* : 453-455, 1986



- Harrison, B. and Zimmerman, S. B.: Stabilisation of T4 polynucleotide kinase by macromolecular crowding. *Nucleic Acids Res.* 14 : 1863-1870, 1986
- Hildemann, W. H., Johnson, I. S. and Jokiel, P. L.: Immunocompetence in the lowest metazoan phylum: transplantation immunity in sponges. *Science* 204 : 420-422, 1979
- Hoang-Xuan, M., Leveziel, H., Zilber, M. T., Parodi, A. L. and Levy, D.: Immunochemical characterisation of major histocompatibility antigens in cattle. *Immunogenetics* 15 : 207-211, 1982a
- Hoang-Xuan, M., Charron, D., Zilber, M. T. and Levy, D.: Biochemical characterisation of class II bovine major histocompatibility complex antigens using cross-reactive antibodies. *Immunogenetics* 15 : 621-624, 1982b
- Hughes, A. L. and Nei, M.: Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* 335 : 167-170, 1988
- Huynh, T. V., Young, R. A. and Davis, R. W.: Constructing and screening cDNA libraries in gt10 and gt11. In D. M. Glover (ed.): *DNA cloning I. A Practical Approach*, IRL press, Oxford, 1986
- Janeway, Jr, C. A.: Frontiers of the immune system. *Nature* 333 : 804-806, 1988
- Janeway, Jr, C. A., Jones, B. and Hayday, A.: Specificity and function of T cells bearing  $\gamma\delta$  receptors. *Immunol. Today* 9 : 73-76, 1988
- Jelinek, W. R. and Schmid, C. W.: Repetitive sequences in eukaryotic DNA and their expression. *Ann. Rev. Biochem.* 51 : 813-844, 1982
- Jones, P. P., Murphy, D. B. and McDevitt, H. O.: Two-gene control of the expression of a murine Ia antigen. *J. Exp. Med.* 148 : 925-939, 1978

- Kappes, D. and Strominger, J. L.: Human class II major histocompatibility complex genes and proteins. *Ann. Rev. Biochem* 57 : 991-1028, 1988
- Kastern, W.: Characterisation of two class I major histocompatibility rat cDNA clones, one of which contains a premature termination codon. *Gene* 34 : 227-233, 1985
- Kastner, D. L., Rich, R. R. and Shen, F. W.: *Qa-1*- associated antigens. I. Generation of *H-2*- nonrestricted cytotoxic T lymphocytes specific for determinants of the *Qa-1* region. *J. Immunol.* 123 : 1232-1238, 1979a
- Kastner, D. L., Rich, R. R. and Chu, L.: *Qa-1*- associated antigens. II. Evidence for functional differentiation from H-2K and H-2D antigens. *J. Immunol.* 123 : 1239-1244, 1979b
- Katz, D. H., Hamaoka, T. and Benacerraf, B.: Cell interactions between histoincompatible T and B lymphocytes. *J. Exp. Med.* 137 : 1405-1418, 1973
- Kaufman, J. F., Anderson, R. L. and Strominger, J. L.: HLA-DR antigens have polymorphic light chains and invariant heavy chains as assessed by lysine-containing tryptic peptide analysis. *J. Exp. Med.* 152 : 37-53, 1980
- Kaufman, J. F. and Strominger, J. L.: The extracellular region of light chains from human and murine MHC class II antigens consists of two domains. *J. Immunol.* 130 : 808-817, 1983
- Kaufman, J. F., Auffray, C., Korman, A. J., Shackelford, D. A. and Strominger, J.: The class II molecules of the human and murine major histocompatibility complex. *Cell* 36 : 1-13, 1984a
- Kaufman, J. F., Krangel, M. S. and Strominger, J. L.: Cysteines in the transmembrane region of major histocompatibility complex antigens

- are fatty acylated via thioester bonds. *J. Biol. Chem.* 259 : 7230-7238, 1984b
- Kemp, S. J., Spooner, R. L. and Teale, A. J.: A comparative study of major histocompatibility antigens in East African and European breeds. *Anim. Genet.* 19 : 17-30, 1988
- Kieny, M. P., Lathe, R. and Lecocq, J. P.: New versatile cloning and sequencing vectors based on bacteriophage M13. *Gene* 26 : 91-99, 1983
- Kimball, E. S. and Coligan, J. E.: Structure of class I major histocompatibility antigens. *Contemp. Topics Molec. Immunol.* 9 : 1-63, 1983
- Klein, J.: *Biology of the mouse Major Histocompatibility-2 Complex.* Springer-Verlag, New York, 1975
- Klein, J.: An attempt at an interpretation of the mouse *H-2* complex. *Contemp. Top. Immunobiol.* 5 : 297-336, 1976
- Klein, J.: Evolution and function of the major histocompatibility system: facts and speculations. In D. Gotze (ed.): *The major histocompatibility system in man and animals.* Springer-Verlag pp 339-378, 1977
- Klein, J.: The major histocompatibility complex of the mouse. *Science* 203 : 516-521, 1979
- Klein, J.: *Immunology the science of self-nonsel discrimination,* John Wiley and sons, New York, 1982
- Klein, J.: *Natural history of the Major Histocompatibility Complex,* John Wiley and sons, New York, 1986
- Klein, J. and Shreffler, D. C.: Evidence supporting a two-gene model for the *H-2* histoincompatibility system of the mouse. *J. Exp. Med.* 135 : 924-937, 1972

- Klein, J., Juretic, A., Baxevanis, C. N. and Nagy, Z. A.: The traditional and a new version of the mouse *H-2* complex. *Nature* 291 : 455-460, 1981
- Klein, J. and Figueroa, F.: The evolution of class I MHC genes. *Immunol. Today* 7 : 41-44, 1986
- Kobori, J. A., Strauss, E., Minard, K. and Hood, L.: Molecular analysis of the hotspot of recombination in the murine major histocompatibility complex. *Science* 234 : 173-179, 1986
- Koch, G. L. E. and Smith, M. J.: An association between actin and the major histocompatibility antigen H-2. *Nature* 273 : 274-278, 1978
- Koeller, D., Lieberman, R., Miyazaki, J. I., Appella, E., Ozato, K., Mann, D. W. and Forman, J.: Introduction of H-2D<sup>d</sup> determinants into the H-2L<sup>d</sup> antigen by site-directed mutagenesis. *J. Exp. Med.* 166 : 744-760, 1987
- Koller, B. H., Sidwell, B., DeMars, R. and Orr, H. T.: Isolation of *HLA* locus-specific DNA probes from the 3'-untranslated region. *Proc. Natl. Acad. Sci. USA* 81 : 5175-5178, 1984
- Koller, B. H. and Orr, H. T.: Cloning and complete sequence of an *HLA-A2* gene: analysis of two *HLA-A* alleles at the nucleotide level. *J. Immunol.* 134 : 2727-2733, 1985
- Koller, B. H., Geraghty, D. E., Shimizu, Y., DeMars, R. and Orr, H. T.: *HLA-E* A novel *HLA* class I gene expressed in resting T lymphocytes. *J. Immunol.* 141 : 897-904, 1988
- Koller, B. H., Geraghty, D. E., DeMars, R., Duvick, L., Rich, S. S. and Orr, H. T.: Chromosomal organisation of the human major histocompatibility complex class I gene family. *J. Exp. Med.* 169 : 469-480, 1989
- Kottmann, A. H., Seemann, G. H. A., Gussow, H. D. and Roos, M. H.: DNA

- sequence of the coding region of the *HLA-Bw44* gene. *Immunogenetics* 23 : 396-400, 1986
- Kramerov, D. A., Lekakh, I. V., Samarina, O. P. and Ryskov, A. P.: The sequences homologous to major interspersed repeats B1 and B2 of mouse genome are present in mRNA and small cytoplasmic poly(A)<sup>+</sup> RNA. *Nucleic Acids Res.* 10 : 7477-7491, 1982
- Krangel, M. S., Biddison, W. E. and Strominger, J. L.: Comparative structural analysis of HLA-A2 antigens distinguishable by cytotoxic T lymphocytes. *J. Immunol.* 130 : 1856-1862, 1983
- Kress, M., Cosman, D., Khoury, G. and Jay, G.: Secretion of a transplantation-related antigen. *Cell* 34 : 189-196, 1983a
- Kress, M., Glaros, D., Khoury, G. and Jay, G.: Alternative splicing in the expression of the *H-2K* gene. *Nature* 306 : 602-604, 1983b
- Kress, M., Liu, W. Y., Jay, E., Khoury, G. and Jay, G.: Comparison of class I (H-2) gene sequences. *J. Biol. Chem.* 258 : 13929-13936, 1983c
- Kress, M., Barra, Y., Seidman, J. G., Khoury, G. and Jay, G.: Functional insertion of an alu type 2 (B2 SINE) repetitive sequence in murine class I genes. *Science* 226 : 974-977, 1984
- Kvist, S., Roberts, L. and Dobberstein, B.: Mouse histocompatibility genes: structure and organisation of a K<sup>d</sup> gene. *EMBO J.* 2 : 245-254, 1983
- Lalanne, J. L., Bregegere, F., Delarbre, C., Abastado, J. P., Gachelin, G. and Kourilsky, P.: Comparison of nucleotide sequences of mRNAs belonging to the mouse H-2 multigene family. *Nucleic Acids Res.* 10 : 1039-1049, 1982
- Lalanne, J. L., Delarbre, C., Gachelin, G. and Kourilsky, P.: A cDNA clone containing the entire coding sequence of a mouse H-2K<sup>d</sup> histocompatibility antigen. *Nucleic Acids Res.* 11 : 1567-1577,

1983

- Lalanne, J. L., Transy, C., Guerin, S., Darche, S., Meulien, P. and Kourilisky, P.: Expression of class I genes in the major histocompatibility complex: identification of eight distinct mRNAs in DBA/2 mouse liver. *Cell* 41 : 469-478, 1985
- Lalor, P. A., Morrison, W. I., Goddeeris, B. M., Jack, R. M. and Black, S. J.: Monoclonal antibodies identify phenotypically and functionally distinct cell types in the bovine lymphoid system. *Vet. Immunol. Immunopathol.* 13 : 121-140, 1986
- Lamont, S. J., Bolin, C. and Chevillat, N.: Genetic resistance to fowl cholera is linked to the major histocompatibility complex. *Immunogenetics* 25 : 284-289, 1987
- Lathe, R., Vilotte, J. L. and Clark, A. J.: Plasmid and bacteriophage vectors for excision of intact inserts. *Gene* 57 : 193-201, 1987
- Lawlor, D. A., Ward, F. E., Ennis, P. D., Jackson, A. P. and Parham, P.: *HLA-A* and *B* polymorphism predate the divergence of humans and chimpanzees. *Nature* 335 : 268-271, 1988
- Lawrence, S. K., Smith, C., Srivastava, R., Cantor, C. R. and Weissman, S. M.: Megabase-scale mapping of the HLA gene complex by pulsed field gel electrophoresis. *Science* 235 : 1387-1390, 1987
- Lee, J. S., Trowsdale, J. and Bodmer, W. F.: cDNA clones coding for the heavy chain of human HLA-DR antigen. *Proc. Natl. Acad. Sci. USA* 79 : 545-549, 1982
- LeGuern, C., Marche, P. N. and Kindt, T. J.: Molecular evidence for five distinct MHC class II  $\alpha$  genes in the rabbit. *Immunogenetics* 22 : 139-148, 1985
- LeGuern, C., Weissman, J. D., Marche, P. N., Jouvin-Marche, E.,

- Laverriere, A., Bagnato, M. R. and Kindt, T. J.: Sequence determination of a transcribed rabbit class II gene with homology to *HLA-DQ $\alpha$* . *Immunogenetics* 25 : 104-109, 1987
- Le Meur, M., Gerlinger, P., Benoist, C. and Mathis, D.: Correcting an immune-response deficiency by creating *E $\alpha$*  gene transgenic mice. *Nature* 316 : 38-42, 1985
- Letesson, J. J., Coppe, P., Lostrie-Trussart, N. and Depelchin, A.: A bovine Ia-like antigen detected by xenogenic monoclonal antibody. *Anim. Blood Groups and Biochem. Genet.* 14 : 239-250, 1983
- Levine, B. B., Ojeda, A. and Benacerraf, B.: Studies on artificial antigens. *J. Exp. Med.* 118 : 953-957, 1963
- Lew, A. M., Lillehoj, E. P., Cowan, E. P., Maloy, W. L., Van Schravendijk, M. R. and Coligan, J. E.: Class I genes and molecules: an update. *Immunology* 57 : 3-18, 1986a
- Lew, A. M., Valas, R. B., Maloy, W. L. and Coligan, J. E.: A soluble class I molecule analogous to mouse Q10 in the horse and related species. *Immunogenetics* 23 : 277-283, 1986b
- Lewin, H. A., Davis, W. C. and Bernoco, D.: Monoclonal antibodies that recognise bovine T and B lymphocytes. *Vet. Immunol. Immunopathol.* 9 : 87-102, 1985
- Lewin, H. A., Wu, M. C., Stewart, J. A. and Nolan, T. J.: Association between *BoLA* and subclinical bovine leukemia virus infection in a herd of Holstein-Friesian cows. *Immunogenetics* 27 : 338-344, 1988
- Lie, Ø., Solbu, H., Larsen, H. J. and Spooner, R. L.: Possible association of antibody responses to human serum albumin and (T, G)-A-L with the bovine major histocompatibility complex (*BoLA*). *Vet. Immunol.*

