Proteasome Subunits Related to the MHC-Encoded LMP Proteins.

A thesis presented for the degree of Doctor of Philosophy by

Mônica Polidoro Belich

University of London
Department of Immunology
Human Immunogenetics Laboratory
Imperial Cancer Research Fund
44 Lincoln's Inn Fields
London WC2A 3PX
Abstract

The objective of this work was to identify and characterize new 20S proteasome subunits that are homologous to the human β subunits LMP2 and LMP7. The LMP genes are tightly linked to the TAP peptide transport loci in the class II region of the MHC implicating this multi-subunit protease in antigen processing for class I presentation. Nevertheless, studies with .174 mutant cell line which lacks the LMP genes show that they were able to process and present antigens at the cell surface at a similar level to wild type 721 cells.

The results presented in chapter 3 suggest that MB1 and Delta gene subunits are most probably the LMP7 and LMP2 subunit homologues. MB1 gene was discovered by screening cDNA libraries under non-stringent conditions with a LMP7 probe. Delta was cloned by hybridise complete cDNA libraries with a Delta PCR probe. Protein analysis showed that MB1 is 68.1% identical to the LMP7 sequence and Delta 60.8% identical to LMP2.

The data presented in chapter 4 shows the chromosomal localization of MB1 and Delta loci examined by using fluorescence in situ hybridization (FISH) method. MB1 and Delta genes localized to different chromosomes, 14q11.2 and 17p13 respectively.

Antisera raised against the C-termini of MB1 and Delta protein were produced to further characterize these β subunits. Results presented on chapter 5 shows that MB1 and Delta proteins are part of the 20S proteasome complex and that they have reciprocal expression to that of the LMP genes: MB1 and Delta are up-regulated in mutant cells lacking the LMPs genes and down-regulated in the presence of γ-IFN.

The data presented in chapter 6 shows that proteasomes from 721 and .174 cell lines have similar quaternary structure and the immunolocalization of MB1 and Delta subunits into .174 cells supports the proposed role of these subunits in directly substituting the genes LMP2 and LMP7.
Chapter 7 describes the analysis of a large number of B cell and non B cell lines for LMP2, LMP7, MB1 and Delta expression. It shows that these subunits are expressed at different levels and compositions and that 20S proteasomes are formed from many different combinations of the available subunits.
I would like to thank my supervisor, John Trowsdale, for his technical expertise, good judgment, encouragement and moral support throughout the period of my research at the ICRF. I am also particularly grateful to Peter Beverly for his insightful comments, guidance and expert advice. I also wish to thank the Imperial Cancer Research Fund for providing me with financial support which enabled me to work in such distinguished scientific environment.

I thank the members of Trowsdale’s laboratory, Richard, Amanda, Philippe, Louise, Sheetal, Claire, Jethro, Derek, Ulli and Toshihiro for making the laboratory such an exciting and friendly environment. I am grateful to Richard Newman at the Protein Structure laboratory of the ICRF for the electron microscopy analysis, and to Gabrielle Senger at the Human Cytogenetics laboratory for the in situ hybridization of MB1 and Delta genes. I wish to thank my colleagues at the ICRF’s Tissue Antigens, Cell Culture and Genome Analysis laboratories for providing cell lines, cosmid and cDNA libraries. Thanks to P.J. for proof reading my thesis.

I am grateful to all my friendly colleagues at ICRF for making my life happier. I would especially like to thank Vas Ponnambalam for his friendship, support and scientific advice. I wish to thank my closest friends, Aurelia Rughetti, Beatrice Plougastel, Marieka Van Ham and Isabel Correa, for their friendly counsel and the joy they brought to my doctorates studies. It would have been much harder to make it through without them. It is impossible for me to fully express my deep appreciation and the respect I have for them in a few sentences. I know they realize they have been indispensable to me.

Many thanks to my parents, my brother and in special to my sister Beatriz for given me so much love and support during my scientific career.

Michael Korzinski has played a very special role in my life. I thank him for his love and the unlimited support he has given to me during my studies. His bountiful energy and determination have inspired me. I deeply admire his capacity for love, and the profound respect he has for the welfare of all human beings. I thank him for all the things we have learned together, and the happiness and enchantment he has brought to my life.
Abstract ........................................................................................................................2
Acknowledgments .......................................................................................................4
List of Figures ..............................................................................................................9

Chapter 1
Introduction ................................................................................................................13
1.1 History of the human MHC .......................................................................................13
1.2 Genomic organization of the human MHC complex ...............................................15
1.3 Structural organization of the human MHC antigens ..............................................18
   1.3.1 Class I structure and bound peptide .................................................................18
   1.3.2 Class II structure and bound peptide ................................................................24
1.4 Cell biology of MHC class I antigens .......................................................................26
   1.4.1 Class I molecules bind peptides .......................................................................26
   1.4.2 Mutant cells defective in antigen presentation ................................................27
   1.4.3 Entry of proteins into the class I processing pathway ....................................28
   1.4.4 Biosynthesis and assembly of class I molecules .............................................30
   1.4.5 The transporter associated with antigen processing, TAP .............................31
   1.4.6 Chaperones in class I assembly........................................................................34
   1.4.7 Production of peptide for class I presentation .................................................35
   1.4.8 ER proteases involved in antigen processing..................................................36
1.5 The proteasome ...........................................................................................................36
   1.5.1 Characterization of the proteasome .................................................................37
   1.5.2 Structure of the proteasome .............................................................................39
   1.5.3 Proteasome catalytic activities and active sites ...............................................43
   1.5.4 Ubiquitin system and the 26S proteasome ......................................................46
   1.5.5 Regulatory proteins of the proteasome ...........................................................50
   1.5.6 Proteasome location and expression ..............................................................52
   1.5.7 Functions of the proteasome ............................................................................54
   1.5.8 Proteasome subunits .........................................................................................55
   1.5.9 Proteasomes in antigen processing .................................................................57
1.6 Aims of this thesis .....................................................................................................64
Chapter 2

Materials and Methods ................................................................. 65

2.1 Cell lines, bacterial strains and DNA libraries ......................... 65
  2.1.1 Cell lines ............................................................................ 65
  2.1.2 Bacterial strains and libraries ............................................ 67

2.2 Bacterial preparation and manipulation ..................................... 67
  2.2.1 Bacterial cell culture .......................................................... 67
  2.2.2 Preparation of competent *E. coli* cells for transformation ... 68
  2.2.3 Introduction of DNA into cells ......................................... 68

2.3 DNA preparation and manipulation ........................................ 68
  2.3.1 Small scale preparation of plasmid or cosmid DNA ............. 68
  2.3.2 Large scale preparation of plasmid or cosmid DNA ............ 70
  2.3.3 Restriction endonuclease digestion of DNA ....................... 71
  2.3.4 Electrophoresis of DNA ..................................................... 71
  2.3.5 Labelling of DNA fragments using random hexamer priming 71
  2.3.6 Screening of cDNA libraries ............................................. 72
  2.3.7 Hybridization of filters .................................................... 73
  2.3.8 Southern blotting .............................................................. 75
  2.3.9 Sequencing of double stranded plasmid DNA .................. 75
  2.3.10 Denaturing polyacrylamide gel electrophoresis ............... 76

2.4 RNA preparation and manipulation ......................................... 77
  2.4.1 RNA preparation .............................................................. 77
  2.4.2 Electrophoresis of RNA .................................................... 77
  2.4.3 Northern blotting ............................................................. 78
  2.4.4 Hybridisation of blots ....................................................... 78

2.5 Preparation of antisera, immunoprecipitations and western blotting 79
  2.5.1 Production of polyclonal rabbit antisera ............................ 79
  2.5.2 Immunoprecipitation ......................................................... 79
  2.5.3 Pulse chase analysis .......................................................... 81
  2.5.4 Reprecipitation ................................................................. 81
  2.5.5 Polyacrylamide protein gel electrophoresis ....................... 82
  2.5.6 Western blotting ............................................................... 82
  2.5.7 Nuclear extracts ............................................................... 83
  2.5.8 Cell culture synchronization and γ-interferon treatment .... 84
  2.5.9 Large scale cell cultures .................................................. 85
  2.5.10 Immunoaffinity column .................................................. 85
  2.5.11 Ammonium sulphate fractionation method ..................... 85
Chapter 3
Cloning of MB1 and Delta Proteasome Subunits .............................................................. 87
3.1 Introduction .................................................................................................................. 87
3.2 Cloning and sequence analysis of MB1 gene ............................................................... 88
3.3 Cloning and sequence analysis of Delta gene .............................................................. 95
3.4 mRNA expression of MB1 and Delta genes ................................................................. 98
3.5 Analysis of the proteasome $\beta$ subunits .................................................................. 100
3.6 Conclusion ................................................................................................................... 104

Chapter 4
Mapping of MB1 and Delta Genes ................................................................................. 106
4.1 Introduction ................................................................................................................ 106
4.2 Chromosome localization of MB1 gene ..................................................................... 106
4.3 Chromosome localization of Delta gene .................................................................... 107
4.4 Models for the genetic origins of the MB1, Delta and LMP genes ............................. 108
4.5 Conclusion ................................................................................................................... 114

Chapter 5
Analysis of MB1 and Delta Proteasome Components .................................................. 119
5.1 Summary ..................................................................................................................... 119
5.2 Introduction ................................................................................................................ 119
5.3 MB1 is a subunit of the 20S proteasome complex ...................................................... 121
5.4 Delta is a subunit of the 20S proteasome complex ..................................................... 124
5.5 MB1 and Delta are synthesized with an N-terminal leader sequence ....................... 128
5.6 MB1 and Delta are down-regulated by $\gamma$-interferon ............................................. 129
5.7 Levels of MB1 and Delta protein decrease when LMP7 and LMP2
    are up-regulated by interferon-$\gamma$ ............................................................................. 131
5.8 Conclusion ................................................................................................................... 134

Chapter 6
Localization of LMP2, LMP7, MB1, and Delta Subunits in Proteasomes ..................... 136
6.1 Summary ..................................................................................................................... 136
6.2 Introduction ................................................................................................................ 136
6.3 721 and .174 cells have similar 20S proteasome structure ......................................... 137
6.4 721 cells have LMP2, LMP7 and Delta proteasome subunits
    localized in the inner rings ............................................................................................ 139
6.5 .174 cells contain MB1 and Delta proteasome subunits localized
    in the inner rings .......................................................................................................... 141
6.6 Conclusion ................................................................................................................... 144
| Figure 1.1 | Map of the human histocompatibility complex | 16 |
| Figure 1.2 | Primary structure of the class I and class II histocompatibility molecules | 19 |
| Figure 1.3 | The structure of an MHC class I molecule determined by x-ray crystallography | 21 |
| Figure 1.4 | MHC molecule binds peptides within the cleft | 23 |
| Figure 1.5 | The structure of an MHC class II molecule, determined by x-ray crystallography | 25 |
| Figure 1.6 | Crystal structure of the 20S proteasome α- and β-subunits from Archaebacteria *T. acidophilum* | 41 |
| Figure 1.7 | Crystal structure of the 20S proteasome from Archaebacteria *T. acidophilum* | 42 |
| Figure 1.8 | Crystal structure of the 20S proteasome from Archaebacteria *T. acidophilum*, showing the location of the peptidase active site | 45 |
| Figure 1.9 | Structural features of the 20S and 26S proteasomes | 47 |
| Figure 1.10 | The putative proteasome ubiquitin cycle | 49 |
| Figure 1.11 | Phylogenetic tree of proteasome subunits | 58 |
| Figure 3.1 | Sequence of MB1 cDNA clone | 90 |
| Figure 3.2 | Comparison of amino-acid sequence of MB1 and LMP7 human proteasome β-subunits | 91 |
| Figure 3.3 | Similarity of primary structures of various human proteasome β subunits | 94 |
| Figure 3.4 | Sequence of the Delta cDNA clone | 96 |
| Figure 3.5 | Comparison of amino-acid sequence of Delta and LMP2 human proteasome β-subunits | 97 |
| Figure 3.6 | Analysis of mRNA expression in wild-type and LMP-deficient cells | 99 |
| Figure 3.7 | Analysis of the protein sequences of various human proteasome β-subunits | 101 |
| Figure 3.8 | Comparison of amino acid sequence of the published human proteasome β-subunits | 102 |
| Figure 4.1 | Chromosome localization of MB1 gene by fluorescence in situ hybridization (FISH) | 108 |
Figure 4.2 Chromosome localization of Delta gene by fluorescence in situ hybridization (FISH) ................................................................. 110
Figure 4.3 Similarity of primary structures of yeast and human proteasome subunits ...................................................................................... 112
Figure 4.4 Model for the genetic origin of the LMP genes ............................................ 115
Figure 4.5 Model of the 20S proteasome .................................................................. 117
Figure 5.1 Two dimensional PAGE analysis of the proteasome purified from placenta and immunoblot of MB1 subunit .................................. 122
Figure 5.2 Assay of the specificity of anti-MB1 antiserum ........................................ 123
Figure 5.3 Immunoprecipitations of proteins using the anti-Delta antiserum .......................................................... 126
Figure 5.4 Assay of the specificity of anti-Delta antiserum ........................................ 127
Figure 5.5 Delta is processed from a ~23kDa precursor ........................................... 130
Figure 5.6 Analysis of MB1, Delta, LMP2 and LMP7 protein expression in .174 cell lines ........................................................................ 132
Figure 5.7 Analysis of MB1, Delta, LMP2 and LMP7 protein expression in cell lines cultured with and without addition of γ-interferon .................................................................... 133
Figure 6.1 Electron micrographs of negative stained 20S proteasomes from 721 and .174 cell lines ...................................................... 138
Figure 6.2 Electron micrographs of negatively stained IgG labeled 20S proteasomes from 721 cells ..................................................... 140
Figure 6.3 LMP2 and Delta subunits are part of the 20S proteasome complex in Raji cells ........................................................................ 142
Figure 6.4 Electron micrographs of negatively stained IgG labeled 20S proteasomes from .174 cells ..................................................... 143
Figure 7.1 List of the 17 human B cell lines from patients or individuals from families that carry abnormalities ........................................... 149
Figure 7.2 Analysis of LMP2, LMP7, MB1 and Delta protein expression in human B cell lines from patients with Hodgkin's disease ........................................................................ 150
Figure 7.3 Analysis of LMP2, LMP7, MB1 and Delta protein expression in human B cell lines from patients carrying various diseases ........................................................................ 151
Figure 7.4 List of the 26 human cell lines, which include non-B cells, tumour cells, and mutant cells ........................................................................ 153
Figure 7.5 Analysis of LMP2, LMP7, MB1 and Delta protein expression in EJ28, Butler, 59M, GM04286, 180D, GER and HELF human cell lines ........................................................................ 154
Figure 7.6 Analysis of LMP2, LMP7, MB1 and Delta protein expression in A818.7, GM05401, GM7730A, SKOV3, OVCA433, ZR75, SKBR3 and L132 human cell lines ...............155