- Immunopathol. 11* : 335-350, 1986
- Lindberg, P. G. and Andersson, L.: Close association between DNA polymorphism of bovine major histocompatibility complex class I genes and serological BoLA-A specificities. *Anim. Genet. 19* : 245-255, 1988
- Lindsay, S. and Bird, A. P.: Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature 327* : 336-338, 1987
- Longenecker, B. M., Pazderka, F., Gavara, J. S., Spencer, J. L., Stevens, E. A., Witter, R. L. and Ruth, R.: Role of the major histocompatibility locus in resistance to Marek's disease: restriction of the growth of JMV-MV tumour cells in genetically resistant birds. *Adv. Exp. Med. Biol. 88* : 287-298, 1977
- Lopez de Castro, J. A., Orr, H. T., Kostyk, T., Mann, D. L. and Strominger, J. L.: Complete amino acid sequence of a papain-solubilised human histocompatibility antigen, HLA-B7. I. Isolation and amino acid composition of fragments and of tryptic and chymotryptic peptides. *Biochemistry 18* : 5704-5711, 1979
- Lopez de Castro, J. A., Strominger, J. L., Strong, D. M. and Orr, H. T.: Structure of crossreactive human histocompatibility antigens HLA-A28 and HLA-A2: possible implications for the generation of HLA polymorphism. *Proc. Natl. Acad. Sci. USA 79* : 3813-3817, 1982
- Lopez de Castro, J. A., Bragado, R., Strong, D. M. and Strominger, J. L.: Primary structure of papain-solubilised human histocompatibility antigen HLA-B40 (-Bw60). An outline of alloantigenic determinants. *Biochemistry 22* : 3961-3969, 1983
- Malissen, M., Malissen, B. and Jordan, B. R.: Exon/intron organisation



- and complete nucleotide sequence of an *HLA* gene. *Proc. Natl. Acad. Sci. USA* 79 : 893-897, 1982
- Maloy, W. L., Hammerling, G., Nathenson, S. G. and Coligan, J. E.: Comparison of alloantisera and hybridoma antibody for purification of the H-2D<sup>b</sup> murine histocompatibility antigen and preliminary molecular characterisation of this antigen. *J. Immunological Methods* 37 : 287-299, 1980
- Maloy, W. L. and Coligan, J. E.: Primary structure of the H-2D<sup>b</sup> alloantigen. *Immunogenetics* 16 : 11-22, 1982
- Maniatis, T., Fritsch, E. F. and Sambrook, J.: *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York 1982
- Mann, D. W. and Forman, J.: A third class I major histocompatibility complex antigen encoded by a gene in the *D* region of the *H-2<sup>d</sup>* haplotype recognised by cytotoxic T lymphocytes. *Immunogenetics* 28 : 38-45, 1988
- Marche, P. N., Tykocinski, M. L., Max, E. E. and Kindt, T. J.: Structure of a functional rabbit class I MHC gene: similarity to human class I genes. *Immunogenetics* 21 : 71-82, 1985
- Martin, L. H., Calabi, F. and Milstein, C.: Isolation of CD1 genes: a family of major histocompatibility complex-related differentiation antigens. *Proc. Natl. Acad. Sci. USA* 83 : 9154-9158, 1986
- Matis, L. A., Cron, R. and Bluestone, J. A.: Major histocompatibility complex-linked specificity of  $\gamma\delta$  receptor-bearing T lymphocytes. *Nature* 330 : 262-264, 1987
- Mayer, W. E., Jonker, M., Klein, D., Ivanyi, P., Seventer, G. van and Klein,

- J.: Nucleotide sequences of chimpanzee MHC class I alleles: evidence for *trans*-species mode of evolution. *EMBO J.* 7 : 2765-2774, 1988
- McDevitt, H. O. and Sela, M.: Genetic control of the antibody response. *J. Exp. Med.* 122 : 517-531, 1965
- McDevitt, H. O. and Chinitz, A.: Genetic control of the antibody response: relationship between immune response and histocompatibility (H-2) type. *Science* 163 : 1207-1208, 1969
- McGuire, K. L., Duncan, W. R. and Tucker, P. W.: Syrian hamster DNA shows limited polymorphism at class I-like loci. *Immunogenetics* 22 : 257-268, 1985
- McGuire, K. L., Duncan, W. R. and Tucker, P. W.: Structure of a class I gene from syrian hamster. *J. Immunol* 137 : 366-372, 1986
- McKean, D. J., Melvold, R. W. and David, C.: Tryptic peptide comparison of Ia antigen  $\alpha$  and  $\beta$  polypeptides from the *I-A* mutant B6.C-H-2<sup>bm12</sup> and its congenic parental strain B6. *Immunogenetics* 14 : 41-51, 1981
- Mellor, A. L., Weiss, E. H., Ramachandran, K. and Flavell, R. A.: A potential donor gene for the *bm1* gene conversion event in the C57BL mouse. *Nature* 306 : 792-795, 1983
- Mellor, A. L., Weiss, E., Kress, M., Jay, G. and Flavell, R. A.: A nonpolymorphic class I gene in the murine major histocompatibility complex. *Cell* 36 : 139-144, 1984
- Meo, T., Vives, J., Miggiano, V. and Shreffler, D.: A major role for the Ir-1 region of the mouse H-2 complex in the mixed leukocyte reaction. *Transplant. Proc.* 5 : 377-381, 1973
- Michaelson, J., Flaherty, L., Vitetta, E. and Poulik, M. D.: Molecular

- similarities between the Qa-2 alloantigen and other gene products of the 17th chromosome of the mouse. *J. Exp. Med.* 145 : 1066-1070, 1977
- Michaelson, J.: Genetic polymorphism of  $\beta_2$ -microglobulin (*B2m*) maps to the *H-3* region of chromosome 2. *Immunogenetics* 13 : 167-171, 1981
- Misra, D. N., Kunz, H. W., Cortese Hassett, A. L. and Gill III, T. J.: Comparison of rat MHC class I antigens by peptide mapping. *Immunogenetics* 25 : 35-46, 1987
- Mizuno, S., Trapani, J. A., Koller, B. H., Dupont, B. and Yang, S. Y.: Isolation and nucleotide sequence of a cDNA clone encoding a novel HLA class I gene. *J. Immunol.* 141 : 4024-4030, 1988
- Moore, K. W., Sher, B. T., Sun, Y. H., Eakle, K. A. and Hood, L.: DNA sequence of a gene encoding a BALB/c mouse  $L^d$  transplantation antigen. *Science* 215 : 679-682, 1982
- Morello, D., Daniel, F., Baldacci, P., Cayre, Y., Gachelin, G. and Kourilsky, P.: Absence of significant H-2 and  $\beta_2$ -microglobulin mRNA expression by mouse embryonal carcinoma cells. *Nature* 296 : 260-262, 1982
- Morita, T., Delarbre, C., Kress, M., Kourilsky, P. and Gachelin, G.: An *H-2K* gene of the  $t^{w32}$  mutant at the *T/t* complex is a close parent of an *H-2K<sup>q</sup>* gene. *Immunogenetics* 21 : 367-383, 1985
- Morrison, W. I., Goddeeris, B. M., Teale, A. J., Baldwin, C. L., Bensaid, A. and Ellis, J.: Cell-mediated immune responses of cattle to *Theileria parva*. *Immunol. Today* 7 : 211-216, 1986
- Muggli-Cockett, N. E. and Stone, R. T.: Identification of genetic variation in the bovine major histocompatibility complex DR $\beta$ -like

- genes using sequenced bovine genomic probes. *Animal Genetics* 19 : 213-225, 1988
- Muggli-Cockett, N. E. and Stone, R. T.: Nucleotide sequence of a bovine DR $\beta$ -like gene: partial nucleotide sequence of a bovine DR $\beta$ -like gene. *Animal Genetics* in press
- Muller, U., Jongeneel, C. V., Nedospasov, S. A., Fischer-Lindahl, K. and Steinmetz, M.: Tumour necrosis factor and lymphotoxin genes map close to *H-2D* in the mouse major histocompatibility complex. *Nature* 325 : 265-267, 1987
- Muller-Eberhard, H. J.: Molecular organisation and function of the complement system. *Ann. Rev. Biochem.* 57 : 321-347, 1988
- Murphy, D. B.: The I-J puzzle. *Ann. Rev. Immunol.* 5 : 405-427, 1987
- Murray, R., Hutchison III, C. A. and Frelinger, J. A.: Saturation mutagenesis of a major histocompatibility complex protein domain: identification of a single conserved amino acid important for allorecognition. *Proc. Natl. Acad. Sci. USA* 85 : 3535-3539, 1988
- Murre, C., Choi, E., Weis, J., Seidman, J. G., Ozato, K., Liu, L., Burakoff, S. J. and Reiss, C. S.: Dissection of serological and cytolytic T lymphocyte epitopes on murine major histocompatibility antigens by a recombinant H-2 gene separating the first two external domains. *J. Exp. Med.* 160 : 167-178, 1984a
- Murre, C., Reiss, C. S., Bernabeu, C., Chen, L. B., Burakoff, S. J. and Seidman, J. G.: Construction, expression and recognition of an H-2 molecule lacking its carboxyl terminus. *Nature* 307 : 432-436, 1984b
- Nairn, R., Yamaga, K. and Nathenson, S. G.: Biochemistry of the gene products from murine MHC mutants. *Ann. Rev. Genetics* 14 :

241-277, 1980

- Nathenson, S. G., Uerhara, H. and Ewenstein, B. M.: Primary structural analysis of the transplantation antigens of the murine H-2 major histocompatibility complex. *Ann. Rev. Biochem.* 50 : 1025-1052, 1981
- Nathenson, S. G., Geliebter, J., Pfaffenbach, G. M. and Zeff, R. A.: Murine major histocompatibility complex class-I mutants: molecular analysis and structure-function implications. *Ann. Rev. Immunol.* 4 : 471-502, 1986
- N'Guyen, C., Sodoyer, R., Trucy, J., Strachan, T. and Jordan, B. R.: The *HLA-AW24* gene: sequence, surroundings and comparison with the *HLA-A2* and *HLA-A3* genes. *Immunogenetics* 21 : 479-489, 1985
- Nizetic, D., Figueroa, F., Nevo, E. and Klein, J.: Major histocompatibility complex of the mole-rat. II. Restriction fragment polymorphism. *Immunogenetics* 22 : 55-67, 1985
- Nizetic, D., Stevanovic, M., Soldatovic, B., Savic, I. and Crkvenjakov, R.: Limited polymorphism of both classes of MHC genes in four different species of the Balkan mole rat. *Immunogenetics* 28 : 91-98, 1988
- Obata, Y., Chen, Y., Stockert, E. and Old, L. J.: Structural analysis of *TL* genes of the mouse. *Proc. Natl. Acad. Sci. USA* 82 : 5475-5479, 1985
- Obata, Y., Stockert, E., Chen, Y. T., Takahashi, T. and Old, L. J.: Influence of 5' flanking sequences on *TL* and H-2 expression in transfected L cells. *Proc. Natl. Acad. Sci. USA* 85 : 3541-3545, 1988
- Okayama, H. and Berg, P.: High-efficiency cloning of full-length cDNA. *Molecular and Cellular Biol.* 2 : 161-170, 1982
- Old, L. J., Boyse, E. A. and Stockert, E.: Antigenic properties of

- experimental leukemias. I. Serological studies *In Vitro* with spontaneous and radiation-induced leukemias. *J. Natl. Cancer Instit.* 31 : 977-986, 1963
- Oliver, R. A., McCoubrey, C. M., Millar, P., Morgan, A. L. G. and Spooner, R. L.: A genetic study of bovine lymphocyte antigens (BoLA) and their frequency in certain breeds. *Immunogenetics* 13 : 127-132, 1981
- Orr, H. T., Lancet, D., Robb, R. J., Lopez de Castro, J. A. and Strominger, J. L.: The heavy chain of human histocompatibility antigen HLA-B7 contains an immunoglobulin-like domain. *Nature* 282 : 266-270, 1979
- Orr, H. T., Bach, F. H., Ploegh, H. L., Strominger, J. L., Kavathas, P. and DeMars, R.: Use of HLA loss mutants to analyse the structure of the human major histocompatibility complex. *Nature* 296 : 454-456, 1982
- Orr, H. T. and DeMars, R.: Class I-like HLA genes map telomeric to the *HLA-A2* locus in human cells. *Nature* 302 : 534-536, 1983
- Ozato, K., Evans, G. A., Shykind, B., Margulies, D. H. and Seidman, J. G.: Hybrid *H-2* histocompatibility gene products assign domains recognised by alloreactive T cells. *Proc. Natl. Acad. Sci. USA* 80 : 2040-2043, 1983
- Palmer, M., Wettstein, P. J. and Frelinger, J. A.: Evidence for extensive polymorphism of class I genes in the rat major histocompatibility complex (*RTI*). *Proc. Natl. Acad. Sci. USA* 80 : 7616-7620, 1983
- Palmer, M. J. and Frelinger, J. A.: Widespread transcription of a *Qa* region gene in adult mice. *J. Exp. Med.* 166 : 95-109, 1987
- Passmore, H. C. and Shreffler, D. C.: A sex-limited serum protein variant in the mouse: inheritance and association with the *H-2*