Figure 7.7 Analysis of LMP2, LMP7, MB1 and Delta protein expression in GM5131, 1847, ME180, JATP, A431, HaPaC1 and JAR human cell lines .............................................156

Figure 7.8 Analysis of LMP2, LMP7, MB1 and Delta protein expression in 5637, HepG2, MCF7 and W138 human cell lines ..............................................................................157

Figure 7.9 Analysis of LMP2, LMP7, MB1 and Delta protein expression in the cytoplasm and nuclei of the following cell lines: 721, .174 and HeLa-cervix carcinoma ..................160

Figure 7.10 Analysis of the specificity of the anti-MB1 antiserum ..........161

Figure 7.11 Analysis of LMP2, LMP7, MB1 and Delta protein expression in the cytoplasm and nuclei of the following cell lines; JAR-choriocarcinoma placenta, MCF7-breast adenocarcinoma, and WI38-fetal lung ..........164

Figure 7.12: Analysis of LMP2, LMP7, MB1 and Delta protein expression in the cytoplasm and nuclei of the following cell lines; 5637-bladder carcinoma, HepG2-hepatocellular carcinoma liver, GM5131-chromosome aberrant fibroblast ..............................................165

Figure 7.13 Analysis of LMP2, LMP7, MB1 and Delta protein expression in the cytoplasm and nuclei of the WJRo76 LMP-2 mutant B cell line ................................................166

Figure 7.14 Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of 721 cell line cultured with and without addition of γ-interferon ...............169

Figure 7.15 Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of .174 cell line cultured with and without addition of γ-interferon ..........170

Figure 7.16: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of HeLa cell line cultured with and without addition of γ-interferon ..........171

Figure 7.17 Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of HELF cell line cultured with and without addition of γ-interferon ..........173
Figure 7.18  Analysis of MB1 and Delta protein expression in the cytoplasm and nuclei of synchronized 721, Raji and Hela cell lines ........................................................... 174

Figure 7.19  Analysis of LMP2 and LMP7 protein expression in the cytoplasm and nuclei of synchronized 721, Raji and Hela cell lines ........................................................... 176
Chapter 1

Introduction

The recognition of antigen by the immune system is mediated by two sets of highly variable receptor molecules, the immunoglobulins that serve as receptors of B cells, and the antigen-specific receptors of T cells. Immunoglobulins are secreted by activated B cells and can deal with extracellular antigens, such as bacteria, but are not able to eliminate intracellular antigens. T cells do not recognize foreign antigens directly, but instead recognize small fragments of antigens bound to class I or class II molecules of the MHC. The T cells either kill the cell presenting these peptides, deliver help to B cells to generate an antibody response or stimulate inflammation. Peptides that associate with class I molecules are usually generated from proteins synthesized or introduced directly to the cytoplasm. Peptides that associate with class II molecules are usually from extracellular sources and are internalized by antigen presenting cells to be degraded in an endosomal or lysosomal compartment.

Each individual has only a small number of different MHC molecules. However, because each MHC molecule can bind a large number of different peptide antigens, a tremendous diversity of peptide-MHC complexes can be formed, permitting the immune system to respond specifically to a wide variety of antigens. MHC molecules are able to retain their ligands on the cell surface long enough for T cell recognition to occur.

1.1 History of the human MHC

The human major histocompatibility complex (MHC) is a region of the human genome encompassing about 4 million base pairs of DNA on the short arm of chromosome 6, approximately 0.1% of the human genome. Over 150 genes in the MHC region have been identified, many of them encode proteins that regulate interactions among cells of the immune, including the histocompatibility antigens. The histocompatibility antigens are polymorphic or genetically variable antigens expressed on the surface of all nucleated cells that determine the ability of tissues to be grafted successfully or be rejected. A large number of diseases, many of the autoimmune type, have been associated with this region. In addition to its importance
in immunology, it provides a useful model for investigating gene clustering, polymorphism, linkage disequilibrium and recombination.

The histocompatibility phenomenon was first observed at the beginning of this century when mice tumours transplanted into healthy hosts revealed that transplantability of tumours was dependent on the strain of animal, and approximately 15 independent genes were responsible for the tumour rejection (Little and Tyzzer 1916).

The genetics of tumour transplantation and of blood group antigens were first brought together when one blood group antigen, called antigen II, was linked to resistance of tumour transplants in mice (Gorer 1937). Subsequently, skin graft studies on human burn patients showed that graft rejection was an immunological process in which the host became sensitized by the first-set grafts and therefore rejected the second-set grafts more rapidly (Medawar 1946).

To identify individual loci involved in tumour rejection, mouse strains that differed at 15 or more of the resistant genes were crossed and the produced lines, called congenic lines, differed from each other for some of the resistant genes, but not all. These congenic lines tested with the antigen II antisera, from mice that rejected tumour transplants, demonstrated that the resistant factor was probably an allele at the locus coding for the antigen II (Gorer et al. 1948). This locus was named H-2 for histocompatibility-2 and the tumour resistant factors were called histocompatibility (H) genes. Some histocompatibility genes had a stronger (major) effect than others, resulting in a marked difference in their ability to cause rejection. The major genes located in the same region as the H-2 complex, in the mouse chromosome 17. This became known as the major histocompatibility complex (MHC).

The discovery of the human MHC, in the early 1950's, started with studies performed by Jean Dausset with leukopenic patients where he showed that leukocytes could be agglutinated by antibodies the same way as erythrocytes (Dausset 1950). By reproducing these experiments, Rose Payne noticed that leukoagglutinins were not only present in the sera of leukopenic patients, but also in patients that had undergone multiple transfusions and women whom had several pregnancies (Payne and Rolfe 1958). Similar studies performed by van Rood proved that multiple pregnancies immunized mothers against leukocytes leaked from the foetus into the mother's circulation (van Rood et al. 1958). The clinical relevance of the leukocyte antigens in organ transplantation led, during the next decade, to the serological characterization of the complex.
The first evidence that a second kind of locus existed in the MHC came from experiments in which the lymphocytes from different individuals were mixed in culture. It was observed that the cells began to transform into large blasts and divide before they returned to the resting stage again. This phenomenon was called mixed lymphocyte reaction (MLR) (Bach and Amos 1967). By typing H-2 recombinants it was observed that the correlation between serologically defined determinants and MLR determinants was not absolute (Yunis and Amos 1971).

Simultaneous evidence for new loci in the MHC was observed by injecting synthetic polypeptides (PPL) into rabbits and pigs. Not all the animals responded to the peptide, and by crossing different animals it was soon observed that the response to PPL was controlled by a small number of genes. It became known as immune response gene, or Ir. More detailed typing of mouse strains recombinant over the H-2 region showed that the Ir locus was linked to, but distinct from, H-2 (McDevitt et al. 1972).

The classical MHC molecules were then called Class I antigens and the Ir and MRL-defined loci were called Class II antigens. The Class I and class II molecules were shown to consist of two chains, α and β (Walsh and Crumpton 1977). Both chains of the Class II molecules are encoded by the MHC region, while in class I molecules the α chain is MHC encoded, and is associated with a low molecular weight protein known as β2 microglobulin which maps to another chromosome.

The biological role of MHC antigens was determined by the discovery that foreign antigens could only stimulate T lymphocytes when presented in association with MHC antigens (Zinkernagel and Doherthy 1974). The elucidation of the three-dimensional structure of MHC antigens and the discovery of the T cell antigen receptor have enhanced understanding of how MHC antigens function (Bjorkman et al. 1987a). The MHC has been divided in three sub-regions, denoted class I, class II and class III.

1.2 Genomic organization of the human MHC complex

Figure 1.1. shows a detailed map of the human MHC region. A list of individual genes is found in Trowsdale et al. (1991) and Campbell and Trowsdale (1993).
Figure 1.1: Map of the human histocompatibility complex. This figure was compiled from information from a number of different laboratories. It is based upon a map originally prepared by Drs. I. M. Hanson and J. Trowsdale. References for individual genes can be found in Trowsdale et al., 1991 and Campbell and Trowsdale, 1993.
The class I region contains three classic well-characterized loci, HLA-A, HLA-B and HLA-C, each of which encodes the α chain of the class I antigen. Other loci, HLA-E, HLA-F and HLA-G are called non-classical, or class Ib, and have also been characterized in some detail (Geraghty 1993, Stroynowski et al. 1994). Each gene can potentially direct the synthesis of class I proteins and they can contribute to immune response against pathogens. HLA-G product is expressed on extravillous trophoblasts and cell lines derived from placental tissues, and its function may be related to survival of the semi-allogenic foetus (Parham 1995). The class I region also contains a number of other class I genes which are either pseudogenes or fragmented genes.

The class II region contains the α and β chain genes which encode the class II antigens HLA-DR, -DP and -DQ. The HLA-DR gene family consists of one α chain and nine β chain genes (DRB1-9), including pseudogenes (Rollini et al. 1985, Andersson et al. 1987, Gorski et al. 1987, Kawai et al. 1989). The HLA-DQ and -DP families each have two α and two β chain genes, HLA-DQA1, -DQA2, -DQB1, -DQB2 and -DPA1, -DPA2, -DPB1 and -DPB2 (Gustafsson et al. 1987, Jonsson et al. 1987). The HLA-DR, -DP and -DQ α chains primarily associate with β chains from their own loci, although studies have demonstrated that some cross-pairing can take place in vitro (Altmann et al. 1990). This pairing may also occur in vitro, DRA coupling with DQB in some cell lines (Lotteau et al. 1987, Lechler 1988). The HLA-DQA1 and -DQB1 gene products associate to form the HLA-DQ antigens; similarly, the HLA-DPA1 and -DPB1 associate to form the HLA-DP gene products. The HLA-DQA2 and -DQB2 genes have no obvious defects in their coding regions and could be functional, although their mRNAs are not widely expressed (Auffray et al. 1987). Nucleotide sequencing has shown that HLA-DPA2 and -DPB2 are pseudogenes incapable of producing functional products.

Additional α and β chain encoding genes have been found in the class II region. The HLA-DNA and -DOB loci are located a considerable distance from each other and apparently they can associate to form the DO molecule which appears to be expressed only on thymus and B cells. The HLA-DMA and HLA-DMB are also highly related to class II α and β chains, and associate to form a dimer (Sanderson et al. 1995). DM molecules are suggested to facilitate peptide binding to class II molecules within multilamina intracellular compartments (Morris et al 1994).

A group of genes involved in antigen presentation for class I molecules have been found in the class II region of the complex. They are TAP1 and TAP2 for transporters associated with antigen processing, and they are part of the ABC

The class III region contains around 40 genes within a region covering about 680 kb of DNA. A number of these genes encode products with known immunological functions. These include the complement component genes C4, C2 and Bf, and the tumour necrosis factor genes TNF-A and TNF-B. Heat shock protein Hsp70 genes are also located within the class III region, but their potential control in immune response remains controversial. The function of the remaining genes is not known, with the exception of G7a, which encodes valyl-tRNA synthetase, and CYP21B, which encodes the steroid 21-hydroxylase (Carroll et al. 1985, Spies et al. 1989).

The class I and class II genes have a conventional organization into introns and exons which reflects the separations between protein domains (Lee et al. 1982, Malissen et al. 1982, Trowsdale et al. 1985). All class I genes have three separate exons encoding α1, α2 and α3 domains, and the signal sequence plus the first few amino acids of the mature protein in a separate exon. The rest of the mRNA molecule is encoded by a complex arrangement, with the connecting peptide and signal sequence plus part of the cytoplasmic tail on one exon, followed by one or two further exons to encode the rest of the cytoplasmic tail. A low level of alternative splicing occurs in some genes, the relevance of which is not apparent, although in some cases, as in DQ, splicing out the transmembrane region yields a secreted form of the molecule.

1.3 Structural organization of the human MHC antigens

1.3.1 Class I structure and bound peptide

MHC class I molecules are composed of a membrane anchored heavy or α chain (approximately 45kDa) and a soluble noncovalently associated light chain, β2-microglobulin (12kDa) (figure 1.2). The heavy chain contains three distinct extracellular domains, a short transmembrane domain and a cytoplasmic domain (Ploegh et al. 1981). The three extracellular domains are called α1, α2 and α3 and are composed of 90, 92 and 92 amino acids respectively. Following the α3 domain is the transmembrane region of around 25 amino acids and the cytoplasmic tail of around 30 amino acids. The transmembrane segment anchors the protein in the
Figure 1.2: Primary structure of the class I and class II histocompatibility molecules.

(A) Class I molecules are composed of an α-chain and β2-microglobulin. The α-chain is oriented with an extracellular amino-terminus, has three extracellular domains (α1, α2, α3), a single transmembrane domain (TM) of around 25 amino acids, and a carboxy-terminal cytoplasmic domain of around 30 amino acids.

(B) Class II molecules are heterodimers composed of a 33-35 kDa α-chain and a 25-29 kDa β-chain. Both the α- and β-chains are oriented with an extracellular amino-terminus, have two extracellular domains, a single TM domain, and a carboxy-terminal cytoplasmic tail of around 10-15 amino-acids. (Photographs from Immunobiology 1994, Janeway-Travers, Current Biology Ltd./Garland Publishing Inc., pages 4.4 and 4.5)
membrane and may mediate interaction with accessory proteins required for proper assembly and folding of class I molecules.

The light chain β2-microglobulin (β2m), and the α3 domain show sequence similarity to the immunoglobulin constant-region domains, and are therefore members of the Ig gene superfamily, which includes among others TcRs (T cell receptors) and class II histocompatibility antigens (Peterson et al. 1972, Becker and Reeke, 1985). In addition to the class I heavy chain, β2m binds to a number of other class I related molecules, such as CD1 encoded on chromosome 1 in humans and an Fc receptor which is responsible for the uptake of IgG from the intestinal cells of neonatal rats (Calabi and Milstein 1986, Simister and Mostov 1989). β2m is required for the formation of a stable class I complex. The absence of β2m leads to retention of the incompletely assembled complex in the ER resulting in reduced cell surface class I expression (Gomez et al. 1978, Rein et al. 1987, Koller et al. 1990, Zijlstra et al. 1990). However, the requirement of β2m is not absolute and free heavy chains are expressed at the cell surface in the absence of β2m, but in a far less efficient manner (Allen et al. 1986).

Most amino acid substitutions between the products of different class I alleles are localized to the α1 and α2 domains with the α3 domain being relatively conserved, and β2m invariant. It was predicted from sequence analysis that the α1 and α2 domains would comprise the part of the molecule that binds peptides and is contacted by the TcR.

The three-dimensional structure of the class I molecule HLA-A2 determined using X-ray crystallographic techniques was a major advance in the study of MHC structure and function (figure 1.3) (Bjorkman et al. 1987a, Bjorkman et al. 1987b). As anticipated from their primary sequences the α3 and β2m domains were folded into β-sandwich structures similar to the Ig constant region. However, these Ig domains interacted with each other in a novel arrangement not previously seen in any of the known Ig structures. The α1 and α2 domains sit on top of the Ig-like domains forming a platform composed of two α helices supported by an eight-stranded β-pleated sheet. The large cleft between the α-helices was the most likely location for binding a peptide and would be accessible to the TcR. This was also supported by the fact that the HLA-A2 shows electron-dense material in the cleft thought to represent a peptide.
Figure 1.3: The structure of an MHC class I molecule, determined by x-ray crystallography.

(A) Ribbon diagram of the HLA-A2. The MHC class I, shown schematically, is a heterodimer of an α-chain, non-covalently associated with β2-m. The β2m does not span the membrane. The α chain folds in three domains, α1, α2 and α3. The α3 and β2m have homology to immunoglobulin domains, and have a similar folded structure, while the α1 and α2 domains are folded together into a single structure consisting of two segmented α helices lying on a sheet of eight β-strands. The folding of the α1 and α2 domains create a long cleft that is the site at which peptide antigens bind to MHC molecules. The transmembrane region and the short stretch of peptide that connects the external domain are not seen as they have been cleaved off by papain digestion.

(B) Top view of the structure of HLA-A2 molecule. The sites of the cleft are formed from the inner faces of the two α helices, while the β-pleated sheet is formed by the pairing of the α1 and α2 domains creates the floor of the cleft. (Photographs from Immunobiology 1994, Janeway-Travers, Current Biology Ltd./Garland Publishing Inc., page 4.4)
Resolving the structure of HLA-B27 made it possible to see in detail the electron-dense structure within the cleft (figure 1.4) (Madden et al. 1991). As predicted, this showed a short peptide, of nine amino acids in length, some of whose side chains were extending into pockets. The overall domain arrangement and backbone structure of three human class I allelic products were very similar with the amino and carboxyl termini of the peptides that they bind being strongly fixed to the edges of the groove.

In most cases these peptides have a length determined by the MHC allele, varying from 9 to 11 amino acids. Some class I molecules can accommodate even longer peptides, up to 13 a.a., as in HLA-Aw68, and still have the amino and carboxyl termini fixed in their groove. The extra length is accommodated by a bulging of the middle part of the peptide (Madden et al. 1991). Peptides that are too short to reach the end of the peptide groove can still bind, but with only one end fixed into the groove (Bouvier and Wiley, 1994). Some peptides might have one end that hangs out over the end of the groove, as seen for the HLA-B27 molecule, where subsets of peptides that were too long had one or both of their ends exceeding the groove (Urban et al. 1994).

Another major advance in the understanding of the nature and size of peptides bound to the class I molecules was due to the analysis of naturally processed peptides. The peptides were also short, most commonly 8 to 10 amino acids long, showing common amino acids at particular positions in the sequence called anchor residues (Falk et al. 1991). Most peptides that bind to class I have an anchor residue at the carboxy-terminus, and this is almost always a hydrophobic or charged residue. Changing any anchor residue prevents the peptide from binding, but peptides containing the correct anchor residues and length will bind the appropriate molecule independent of the sequence of the peptide at other positions allowing the class I molecules to bind a wide variety of peptides. Besides the anchor at the carboxyl terminus, another anchor of class I ligands was found at position 2 (P2) of the peptide, and in many cases at P3, P5 or P7 (reviewed in Elliot et al. 1993). The amino acids side chains of the peptides are accommodated into pockets along the groove, designated A through F (Matsumura et al. 1992). Class I binding peptides have one common denominator, they all anchor in the peptide binding cleft with their C-terminal amino acid side chain in the F-pocket.
Figure 1.4: MHC molecules binds peptides within the cleft.
(A) Crystal structure of HLA-B27 molecule containing additional material lying within
the peptide binding cleft. (Photograph from Madden et al. 1991)
(B) Peptide bound into the groove shows that it lies in an elongated conformation along
the groove, with both ends tightly bound at either end of the cleft. (Photograph from
Madden et al. 1991)
(C) The upper surface of the MHC molecule, shown in white, is the surface recognized
by the T cell molecule. The peptide is shown in red. (Photographs from
Immunobiology 1994, Janeway-Travers, Current Biology Ltd./Garland Publishing
Inc., page 4.6)
1.3.2 Class II structure and bound peptide

Like class I molecules, the HLA class II molecules are transmembrane glycoproteins. However, two transmembrane chains form the class II heterodimer, an α chain (30-34 kD) and a β chain (26-29kDa) (figure 1.2). Each chain has two extracellular domains of approximately 90 amino acids each (the α1, α2, β1 and β2 domains), followed by a short peptide connecting to a transmembrane region of about 30 amino acids and a short cytoplasmic tail (Auffray and Strominger 1985). The α2 and β2 domains, like the α3 and β2-microglobulin domains of the MHC class I molecules, have amino acids sequence similarities to immunoglobulin constant domains and each one contains an internal disulphide bond. The membrane distal domains, α1 and β1, associate to form a structure that constitutes the peptide binding region of the molecule. This structure closely resembles the peptide binding structure of the class I molecule, which in class I is generated by the two N-terminal domains (α1 and α2).

The structure of the peptide binding cleft was predicted on the basis of sequence similarities and the location of conserved and polymorphic residues between the α1 and α2 domains of class I and the α1 and β1 domains of class II (Brown et al. 1988). Recently the x-ray crystallographic analysis of the HLA-DR has confirmed the major features of the prediction (Brown et al. 1993) (figure 1.5). The structure consists of eight strands of anti-parallel β-sheet with two anti parallel α-helical regions overlaying them. There is a deep cleft between the α-helices which accommodates the bound peptide. Four of the β-strands and one helical region are derived from each subunit.

There are striking differences in the sets of peptides bound by class I and class II molecules. Peptides associated with class II molecules are longer than the class I ones, usually 13-25 residues (Rudensky et al. 1991, Newcomb and Cresswell 1993). The difference in length preference are due to small structural differences between the two binding sites, and the positioning of key residues that make hydrogen bonds to the bound peptide. Structural analysis of class II molecules has confirmed earlier predictions that a central stretch of the peptide is held in the MHC cleft, whereas both the amino and carboxyl termini protrude out of the end of the groove, unlike class I. The HLA-DR molecule, characterized by x-ray crystallography, has a single nonpolar pocket near one end of the groove, which accommodates a hydrophobic amino acid side chain of the bound peptide. Other conserved residues in the groove bind to the extended polypeptide backbone, independent of the individual amino acid side chains (Brown et al. 1993).
Figure 1.5: The structure of an MHC class II molecule, determined by x-ray crystallography.
(A) Ribbon diagram of HLA-DR1. MHC class II, shown schematically, is a heterodimer of a α-chain, non-covalently associated with a β-chain. Each chain has two domains, and the two together form a compact four-domain structure similar to class I. The peptide binding-site is between the α-helices contributed by the α1- and β1- domains. The class I molecules have differences in the structure of the α-helices, compared to class II, and the peptide-binding cleft is open at both ends. The α2- and β2- domains have homology to the immunoglobulin domains.
(B) Schematic representation of the top view of the class II molecule. The sites of the cleft are formed from the inner faces of the two α helices, while the β-pleated sheets is formed by the pairing of the α1 and β1 domains creates the floor of the cleft.
(C) The upper surface of the MHC class II molecule, shown in white, is the surface recognized by the T cell molecule. The peptide is shown in red. The more open ends of the binding groove allow the peptide to protrude at either end of the cleft (Photographs from Immunobiology 1994, Janeway-Travers, Current Biology Ltd./Garland Publishing Inc., pages 4.5 and 4.6)
Comparisons between the class II α2 domain and the class I β2m subunit show that they are folded and oriented in a similar manner, probably due their similar interaction with the T cell receptor.

1.4 Cell biology of MHC class I antigens

Class I molecules present peptides originating in the cytosol, whereas class II molecules present peptides from extracellular origin. In both cases, the generation of peptides involves processing of native proteins in order to be recognized by MHC molecules. In the last few years new discoveries in the mechanism of antigen processing and presentation has led to a better understanding of class I and class II antigen processing pathways.

1.4.1 Class I molecules bind peptides

The classical studies from Zinkernagel and Doherty on T cell recognition of lymphocytic choriomeningitis virus, indicated that the T cells could only lyse infected cells that had the same MHC haplotype (Zinkernagel and Doherthy 1974). This phenomenon was called "MHC restriction". The MHC class I-restricted nature of cytotoxic T lymphocytes (CTL) recognition of virus represented the first function of the class I products. It was first proposed that the viral antigens were membrane glycoproteins, but at that time it was not possible to identify the nature of the antigen recognized on virus-infected cells.

The first indication that CTL recognize viral peptides, instead of whole viral proteins, came from enzymatically and chemically cleaved Sendai virus. Peptides smaller than 5kDa could stimulate Sendai virus-specific CTL cells (Guertin and Fan, 1980). These findings were confirmed by the demonstration that cells incubated with synthetic peptides of 14 or more residues, corresponding to sequences in influenza virus nucleoprotein-sensitized target cells for recognition by CTL (Townsend et al. 1986). The use of synthetic peptides made it possible to define the determinants recognized by CTLs specific for different proteins. However, synthetic peptides could only give a rough approximation of the natural epitopes that are actually recognized by the T cells in vivo.

A great advance in the understanding of size, motifs and nature of peptides was possible by using an acid extraction method for elution of MHC-bound peptides (Rötzschke et al. 1990a, Rötzschke et al. 1990b, Falk et al. 1990, Falk et al. 1991,
Van Bleek and Nathenson 1990). Analysis of naturally processed peptides was the first precise information on the content of MHC and provided new insights into mechanisms of peptide generation and presentation. The fact that most peptides were soluble under acidic conditions was used to develop two different methods for the isolation of naturally-processed peptides. One method analyses the complete cellular peptide pool, whereas the other one, after immunoprecipitating MHC molecules, analyses only peptide antigens that are actually presented. The final step in both isolation procedures is the separation of biological material by reverse-phase high performance liquid chromatography (HPLC).

Antigenic peptides could be identified in HPLC fractions by their capacity to sensitize target cells for lysis by antigen-specific CTL. Fractions containing the naturally-processed antigens could be further characterized, by comparing their elution behaviour with that of synthetic peptides (Rötzschke et al. 1990a, Van Bleek and Nathenson 1990, Falk et al. 1991).

The higher purity of fractions obtained by the precipitation method made it possible to sequence a few abundant peptides (Van Bleek and Nathenson 1990, Falk et al. 1991). The sequence information on naturally-presented viral antigens and self peptides suggested a uniform length and allele-specific sequence characteristics for MHC-presented peptides (Falk et al. 1991). However, the consensus motif is merely a minimal requirement for a peptide to be selected for presentation further conditions must be met, such as intracellular location, concentration, access to proteolytic degradation of the parent protein, structure of the protein and affinity to MHC molecule.

The first direct structural examination of the peptide with class I molecules was visualised by X-ray crystallographic studies of the human HLA-class I molecule HLA-A2.1 and HLA B-27 which showed the potential sites for interaction with the T cell receptor (Bjorkman et al. 1987a, Jardetzky et al. 1991). The question now was how cytosolic proteins were processed into peptides, transported into the ER to bind the MHC molecule and transported to the cell surface.

1.4.2 Mutant cells defective in antigen presentation

Understanding the different steps involved in antigen processing and presentation was greatly enhanced by using a number of mutant cell lines that fail in one or more steps of the pathway. RMA was the first cell line used that produced mutants deficient in antigen processing. RMA was derived from a chemically
The human mutant cell line LCL 721.174 (called .174) has a similar phenotype to murine RMA-S. The .174 line was produced by γ-irradiating Epstein-Barr virus-transformed 721 B lymphoblastoid cell lines and selecting for mutants (DeMars et al. 1984). One derivative of the mutageneses of 721 cells, called LCL .45.1, had a deletion spanning the entire HLA complex in one copy of chromosome 6. LCL 45.1 cells were then re-irradiated and selected for class II products. The .174 line resulting from this treatment had a large deletion in the class II region of the other copy of the chromosome 6 (Erlich et al. 1986). The class I gene products encoded by this chromosome were normal, but their assembly with β2m and transport to the cell surface were less efficient than in normal cells. When .174 were fused with cells with normal MHC complexes the surface expression of class I molecules encoded by .174 was restored (DeMars, 1985).