- region. *Biochem. Genetics* 4 : 351-365, 1970
- Passmore, H. C. and Shreffler, D. C.: A sex-limited protein variant in the mouse: hormonal control of phenotypic expression. *Biochem. Genetics* 5 : 201-209, 1971
- Pazderka, F., Longenecker, B. M., Law, G. R. J. and Ruth, R. F.: The major histocompatibility complex of the chicken. *Immunogenetics* 2 : 101-130, 1975
- Pease, L. R., Nathenson, S. G. and Leinwand, L. A.: Mapping class I gene sequences in the major histocompatibility complex. *Nature* 298 : 382-385, 1982
- Pease, L. R., Schulze, D. H., Pfaffenbach, G. M. and Nathenson, S. G.: Spontaneous *H-2* mutants provide evidence that a copy mechanism analogous to gene conversion generates polymorphism in the major histocompatibility complex. *Proc. Natl. Acad. Sci. USA* 80 : 242-246, 1983
- Peterson, P. A., Cunningham, B. A., Berggard, I. and Edelman, G. M.:  $\beta_2$ -microglobulin- a free immunoglobulin domain. *Proc. Natl. Acad. Sci. U. S. A.* 69 : 1697-1701, 1972
- Pink, J. R. L., Droege, W., Hala, K., Miggiano, V. C. and Ziegler, A.: A three-locus model for the chicken major histocompatibility complex. *Immunogenetics* 5 : 203-216, 1977
- Pinkert, C. A., Widera, G., Cowing, C., Heber-Katz, E., Palmiter, R. D., Flavell, R. A. and Brinster, R. L.: Tissue-specific, inducible and functional expression of the  $E_{\alpha}^d$  MHC class II gene in transgenic mice. *EMBO J.* 4 : 2225-2230, 1985
- Ploegh, H. L., Orr, H. T. and Strominger, J. L.: Molecular cloning of a human histocompatibility antigen cDNA fragment. *Proc. Natl. Acad.*

*Sci. USA* 77 : 6081-6085, 1980

Ploegh, H. L., Orr, H. T. and Strominger, J. L.: Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell* 24 : 287-299, 1981

Pober, J. S. and Strominger, J. L.: Transglutaminase modifies the carboxy-terminal intracellular region of HLA-A and -B antigens. *Nature* 289 : 819-821, 1981

Potter, T. A., Bluestone, J. A. and Rajan, T. V.: A single amino acid substitution in the  $\alpha 3$  domain of an H-2 class I molecule abrogates reactivity with CTL. *J. Exp. Med.* 166 : 956-966, 1987

Potter, T. A., Rajan, T. V., Dick III, R. F. and Bluestone, J. A.: Substitution at residue 227 of H-2 class I molecules abrogates recognition by CD8-dependent, but not CD8-independent, cytotoxic T lymphocytes. *Nature* 337 : 73-75, 1989

Rask, L., Gustafsson, K., Larhammar, D., Ronne, H. and Peterson, P. A.: Generation of class II antigen polymorphism. *Immunol. Rev.* 84 : 59-83, 1985

Rebiere, M. C., Marche, P. N. and Kindt, T. J.: A rabbit class I major histocompatibility complex gene with a T cell-specific expression pattern. *J. Immunol.* 139 : 2066-2074, 1987

Reiss, C. S., Evans, G. A., Margulies, D., Seidman, J. G. and Burakoff, S. J.: Allospecific and virus-specific cytolytic T lymphocytes are restricted to the N or C1 domain of H-2 antigens expressed on L cells after DNA-mediated gene transfer. *Proc. Natl. Acad. Sci. USA* 80 : 2709-2712, 1983

Reyes, A. A., Schold, M., and Wallace, R. B.: The complete amino acid sequence of the murine transplantation antigen H-2D<sup>b</sup> as deduced by



- molecular cloning. *Immunogenetics* 16 : 1-9, 1982
- Robb, R. J., Terhrst, C. and Strominger, J. L.: Sequence of the COOH-terminal hydrophilic region of histocompatibility antigens HLA-A2 and HLA-B7. *J. Biol. Chem.* 253 : 5319-5324, 1978
- Robinson, P. J., Bevec, D., Mellor, A. L. and Weiss, E. H.: Sequence of the mouse Q4 class I gene and characterisation of the gene product. *Immunogenetics* 27 : 79-86, 1988
- Rogers, J. H.: Family organisation of mouse H-2 class genes. *Immunogenetics* 21 : 343-353, 1985a
- Rogers, J. H.: Mouse histocompatibility-related genes are not conserved in other mammals. *EMBO J.* 4 : 749-753, 1985b
- Roos, M. H. and Demant, P.: Murine complement factor B (BF): sexual dimorphism and H-2 linked polymorphism. *Immunogenetics* 15 : 23-30, 1982
- Rothbard, J. B., Hopp, T. P., Edelman, G. M. and Cunningham, B. A.: Structure of the heavy chain of the H-2K<sup>k</sup> histocompatibility antigen. *Proc. Natl. Acad. Sci. USA* 77 : 4239-4243, 1980
- Sachs, D. H., Germana, S., El-Gamil, M., Gustafsson, K., Hirsch, F. and Pratt, K.: Class II genes of miniature swine. *Immunogenetics* 28 : 22-29, 1988
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N.: Enzymatic amplification of  $\beta$ -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230 : 1350-1354, 1985
- Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. and Erlich, H. A.: Analysis of enzymatically amplified  $\beta$ -globin and HLA-DQ $\alpha$  DNA with allele-specific oligonucleotide probes. *Nature* 324 : 163-166,

1986

- Sanger, F., Nicklen, S. and Coulson, A. R.: DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74 : 5463-5467, 1977
- Sarmiento, U. M. and Storb, R. F.: Restriction fragment length polymorphism of the major histocompatibility complex of the dog. *Immunogenetics* 28 : 117-124, 1988
- Satz, M. L., Wang, L., Singer, D. S. and Rudikoff, S.: Structure and expression of two porcine genomic clones encoding class I MHC antigens. *J. Immunol.* 135 : 2167-2175, 1985
- Scott, P. C., Choi, C. L. and Brandon, M. R.: Genetic organisation of the ovine MHC class II region. *Immunogenetics* 25 : 116-122, 1987
- Seemann, G. H. A., Rein, R. S., Brown, C. S. and Ploegh, H. L.: Gene conversion like mechanisms may generate polymorphism in human class I genes. *EMBO J.* 5 : 47-52, 1986
- Shaw, S., Johnson, A. H. and Shearer, G. M.: Evidence for a new segregant series of B cell antigens that are encoded in the HLA-D region and that stimulate secondary allogeneic proliferative and cytotoxic responses. *J. Exp. Med.* 152 : 565-580, 1980
- Sher, B. T., Nairn, R., Coligan, J. E. and Hood, L. E.: DNA sequence of the mouse H-2D<sup>d</sup> transplantation antigen gene. *Proc. Natl. Acad. Sci. USA* 82 : 1175-1179, 1985
- Shimizu, Y., Geraghty, D. E., Koller, B. H., Orr, H. T. and DeMars, R.: Transfer and expression of three cloned human non-*HLA-A,B,C* class I major histocompatibility complex genes in mutant lymphoblastoid cells. *Proc. Natl. Acad. Sci. USA* 85 : 227-231, 1988
- Shiroishi, T., Evans, G. A., Appella, E. and Ozato, K.: Role of a disulphide bridge in the immune function of major histocompatibility class I

antigen as studied by *in-vitro* mutagenesis. *Proc. Natl. Acad. Sci. USA* 81 : 7544-7548, 1984

Shreffler, D. C.: A serologically detected variant in mouse serum: further evidence for genetic control by the histocompatibility-2 locus. *Genetics* 49 : 973-978, 1964

Shreffler, D. C.: *MHC-linked complement components*, in ed. P. Parham and J. L. Strominger *Structure and function . Receptors and recognition Series B*. Chapman Hall, pp187-219, 1982

Shreffler, D. C. and Owen, R. D.: A serologically detected variant in mouse serum: inheritance and association with the histoincompatibility-2 locus. *Genetics* 48 : 9-25, 1963

Shreffler, D. C. and David, C. S.: Studies on recombination within the mouse H-2 complex. *Tissue Antigens* 2 : 232-240, 1972

Shreffler, D. C. and David, C. S.: The *H-2* major histocompatibility complex and the *I* immune response region: genetic variation, function and organisation. *Advan. Immunol.* 20 : 125-195, 1975

Silver, J., Swain, S. L. and Hubert, J. J.: Small subunit of I-A subregion antigens determines the allospecificity recognised by a monoclonal antibody. *Nature* 286 : 272-274, 1980

Singer, D. S., Camerini-Otero, R. D., Satz, M. L., Osborne, B., Sachs, D. and Rudikoff, S.: Characterisation of a porcine genomic clone encoding a major histocompatibility antigen: expression in mouse L cells. *Proc. Natl. Acad. Sci. USA* 79 : 1403-1407, 1982

Sinha, A. A., Brautbar, C., Szafer, F., Friedmann, A., Tzfon, E., Todd, J. A., Steinman, L. and McDevitt, H. O.: A newly characterised HLA DQ $\beta$  allele associated with pemphigus vulgaris. *Science* 239 : 1026-1029, 1988

- Sittisombut, N. and Knight, K. L.: Rabbit major histocompatibility complex. 1. Isolation and characterisation of three subregions of class II genes. *J. Immunol.* 136 : 1871-1875, 1986
- Smithies, O. and Poulik, M. D.: Initiation of protein synthesis at an unusual position in an immunoglobulin gene. *Science* 175 : 187-189, 1972
- Snell, G. D.: Methods for the study of histocompatibility genes. *J. Genetics* 49 : 87-108, 1948
- Snell, G. D.: Histocompatibility genes of the mouse. II. Production and analysis of isogenic resistant lines. *J. Natl. Cancer Instit.* 21 : 843-877, 1958
- Snell, G. D., Demant, P. and Cherry, M.: Haemagglutination and cytotoxic studies of H-2 V. The anti-27, 28, 29 family of antibodies. *Folia Biologica (Praha)* 20 : 145-160, 1974
- Sodoyer, R., Damotte, M., Delovitch, T. L., Trucy, J., Jordan, B. R. and Strachan, T.: Complete nucleotide sequence of a gene encoding a functional human class I histocompatibility antigen (HLA-CW3). *EMBO J.* 3 : 879-885, 1984
- Soloski, M. J., Uhr, J. W., Flaherty, L. and Vitetta, E. S.: Qa-2, H-2K and H-2D alloantigens evolved from a common ancestral gene. *J. Exp. Med.* 153 : 1080-1093, 1981
- Soloski, M. J., Hood, L. and Stroynowski, I.: Qa-region class I gene expression: identification of a second class I gene, Q9, encoding a Qa-2 polypeptide. *Proc. Natl. Acad. Sci. USA* 85 : 3100-3104, 1988
- Sood, A. K., Pereira, D. and Weissman, S. M.: Isolation and partial nucleotide sequence of a cDNA clone for human histocompatibility antigen HLA-B7 by use of an oligodeoxynucleotide primer. *Proc. Natl. Acad. Sci. USA* 78 : 616-620, 1981

- Sood, A. K., Pan, J., Biro, P. A., Pereira, D., Srivastava, R., Reddy, V. B., Duceaman, B. W. and Weissman, S. M.: Structure and polymorphism of class I MHC antigen mRNA. *Immunogenetics* 22 : 101-121, 1985
- Southern, E. M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98 : 503-517, 1975
- Spencer, S. C. and Fabre, J. W.: Water-soluble form of RT1.A class I MHC molecules in the kidney and liver of the rat. *Immunogenetics* 25 : 91-98, 1987
- Spies, T., Morton, C. C., Nedospasov, S. A., Fiers, W., Pious, D. and Strominger, J. L.: Genes for the tumour necrosis factor  $\alpha$  and  $\beta$  are linked to the human major histocompatibility complex. *Proc. Natl. Acad. Sci. USA* 83 : 8699-8702, 1986
- Spies, T., Blanck, G., Bresnahan, M., Sands, J. and Strominger, J. L.: A new cluster of genes within the human major histocompatibility complex. *Science* 243 : 214-217, 1989
- Spooner, R. L.: The bovine major histocompatibility complex. In W. I. Morrison (ed.) *The ruminant immune system in health and disease*, Cambridge university press, Cambridge, pp133-151, 1986
- Spooner, R. L., Leveziel, H., Grosclaude, F., Oliver, R. A. and Vaiman, M.: Evidence for a possible major histocompatibility complex (BLA) in cattle. *J. Immunogenetics* 5 : 335-346, 1978
- Spooner, R. L. and Morgan, A. L. G.: Analysis of BoLA w6. Evidence for multiple subgroups. *Tissue Antigens* 17 : 178-188, 1981
- Spooner, R. L. and Pinder, M.: Monoclonal antibodies potentially detecting bovine MHC products. *Vet. Immunol. Immunopath.* 4 : 453-458, 1983
- Spooner, R. L., Innes, E. A., Millar, P., Webster, J. and Teale, A. J.:

- Bovine alloreactive cytotoxic cells generated *in-vitro* detect *BoLA* w6 subgroups. *Immunology* 61 : 85-91, 1987a
- Spooner, R. L., Leveziel, H., Queval, R. and Hoste, C.: Studies on the histocompatibility complex of indigenous cattle on the Ivory coast. *Vet. Immunol. Immunopathol.* 15 : 377-348, 1987b
- Srivastava, R., Duceaman, B. W., Biro, P. A., Sood, A. K. and Weissman, S. M.: Molecular organisation of the class I genes of human major histocompatibility complex. *Immunol. Rev.* 84 : 93-121, 1985
- Srivastava, R., Chorney, M. J., Lawrance, S. K., Pan, J., Smith, Z., Smith, C. and Weissman, S. M.: Structure, expression and molecular mapping of a divergent member of the class I *HLA* gene family. *Proc. Natl. Acad. Sci. USA* 84 : 4224-4228, 1987
- Stavenhagen, J. B. and Robins, D. M.: An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. *Cell* 55 : 247-254, 1988
- Steinmetz, M., Frelinger, J. G., Fisher, D., Hunkapiller, T., Pereira, D., Weissman, S. M., Uehara, H., Nathenson, S. and Hood, L.: Three cDNA clones encoding mouse transplantation antigens: homology to immunoglobulin genes. *Cell* 24 : 125-134, 1981a
- Steinmetz, M., Moore, K. W., Frelinger, J. G., Sher, B. T., Shen, F. W., Boyse, E. A. and Hood, L.: A pseudogene homologous to mouse transplantation antigens: transplantation antigens are encoded by eight exons that correlate with protein domains. *Cell* 25 : 683-692, 1981b
- Steinmetz, M., Minard, K., Horvath, S., McNicholas, J., Srelinger, J., Wake, C., Long, E., Mach, B. and Hood, L.: A molecular map of the immune response region from the major histocompatibility complex of the mouse. *Nature* 300 : 35-42, 1982a

- Steinmetz, M., Winoto, A., Minard, K. and Hood, L.: Clusters of genes encoding mouse transplantation antigens. *Cell* 28 : 489-498, 1982b
- Steinmetz, M., Malissen, M., Hood, L., Orn, A., Maki, R. A., Reza, G., Dastoornikoo, G. R., Stephan, D., Gibb, E. and Romaniuk, R.: Tracts of high or low sequence divergence in the mouse major histocompatibility complex. *EMBO J.* 3 : 2995-3003, 1984
- Steinmetz, M., Stephan, D. and Lindahl, K. F.: Gene organisation and recombinational hotspots in the murine major histocompatibility complex. *Cell* 44 : 895-904, 1986
- Stephan, D., Sun, H., Lindahl, K. F., Meyer, E., Hammerling, G., Hood, L. and Steinmetz, M.: Organisation and evolution of D region class I genes in the mouse major histocompatibility complex. *J. Exp. Med.* 163 : 1227-1244, 1986
- Stock, W. and Gunther, E.: Serologic and cellular characterisation of products of a new major histocompatibility gene region, RT1.C of the rat; possible homology to mouse H-2 Qa. *J. Immunol.* 128 : 1923-1928, 1982
- Strachan, T., Sodoyer, R., Damotte, M. and Jordan, B. R.: Complete nucleotide sequence of a functional class I HLA gene, *HLA-A3* : implications for the evolution of HLA genes. *EMBO J.* 3 : 887-894, 1984
- Stroynowski, I., Soloski, M., Low, M. G. and Hood, L.: A single gene encodes soluble and membrane bound forms of the major histocompatibility Qa-2 antigen: anchoring of the product by a phospholipid tail. *Cell* 50 : 759-768, 1987
- Szots, H., Riethmuller, G., Weiss, E. and Meo, T.: Complete sequence of HLA-B27 cDNA identified through the characterisation of

structural markers unique to the *HLA-A*, *-B* and *-C* allelic series.

*Proc. Natl. Acad. Sci. USA* 83 : 1428-1432, 1986

Teale, A. J., Morrison, W. I., Goddeeris, B. M., Grocock, C. M., Stagg, D. A. and Spooner, R. L.: Bovine alloreactive cytotoxic cells generated *in-vitro* : target specificity in relation to *BoLA* phenotype.

*Immunology* 55 : 355-362, 1985

Teale, A. J., Morrison, W. I., Spooner, R. L., Goddeeris, B. M., Grocock, C. M. and Stagg, D. A.: Bovine alloreactive cytotoxic T-cells. *In The ruminant immune system in health and disease* . ed W. I. Morrison, Cambridge University press 322-345, 1986a

Teale, A. J., Baldwin, C. L., Ellis, J. A., Newson, J., Goddeeris, B. M. and Morrison, W. I.: Alloreactive bovine T lymphocyte clones: An analysis of function, phenotype and specificity. *J. Immunol.* 136 : 4392-4398, 1986b

Teale, A. J. and Kemp, S. J.: A study of *BoLA* class II antigens with *Bo* T4+ T lymphocyte clones. *Anim. Genet.* 18 : 17-28, 1987

Todd, J. A., Bell, J. I. and McDevitt, H. O.: *HLA-DQ $\beta$*  gene contributes to susceptibility and resistance to insulin-dependant diabetes mellitus. *Nature* 329 : 599-604, 1987

Todd, J. A., Acha-Orbea, H., Bell, J. I., Chao, N., Fronck, Z., Jacob, C. O., McDermott, M., Sinha, A. A., Timmerman, T., Steinman, L. and McDevitt, H. O.: A molecular basis for MHC class II-associated autoimmunity. *Science* 240 : 1003-1009, 1988

Tonegawa, S.: Somatic generation of antibody diversity. *Nature* 302 : 575-581, 1983

Tonnelle, C., DeMars, R. and Long, E. O.: *DO $\beta$* : a new  $\beta$  chain gene in *HLA-D* with a distinct regulation of expression. *EMBO J.* 4 :



2839-2847, 1985

- Tosi, R., Tanigaki, N., Centis, D., Ferrara, G. B. and Pressman, D.: Immunological dissection of human Ia molecules. *J. Exp. Med.* 148 : 1592-1611, 1978
- Toyonaga, B. and Mak, T. W.: Genes of the T-cell antigen receptor in normal and malignant T cells. *Ann. Rev. Immunol.* 5 : 585-620, 1987
- Trowsdale, J. and Kelly, A.: The human HLA class II chain gene DZ $\alpha$  is distinct from genes in the DP, DQ and DR subregions. *EMBO J.* 4 : 2231-2237, 1985
- Trowsdale, J., Young, J. A. T., Kelly, A. P., Austin, P. J., Carson, S., Meunier, H., So, A., Ehrlich, H. A., Spielman, R. S., Bodmer, J. and Bodmer, W. F.: Structure, sequence and polymorphism in the HLA-D region. *Immunol. Rev.* 85 : 5-43, 1985
- Tykocinski, M. L., Marche, P. N., Max, E. E. and Kindt, T. J.: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expressed gene and a putative pseudogene. *J. Immunol.* 133 : 2261-2268, 1984
- Usinger, W. R., Curie-Cohen, M. and Stone, W. H.: Lymphocyte-defined loci in cattle. *Science* 196 : 1017-1018, 1977
- Usinger, W. R., Curie-Cohen, M., Benforado, K., Pringnitz, D., Rowe, R., Splitter, G. A. and Stone, W. H.: The bovine major histocompatibility complex (*BoLA*): close linkage of the genes controlling serologically defined antigens and mixed lymphocyte reactivity. *Immunogenetics* 14 : 423-428, 1981
- Vaiman, M., Chardon, P. and Cohen, D.: DNA polymorphism in the major histocompatibility complex of man and various farm animals. *Anim.*

*Genet. 17* : 113-133, 1986

Van Someron, H., Westerveld, A., Hagemeyer, A., Mees, J. R., Khan, P. M. and Zaalberg, O. B.: Human antigen and enzyme markers in man-chinese hamster somatic cell hybrids: evidence for syntenicity between the *HL-A*, *PGM<sub>3</sub>*, *ME<sub>1</sub>*, and *IPO-B* loci. *Proc. Natl. Acad. Sci. U. S. A.* *71* : 962-965, 1974

Vega, M. A., Bradago, R., Ezquerra, A. and Lopez de Castro, J. A.: Variability and conformation of HLA class I antigens: a predictive approach to the spatial arrangement of polymorphic regions. *Biochemistry* *23* : 823-831, 1984

Vega, M. A., Ezquerra, A., Rojo, S., Aparicio, P., Bradado, R. and Lopez de Castro, J. A.: Structural analysis of an HLA-B27 functional variant: identification of residues that contribute to the specificity of recognition by cytolytic T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* *82* : 7394-7398, 1985

Von Heijne, G.: Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* *133* : 17-21, 1983

Wahl, G. M., Stern, M. and Stark, G. R.: Efficient transfer of large DNA fragments from agarose gels to diazobenzylomethyl-paper and rapid hybridisation by using dextran sulphate. *Proc. Natl. Acad. Sci. U. S. A.* *76* : 3683-3687, 1979

Wake, C. T., Long, E. O. and Mach, B.: Allelic polymorphism and complexity of the genes for HLA-DR  $\beta$ -chains - direct analysis by DNA-DNA hybridisation. *Nature* *300* : 372-374, 1982

Wake, C. T. and Flavell, R. A.: Multiple mechanisms regulate the expression of murine immune response genes. *Cell* *42* : 623-628, 1985

- Walsh, F. S. and Crumpton, M. J.: Orientation of cell-surface antigens in the lipid bilayer of lymphocyte plasma membrane. *Nature* 269 : 307-311, 1977
- Walter, P., Gilmore, R. and Blobel, G.: Protein translocation across the endoplasmic reticulum. *Cell* 38 : 5-8, 1984
- Waneck, G. L., Sherman, D. H., Calvin, S., Allen, H. and Flavell, R. A.: Tissue-specific expression of cell-surface Qa-2 antigen from a transfected *Q7<sup>b</sup>* gene of C57BL/10 mice. *J. Exp. Med.* 165 : 1358-1370, 1987
- Warner, C. M.: Genetic manipulation of the major histocompatibility complex. *J. Anim. Sci.* 63 : 279-287, 1986
- Watson, C. J. and Jackson, J. F.: An alternative procedure for the synthesis of double-stranded cDNA for cloning in phage and plasmid vectors, *In* D. M. Glover (ed.): *DNA cloning I. A Practical Approach*, IRL press, Oxford, 1986
- Ways, J. P., Coppin, H. L. and Parham, P.: The complete primary structure of HLA-Bw58. *J. Biol. Chem.* 260 : 11924-11933, 1985
- Weiss, E., Golden, L., Zakut, R., Mellor, A., Fahrner, K., Kvist, S. and Flavell, R. A.: The DNA sequence of the H-2K<sup>b</sup> gene: evidence for gene conversion as a mechanism for the generation of polymorphism in histocompatibility antigens. *EMBO J.* 2 : 453-462, 1983a
- Weiss, E., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. and Flavell, R. A.: The structure of a mutant *H-2* gene suggests that the generation of polymorphism in *H-2* genes may occur by gene conversion-like events. *Nature* 301 : 671-674, 1983b
- Weiss, E. H., Golden, L., Fahrner, K., Mellor, A. L., Devlin, J. J., Bullman, H., Tiddens, H., Bud, H. and Flavell, R. A.: Organisation and evolution

- of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse. *Nature* 310 : 650-655, 1984
- White, P. C., Chaplin, D.D., Weis, J. H., Dupont, B., New, M. I. and Seidman, J. G.: Two steroid 21-hydroxylase genes are located in the murine S region. *Nature* 312 : 465-467, 1984a
- White, P. C., New, M. I. and Dupont, B.: HLA-linked congenital adrenal hyperplasia results from a defective gene encoding a cytochrome P-450 specific for steroid 21-hydroxylation. *Proc. Natl. Acad. Sci. USA* 81 : 7505-7509, 1984b
- Widera, G. and Flavell, R. A.: The nucleotide sequence of the murine I-E $\beta^b$  immune response gene: evidence for gene conversion events in class II genes of the major histocompatibility complex. *EMBO J.* 3 : 1221-1225, 1984
- Wilmot, I., Sales, D. I. and Ashworth, C. J.: The influence of variation in embryo stage and maternal hormone profiles on embryo survival in farm animals. *Theriogenology* 23 : 107-119, 1985
- Winchester, R. J. and Kunkel, H. G.: The human Ia system. *Advan. Immunol.* 28 : 221-292, 1979
- Winoto, A., Steinmetz, M. and Hood, L.: Genetic mapping in the major histocompatibility complex by restriction enzyme site polymorphisms: most mouse class I genes map to the *Tla* complex. *Proc. Natl. Acad. Sci. USA* 80 : 3425-3429, 1983
- Yamamura, K. I., Kikutani, H., Folsom, V., Clayton, L. K., Kimoto, M., Akira, S., Kashiwamura, S., Tonegawa, S. and Kishimoto, T.: Functional expression of a microinjected  $E_{\alpha}^d$  gene in C57BL/6 transgenic mice. *Nature* 316 : 67-69, 1985

- Yanisch-Perron, C., Vieira, J. and Messing, J.: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33 : 103-119, 1985
- Yokoyama, K. and Nathenson, S. G.: Intramolecular organisation of class I H-2 MHC antigens; localisation of the alloantigenic determinants and the  $\beta_2m$  binding site to different regions of the H-2K<sup>b</sup> glycoprotein. *J. Immunol.* 130 : 1419-1425, 1983
- Yuhki, N. and O'Brien, S. J.: Molecular characterisation and genetic mapping of class I and class II MHC genes of the domestic cat. *Immunogenetics* 27 : 414-425, 1988
- Yunis, E. J. and Amos, D. B.: Three closely linked genetic systems relevant to transplantation. *Proc. Natl. Acad. Sci. USA* 68 : 3031-3035, 1971
- Zinkernagel, R. M. and Doherty, P. C.: Restriction of *in vitro* T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature* 248 : 701-702, 1974a
- Zinkernagel, R. M. and Doherty, P. C.: Immunological surveillance against self components by sensitised T lymphocytes in lymphocytic choriomeningitis. *Nature* 251 : 547-548, 1974b
- Zinkernagel, R. M. and Doherty, P. C.: MHC-restricted cytotoxic T cells: studies on the biological role of polymorphic major transplantation antigens determining T-cell restriction-specificity, function and responsiveness. *Adv. Immunol.* 27 : 51-142, 1979

Addendum.