LCL .174 fused to a T cell lymphoblastoid line CEMR produced a cell line called T1, which again restored the surface expression of the class I gene products of .174. T2, a subclone of T1, had both copies of chromosome 6 derived from CEMR deleted, but class I expression was lost (Salter and Cresswell 1986). Both the T2 and .174 lines presented synthetic peptides, but not endogenously synthesized antigens, to CTL (Cerundolo et al. 1990, Hosken and Bevan, 1990).

All these mutant cells, human and murine, could synthesize normal amounts of class I α chain and β2m, but failed to assemble class I molecules and transport them to the cell surface efficiently.

1.4.3 Entry of proteins into the class I processing pathway

After the discovery that MHC molecules are recognized by the T cell in association with processed proteins, the next problem was how proteins would gain entry into the class I pathway. It was demonstrated that extracellular proteins were not efficiently processed by cells for association with class I antigens, by noticing that immunizations with protein antigens rarely induced an antigen-specific CTL responses. However, proteins produced during viral infection efficiently elicited CTL
response. It was proposed that either biosynthesized proteins or endogenous proteins would enter the class I processing pathway.

The first evidence that the cytosolic location of newly synthesized proteins was critical to their entry into the class I processing pathway came from experiments where inactivated influenza virus could not sensitize target cells for recognition by CTL specific for various influenza virus proteins (Hosaka et al. 1985, 1988, Yedwell et al. 1988). The delivery of proteins to the cytosol was directly demonstrated by immunofluorescence localization of viral structural proteins in the nucleus. Subsequently, it was shown that presentation of endogenous proteins could be inhibited by the drug Brefeldin A (BFA), which blocked presentation of processed proteins but not antigenic peptides to CTL (Yewdell and Bennink 1989). However, there is evidence that exogenous viral-produced proteins also enter in the class I pathway. Experiments with hemagglutinin (a viral membrane protein) showed that human cytomegalovirus virions and hepatitis B virus surface antigens could present peptides for class I presentation (Riddell et al. 1991, Jin et al. 1990). It appears that viral surface antigens that enter in the class I processing pathway are produced in the cytosol.

The first indication that cytosolic proteins efficiently entered in the class I processing pathway came from studies with T antigen, where cells expressing T antigen by DNA transfection were recognized by CTL induced by immunization with simian virus 40 transformed cells (Gooding and O'Connell 1983). At the same time, cells transfected with viral nucleoprotein DNA or recombinant vaccinia expressing the nucleoprotein gene, could be recognized by specific CTL cells (Townsend et al. 1984, Yewdell et al. 1984). Some of the influenza nucleoprotein fragments presented to CTL were degraded rapidly after their synthesis and were difficult to detect other than by CTL recognition.

It was proposed that nucleoprotein was degraded to peptides in the cytosol, and these peptides were transported to class I molecules (Townsend et al. 1985). In agreement with this, nearly all identified HLA class I eluted peptides were shown to be derived from proteins located in the cytosol or nucleus (Jardetzky et al. 1991).

A small proportion of class I molecules present on phagocytic macrophages can apparently bind peptides derived from endocytic compartments, in a similar way as the class II molecule. The invariant chain (Ii), a polypeptide normally found associated with class II molecules and believed to target class II complexes to
the endosomes, has been shown to associate with a subset of class I molecules (Cerundolo et al. 1992, Sugita and Brenner 1995). Class I molecules associated with Ii could enter the endocytic pathway and may be loaded with exogenously derived peptides for class I restricted presentation or processed into peptides themselves for binding to and presentation by MHC class II molecules (Chicz et al. 1992, Chen et al. 1990). But the majority of antigens can enter the class I pathway only via the cytosol.

1.4.4 Biosynthesis and assembly of class I molecules

One of the first questions asked after the discovery that peptide antigens were in the cytosol, was if the antigens would associate with class I prior to or after their arrival at the cell surface. Some evidence of where the assembly was taking place came from the discovery that cells treated with the antibiotic BFA could block the export of virus glycoproteins to the cell surface. BFA blocked presentation of viral proteins without interfering with their synthesis. However, BFA treated cells could present synthetic peptides containing determinants recognized by matrix- or nucleoprotein specific CTL. It was then concluded that exogenously added synthetic peptides associate with class I molecules that have already reached the cell surface, and that peptides from cytosolic proteins associated with class I molecules in an intracellular compartment. It was subsequently discovered that BFA inhibits transport of molecules into the cis-Golgi, resulting in accumulation of Golgi enzymes in the ER and collapse of the Golgi apparatus (Lippincott-Schwartz et al. 1990).

Further information about the intracellular location of class I assembly came from studies of the adenovirus E19 protein. E19 is a non-structural protein, deletion of which was known to reduce adenovirus infection in vitro. E19 protein is retained in the ER and cells infected with the virus or transfected with the E19 gene show reduced surface expression of the restricting class I molecule (Burget et al. 1987, Andersson et al. 1987). Further work revealed that the luminal domain of the E19 protein mediates the binding to class I and the transmembrane domain of E19 protein contains a retention signal in the cytoplasmic portion (Nilsson et al. 1990). This data demonstrate that antigen association occurs in the ER.

Studies in mutant cells showed that the assembly of α-chain and β2m in the ER had conformational effects on the class I molecule. By using the Daudi cell line, a mutant that lacks expression of β2m, it was shown that the class I heavy chain was malfolded, never exported from the ER and later degraded (Krangel et al. 1979). Additional experiments using RMA-S mutants, showed that the peptide has a role in stabilizing the class-I structure. In these cells the processed peptide is not able to
reach the compartment were newly synthesized class I molecules are located, producing very unstable class I molecules. However, these molecules can leave the ER at 26°C where they can bind exogenously added peptides, resulting in a conformational stabilization and increase in cell surface expression. The addition of peptides to RMA-S cells was able to stabilise class I when the cell was shifted to 37°C (Ljunggren et al. 1990).

Class I molecules from cell lysates of RMA-S cells could also be stabilised by adding peptides to the lysate, showing conformational changes and an increase in affinity for β2m (Townsend et al. 1989, Ljunggren et al. 1990, Townsend et al. 1990). Empty class I was also detected in the parental RMS cell line which had no defect in class I presentation (Ljunggren et al. 1990, Townsend et al. 1990). Suggesting that the availability of peptides in the ER was a limiting factor in class I expression at cell surface.

Experiments using 174 also showed that the mutant was unable to present antigen from infecting virus, but was able to present synthetic peptide epitopes when added exogenously. In fact, 174 could present epitopes better than the wild type parental cell line 721. Surface class I expression was also induced by adding peptide epitopes (Cerundolo et al. 1990). Similarly, the derivative of 174, T2, was unable to present endogenously synthesized antigen but could present added peptide epitope (Hosken and Bevan 1990).

For all the class I molecules, the optimal conformation required for the ER export seems to be dependent on the co-assembly of the α-chain with β2m and appropriate peptide.

1.4.5 The transporter associated with antigen processing, TAP.

After the demonstration that proteins lacking signals for access to the secretory compartment could contribute peptide epitopes for presentation via class I molecules at the cell surface, the existence of an auxiliary molecule that could be involved in the transport of cytosolic peptides into the ER was proposed.

A first suggestion that transporter proteins existed was provided by studying the class I-deficient mutant cell line, RMA-S, which could present exogenous peptides but not cytosolic ones. A similar phenotype was observed for several other mutants with defects in class I assembly. It was proposed that they could be deficient in a specific mechanism for translocation of peptides from the
cytosol to the ER (Townsend et al. 1989). Furthermore, the majority of class I molecules expressed in the mutants .174 and T2 contained either no peptide, or a limited set of peptides derived from signal sequences that are cleaved in the ER (Wei et al. 1992, Arnold et al. 1992).

Analysis of the .174 hybrid with normal cells established that the deletion in the MHC on chromosome 6 of .174 was responsible for the defect in antigen presentation and the class I assembly defect (Salter et al. 1986, Erlich et al. 1986, Cerundolo et al. 1990). It was suggested that gene or genes encoding the peptide transport should exist within this deletion (Cerundolo et al. 1990).

In 1990 four laboratories localized genes in the MHC encoding proteins very similar to the ABC family of transporters (Spies et al. 1990, Trowsdale et al. 1990, Monaco et al. 1990, Deverson et al. 1990). The transporter genes were then called TAP1 and TAP2, for transporters associated with antigen processing. The human genes were cloned by analysing the mutant .174 cell line which has a large deletion in the class II region of the MHC.

Immunoelectron microscopy analysis showed that TAP1 is localized in the ER membrane with the ATP binding domain most probably oriented to the cytosol (Kleijmeer et al. 1992). The TAP1 and TAP2 genes in humans are not highly polymorphic, and until now, 3 alleles of TAP1 and 5 alleles of TAP2, have been identified (Powis et al. 1992a). However, the rat TAP2 locus is highly polymorphic, probably to compensate the low number of class I alleles in this species. In rat the TAP genes have probably evolved in a way to increase the diversity of peptides presented by the limited range of class I available.

In order to study the function of the TAP genes, mutant cell lines that lacked class I at the cell surface were transfected with either TAP1 or TAP2 genes or both. The mutant line .134 was transfected with the TAP1 gene, RMA-S was transfected with the TAP2 gene and .174 with both TAP1 and TAP2 (Spies and DeMars 1991, Spies et al. 1992, Powis et al. 1991, Attaya et al. 1992, Kelly et al. 1992, Arnold et al. 1992, Momburg et al. 1992). In each case class I expression was restored to 50-100% of wild type, and the mutant phenotype was a result of mutations of either TAP1 or TAP2. In addition, it was possible to reconstitute antigen presentation in the mutant cell with TAP genes from different species (Momburg et al. 1992, Powis et al. 1991). In the murine mutant RMA-S the TAP2 deficiency could be restored for presentation of influenza antigens to murine T cells by transfection with the rat TAP2 sequence (Powis et al. 1991).
The role of TAP in peptide transport was questioned by experiments showing the existence of ATP- and TAP-independent peptide transport into microsomal vesicles (Koppelman et al. 1992, Levy et al. 1991). Later, a number of studies described the development of in vitro peptide transport assays that showed that peptide translocation was ATP- and TAP-dependent (Dobberstein 1992, Shepherd et al. 1993, Neefjes et al. 1993, Androlewicz et al. 1993).

Two further independent assays using lymphoblastoid cells permeabilized by streptolysin O showed accumulation of test peptides in the ER in an ATP- and TAP-dependent manner (Neefjes et al. 1993, Androlewicz et al. 1993, Shepperd et al. 1993). In the one assay the retention of peptides in the ER was achieved by including a consensus N-glycosylation site in the test peptide, enabling ER-translocated peptides to be recovered later on a ConA Sepharose column (Neefjes et al. 1993). In the second assay, peptide accumulation in the ER was observed by measuring peptides bound and retained by the nascent class I molecule of different cells (Androlewicz et al. 1993, Shepperd et al. 1993).

To analyse in detail the conditions for transport of peptides by TAP in vitro assays were developed. Because class I molecules bind preferentially peptides 8-10 amino acids in length the possibility of TAP having any additional function by modifying peptides to be loaded in the ER, or if any additional proteolytic activity in the ER lumen would cleave peptides from longer precursors derived from the cytosol. In most assays the TAPs showed preference for the length of the peptide transported, which corresponded to the peptide length preferred by the class I molecule (Shepherd et al. 1993, Androlewicz et al. 1993, Momburg et al. 1994b 1994c, Androlewicz and Cresswell 1994a). Cells permeabilized with streptolysin O could transport iodinated 16-40 mer peptides, which ones were first proteolytically processed into shorter peptides prior their transportation (Momburg et al. 1994b).

It was recently proposed that peptides longer than 8-10 amino acids long could also be transported into the ER, and be loaded onto class I molecules more frequently than previously supposed. Loaded longer peptides can have the amino- or carboxy-terminal amino acids located in their complementary pockets in the class I groove with the extra length accommodated by bulging (Collins et al. 1994). It is also possible that longer peptides do not have an ideal binding site and they have one or both termini left hanging loose (Collins et al. 1994, Urban et al. 1994). It has been shown that the identity of the peptide C-terminal residue essentially governs the species-specific substrate specificity of TAP (Neefjes et al. 1995). Rat and Mouse u TAP alleles preferentially transport peptides with hydrophobic C-terminal residues.
No such selection was reported for human TAP or for rat TAP a allele (Momburg et al 1996).

To identify the peptide binding site in TAP proteins, photoactivable peptide cross-linkers were used, revealing that TAP has a combinatorial binding site that consists of both TAP chains (Androlewicz et al 1994b). It was also observed that peptides bind to TAPs in an ATP independent manner. ATP hydrolysis is necessary for peptide translocation. Expression of TAPI and TAP2 in insect cells has shown that they are the minimal components necessary for peptide translocation across the ER membrane, in conjunction with ATP binding to the COOH-terminal of murine TAP1 and TAP2 (Meyer et al. 1994, van Endert et al. 1994, Wang et al. 1994).

It was shown that the TAP complex interacts with MHC molecules, and dissociation of the transporter from class I molecules coincides with exit of the complexes from the ER (Suh et al. 1994, Ortmann et al. 1994). However, it was not known if this association was direct or mediated by another protein. By studying a new mutant cell line 721.220, it was observed that the expressed functional class I and TAP molecules did not associate in the ER (Greenwood et al. 1994). This cell line showed a 80% reduction of class I surface expression that could be restored by introducing an unidentified 48kd MHC-linked gene, suggesting that additional proteins may be involved in the assembly of class I molecules (Grandea et al. 1995).

1.4.6 Chaperones in class I assembly

Class I molecules are normally retained in the ER until they are released by binding peptides. And like other nascent glycoproteins, class I molecules were found associated with an ER resident chaperone, called calnexin (Degen et al. 1991, Rajagopalan et al. 1994). Calnexin mediates disulphide bond formation in class I heavy chain and may promote dimerization of heavy chain with β2m (Tector et al. 1995, Sugita et al. 1994). It has been shown that murine and human calnexin-associated class I heavy chain associates with β2m prior to TAP association (Suh et al. 1994, Sugita et al. 1994, Nößner et al. 1995, Carreno et al. 1995). These data suggest that β2m is a prerequisite for class I heavy chain binding to TAP. In addition, it has been shown that in the absence of heavy chain, β2m can itself bind to TAP (Carreno et al. 1995). However, analysis of calnexin deficient cells showed that calnexin is not essential for class I assembly, and calnexin deficient cells were able to assemble mature and functional class I molecules (Scott et al. 1995). The association of class I molecules with TAPs is released in the presence of peptide, and assembled MHC class I molecules are transported to the cell surface (Carreno et al. 1995).
However, it is not known whether peptide binding is sufficient for correct class I folding, or whether TAP, or other chaperones, are required for class I folding (Day et al. 1995).

1.4.7 Production of peptide for class I presentation

As detailed before, the cytosol is responsible for the entrance of proteins to the antigen processing pathway, and it was likely that proteolysis occurs in this compartment. The existence of a cytosolic degradation pathway was first suggested by the inability of agents that inhibit lysosomal proteolysis to completely inhibit protein turnover in the cell. In addition, an ATP-dependent cytosolic proteolytic system could be reconstituted from cytosolic extracts. This system was shown to use ubiquitin to target proteins to a cytosolic protease. The ubiquitinated proteins could be degraded in an ATP dependent manner. This protease, recently called proteasome, has been implicated as the major protease in the class I processing pathway (further discussed in section 1.5).

Indications that proteins were degraded in the cytosol were first observed by using recombinant vaccinia virus that expressed "minigenes" in the cytosol. Nucleoprotein peptides of 30 amino acids or less containing the "naturally processed" nonamer residues were expressed in cells and further recognized by nucleoprotein specific CTL (Gould et al. 1989). The lysis of sensitized cells by nucleoprotein-specific CTL could be blocked by Brefeldin A, indicating that peptides were processed as cytosolic proteins (Eisenlohr et al. 1992a). Similar results were found for HLA-A2 restricted influenza virus matrix CTL (Anderson et al. 1991). These findings demonstrated that proteolysis most probably occurred in the cytosol.

A selectivity in peptide sequences for proteolysis was observed by using vaccinia virus expressing short peptide sequences from nucleoprotein or cytomegalovirus containing altered flanking regions. When 13 residues of HLA class I was expressed in the middle of nucleoprotein it was shown that this segment could be presented to Kd-restricted HLA-specific CTL (Chemini et al. 1989). Similar results were obtained with vaccinia virus expressing influenza hemagglutinin molecules. In contrast, short nonamer cytomegalovirus sequences inserted in the amino terminus of hepatitis B virus could not sensitize specific CTL cells (del Val et al. 1991a). However, the insertion of five alanine residues on each side of the determinant rescue presentation (del Val et al. 1991b). Flanking sequences also influence the presentation of oligopeptides expressed by recombinant vaccinia virus.
nucleoprotein, and the addition of two amino acids diminished the antigenicity in vitro and in vivo (Eisenlohr et al. 1992b).

1.4.8 ER proteases involved in antigen processing

It has been proposed that longer peptides transported by the TAPs required further trimming to be presented, suggesting additional peptide trimming in the ER (Falk et al. 1990).

To test this possibility, long intracellular peptides were expressed by recombinant vaccinia virus and expression was greatly enhanced by simultaneous expression of a membrane protease targeted to the ER (Eisenlohr et al. 1992a). However, the only characterized ER resident protein is the signal peptidase, responsible for cleaving signal sequences from secreted type I transmembrane proteins. In at least one case, proteolysis by the ER-resident signal peptidase leads to the generation of presentable peptides. Peptides derived from ER signal sequences have been isolated from HLA-A2 class I molecules (Henderson et al. 1992, Wei et al. 1992). The signal sequence targets the protein into the ER and the signal peptidase enzyme cleaves off the signal peptide, which is free to bind class I. Binding of these peptides is independent of the expression of TAP peptide transporters confirming that these peptides are generated within the lumen of the ER itself.

Recently, it has been confirmed that processing of peptides can occur at both amino and carboxyl termini (Snyder et al. 1994, Roelse et al. 1994, Elliot et al. 1995). Experiments using large vaccinia virus recombinants (150 amino acids long) containing influenza nucleoprotein epitopes and ER signal sequence that facilitate the entry into the secretory pathway, showed that the large protein was processed in the ER to the appropriate 9-residue epitope recognized by influenza specific CTL cells (Elliot et al. 1995). It appears that ER processing can occur, but its efficiency and physiological importance are uncertain.

1.5 The proteasome

All proteins within a cell have a finite lifetime, therefore their degradation and synthesis must be closely coupled and regulated. One protein complex that has been implicated in the turnover of many different proteins is the proteasome (Rivett 1993, Goldberg and Rock 1992b). Proteasomes are found in the nucleus, cytoplasm and ER and they are the major non-lysosomal complex responsible for the breakdown of self and non-self intracellular proteins. The proteasome consists of about 14
different subunits and displays three types of catalytic activity that enable them to cleave proteins at the C-termini of hydrophobic, basic and acidic residues (Goldberg and Rock 1992b). The proteasome can be isolated as a 26S sedimenting particle or a smaller 20S form which is also part of the 26S complex. The 26S complex is involved in the ATP dependent degradation of cellular proteins that have been targeted for destruction by multiple additions of ubiquitin (Hershko and Ciechanover 1992).

During the last few years the proteasome has become the subject of extensive research in many fields. This is on account of their unique structural and catalytic properties and evidence of their essential role in several fundamental cellular processes. These include the general housekeeping role in extralysosomal protein degradation, cell cycle control and transcriptional regulation by the degradation of regulatory proteins conjugated with ubiquitin such as cyclins, transcription factors and oncoproteins.

Functional studies have demonstrated that the proteasome is not only involved in the protein turnover inside the cell, but is also involved in the production of most antigenic peptides presented to the immune system by the class I pathway. The first strong evidence that the proteasome was involved in antigen presentation came from the cloning of two genes within the MHC class II region, LMP2 and LMP7, whose expression is upregulated by γ-interferon and whose sequences are highly homologous to proteasome subunits. Since then, many studies have been performed to elucidate the properties of the proteasome in antigen presentation.

1.5.1 Characterization of the proteasome

The first evidence that proteolysis for class I presentation could occur in the cytoplasm came when cells treated with lysosomal inhibitor did not inhibit protein turnover in cells showing that the degradation of antigens for class I molecules was not blocked (Morrison et al. 1986). Previously, a non-lysosomal protease had been characterized in cytosolic extract of reticulocytes when abnormal haemoglobin was rapidly degraded by a process that was sensitive to ATP depletion. The proteolytic activity could be assayed in a cell free system and was activated by ATP (Etlinger and Goldberg 1977). Fractionation of homogenates from rat and mice showed that the protease had a molecular weight of around 550kDa (deMartino and Goldberg 1979, Rose et al. 1979).

A macromolecular particle was first identified using electron microscopy by Harris, in 1968, when studying protein components in human erythrocytes (Harris
1968). The first functional characterization of this macromolecule appeared in 1980, when Wilk and Orlowski described a complex of about 700 kDa with multicatalytic activity isolated from pituitary (Wilk and Orlowski 1980). At the same time Hase et al. described a tube-shaped protease complex of about 600 kDa consisting of four rings with an overall length of 15 nm and a diameter of 10 nm (Hase et al. 1980). These protein complexes turned out to be the same protease as the hollow cylinder complex which had previously been discovered by electron microscopy of erythrocytes.

Multicatalytic proteases were then isolated from various organisms, such as bovine lens (Ray et al. 1985, Blow et al. 1975), rat liver (Rivett 1992), muscle (Dahlmann et al. 1985), human lung (Zolfaghari et al. 1987b), and kidney (Zolfaghari et al. 1987a). The enzyme had similar properties to the ATP-dependent proteolytic systems isolated earlier (deMartino and Goldberg 1979, Rose et al. 1979). Proteolytic studies using rat liver enzyme showed that cleavage of peptide occur at five positions, different from those seen with other types of proteases (lysosomal cathepsins, calpain I, calpain II, papain, insulin protease, metalloproteases, serine proteases, chymotripsin, cathepsin C and elastase) (Rivett 1993a).

Today, this non-lysosomal protein complex is referred to as the 20S proteasome, multicatalytic proteinase (MCP), the 20S cylinder particle, proteasome and prosome. The MCP name was used due to its unusual ability to cleave peptide bonds carboxy-terminal to basic, hydrophobic or acid residues. These activities are referred to as trypsin-like, chymotrypsin-like and peptidyl-glutamyl hydrolysing activity, respectively. In addition to these three activities there is evidence that other peptidase activities exist in the proteasome (Rivett 1993a). The proteasome complex has been found both in the nucleus and cytosol of all eukaryotic cells (Arrigo et al. 1988, Tanaka et al. 1989, Orlowski et al. 1991). In each case these particles were shown to be 700kDa protein complexes, with sedimentation coefficient of 20S, composed of approximately 12-15 different polypeptide subunits in the molecular weight range from 19 to 36kDa.

It has been recently shown that the 20S core complex or 20S proteasome, is part of a large complex called 26S, the ATP-dependent form of proteasome complex. The 26S proteasome was discovered by the groups of Goldberg, Rechesteiner and Hershko, and it is the central protease of the ubiquitin dependent pathway of protein degradation (Waxman et al. 1987, Hough et al. 1987, Ganoth et al. 1988, Hershko et al. 1992). It is formed by a central core complex and by two 19S
cap complexes that associate with the two termini of the core. The 26S complex is
described in the section 1.5.5

1.5.2 Structure of the proteasome

Archaebacterium proteasomes isolated from *Thermoplasma acidophilum*
have been shown to have a similar subunit composition. They are composed of two
different polypeptides only, called α and β subunits, with molecular masses of 25.8
*Thermoplasma acidophilum* 20S complex is made of multiple copies of α and β
subunits showing a stoichiometry of α14β14 and a calculated total molecular mass of
673.4 kDa. The seven fold symmetry of the complex, revealed by electron microscopy,
has been recently corroborated by X-ray crystallography (Löwe et al. 1995).
Eukaryotic 20S proteasome complexes are made up formed of 14 different subunits
(Hendil et al. 1995). Each proteasome subunit can be grouped into two families
based on homology to either the α or β subunit of the *Thermoplasma acidophilum*
(Löwe et al. 1995).

The location of the α- and β-subunits in the *Thermoplasma* proteasome
determined by immunoelectron microscopy showed that α-subunits localize to the
outer rings, and the β-subunits to the inner rings (Grziwa et al. 1991). By using the
same technique it was shown that the proteasome is composed of two identical halves
related by C2 symmetry. This was also confirmed by immunoelectron microscopy of
eukaryotic proteasomes (Schauer et al. 1993, Koop et al. 1993).

Recombinant α- and β-subunits of *Thermoplasma* expressed in
*Escherichia coli* showed that α subunits seem to be necessary for 20S proteasome
assembly (Zwickl et al. 1992). Mature β-subunits formed irregular aggregates when
expressed separately, however, α-subunits formed seven-subunit rings that associated
in double rings. Recombinant proteins assembled normally and were fully functional
when both subunits were present (Zwickl et al. 1992).

The 20S proteasome appears either, as a ring-shape or rectangular-shape
when observed using electron microscopy. The ring-shaped appearance represents a
two-dimensional projection along the cylindrical axis (top-view). While the
rectangular appearance result from a projection perpendicular to the cylindrical axis
(side-view) (Koster et al. 1995). The eukaryotic and archaebacterial proteasomes
were almost identical in size and shape at about 1.5nm resolution.
The first evidence that the site of proteolysis was located in the interior of the 20S barrel was obtained by mixing nanogold-labelled insulin with the 20S proteasome. Because the size of the nanogold prevents the insulin from entering the proteasome and getting degraded, the nanogold accumulates at the center of the α-rings (Wenzel and Baumeister, 1995). The ability of the α-subunits to form seven-membered rings was suggested to be determined by their N-termini. This region appears to be very conserved between the Thermoplasm and eukaryotic subunits. Mutation in these regions leads to folded, stable but monomeric subunits (Zwickl et al. 1994).

The crystal structure of the 20S proteasome from Thermoplasma acidophilum has been elucidated by X-ray crystallographic analysis at 3.4 Å resolution (Lowe et al. 1995). Many of the proteasome characteristics previously described by electron microscopy have been confirmed by X-ray analysis. The proteasome complex consisted of 14 copies of two different subunits, α and β, forming a barrel-shaped structure of four stacked rings. The two inner rings consisted of seven β subunits each, and the two outer rings consisted of seven α subunits each. Three inner compartments were observed with a narrow channel controlling the access to these compartments (figure 1.7 D). The structure of the α and the β subunits were very similar, consisting of a core of two antiparallel β sheet flanked by α helices on both sides (figure 1.6).

The structure of the α subunit is a sandwich of two five stranded antiparallel β sheets with one side open, where the NH2-terminal α helix (OH) fills the cleft (figure 1.6). The opposite side of the β sandwich is closed by four hairpin loops connecting β strands. The β sandwich is flanked by three α helices on top and by two α helices at the bottom.