- Chen, P., B. and Parham, P.: Direct binding of influenza peptides to class I HLA molecules. *Nature 337* : 743-745, 1989
- Kane, K. P., Vitiello, A., Sherman, L. A. and Mescher, M. F.: Cytolytic T-lymphocytes response to isolated class I H-2 proteins and influenza peptides. *Nature 340* : 157-159, 1989
- Kisielow, P., Bluthmann, H., Staerz, U. D., Steinmetz, M. and von Boehmer, H.: Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature CD4<sup>+</sup>8<sup>+</sup> thymocytes. *Nature 333* : 742-746, 1988
- Kourilsky, P. and Claverie, J. M.: MHC restriction, alloreactivity and thymic education: a common link? *Cell 56* : 327-329, 1989
- Sha, W. C., Nelson, C. A., Newberry, R. D., Kranz, D. M., Russell, J. H. and Loh, D. Y.: Positive and negative selection on T cells in transgenic mice. *Nature 336* : 73-76, 1988
- Sprent, J., Lo, D., Gao, E. K. and Ron, Y.: T cell selection in the thymus. *Immunol. Rev. 101* : 173-190, 1988
- Townsend A. R. M., Rothbard, J., Gotch, F. M., Bahadur, G., Wraith, D. and McMichael, A. J.: The epitopes of influenza nucleoprotein recognised by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell 44* : 959-968, 1986
- Townsend, A. R. M., Ohlen, C., Bastin, J., Ljunggren, H. G., Foster, L. and Karre, K.: Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature 340* : 443-448, 1989

# The analyses of antigen and DNA polymorphism within the bovine major histocompatibility complex: 1. The class I antigens

R. A. OLIVER, P. BROWN, R. L. SPOONER, I. JOOSTEN\* & J. L. WILLIAMS

*AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, Edinburgh, UK and \*Department of Immunology, Faculty of Veterinary Medicine, University of Utrecht, The Netherlands*

**Summary.** Serology, isoelectric focusing (IEF) of expressed antigens, and restriction fragment length polymorphisms (RFLP) were compared for the identification of BoLA class I haplotypes. Expressed antigens identified as bands by IEF correlated well with serological definition confirming and extending our earlier findings (Joosten *et al.* 1988). Comparison of serology and isoelectric focusing bands with restriction fragments was more complicated; fragments were found which correlated both with broadly reacting and antigen specific sera. We also found correlation of fragments with two or more sera which showed no cross-reactivity. Fragments unique to particular haplotypes were also observed.

Serology remains the simplest method of typing BoLA class I antigens. Isoelectric focusing generally agrees with serological definition of antigens and detects antigens not yet defined by serology. It may also be useful in defining the products of other expressed BoLA class I loci. In order to identify RFLPs which could be used for typing, comparison with serology or IEF is essential. Haplotype specific RFLPs could be useful in identifying genes linked to the MHC.

*Keywords:* BoLA class I, IEF, RFLP

## Introduction

The major histocompatibility complex (MHC) class I antigens are cell surface glycoproteins expressed on most cell types. The BoLA-A locus encoded MHC class I antigens were identified by classical serology in 1978 (Amorena & Stone 1978; Spooner *et al.* 1978). Over the last decade, three international workshops have identified more than 30 BoLA class I alleles (Spooner *et al.* 1979; Anon. 1982; Bull *et al.* 1989). These behave in both population and family studies as products of a single

Correspondence: R.A. Oliver, AFRC Institute of Animal Physiology and Genetics Research, Edinburgh Research Station, West Mains Road, Edinburgh EH9 3JQ, Scotland.

Accepted 12 September 1988

locus (Oliver *et al.* 1981). In contrast to the three serologically defined MHC class I loci in man (HLA A, B and C) and mouse (K, D and L) (Klein 1986) there is no conclusive evidence for additional expressed BoLA class I loci, although at least one further expressed BoLA class I locus may exist (Bensaid *et al.* 1988; Bull *et al.* 1989).

One- and two-dimensional isoelectric focusing (IEF) have been used to characterize products of the human MHC (van der Poel *et al.* 1983; Neefjes *et al.* 1986). We have shown that one-dimensional IEF can be used to identify MHC class I haplotypes in cattle, the technique being particularly useful in that it is possible to identify BoLA class I products not yet identified by classical serology (Joosten *et al.* 1988).

At the DNA level a large number of MHC class I like sequences have been found in man (Cohen *et al.* 1985; Wake 1986) and mouse (Mellor 1986). Only three to five of these encode the classical MHC class I antigens, identified by one or more polymorphic and a few constant fragments by restriction enzyme analysis. The remainder encode the non-polymorphic Qa, Tl antigens or are pseudogenes. Restriction fragment analysis of DNA from various mammals revealed a few polymorphic and many non-polymorphic MHC class I fragments (Palmer *et al.* 1983; Chardon *et al.* 1985; McGuire *et al.* 1985). In cattle four to five polymorphic fragments and up to 10 non-polymorphic fragments are revealed using several restriction enzymes and a bovine class I cDNA probe (Brown *et al.* 1988).

Restriction fragment length polymorphisms (RFLPs) have been used to identify HLA class I alleles (Steinmetz *et al.* 1986). However, DNA sequences encoding the same expressed antigen gave rise to more than one set of polymorphic fragments. The relationship between DNA polymorphisms and expressed molecules was only solved by comparison with serological definition (Cohen *et al.* 1985).

In this study we have compared serology, IEF and RFLP for BoLA class I typing, using a selected panel of animals including the most frequent serologically defined alleles in the British Friesian breed. Results from each method were analysed to study the relationship between polymorphism of the expressed products (identified by serology and IEF) with that shown at the DNA level (RFLP).

## Materials and methods

### *Animals*

The cattle used in this study were from the AFRC British Friesian herd in Edinburgh. Unrelated animals were selected to represent the common BoLA serological specificities.

### *Serology*

BoLA class I antigens were defined by microlymphocytotoxicity testing as described by Spooner *et al.* (1979) with minor modifications. The lymphocytes were incubated with alloantisera at room temperature for 30 min before the addition of 5  $\mu$ l of neat rabbit complement and subsequent incubation at room temperature for 1 h.



Sera detecting most of the internationally recognized workshop specificities and additional local specificities were used.

#### *DNA preparation*

DNA was prepared from 25 ml of blood collected into heparinized vacutainers (Becton Dickinson). The blood was diluted 5:1 with 0.9% saline then centrifuged at 500xg for 15 min and the buffy coat recovered. Erythrocytes were lysed by hypotonic shock and the peripheral blood mononuclear cells (PBM) collected then lysed in 0.15 M NaCl, 10 mM Tris, 1% SDS, 1 mM EDTA pH 8.0. The lysate was digested with 200 µg/ml proteinase K for 3 h at 37°C, followed by two extractions with phenol, chloroform, isoamyl alcohol (25:24:1) and one extraction with chloroform. DNA was precipitated with ethanol and collected by spooling on to a glass rod, air dried and dissolved in TE (10 mM Tris, 1 mM EDTA pH 8.0).

#### *Restriction enzyme digestion and Southern blotting*

Ten µg of DNA were digested with 40 units of restriction enzyme (*Bam*HI, *Pvu*II, *Pst*I or *Taq*I) for 4 h according to the manufacturers' recommendations. Digested DNA was separated by electrophoresis through 0.75% agarose gels in TBE buffer (Maniatis *et al.* 1982) at 2.5 V/cm for 17 h. DNA was denatured by soaking the gels in 0.5 M NaOH, 1.5 M NaCl for 30 min, then in 0.5 M Tris, 1.5 M NaCl pH 7.5 for 30 min and transferred to nylon membranes (Hybond N, Amersham) by capillary blotting (Southern 1975).

#### *Hybridization*

The DNA probe used in this study was a BoLA class I cDNA cloned in this laboratory (pBoLA1). This has been sequenced and shows up to 80% similarity to HLA class I sequences (Brown *et al.* 1988). An *Eco*RI/*Sac*I fragment containing the 5' region of the cDNA (corresponding to the 1st domain of the protein) was labelled with <sup>32</sup>P dCTP using a multiprime kit (Amersham) producing probes with a specific activity of 10<sup>9</sup>–10<sup>10</sup> dpm/µg. Filters were pre-hybridized for 30 min at 65°C in 0.5 M phosphate buffer pH 7.2, 7% SDS, 1 mM EDTA, 10 µg/ml herring sperm DNA (Church & Gilbert 1984). The denatured probe was then added in fresh buffer and hybridization allowed to take place for 17 h. Filters were washed twice in 0.5% phosphate buffer pH 7.2, 0.1% SDS, 1 mM EDTA at 65°C for 30 min then twice in 0.5x SSC, 0.1% SDS at 65°C for 30 min and exposed to Curix RP1 film for 5–20 h at –70°C.

#### *IEF*

One-dimensional isoelectric focusing and immunoprecipitation of BoLA class I antigens was by the method of Joosten *et al.* (1988) using the monoclonal antibody W6/32 (Serotec, Oxford, UK).

None of the bands seen following precipitation with W6/32 are found following precipitation with a monoclonal antibody recognising rabbit uteroglobulin (Joosten *et al.* 1988).

**Results**

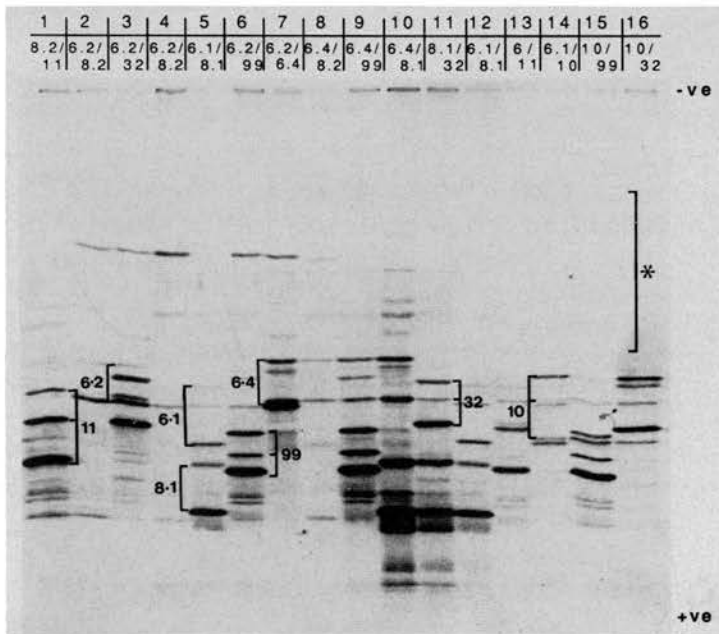
Sixteen animals with common serological haplotypes were selected at random from unrelated cattle. These were analysed using an extensive panel of alloantisera, by isoelectric focusing of immunoprecipitated BoLA class I antigens and by restriction fragment length analysis.

*Serology*

The antisera used fall into three categories:

- (1) those which identify a common epitope on a family of antigens, for example Ed085 which recognizes w4, w7 and w10;
- (2) sera identifying specific (unique) antigens, e.g. Ed069 which defines w10;
- (3) sera considered operationally monospecific and detecting antigens not as yet defined internationally, e.g. Ed099.