The structure of the α and β subunits were very similar, and the Cα atoms of the two subunits could be superimposed. The structure of the β subunit was slightly different in the NH2-terminal part compared to the α subunit, it lacks 35 amino acid residues, which provides access to the interior of the β sandwich (figure 1.6). The NH2-terminal region of the α subunits has an extra α helix that may be critical for the interaction with the regulatory complexes such as the 19S complex.
Figure 1.6: Crystal structure of the 20S proteasome α- and β-subunits from Archaebacteria *T. acidophilum*.

(A) Ribbon diagram of the α-subunit; α helices are labelled by H, β strands by S. The NH2 terminal helix H0 represents the front side of the hole proteasome. The identically oriented subunit within the complex is indicated (shown as a yellow sphere) at the top left corner; each subunit is represented by a sphere around its center of gravity.

(B) Ribbon diagram of the β-subunit in the same orientation. (Photographs from Löwe *et al.* 1995)
Figure 1.7: Crystal structure of the 20S proteasome from Archaeabacteria *T. acidophilum*.

(A) Cα drawing of two α(red) and two β subunits (blue) showing the major contacts between the subunits.

(B) Top view of the 20S proteasome showing Cα atoms only.

(C) Side view of the 20S proteasome. The overall dimension are 148Å in length, 113Å maximum diameter and 75Å minimum diameter.

(D) View of the proteasome cut open along the sevenfold axis; the three compartments and the gates are clearly visible. Dimension: entrance to the channel: 13Å, bottlenecks in the center: 22Å, left and right cavities: 50Å and length 40Å, central cavity: 53Å and length 38Å.

(Photographs from Löwe *et al.* 1995)
The overall dimensions of the 20S proteasome showed an elongated cylinder with a penetrating channel, large cavities and narrow constrictions. The overall dimensions of the complex are 148 Å in length and 113 Å in diameter (figure 1.7). The maximum diameter of the cavity is 53 Å. The central channel has three large cavities. Two of them located at the interface between the α and β rings, the third is formed by the β rings in the center of the molecule. The access to these compartments is controlled by four narrow gates.

1.5.3 Proteasome catalytic activities and active sites

Most of the investigation on the catalytic action of the proteasome has been performed using model oligopeptide substrates and protease inhibitors. The three best characterized activities are responsible for cleavage on the carboxyl side of basic (usually Arg), hydrophobic (Leu, Tyr, Phe) and acidic (Glu) residues. Referred to as trypsin-like, chymotrypsin-like and peptidylglutamyl-peptide bond hydrolase activities, respectively, these activities have been widely used to characterize the activities of the proteasomes isolated from different sources (reviewed in Rivett 1993a).

Evidence that the trypsin-like, chymotrypsin-like, and peptidylglutamyl peptide-hydrolyzing activities was expressed by distinct active sites was first provided by experiments showing that cleavage after basic residues could be inhibited by leupeptin, and cleavage after hydrophobic residues could be inhibited by peptidyl aldehyde Cbz-Gly-Gly-Leu-CHO (Wilk et al. 1980). Similar evidence for the three activities was found by treating proteasomes with a serine inhibitor called 3,4-dichloroisocoumarin (DCI) (Orlowski et al. 1989, Harper et al. 1985). Further studies with inhibitors, using rat liver and bovine pituitary proteasomes, revealed that there are at least two other distinct peptidase activities. Mammalian proteasomes can also catalyse cleavage on the carboxyl side of branched chain amino acid residues (such as Val, Leu and Ile), and between small neutral amino acids (such as Ala or Gly) (Orlowski et al. 1993, Djaballah et al. 1992, Savory et al. 1993).

Recently, the catalytic mechanism of the 20S proteasome from the archaeabacterium Thermoplasma acidophilum has been analyzed by site-directed mutagenesis of the beta subunit and by inhibitor studies (Seemüller et al. 1995b). Deletion of the amino-terminal threonine or its mutation to alanine led to inactivation of the enzyme.
Similarly, analysis of the crystal structure of proteasome inhibitors complexes have shown that nucleophilic attack is mediated by the amino-terminal threonine of processed β subunits (Löwe et al. 1995). Proteasome crystals formed in the presence of calpain inhibitor I showed difference in electron density in a tripeptide at the N-terminus. The tripeptide is situated at the terminus of the β subunit in a hydrophobic cleft inside the central cavity of the complex (figure 1.8). It was proposed that the terminal amino group and the tripeptide form a catalytic tetrad. All the β subunits in archaeobacterial proteasomes are identical and probably active, however analysis of human proteasomes showed that not all β subunits are proteolytic active, only six of the ten β subunits have a Thr1 residue, called LMP2, LMP7, MB1, Delta, Z and MECL1.

Further proteasome inhibition studies in mammalian cells using lactacystin, a streptomyces metabolite that inhibits cell cycle progression, identified the 20S proteasome as its specific cellular target (Fenteany et al. 1995). It showed that lactacystin covalently modified the Thr1 residue of one specific proteasome subunit, called MB1. Three distinct peptidase activities of this enzyme complex (trypsin-like, chymotrypsin-like, and peptidylglutamyl-peptide hydrolyzing activities) were inhibited by lactacystin, the first two irreversibly.

It was also shown that proteolysis requires a second conserved residue that carries a nucleophilic group (Seemüller et al. 1995b). Studies performed in Thermoplasma acidophilum proposed two possible residues: Lys33 and the amino-group of Thr1 (Seemüller et al. 1995b). Mutation of Lys33 showed that the proteasome became inactive but could fold and assemble correctly, demonstrating that Lys33 was essential for proteasome function. However, the results did not show whether Lys33 is directly involved in proton transfer or whether it served to polarise the Thr1 amino-group. On the basis of the crystal structure both roles appear to be possible, however the Thr1 amino group was favored as the primary acceptor-donor (Seemüller et al. 1995b). Two other conserved regions in the amino-acid sequence have been proposed to be involved in the active sites of the β-subunits, the pattern Gly128SerGly and Ser169GlyGly, are in direct proximity to the active-sites residues and are conserved in active subunits (Seemüller et al. 1995b).
Figure 1.8: Crystal structure of the 20S proteasome from Archaebacteria *T. acidophilum* showing the location of the peptidase active site.

(A) Fourteen fold difference in electron density showing seven inhibitors sitting around the central cavity formed by β-subunits. Distance from active sites within one β-ring are 28Å, 51Å, and 64Å.

(B) Stereo drawing of the accessible surface of the central cavity built by two rings of β-subunit cut open along the sevenfold axis. Superimposed is a space-filling model of the LLnL inhibitors. (Photographs from Löwe *et al.* 1995)
1.5.4 Ubiquitin system and the 26S proteasome

Since the beginning of the 1970s, several reports had suggested that the biochemical mechanism for intracellular degradation was ATP and ubiquitin dependent. It was shown that ATP stimulated the degradation of abnormal proteins in crude soluble extracts of reticulocytes (Etlinger and Goldberg 1977). Degradation of protein was also observed by its ligation to ubiquitin, a highly conserved 76-amino acid polypeptide that was essential for an ATP-dependent proteolytic degradation in reticulocytes (Ciachanover et al. 1980, Wilkinson et al. 1980).

Proteasomes were shown to be involved in ubiquitin-dependent degradation when a large enzymatic complex, similar to the proteasome, was isolated by its ability to degrade ubiquitinated proteins (Hough et al. 1986 and 1987). It was further demonstrated that the 20S proteasome is part of a large complex called 26S, the ATP-depend form of proteasome complex. The 26S form is involved in the ubiquitin dependent pathway of protein degradation. It is formed by a central 20S core complex and by two 19S cap complexes that associate with the two termini of the core (Waxman et al. 1987, Hough et al. 1987, Ganoth et al. 1988, Hershko et al. 1992) (figure 1.9). Negative-staining images of the purified 26S complex confirmed the two different forms of the complex, one consisting of the 20S barrel-shape proteasome with a rombohedral base which contains the components that allow recognition of ubiquitinated proteins, and a larger form in which the structures contained a cap at the end of the proteasome (Heuser 1989).

A number of recent investigations suggested that the 26S proteasome needs ATP for proteolysis (Goldberg and Rock 1992b). ATP interacts with multiple components of the large complex. The 26S particle has significant ATPase activity that can be activated by protein and peptide substrates. The 20S particle isolated from muscle also needs ATP for proteolysis and is probably required in its native state. The latent 20S and the 26S particle are temporarily activated by substrates, like ubiquitin, and linkage to ATP hydrolysis. The latent state explains how these structures can exist in the nucleus and the cytosol without damaging cell proteins.

Many substrates for the ubiquitin system have been identified in the past few years, but structural features for degradation signals have not been identified yet (reviewed in Hoshstrasser 1995). However, it was shown that the ATP-dependent degradation of ornithine decarboxylase was not ubiquitinated (Murakami et al. 1992). The degradation of ornithine decarboxylase was dependent on its association with another small protein, called antizyme the function of which is unknown.
Figure 1.9: Structural feature of the 20S and 26S proteasomes. Left, contour plot of the 26S proteasome derived from electron microscopy and image analysis. Top right, schematic cross-section of the 20S proteasome, showing the location of α-subunits (red), β-subunits (blue), and peptidase active sites (yellow). Bottom right, schematic view of subunit arrangement in the proteasome. (Photograph from Current Biology 1995, 5; 855).
Protein ubiquitination and deubiquitination are both mediated by a large family of enzymes (reviewed in Hochstrasser 1995). Ubiquitin is joined reversibly to other proteins via linkage between the α-carboxyl group of ubiquitin and lysine ε-amino groups of the acceptor protein (isopeptide bonds). The ubiquitin pathway is highly conserved in eukaryotes and it involves three enzymes, called E1, E2 and E3. In summary, in an energy dependent reaction, ubiquitin is first activated by an E1 enzyme, to which it becomes linked by a high energythiol ester bond. Ubiquitin then forms a thiolester bond with a second protein, an ubiquitin-conjugating enzyme E2. The E2, often with the additional factor E3, catalyses isopeptide bond formation between ubiquitin and the substrate. For substrates destined for proteolysis, additional ubiquitin molecules are added to the substrate by the same enzyme cascade to form a chain or chains of ubiquitin molecules in isopeptide linkage to one another. Degradation of multi-ubiquitinated proteins occurs on the 26S proteasome complex. A putative proteasome cycle is shown in figure 1.10.

In the ubiquitin pathway, the primary determinant of substrate specificity is the E2-E3 ubiquitin-protein ligase complex. There are at least 12 different types of E2 enzymes in yeast, suggesting that E2 proteins have different substrates specificities. For example, ubiquitination of the yeast MATα2 transcription factor needs at least four different E2 enzymes for degradation in vivo (Chen et al. 1993). It was also suggested that there may be different kinds of E3 proteins, that may not simply be adapter molecules that bring E2 proteins to their correct targets but may participate as intermediates in ubiquitin transfer reactions (Hochstrasser 1995).

Ubiquitin chains that assemble on various proteins are highly dynamic, with rapid addition and removal of ubiquitin units. The enzyme responsible for removing ubiquitin from substrates, known as ubiquitin carboxyl-terminal hydrolases, ubiquitin isopeptidases, or deubiquitinating enzymes, fall into two distinct families: a set of small proteins that preferentially cleave ubiquitin from small molecules, such as peptides, lysine or nucleophile glutathione, and a group of larger proteins that can generally cleave ubiquitin from a range of protein substrates in vitro (Mayer et al. 1989). This large family of enzyme is called Ubp class (Ubiquitin-specific processing protease) and all members contain several short conserved sequences that probably form the active site (Baker et al. 1992, Papa et al. 1993).

The Ubp family is large, with at least fifteen enzymes in yeast. It was suggested that changing the rate of ubiquitin removal from a substrate will alter the multi-ubiquitinated intermediate being recognized by the 26S proteasome complex, and the high number of deubiquitination enzymes make possible that specific protein
Figure 1.10: The putative ubiquitin-proteasome cycle. The proteolytic substrate is shown in red, ubiquitin in green. The peptidase sites are shown as blue scissors. The 26S is shown with a single 19S particle; the 19S particle is shown without the protein mass that occupies the mouth of the wedge. Degradation of the substrate to oligopeptides is accompanied by the regeneration of free ubiquitin molecules. Ubiquitin is activated at its carboxyl terminus by an E1 enzyme, then passed to an E2 enzyme, and finally to an E3 enzyme (ubiquitin-protein ligase). All three enzymes form thiolester bonds to ubiquitin, although in some cases the E2 enzyme might directly donate ubiquitin to target proteins. The proteolytic substrate is linked to ubiquitin by an isopeptide bond. Multiple ubiquitin groups are added to the substrate sequentially to form a multiubiquitin chain. (Photograph from Current Biology 1995, 5; 857).
turnovers rates can be differentially regulated by these enzymes. Deubiquitination enzymes appear to play a central regulatory role in the growth and development of eukaryotic organisms (Horchstrasser 1995).

### 1.5.5 Regulatory proteins of the proteasome

The proteasome regulatory proteins can be classified into three groups (Reviewed in Tanaka 1995, Dubiel et al. 1995); 1) Activators are proteins that can directly increase one or more of the various proteasome activities. 2) Inhibitors are proteins that can directly decrease one or more of the proteasome activities. 3) Modulators are proteins that do not have direct effects on the proteasome activities, but affect the action of other regulatory proteins. A large number of proteins involved in proteasome regulation have been identified, but a more detailed biochemical and molecular mechanism need to be defined. The better defined proteasome regulators are two proteasome activators called, PA28 and PA700, which have similar general characteristics (Ma et al. 1992, 1994).

The PA28 was identified as an activator of the proteasome peptidase activities (Dubiel et al. 1992b, Chu-Ping et al. 1992). It stimulates at least three hydrolytic activities by using small synthetic peptides as substrates. PA28 was shown to increase up to 500-fold the maximal velocity of hydrolysis of some substrates, and the concentration of substrate required for a half-maximal reaction velocity were decreased (Ma et al. 1992).

The protein has been found in a variety of higher eukaryotic species and tissues in a similar distribution to the proteasome (Ma et al. 1992, Dubiel et al. 1992b). It consists of two subunits of predicted molecular weight of 28.7 kDa for PA28a and 27.1 kDa for PA28β subunit which are both strongly inducible by γ-interferon (Ahn et al. 1995). The PA28 complex has a molecular weight of 200kDa and it consists of 6 to 7 PA28α and PA28β subunits arranged as a ring particle with the size of the ring similar to that of the terminal rings of the proteasome (Gray et al. 1994). The PA28 complex was shown to associate with the terminal α-rings of the proteasome (Gray et al. 1994).

Although PA28 could be detected in many types of tissues it was active only in some of them (Ma et al. 1993). This was due to the presence of a proteolytic enzyme, called carboxypeptidase which inactivates the PA28 by cleavage of the C-terminal residues. The inactive and native forms of PA28 were compared by interactions with the proteasome, and the inactive PA28 form fails to bind proteasome
indicating that the carboxyterminus contains a critical domain for the binding of PA28 to the proteasome (Ma et al. 1993).

Binding of PA28 to the 20S proteasome was shown to alter the quality and quantity of cleavage products generated from 25 mer polypeptides in vitro (Groetttrup et al. 1995). These results suggest that the 11S regulator increases the spectrum of peptides which can be generated in antigen presenting cells. However, the mechanism by which PA28 binds affects the cleavage process by the 20S proteasome remains to be elucidated.

The work on PA28 led to the discovery of PA700 (Ma et al. 1994). The PA700 appears to be similar to 19S cap proteins, and it is a 700kDa protein complex with at least 16 subunits that varies from 20 to 112 kDa in molecular weight. It activates the proteasome peptidase activities, but does not activate the degradation of non-ubiquitinated proteins, such as casein, lysozyme or albumin. PA700 has a similar effect to PA28 on proteasome activity. In contrast, PA700 requires hydrolysis of ATP to bind the proteasome during incubation of both proteins (DeMartino and Slaughter 1993).

The role of ATP in proteasome activation by PA700 is not well understood, and detailed information on the 16 subunits of PA700 is required to further understand the function of the complex. Seven of these subunits have been characterized. Three subunits have no homology with any previously characterized protein. One is similar to the tumour antigen P91A (Lurquin et al. 1989). Three others are homologous to one another, and to previously described protein members of a large family that contains an ATPase module (called AAA family) (Reviewed in Confalonieri and Duguet 1995). The AAA family members have a variety of cellular functions, such as regulation of transcription, cell cycle, and vesicle formation. The relationship of such functions to the regulation of the proteasome is unclear.

There is sufficient data, from purified proteasomes, which suggested that the 26S proteasome contains the PA700 complex as its principal component (Rechesteiner et at. 1993, Goldberg et al. 1992a). In addition many purified PA700 and 26S particles contain some of the same proteins as determined by sequence analysis, such as S4, p56, S6 and p48 members of the AAA family (Dubiel et al. 1992a, 1994). It was later found that a specific subunit of the PA700 particle, called S5, appears to mediate recognition of ubiquitin-protein conjugates (Deveraux et al. 1994). The S5 subunit binds polymers of ubiquitin even when they have not been linked to any substrate.
The 26S and the PA700 particles have similar regulatory properties and structure as defined by electron microscopy (Peters et al. 1993). PA700 binds to the terminal rings of the proteasome, suggesting that 26S particle is formed at least by the PA700 and the 20S proteasome. But it is not known if the 26S isolated in many laboratories contains only these two complexes or if there are additional components.

### 1.5.6 Proteasome location and expression

Early studies on proteasome location were performed in various cell types using electron microscopy as well as biochemical, and immunochemical techniques. It has been shown that proteasomes localize in the nucleus and in the cytoplasm of cells of vertebrates and non-vertebrates (Tanaka et al. 1990). Proteasomes have also been associated with intermediate filaments of the cytokeratin type and at the periphery of heterochromatin. A small fraction of proteasomes of certain cell types is located close to the endoplasmic reticulum, to specific membrane areas, to the nuclear membrane and to the nucleolus (Rivett et al. 1993b).

Much evidence suggested that there are clear differences in the relative amounts of proteasomes present in the nucleus and cytoplasm (Rivett et al. 1993b). In crustaceans muscle proteasomes are almost entirely cytoplasmic and they are apparently absent from nuclei. When proteasomes from leukaemia cells, renal cancer cells and T cells were analysed using antibodies against only one single subunit, they have been found to be predominantly nuclear (Shimbara et al. 1992).

More detailed localization studies using polyclonal antibodies and immunogold electron microscopic techniques have made it possible to quantitate proteasomes associated with different cellular compartments. In cultured human embryo lung, 50% was found associated in the nucleus, 3.5% associated with the endoplasmic reticulum, and the remainder in the cytoplasmic matrix. (Rivett et al. 1992). Other studies have shown that proteasomes might change from nucleus to cytoplasm. In lower eukaryotic organisms, such as the sea urchin, axlot, newt and Drosophila, it was observed that the distribution of proteasomes in the two subcellular compartments was interchangeable during oogenesis, embryogenesis and development (Tanaka et al. 1990). The localization of proteasomes was shown to vary within a single cell type when cells are induced to differentiate (Shimbara et al. 1992). Induced immature leukaemia T-cell (HL60) increased its nuclear fraction of proteasomes and a rapid movement of proteasomes between nucleus and cytoplasm was noticed (Shimbara et al. 1992).
Similarly, many studies have shown that proteasome location changes during the course of the cell cycle. Immunofluorescence studies with human embryo lung cells showed different patterns of distribution in different cells within the same culture (Rivett et al. 1992, Knecht et al. 1991, Palmer et al. 1994). Immunocytochemical analysis demonstrated that depending on the cell division cycle of the ascidian *Halocynthia roretzi* the proteasome undergoes changes in its subcellular distribution (Kawahara et al. 1992b). In interphase, proteasomes were localized in small clusters throughout the cytoplasm and nuclear matrix. In prophase, proteasomes disappear from the nucleoplasm and were observed in the perichromosomal area, suggesting that proteasomes can be rapidly transported from the cytoplasm to the nucleus. Similar results were also observed in an ovarian granuloma cell line (Amsterdam et al. 1993). The distribution of proteasomes during mitosis was shown to be similar to the behavior of cyclins. Subsequent experiments showed that the 26S proteasome was activated at two points of the cell cycle, prophase and metaphase, and it was involved in the degradation of cyclin and cyclin-dependent kinase inhibitor p27 (Kawahara et al. 1992a, Glotzer et al. 1991, Pagano et al. 1995).

It has been also shown that the proteasome can change its cell localization according to stages of embryonic development. Evidence that proteasome activity was under developmental control was obtained by showing that proteasome subunit composition varies considerably during fly development (Haass and Kloetzel 1989, Klein et al. 1990). Analysis of proteasomes from chick embryos at different developmental stages revealed that proteasome subunits are differentially synthesized at various stages during development (Pal et al. 1994). During early embryogenesis some proteasome subunits localized in the nucleus while at later stages they are also present in the cytoplasm. This data suggests that there are subcomponents of proteasomes which are synthesized as well as distributed in an independent manner during development, possibly reflecting subcomponent-specific multiple functions of these particles.

Sequence analysis of several proteasome subunits, from yeast, *Drosophila* and humans, suggested a probable consensus sequence for nuclear protein translocation (Tanaka et al. 1990). A unique cluster of amino acid residues containing a potential tyrosine phosphorylation site was present in the polypeptide subunits that have the nuclear localization signal sequence (NLS). This implies that the nuclear location signal sequence and the tyrosine phosphorylation site may play important roles in nuclear translocation of proteasomes, which may be responsible for their specific nuclear and cytoplasmic function (Tanaka et al. 1990). This observation
agrees with a recent study were tyrosine phosphorylation of the cytosolic form of the
RRC3 α-subunit rapidly triggers nuclear import, which is in turn quickly followed by
conversion to the major form associated with the nuclear scaffold (Benedict et al.
1995). However, it is not known if RRC3 is imported within the proteasome or by
itself. This α subunit, RRC3, is the predominant subunit expressed in hepatocytes,
and it appears to have both cytoplasmic and nuclear localization signal sequences
(Benedict et al. 1995).

1.5.7 Functions of the proteasome

The proteasome has long been known to be important in the degradation
of damaged or mutated cellular proteins, and more recently it has been shown to
degrade a variety of other substrates allowing insight into the different cellular
functions of the complex. These include, regulation of metabolic adaptation, cell
differentiation, cell-cycle control and a role in stress response (Goldberg 1992b,
1995). The rapid removal of these proteins is necessary for the control of cell growth
and metabolism. The continual turnover of cellular proteins by the ubiquitin-
proteasome pathway is also used by the immune system to screen for the presence of
abnormal intracellular proteins (Goldberg and Rock 1992b).

Early functional studies using yeast proteasome have shown, with one
exception, that individual chromosomal deletions of each of the known 14
proteasome are lethal. This suggests that deletion of one of the 20S proteasome genes
leads to disassembly of the 20S and 26S proteasome complex and complete loss of all
proteasome functions. Yeast mutants have been shown to affect the proteolysis,
specifically the chymotrypsin-like activity (PRE1 and PRE2 mutants) and the
peptidylglutamyl peptide hydrolyzing activity (PRE3 and PRE4 mutants)
cap genes CIM3, CIM5, YTA1 or YTA2 have also been shown to be lethal (Ghislain et
al. 1993, Schnall et al. 1994)

The eukariotic cell-division cycle is controlled by cyclin-dependent
protein kinases. The appearance and disappearance of particular active kinase
complexes during different phases of the cell cycle is regulated by synthesis and
degradation of specific cyclins. Many reports indicated that cell cycle regulatory
proteins are degraded by an energy- and ubiquitin- dependent pathway. Cyclin was
shown to be degraded by the ubiquitin pathway (Glotzer et al. 1990). This has also
been shown for c-Myc, c-Jun, c-Mos, tumour suppressor p53, Cyclin B, yeast Clb2,
p53 and p27 (Ichiara and Tanaka 1995). p27 is a mammalian cell cycle protein and it
functions as an inhibitor of cyclin-dependent kinases protein (Pagano et al. 1995). Mutations that prevent proteasome degradation of the Jun oncoprotein lead to tumours (Treir et al. 1994). It has been shown that papilloma virus induces tumours by stimulating the ubiquitin-mediated degradation of the tumour suppressor p53 gene (Scheffner et al. 1990).

The proteasome has also been involved in the control of yeast mating differentiation. The degradation of the regulatory protein MATα2 repressor was shown to be mediated by the proteasome complex (Richter-Ruoff et al. 1994). The MATα2 repressor is a transcriptional regulator protein which is required for mating type differentiation in yeast S. cerevisiae.

In addition the proteasome has an essential role in the proteolytic process required for activation of the transcription factor NF-κB (Palombella et al. 1994). NF-κB is a transcriptional regulator of a multitude of genes involved in the immune and inflammatory responses, such as chemotactic cytokines, hematopoietic growth factors, cell adhesion molecules, antibodies, class I molecules and cytokine receptors. It was shown that the p105 precursor of the p50 subunit of NF-κB is processed in vitro by an ATP-dependent process that requires proteasome and ubiquitin conjugation. The C-terminal region of p105 was rapidly degraded, leaving the N-terminal p50 domain. These results demonstrate that proteasomes are not only responsible for complete degradation of proteins, but are also required for activation of proteins by processing of inactive precursors.

The continual turnover of intracellular proteins by the proteasome is utilised by the immune system to screen for the presence of abnormal intracellular proteins. In this process, lymphocytes continually monitor small fragments of cell protein that are presented on class I molecules. The proteasome has been implicated in the class I presentation pathway by the identification of LMP2 and LMP7 proteasome subunits closely linked to other genes involved in class I presentation. Recently, many groups have suggested that the proteasome is the main protease that generates intracellular peptides for class I presentation. This is discussed in section 1.5.9.

1.5.8 Proteasome subunits

Many studies have shown that the structure of the proteasome is remarkably conserved from Thermoplasma to higher eukaryotes. Thermoplasma proteasome is composed of two different polypeptides, the α and the β subunits and it
is the simplest that has been described to date (Dahmann et al. 1989). It is often described as the ancestral proteasome, since attempts to isolate proteasome from many different species or strains of archaeabacterium, other than Thermoplasma, have failed. This suggests that proteasomes occur exclusively in the genus Thermoplasma (Pühler et al. 1994).

A sequence alignment of the α and β subunits of the Thermoplasma proteasome shows significant similarity (24%). This indicates that the two genes could have arisen from a common ancestor. A monomeric subunit would not necessarily have had a role in protein degradation.

The α-type subunits are distinguished from β-type subunits by a highly conserved N-terminal extension of approximately 35 residues. This region encompasses a α helix that is highly conserved between eukaryotic and Thermoplasma α sequences, but less conserved in eubacteria subunits. This helix lies on the outer surface of α rings and probably interacts with regulatory factors such as PA700 and PA28. Another conserved region in α-type subunits contains an invariant RPXG motif (where X represents a large hydrophobic residue). This sequence is a part of the loop that constricts the central pore through which unfolded polypeptides penetrate into the inner proteolytic compartment (Löwe et al. 1995, Wenzel and Baumeister 1995).

The β subunits have several residues that are conserved near the N-terminus, which are critical for the shape and nature of the active site. The most important is the Thr1 of mature subunits, which provides the catalytic nucleophile (Löwe et al. 1995, Seemüller et al. 1995b). The amino-terminal group of this residue is exposed by cleavage of pro-peptide in three of seven eukaryotic β-subunit types and in Thermoplasma during assembly of the complex. The four other eukaryotic β-subunit types do not contain this threonine and/or are not cleaved at this site and they might be proteolytically inactive (Seemüller et al. 1995b).