Serum reactions were scored as in the 3rd BoLA Workshop (Bull *et al.* 1989). The BoLA class I workshop types of the animals are shown in Fig. 1 and in Table 1. The animals also reacted with several local sera. One such serum, Ed074, has been



**Figure 1.** An autoradiograph of an ID-IEF gel with the animal number and BoLA class I serological type above each track. The standard patterns of eight of the nine serologically defined alleles used in this study is shown. No bands corresponding to w15 are seen.

**Table 1.** The comparison of serological specificities (bold type) (identified by the code used in the 3rd International Workshop) and isoelectric focusing bands (normal type) with restriction fragments (*italics*). IEF bands are numbered according to the serological specificity to which they correlate, hence F11 refers to the IEF bands seen with BoLA w11. Restriction fragments are identified by enzyme (P = *Pst*I, V = *Pvu*II, T = *Taq*I, B = *Bam*HI) and size, i.e. V6.2 is a 6.2-kb fragment from a *Pvu*II digest. The serological typing is shown beside the animal number

BoLA type	V	W	F	T	S	W	F	T	V	T	T	S	W	T	W	W	F	V	S	W	F	E	B	T	E	F
	6	3	3	7	6	1	1	6	6	2	5	3	3	5	8	1	1	3	7	1	1	1	6	0	9	9
	2	2	2	1	1	1	1	1	1	1	1	1	0	1	1	4	4	1	1	0	0	2	1	1	9	9
	2		6	6				5	7	5	9	4	6				6	5			5	0	6			
1 w15 w11				+++		+++							++++													
2 w18 w15													++++													
3 w18 w32	+	+	+	+++																						
4 w18 w15													++++													
5 w17 w14													++++													
6 w18 Ed99													++												++	+
7 w18 w19																										
8 w19 w15													++++													
9 w19 Ed99													++												++	+
10 w19 w14													++++													
11 w14 w32	+	+	+	+++									++++													
12 w17 w14													++++													
13 w6 w11				+++		+++		+++		+++																
14 w17 w10				+														+++	+	+	+	+	+++			
15 w10 Ed99				+									+++	+	+			+++	+						+	+
16 w10 w32	+	+	+	+++														+++	+	+	+	+++				

W = workshop defined sera; E = Edinburgh sera; F = isoelectric focusing

RFLP fragments V = *Pvu*II S = *Pst*I T = *Taq*I

included in a group of sera defining w30 in the 3rd BoLA workshop (Bull *et al.* 1989); however, it is often found to react in addition to two other well-defined workshop specificities. These and other data (R.A. Oliver, unpublished) suggest that the antigen detected by Ed074 may be the product of a second BoLA class I locus.

### Isoelectric focusing

Some of the MHC class I antigens precipitated by W6/32, identified as bands following autoradiography of IEF gels, were readily assigned to serologically defined haplotypes (Fig. 1). No band was identified which correlated with w15. The number of bands assigned to a haplotype varied. For example, only two bands on IEF gels could be assigned to w18, w19 and w14, whereas three bands could be assigned to w10, w11 and w32. In the majority of cases, the animals with a particular workshop-defined specificity gave rise to identical IEF patterns. In addition to the

bands which correlated with the serological specificities, other bands could be seen, which were present in several animals. It is not clear whether these are products of other loci or arise from alternative processing.

To allow further analysis bands were identified as F for focusing and given the workshop specificity with which they were associated.

#### *Restriction fragment length polymorphism*

A total of 40 polymorphic patterns were found in the 16 animals used in this study, following digestion with four restriction enzymes *Bam*HI (B), *Pst*I (S), *Pvu*II (V) and *Taq*I (T) and using the 5' region of the BoLA cDNA probe. In general there were 8–10 polymorphic bands per heterozygous animal and 5–8 non-polymorphic bands (Fig. 2). When the complete BoLA cDNA insert was used the same polymorphic fragments were seen but 10–12 invariant fragments were also identified (data not shown).

The only obvious correlation of fragments with serological workshop-defined specificities were the *Pst*I 7.5-kb fragment with w10 and the *Taq*I 6.5-kb fragment with w11 (Fig. 2). Although other groups of polymorphic fragments were seen, it was not immediately possible to correlate these with the serology.

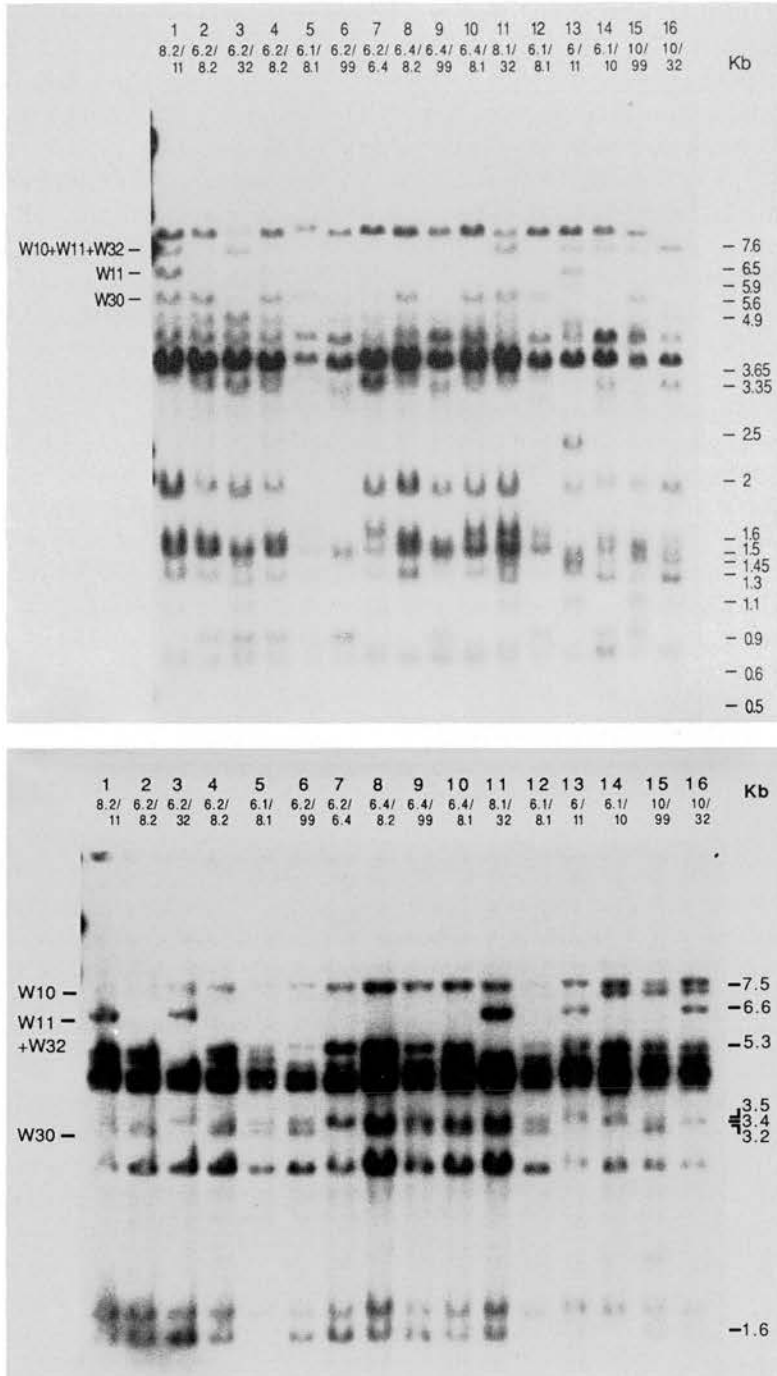
For further analysis the bands observed on the autoradiographs were reduced to a code identifying polymorphic fragments by enzyme and size e.g. V6.2 identifies a *Pvu*II, 6.2-kb fragment.

#### *Analysis*

The results of the RFLP and the IEF patterns along with those from the classical serology were analysed using the 'BoLA' programme (Kemp 1985) and a set of correlation coefficients and similarity indices derived. These results enabled us to identify the following relationships, presented in Table 1.

Bands found by the IEF technique and assigned to the w32 specificity exactly matched the serological typing, as did a 6.2-kb *Pvu*II restriction fragment. The IEF pattern assigned to w11 and a 6.5-kb *Taq*I fragment also showed a perfect match with serology, while a 6.7-kb *Pvu*II fragment, 2.5-kb and 5.9-kb *Taq*I fragments subdivided the w11 animals. A 6.6-kb *Pst*I fragment was present in all of the w32 and the w11 animals whereas a 7.6-kb *Taq*I fragment was found in the same animals and in addition two w10 animals.

A 3.2-kb *Pst*I fragment was seen in all the animals defined as w30; these included all the w8 animals plus several others. It is possible that this fragment arises from a second locus gene (see serology, above). A 5.6-kb *Taq*I fragment on the other hand agreed with the broad w8 specificity but was also found in one extra non-w8 animal. No RFLP fragments were found which corresponded to the w8 subgroups. No IEF bands correlated with the broad w8 specificity but IEF bands did correlate exactly with the w14 specificity. The sera defining w10 correlated perfectly with *Pvu*II 3.6-kb and *Pst*I 7.5-kb fragments and with the F10 IEF bands. Ed125 reacted with



**Figure 2.** Autoradiographs of genomic Southern blots following (A) *TaqI* and (B) *PstI* digestion and hybridization with the 5' region of pBoLA1. Samples are in the same order as in Fig. 1, with serological typing shown above the track. Fragments correlating with serotypes are indicated.

two out of the three w10 animals and correlated with *Bam*HI 6.0-kb and *Taq*I 0.6-kb DNA fragments. No IEF bands correlated with Ed125.

The serum Ed099 has not been accepted by an international workshop, that is no sera identifying this specificity have been produced in other laboratories. However, there was complete agreement with a set of IEF bands suggesting that it does define a specific BoLA haplotype. There was no correlation with any RFLP fragment.

In summary, isoelectric focusing bands correlated extremely well with internationally agreed serological definition of antigens. We previously showed that two IEF bands correlated with w11 (Joosten *et al.* 1988); however, the animals used in this study shared three IEF bands. Additional bands were occasionally missing or in addition to the standard pattern possibly arising from other class I loci. The most obvious being the upper band present in animals 2-4, 6 and 7 which correlates with w18 in the animals used in this study; however, this band was also observed with other serological types (data not shown). The results for RFLP were more difficult to interpret. Fragments could be identified which correlated with the workshop specificities, e.g. the 6.2-kb *Pvu*II fragment with w32. Other fragments occurred with two international specificities which were thought to be totally unrelated, e.g. 6.6-kb *Pst*I with w32 and w11. A third category of fragments were seen which correlated with the broadly reacting sera, e.g. 5.6-kb *Taq*I with broad w8. Finally, fragments were seen which did not arise from the regions encoding expressed antigens and were haplotype specific.

## Discussion

In this study, we have drawn together data on expressed MHC class I antigens and DNA polymorphisms in cattle. Initially, we made comparisons between RFLP data and serology in family groups and found an 11-kb *Pvu*II fragment that correlated with w14 and a 2.9-kb *Pvu*II fragment with w10 (data not shown). These correlations were not observed in this study of unrelated animals from a herd derived from a large panel of AI bulls. The fragments were presumably polymorphisms linked to, but not arising from, the sequence coding for w14 or w10 in the family studied.

Comparison of serology with the predominant bands seen by isoelectric focusing gave good correlation. It is difficult to explain the large number of bands seen on the IEF gels. The major bands could represent the product of a single highly expressed gene at various stages of processing. Alleles showing different modification sites may give rise to the variable number of bands. Alternatively, the different bands may be the products of two or more very tightly linked loci, the variable number of bands arising from differing affinities of W6/32 for the different products. It is possible in the case of w15, where no bands were seen, that the products may have very low affinity for W6/32. Preliminary results from pulse chase experiments presented by Joosten *et al.* (1988) suggest that there is more than one BoLA class I product. Iodination of surface antigens, rather than metabolic labelling, may also help to resolve the identity of the different bands. Van der Poel *et al.* (1986) and Yang *et al.*

(1984) showed IEF patterns corresponding to established serological types in man. In addition, variant patterns were observed for certain serologically defined alleles.

Several minor IEF bands were also observed; these could be products of other loci with lower levels of expression or alternatively they could represent degradation products. The latter is unlikely as only a single 45-Kd band was observed on SDS gels (data not shown).

The comparison of serological typing with RFLP fragments was particularly interesting. Restriction fragments arise from polymorphisms within and outwith the coding region of genes and other MHC class I like sequences including pseudogenes. Many of the polymorphisms are irrelevant to the expressed antigens; nevertheless, there was excellent correlation of several RFLP fragments with various serological and focusing types. The best examples of these correlated with w32 and w11. A 6.6-kb *Pst*I fragment was found in animals with either the w32 or the w11 antigens; likewise a 5.6-kb *Taq*I fragment was found in w8 and non-w8 animals. These may represent different fragments fortuitously of the same size or that the respective antigens arose from the same progenitor genes, although no serological cross-reactivity is observed in either case.

In general, restriction fragment length polymorphisms in the MHC regions segregated with the serology in families. Many of the RFLPs observed probably did not arise directly from the region coding for the MHC antigens but from linked non-expressed sequences, identifying haplotypes. Such RFLPs could be used for parentage determination, or as markers for genes linked to the MHC. In an outbred population fragments occurring consistently with serological specificities are more likely to arise directly from the region coding for particular MHC class I antigens. There are few published data on MHC class I RFLPs in domestic animals. Alexander *et al.* (1987) showed several correlations of RFLP fragments with equine class I serology. This work was carried out in unrelated homozygous horses and correlations could only be made for two of the six antigens tested. However, in order to correlate serology with restriction fragments, the choice of enzymes is crucial.

In conclusion, in order to identify BoLA class I antigens, serology remains the simplest technique with accurate results achieved quickly. Isoelectric focusing can be used to confirm serological results with the added advantage that serologically undefined antigens can also be identified. It should also be possible to identify products from other class I loci. Different information is obtained from RFLP studies. While some fragments may correlate with the antigens identified serologically other polymorphisms are also seen. These may arise from other expressed BoLA class I loci but may equally arise from non-coding regions of DNA: such RFLPs, identifying haplotypes, can be used as markers in studies of genetic linkage.

### **Acknowledgements**

We wish to thank A.L.G. Morgan for BoLA serology, B.E. Harrower for the RFLP gels, the staff of the AFRC farms for bleeding the animals and E. Hensen for valuable discussion. This study was partly funded by a UK Overseas Development Administration research grant (Project R-4153A).

## References

- Anon. (1982) Proceedings of the 2nd bovine lymphocyte antigen (BoLA) workshop. *Animal Blood Groups and Biochemical Genetics* **13**, 33–53.
- Alexander J.A., Bailey E. & Woodward J.G. (1987) Analysis of equine lymphocyte antigen system by Southern blot hybridization. *Immunogenetics* **25**, 47–54.
- Amorena B. & Stone W.H. (1978) Serological defined (SD) locus in cattle. *Science* **201**, 159–60.
- Bensaid A., Naessens J., Kemp S.J., Black S.J., Shapiro S.Z. & Teale A.J. (1988) An immunochemical analysis of class I (BoLA) molecules on the surface of bovine cells. *Immunogenetics* **27**, 139–44.
- Brown P., Spooner R.L. & Clark A.J. (1988) Cloning and characterization of a BoLA class I cDNA clone. *Immunogenetics* **29**, 58–60.
- Bull R.W., Lewin H.A., Wu M.C., Peterbaugh K., Antczak D., Bernoco D., Cwik S., Dam L., Davies C., Dawkins R.L., Dufty J., Gerlach J., Hines C., Lazary S., Leibold W., Leveziel H., Lie Ø., Lindberg P.G., Meggiolaro D., Meyer E., Oliver R., Ross M., Simon M., Spooner R.L., Stear M., Teale A. & Templeton J. (1989) Joint Report of the Third International Bovine Lymphocyte Antigen (BoLA) Workshop. *Animal Genetics* **20**, 109–32.
- Chardon P., Vaiman M., Kirzenbaum M., Geffrotin D., Renard C. & Cohen D. (1985) Restriction fragment length polymorphism of the major histocompatibility complex of the pig. *Immunogenetics* **21**, 161–71.
- Church G.M. & Gilbert W. (1984) Genomic sequencing. *Proceedings of the National Academy of Science* **81**, 1991–5.
- Cohen D., Paul P., Le Gall I., Marcadet A., Font M-P., Cohen-Haguenaer O., Sayagh B., Cann H., Lalouel J-M. & Dausset J. (1985) DNA polymorphism of HLA class I and class II regions. *Immunological Reviews* **85**, 87–105.
- Joosten I., Oliver R.A., Spooner R.L., Williams J.L., Hepkema B.G., Sanders M.F. & Hensen E. (1988) Characterization of class I bovine lymphocyte antigens (BoLA) by one-dimensional isoelectric focussing. *Animal Genetics* **19**, 103–13.
- Kemp S.J. (1985) *The major histocompatibility complex of African cattle*. PhD thesis, University of Edinburgh.
- Klein J (1986) *Natural History of the Major Histocompatibility Complex*. John Wiley and Son, New York.
- Maniatis T., Fritsch E.F. & Sambrook J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, New York.
- McGuire K.L., Duncan W.R. & Tucker P.W. (1985) Syrian Hamster DNA shows limited polymorphism at class I-like loci. *Immunogenetics* **22**, 257–68.
- Mellor A. (1986) The class I MHC gene family in mice. *Immunology Today* **7**, 19–24.
- Neeffjes J.J., Breur-Vriesendorp B.S., van Seventer G.A., Ivannyi P. & Ploegh H. (1986) An improved biochemical method for the analysis of HLA class I antigens. Definition of new HLA class I subtypes. *Human Immunology* **16**, 169–81.
- Oliver R.A., McCoubrey C.M., Millar P., Morgan A.L.G. & Spooner R.L. (1981) A genetic study of bovine lymphocyte antigens (BoLA) and their frequency in several breeds. *Immunogenetics* **13**, 127–32.
- Palmer M., Wettstein P.J. & Frelinger J.A. (1983) Evidence for extensive polymorphism of class I genes in the rat major histocompatibility complex (RTI). *Proceedings of the National Academy of Sciences* **80**, 7616–20.
- Southern E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–17.
- Spooner R.L., Leveziel h., Grosclaude F., Oliver R.A. & Vaiman M. (1978) Evidence for a major histocompatibility complex (BLA) in cattle. *Journal of Immunogenetics* **5**, 335–46.
- Spooner R.L., Oliver R.A., Sales D.I., McCoubrey C.M., Millar P., Morgan A.G., Amorena B., Bailey E., Bernoco D., Brandon M., Bull R.W., Caldwell J., Cwik S., van Dam R.H., Dodd J., Gahne B., Grosclaude F., Hall J.G., Hines H., Leveziel H., Newman M.J., Stear M.H., Stone W.H. & Vaiman M. (1979) Analysis of alloantisera against bovine lymphocytes. Joint report of the 1st International bovine lymphocyte antigen workshop. *Animal Blood Groups and Biochemical Genetics* **10**, 63–86.



- Steinmetz M., Stephan D. & Fischer-Lindahl K (1986) Gene organization and recombination hotspots in the murine major histocompatibility complex. *Cell* **44**, 895-904.
- Van der Poel J.J., Molders H.H., Thompson A. & Ploegh H. (1983) Definition of four HLA-A2 subtypes by CML typing and biochemical analysis. *Immunogenetics* **17**, 609-21.
- Wake C.T. (1986) Molecular biology of HLA class I and class II genes. *Molecular and Biological Medicine* **3**, 1-11.
- Yang S.Y., Morishima Y., Collins N.H., Alton T., Pollack M.S., Yunis E.J. & Dupont B. (1984) Comparison of one-dimensional IEF patterns for serologically detectable HLA-A and B allotypes. *Immunogenetics* **19**, 217-31.

## Gene-antigen register

# Cloning and characterization of a *BoLA* class I cDNA clone

P. Brown, R. L. Spooner, and A. J. Clark

A.F.R.C. Institute of Animal Physiology and Genetics Research, Kings Buildings, West Mains Road, Edinburgh EH9 3JQ, UK

The molecular organization of *BoLA* (Bovine Mhc), particularly the class II genes, has been investigated using human DNA probes in Southern blotting analysis of bovine genomic DNA (Andersson et al. 1986a, b, Andersson and Rask 1988, Andersson et al. 1988). Class I genes in the *BoLA* complex have not been studied as extensively, mainly because the use of heterologous probes is difficult for technical reasons and causes problems in interpretation of results. To circumvent these problems we isolated a class I cDNA clone (pBoLA-1) from a bovine liver cDNA library. Here we present the cloning and characterization of pBoLA-1. pBoLA-1 is 1263 bp in size and codes for a complete mature class I protein. Comparing the protein encoded by pBoLA-1 to other class I proteins from man, rabbit, mouse and rat shows that all the features attributed to functional class I proteins are present. The only feature unique to pBoLA-1 is an amino acid substitution at position 158 where a conserved alanine is replaced by an asparagine.

A bovine liver cDNA library of  $2.5 \times 10^5$  recombinants was constructed in the lambda immunity insertion vector gt10 (Huynh et al. 1986). The unamplified library was screened with the HLA class I cDNA clone pHLA-27 (kindly supplied by Dr. S. Paabo) using standard hybridization conditions as outlined in Maniatis and co-workers (1982), and washed to low stringency in  $2 \times$  SSC, 0.1% SDS. Two positive clones were obtained, plaque purified, and restricted with *EcoRI*. The clone containing the larger insert (pBoLA-1) was chosen for further investigation. The complete nucleotide sequence with the restriction map and sequencing strategy are presented in Figure 1. The cDNA clone is 1263 bp in size and the longest open reading frame stretches from nucleotide 2 to nucleotide 1090. Following this there is 173 bp of 3' untranslated sequence. The cDNA clone is incomplete, lacking 5' untranslated (UT) sequence, an ATG start codon, and some 3' UT sequence. Examination of the translated sequence

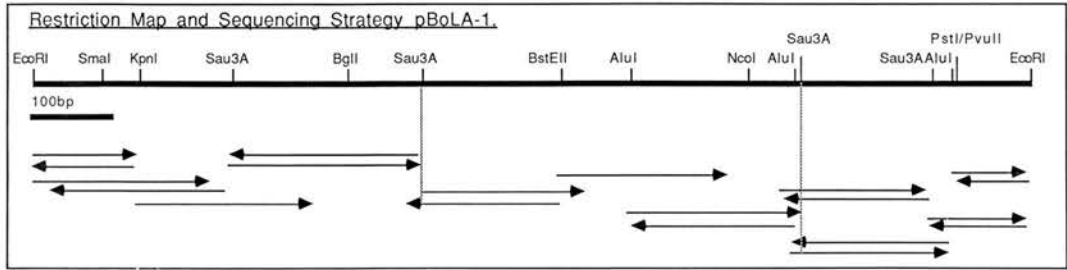
of pBoLA-1 in Figure 1 shows that all the conserved features typical of class I proteins are present. There is a classical hydrophobic sequence capable of acting as a signal peptide in pBoLA-1. The amino acid sequence of the predicted mature protein is similar to those reported for other class I proteins comprising three extracellular domains, a hydrophobic transmembrane region, and a hydrophilic cytoplasmic region.

As for other species in the first extracellular domain, there is a potential N-linked glycosylation site at amino acid 86 (Ploegh et al. 1981). This is the only putative glycosylation site in the pBoLA-1 protein and the only site completely conserved across species (Tykocinski et al. 1984). Cysteines implicated in disulfide loop formation (Kimball and Coligan 1983) are located at amino acids 101/164 in the second domain and at 203/259 in the third domain. The putative transmembrane region of the pBoLA-1 protein is 25 amino acids long and is composed of hydrophobic residues which is the requirement for integral membrane proteins (Ploegh et al. 1981). The cytoplasmic region is 32 amino acids long, has a high proportion of hydrophilic residues, and contains a serine phosphorylation site consensus sequence ser-asp/glu-x-ser(P)-leu. This phosphorylation site has not been shown to be of functional significance in class I proteins but it is very highly conserved in man and mouse (Guild and Strominger 1984). Since these features are conserved in pBoLA-1 protein, it suggests that the clone encodes a functional class I protein.

Extensive sequence comparison of pBoLA-1 protein to a variety of class I molecules reveals that the pBoLA-1 protein is most similar to class Ia molecules. The amino acid sequence of pBoLA-1 is shown aligned with human, rabbit, mouse, and rat class Ia amino acid sequences in Figure 2. The comparison illustrates that the three extracellular domains have very high sequence similarity to corresponding domains of man, rabbit, mouse, and rat. The transmembrane and cytoplasmic regions of pBoLA-1 protein are clearly related to those of other species, but

```

10          30          50          70          90          110
GCCTCTGAGTCTGGGAAGAACAGAGTCTGCTGCTCTCGGGGGCTGGTCTGACCAGACCCGGGCTGGCTCCCCTCGATGAGGTATTTTCAGCACCCGGCTGCCCGCCGGCCT
ProLeuSerLeuGlyArgAsnArgValLeuLeuSerGlyValLeuValLeuThrGluThrArgAlaGlySerHisSerMetArgTyrPheSerThrAlaValSerArgProGlyLe
130          150          170          190          210          230
CGGGAGCCCGGTTACCTGGAAGTCCGGTACGTGGACGACACCGAGTTCGTGCGGTTTCGACGCGCCCGAATCCGAGGATGGAGCCCGGGCGGGTGGGTGGAGCAGGAGGGGGCC
uGlyGluProArgTyrLeuGluValGlyTyrValAspAspThrGlnPheValArgPheAspSerAspAlaProAsnProArgMetGluProArgAlaArgTrpValGluGlnGluGlyPr
250          270          290          310          330          350
CGAGTATTGGGATCAGGAGACGCGAAAGGCCAAGGGCACCGCACAGACTTCCGGGCAACCTGAAACATCGCACTCGGCTACTACAACAGAGCGAGGCGGGTCTCACACCTCCAGTG
cGluTyrTrpAspGlnGluThrArgLysAlaLysGlyThrAlaGlnThrPheArgAlaAsnLeuAsnIleAlaLeuGlyTyrTyrAsnGlnSerGluAlaGlySerHisThrPheGlnTr
370          390          410          430          450          470
GATGTACGGCTGCGACGTGGGGCCGGCGGGCTCTCCGCGCGGGTTCATGCAGTACGGCTACGACGCGCAGAGATTACATCGCCCTGAACGAGGACCTCGCTCTCGCCGCGGGGA
pMetTyrGlyCysAspValGlyProAspGlyArgLeuArgArgGlyPheMetGlnTyrGlyTyrAspGlyArgAspTyrIleAlaLeuAsnGluAspLeuArgSerTrpThrAlaAlaAs
490          510          530          550          570          590
CACGGCGGCTCAGATCACCAAGCGCAAGTGGGAGGGCGGAGGTGAGCGGGAGAGACAGAGAACTACTCGGAGGCGACGTGCGTGGAGTGGCTCCCGCAGATCTCGGAGCCGGGAAGGA
pThrAlaAlaGlnIleThrLysArgLysTrpGluAlaAlaGlyGluAlaGluArgGlnArgAsnTyrLeuGluGlyThrCysValGluTrpLeuArgArgTyrLeuGluThrGlyLysAs
610          630          650          670          690          710
CACGCTGTGCGCGCAGACCTCCAAAGGCACATGTGACCCATCACTCTCTGGTCATGAGGTCAACCTGAGGTGCTGGGCGTGGGCTTCTACCCTGAGGACATCTCACTGACCTG
pThrLeuLeuArgAlaAspProProLysAlaHisValThrHisHisSerIleSerGlyHisGluValThrLeuArgCysTrpAlaLeuGlyPheTyrProGluAspIleSerLeuThrTr
730          750          770          790          810          830
GCAGCGCAATGGGAGGACCAGCCAGGACATGGAGCTTGTGGAGACCAGGCGCTTCAGGGACGGAAACTCCAGAAGTGGGCGGCCCTGGTGGTGCCTTCTGGAGAGGAGCAGAAATA
pGlnArgAsnGlyGluAspGlnThrGlnAspMetGluLeuValGluThrArgProSerGlyAspGlyAsnPheGlnLysTrpAlaAlaLeuValValProSerGlyGluGlnGlnLysTy
850          870          890          910          930          950
CACATGCCGATGTCAGCAGCAGGGGGCTTCAGGAGCCCTCACCTGAAATGGGAACCTCCTCAGCCCTCTCTCCACATGGGCATCATTGTTGGCCGTGTTCTCTCGTGGTCACTGG
rThrCysArgValGlnHisGluGlyLeuGlnGluProLeuThrLeuLysTrpGluProProGlnProSerPheLeuThrMetGlyIleIleValGlyLeuValLeuLeuValValThrGl
970          990          1010          1030          1050          1070
AGCTGGTGGTGGAGTGTGTATCTGATCAGAGAGCGCTCAGGTGAAAACAGGGGACTTATATCCAGGCTCAAGCAGTGACAGTGCCAGGGCTCTGATGTGTCTTCCAGGGTCC
yAlaValValAlaGlyValValIleCysMetLysLysArgSerGlyGluLysArgGlyThrTyrIleGlnAlaSerSerSerAspSerAlaGlnGlySerAspValSerLeuThrValPr
1090          1110          1130          1150          1170          1190
TAAAGTGTGAGACCCGGCTTGGGGGACTGAGTGATGCTTCATCCGGCTATGTGACATCAGATCCCCAGACCCTCTTCTGCAGCTGCATGTGAATGTGTGTGCCCTATTCCG
oLysValEnd
1210          1230          1250
ATAACGTAGGAGTTAGGGAGACTGGCCACCCTATGCCACTGTGCCCTCCCCACCTGACC
    
```



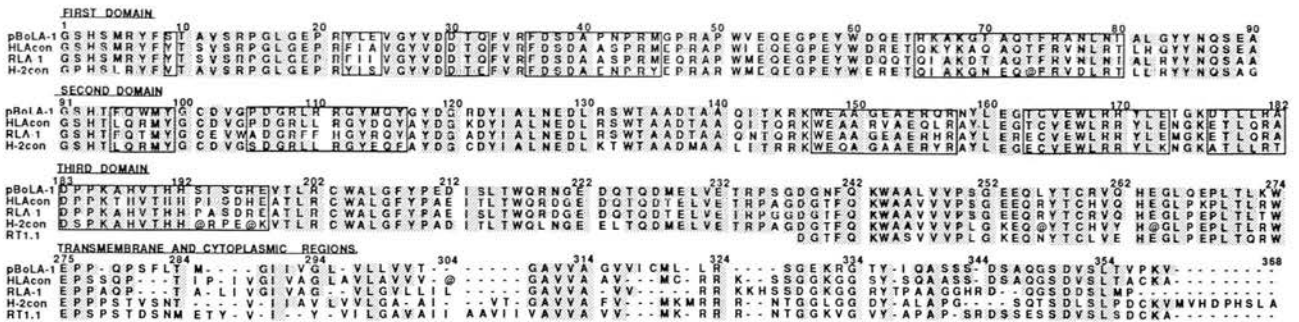
**Fig. 1.** Nucleotide sequence of pBoLA-1 and restriction map. Sequence of pBoLA-1 with deduced amino acid sequence, shown with restriction map and sequencing strategy used. The putative signal peptide is shown underlined as is the serine phosphorylation site consensus sequence. All cysteines proposed to be involved in disulfide bridge formation are indicated in **bold type**, the possible glycosylation site at amino acid 86 is also **highlighted**. pBoLA-1 was subcloned into the *Eco* RI site of plasmid vector pPoly1 (Lathe et al. 1987) and then mapped by single and double restriction enzyme digests. Individual fragments were subcloned into M13 vectors tg130 and tg131 (Kieny et al. 1983) for sequencing by the Sanger dideoxy chain termination method (Sanger et al. 1977)

the degree of similarity is lower than that found in the extracellular domains. Others have also remarked on the strong similarity of the extracellular domains of class I proteins across species (Kimball and Coligan 1983, Klein and Figueroa 1986, Klein 1986). This suggests evolutionary constraints on the extracellular structure of class I proteins, while the transmembrane and cytoplasmic regions are free to diverge, providing that they remain hydrophobic and hydrophilic, respectively (Ploegh et al. 1981, Kimball and Coligan 1983).

The extracellular domains of class I proteins are known to encode all the functional variation of class I molecules. The clusters of variation thus defined are boxed on Figure 2. Confirmation of the importance of these sites of variability came from X-ray crystallographic studies of *HLA-A2*, where they were shown to map to accessible sites on the protein (Bjorkman et al. 1987a, b). Some of the residues formed an  $\alpha$ -helix secondary struc-

ture which combined to form the putative antigen binding site of class I molecules (Bjorkman et al. 1987a). Similar, but not identical clusters of variation are observed when pBoLA-1 is compared to other species, suggesting that it may function as a restriction element.

Gussow and co-workers (1987) have noted that every third or fourth residue of the  $\alpha$ -helical region is conserved, and Bjorkman and co-workers (1987b) have shown that this corresponds to one face of the  $\alpha$ -helix. A similar pattern of amino acid conservation is seen in pBoLA-1, except for the presence of asparagine at position 158 (all other functional class I molecules have alanine at this position). The functional significance of this difference is not clear and could be resolved by experimentally testing for expression of pBoLA-1. Since pBoLA-1 conforms with class Ia molecules in all other respects, it may simply represent a species specific residue (Klein and Figueroa 1986, Klein 1986).



**Fig. 2.** Comparison of pBoLA 1 with human, rabbit, mouse, and rat protein sequences. Alignment of pBoLA 1 deduced amino acid sequence with an HLA consensus sequence; RLA-1 rabbit class I cDNA clone deduced amino acid sequence (Tykocinski et al. 1984); H-2 consensus sequence; RTI.1 rat class I cDNA clone deduced amino acid sequence (Kastern 1985). Residues held in common with pBoLA-1 have been shaded, clusters of variation are shown boxed. Maximum alignment has been achieved by introducing gaps indicated by (-) where necessary. The criteria of maximum parsimony was used to provide a consensus. When there was no consensus due to extreme variability at that position an @ is used instead of the single letter amino acid code. The consensus sequences were derived using UWGCG program on a VAX network. The consensus for HLA was derived from protein sequences which were presented in Figure 5 from Gussow and co-workers (1987). The consensus for H-2 was derived from the mouse class Ia sequences presented in Klein and Figueroa (1986)

We have described the isolation and characterization of a class I cDNA clone. The high similarity of the deduced protein sequence of this clone with other class I proteins is evidence that the clone may encode a functional transcript.

**Acknowledgments.** We wish to thank Dr. S. Harris, Dr. J. P. Simons, and Dr. E. J. Glass for their critical reading of the manuscript. P. Brown would like to thank Dr. R. Lathe for useful discussion and is grateful for help and encouragement received from members of Dr. N. Hastie's laboratory (C.A.P.C.U., Edinburgh), particularly Dr. R. Hill. P. Brown was supported by an A.F.R.C. studentship.

## References

- Andersson, L. and Rask, L.: Characterization of the MHC class II region in cattle. The number of *DQ* genes varies between haplotypes. *Immunogenetics* 27: 110-120, 1988
- Andersson, L., Bohme, J., Rask, L., and Peterson, P. A.: Genomic hybridization of bovine class II major histocompatibility genes: I. Extensive polymorphism of *DQ $\alpha$*  and *DQ $\beta$*  genes. *Anim Genet* 17: 95-112, 1986a
- Andersson, L., Bohme, J., Peterson, P. A., and Rask, L.: Genomic hybridization of bovine major histocompatibility genes. II. Polymorphism of *DR* genes and linkage disequilibrium in the *DQ-DR* region. *Anim Genet* 17: 195-304, 1986b
- Andersson, L., Lunden, A., Sigurdardottir, S., Davies, C. J., and Rask, L.: Linkage relationships in the bovine MHC region. High recombination frequency between class II subregions. *Immunogenetics* 27: 273-280, 1988
- Bjorkman, P. J., Saper, M. A., Samaraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C.: Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329: 506-512, 1987a
- Bjorkman, P. J., Saper, M. A., Samaraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C.: The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329: 512-518, 1987b
- Guild, B. C. and Strominger, J. L.: Human and murine class I MHC antigens share conserved serine 335, the site of HLA phosphorylation *in vivo*. *J Biol Chem* 259: 9235-9240, 1984
- Gussow, D., Rein, R. S., Meijer, I., Hoog, W., Seeman, G. H. A., Hochstenbach, F. M., and Ploegh, H. L.: Isolation, expression and the primary structure of *HLA-Cw1* and *HLA-Cw2* genes: Evolutionary aspects. *Immunogenetics* 25: 313-322, 1987
- Huynh, T. V., Young, R. A., and Davis, R. W.: Constructing and screening cDNA libraries in gt10 and gt11. In D. M. Glover (ed.): *DNA Cloning I. A Practical Approach*, IRL Press, Oxford, 1986
- Kastern, W.: Characterization of two class I major histocompatibility rat cDNA clones, one of which contains a premature termination codon. *Gene* 34: 227-233, 1985
- Kieny, M. P., Lathe, R., and Lecocq, J. P.: New versatile cloning and sequencing vectors based on bacteriophage M13. *Gene* 26: 91-99, 1983
- Kimball, E. S. and Coligan, J. E.: Structure of class I major histocompatibility antigens. In F. P. Inman and T. J. Kindt (eds.): *Contemporary Topics in Molecular Immunology* 9, pp. 1-63, Plenum, New York, 1983
- Klein, J.: *Natural History of the Major Histocompatibility Complex*, John Wiley & Sons, New York, 1986
- Klein, J. and Figueroa, F.: The evolution of class I MHC genes. *Immunol Today* 7: 41-44, 1986
- Lathe, R., Vilotte, J. L., and Clark, A. J.: Plasmid and bacteriophage vectors for excision of intact inserts. *Gene* 57: 193-201, 1987
- Maniatis, T., Fritsch, E. F., and Sambrook, J.: *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York 1982
- Ploegh, H. L., Orr, H. T., and Strominger, J. L.: Major histocompatibility antigens: the human (HLA-A, -B, -C) and murine (H-2K, H-2D) class I molecules. *Cell* 24: 287-299, 1981
- Sanger, F., Nicklen, S., and Coulson, A. R.: DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467, 1977
- Tykocinski, M. L., Marche, P. N., Max, E. E., and Kindt, T. J.: Rabbit class I MHC genes: cDNA clones define full-length transcripts of an expressed gene and a putative pseudogene. *J Immunol* 133: 2261-2269, 1984

Received June 21, 1988; revised version received August 22, 1988

**Note added in proof:** Since submission of this article, Ennis and co-workers (*J Immunol* 141: 642-651, 1988) have described the cloning and characterization of two other bovine class I cDNA clones.