The first eubacterial 20S proteasome complex was recently characterized showing that the proteasome is an ancestral particle of universal distribution (Tamura et al. 1995). Purified preparations of 20S complexes from Rhodococcus sp. revealed the existence of four subunits clearly related to α-type or β-type proteasome subunits. There are two α-subunits (α1 and α2), and two β-subunits (β1 and β2), but it is not known how these subunits are incorporated into the complex. The complex efficiently degrades chymotryptic substrates.
A large number of 20S subunits (over 60) have been identified in various eukaryotic organisms such as human, rat, mice, chicken, Drosophila and Xenopus (Lupas et al. 1993). Sequences of the eukaryotic 20S proteasome subunits determined so far show strong similarity to each other without any detectable sequence similarity to other known proteins. Two eukaryotic subfamilies can be defined by their similarity to the α- and β-subunits of the Thermoplasma, and each family can be divided into seven branches by phylogenetic methods (figure 1.11). It is likely that each of these branches contributes one subunit to the seven membered α- and β-rings.

Human proteasome subunits have been extensively characterized and it seems that there are at least 7 α- (C2a, C2b, C3, C8, C9, Iota and Zeta) and 10 β-subunits (HN3, HC5, HC7, HC10, Delta, MB1, MECL1, Z, LMP2 and LMP7). Three β-type subunits, LMP2, LMP7 and MECL1, were shown to have their expression induced by γ-interferon.

1.5.9 Proteasomes in antigen processing

The first evidence that implicated the proteasome in antigen processing for class I presentation came from studies that identified two new genes in the class II region of the MHC, called LMP2 and LMP7 (Glynne et al. 1991, Kelly et al. 1991, Ortiz-Navarrete et al. 1991, Martinez et al. 1991). Analysis of DNA sequence showed that LMP genes were very similar to those of known proteasomal subunits. This evidence was supported by the fact that the LMP genes were closely associated to the TAP1 and TAP2 genes in the same region of the MHC.

The idea that the LMPs could be involved in antigen processing had been proposed when a complex of Low-Molecular-mass Proteins, "LMP", was precipitated with alloantisera to recombinant allotypes. The antisera were raised in congenic mouse strains against different H-2-encoded antigens. An anti-H-2d serum was used to immunoprecipitate proteins from H-2d cell lines. The proteins were separated by 2D electrophoresis. 45kDa H-2k and H-2d products were seen, as well as a group of 16 proteins of molecular weight 15-30 kDa with a wide range of isoelectric points (pI 4 to 8). These were called Low Molecular weight Proteins, or LMPs. The LMP determinants were expressed on monocyte and macrophage cell lines, as well as normal macrophages and spleen cells (Monaco and McDevitt 1982). The LMP

Early experiments were performed on mouse lines which were assayed for the presence of the LMP complex determinants by testing the ability of spleen cells to
Figure 1.11: Dendogram showing the relationships among known eukaryotic 20S proteasome protein sequences. The dendogram is from Hilt and Wolf, 1996. The 14 branches are numbered, abbreviations for species are: at, Arabidopsis thaliana; dd, Dictyostelium discoideum; dm, Drosophila melanogaster; gg, Gallus gallus; hs, Homo sapiens; mm, Mus musculus; rn, Rattus norvegicus; sc, Saccharomyces cerevisiae; sp, Schizosaccharomices pombe; xl, Xenopus laevis.
inhibit precipitations with the anti-H-2\textsuperscript{d} serum by preabsorption and by direct precipitation. Using mice recombinant within the H-2 region, the polymorphic determinant responsible for the precipitation of the LMPs was mapped between \textit{Pb} and \textit{Ob} genes (Monaco and McDevitt 1986).

Further circumstantial evidence for the involvement of the LMP genes in antigen processing was shown when the expression of LMP2 and LMP7 genes could be induced by \(\gamma\)-interferon, a property shared by other genes involved in the immune response, including the TAPs and HLA class I (Yang \textit{et al.} 1992).

However, evidence that the LMP's were not necessary for processing of antigens arose when mutant cells lacking the LMP genes expressed class I proteins at the cell surface at almost normal levels and could present intracellular antigens (Arnold \textit{et al.} 1992, Momburg \textit{et al.} 1992). .174 cells lacking TAPs and the LMPs could restore the wild type phenotype by transfection of TAP1 and TAP2 genes alone (Momburg \textit{et al.} 1992, Arnold \textit{et al.} 1992). More recently, the kinetics of presentation were shown to be similar for LMP-expressing and non-expressing cells (Yewldell \textit{et al.} 1994). These findings, cast further doubts on the role of LMP2 and LMP7 genes in processing, indicating that the LMP subunits were not essential for antigen presentation.

Several groups have now suggested that the LMP genes are not necessary for antigen presentation, but they play a specialized role by modifying proteasomes to produce peptides more suitable for class I binding. Three groups, Gaczynska \textit{et al.}, Driscoll \textit{et al.} and Boes \textit{et al.}, have analysed peptide hydrolysis by proteasomes from 721, .45 , and .174 cell lines, or cells treated with \(\gamma\)-interferon (Gaczynska \textit{et al.} 1993, Driscoll \textit{et al.} 1993, Boes \textit{et al.} 1994). Gaczynska and Driscoll showed that .174 cells have decreased rates of cleavage after hydrophobic and basic residues. They also showed the opposite effect on \(\gamma\)-interferon treated cells, which affected the endopeptidase activity of the enzyme by increasing the capacity to cleave small hydrophobic and basic residues, favoring the production of the type of peptides with higher affinity to bind class I. However, only Gaczynska showed a decreased rate of cleavage after acidic residues. Boes results were also controversial, showing that \(\gamma\)-interferon only decreases the capacity to cleave small hydrophobic peptides.

It was suggested that the phenotype of the LMP-deficient cell line, .174, might reflect the expression of alternative proteasome subunits, probably highly homologous to the LMPs. In this case, \(\gamma\)-interferon treated cells would up-regulate
the LMP subunits and down-regulate the LMP-related ones. Moreover, the .174 mutant cells would have the LMP-related subunits substituted within the proteasome complex.

The differences observed in proteasome cleavage under \(\gamma\)-interferon induction could be also explained by further discoveries. Two other genes were shown to be up-regulated by \(\gamma\)-interferon, PA28\(\alpha\) and \(\beta\). PA28 genes are part of proteasome 11 S regulator, and are potent activators of the proteasome (Realini et al. 1994, Mott et al. 1994). The binding of the 11S regulator to any proteasome preparation was shown to markedly change both quality and quantity of peptides produced (Groettrup et al. 1995). Thus, it is reasonable to suggest that \(\gamma\)-interferon induction of LMPs and PA28 serve to optimize antigen processing in proteasome-mediated protein degradation. In addition, the recently cloned MECL1 \(\beta\) subunit was also shown to be induced by this cytokine (Groettrup et al. 1996). Preliminary studies indicate that MECL1 has a different pattern of expression in .174 compared to wild type cell lines (I. Correa, personal communication). As described in section 1.5.3, MECL1 is an active \(\beta\) subunit and its differential expression, according to cell line, could explain some of the controversial findings in the various peptide cleavage experiments.

Further experiments with LMP2 and LMP7 mutant mice have shown that these \(\beta\) subunits influence antigen presentation (Van Kaer et al. 1994, Fehling et al. 1994). The altered peptidase activity in the LMP2 mutant resembled those observed in the .174 cell line. Proteasomes isolated from mutant spleen and liver degraded hydrophobic substrates and basic substrates at significantly lower levels, whereas degradation of the acidic substrates was almost two-fold higher than those from wild type mice. In addition, these mutant mice showed reduced capacity to stimulate specific T cells and reduced levels of CD8+ lymphocytes. Interestingly, the loss of LMP2 was associated with differences in the expression of other proteasome subunits, not only up regulation of Delta, indicating that the absence of LMP2 affects the overall composition of the proteasome. LMP7 mutant mice showed inefficient presentation of HY peptide, but in contrasts to the LMP2 mutant mice, it showed reduced class I expression at cell surface. Therefore, these data strongly implicate the LMPs and the proteasome in antigen presentation.

The observation that the absence of LMP2 in the mutant mice affects the overall composition of the proteasome raises interesting questions about the function of such subunits variability. In fact, this is most probably a way to increase the ability of the proteasome to generate a large diversity of peptides. It is possible that some
proteasome subunits have specific functions according to cell type, embryonic developmental stage, and cell cycle stage.

In order to investigate the specific roles of LMP2 and LMP7 subunits in the generation of peptides, Gaczynska et al. and Kuckelkorn et al. transfected HeLa, .174 and T2 cells with either one or both LMP genes and tested the ability of the transfected proteasomes to cleave fluorogenic peptides (Gaczynska et al. 1994, Kuckelkorn et al. 1995). Both groups used this approach, rather than γ-interferon induction, to avoid the influences that this cytokine has in other proteasome components or other unknown genes involved in processing of proteins.

Gaczynska's experiments showed that LMP7-transfected HeLa and .174 cells increased their capacity to cleave peptides after hydrophobic and basic residues without affecting the hydrolysis after acidic residues. These changes were also dependent on the amount of LMP7 incorporated into proteasomes. Transfections with LMP2 reduced cleavage of peptides after acid residues, increased hydrolysis after basic residues and did not affected hydrolysis after hydrophobic residues. This suggests that the expression of LMP2 and LMP7 by transfection mimics the proteasome activity observed in cells treated with γ-interferon.

Kuckelkorn's experiments in T2 transfected LMP2 and/or LMP7 cells showed different results from those of Gaczynska. In Kuckelkorn's work neither LMP2 nor LMP7 transfected cells showed significant changes in hydrolysis after hydrophobic residues. However, double transfectants showed a decreased hydrolysis against hydrophobic substrates, revealing enhanced positive cooperation between the proteasome subunits upon LMP2 and LMP7 incorporation. This effect was not observed when assaying hydrolysis after basic residues, whereas moderate increase of hydrolysis was detectable in the presence of LMP2 or LMP7 alone. Most interesting was the analysis of the hydrolysis against acidic substrates, each subunits induced opposite changes when expressed individually. LMP2-T2 proteasomes showed reduced ability to cleave acidic substrates, whereas LMP7-T2 increased activity to hydrolyse acidic substrates. Nevertheless, both studies indicated that LMP2 and LMP7 subunits induced differences in the quality of peptides generated compared to the untransfected LMP2 and LMP7 cells.

The work of Kuckelkorn and Gaczynska demonstrate that incorporation of either one of the LMP subunits cannot be easily correlated with activation and inactivation of a single type of peptide hydrolytic activity. Depending of the cell type and physiological situation, LMP2 and LMP7 subunits affect the incorporation and
the stoichiometry of at least one other proteasome subunit. It appears that the organization of the 20S proteasome subunits and its basic functional properties are controlled at the assembly level by affinities of the different α and β type proteasome subunit for each other. Incorporation of either one of the LMP subunits cannot be easily correlated with the activation or inactivation of a single type of peptide hydrolytic activity. The differential effects in hydrolysis of fluorogenic peptides reported from many groups are not entirely clear. Moreover, there are difficulties encountered when analyzing different cell types, different preparative protocols, and using artificial synthetic peptide substrates to measure different peptidase activities.

The importance of LMP2 and LMP7 was subsequently confirmed by two other groups (Sibille et al. 1995, Cerundolo et al. 1994). Both groups demonstrated that LMP2 and LMP7 can directly influence both MHC class I-restricted antigen presentation and class I surface expression, by identifying several viral epitopes that required LMP2 and/or LMP7 genes for their presentation.

To understand the function of the proteasome and its role in antigen presentation, Rock et al. identified potent reagents that can block the activities of purified 20S and 26S proteasomes and reduced proteolysis in intact cells (Rock et al. 1994). The class of compounds identified as proteasome inhibitors are peptide aldehydes that reversibly inhibited the chymotryptic and peptidylglutamyl activities of purified 20S and 26S proteasomes, and also blocked the degradation of casein and ubiquitin conjugated lysozyme by the 26S complex. Moreover, when ovalbumin was introduced into the cytosol these reagents blocked the class I presentation of an OVA peptide. However, these reagents did not reduce class I expression if cells were injected with OVA peptide. Thus, the inhibitors act on the generation of antigenic peptides without affecting the capacity of the cell to transport peptide into the ER. In addition, cells treated with inhibitors showed a marked reduction in the generation of class I heterodimers, indicating that peptides generated by the proteasome are the main source of peptide supply for class I presentation. Similar observations from Harding et al. using novel dipeptide aldehydes confirmed these findings (Harding et al. 1995).

Fenteany et al. showed that the most selective proteasome inhibitor presently known is Lactacystin which binds primarily to the MB1 subunit leading to inhibition of multiple peptidase activities of the proteasome (Fenteany et al. 1995). As the inhibition of the chymotrypsin-like and trypsin-like activities appear to be irreversible, the covalent modification of the MB1 subunit may be responsible for the inhibition of these activities. This dual inhibition could also be due to formation of
the lactacystin-protein adduct at one side, causing a change in conformation that is transmitted to the different types of active sites. It is also possible that MB1 may be a component of both chymotrypsin-like and trypsin like active sites. This finding implies that the MB1 subunit may play an important catalytic role in the proteasome.

In conclusion, additional studies performed by Grant et al. and Yang et al., have helped to clarify the role of proteasome in antigen degradation. Grant et al. demonstrated that the rate of antigen degradation by the ubiquitin-proteasome pathway influences class I presentation (Grant et al. 1995). When β-galactosidase was modified by a destabilizing amino-terminal residue the rate of class I presentation was enhanced. This enhanced presentation was then inhibited by blocking ubiquitin sites on the protein and by inhibition of the proteasome with peptide aldehydes. The rate of degradation of β-galactosidase was shown to require ATP and ubiquitin and was blocked by the inhibitors. The results showed that the rate of degradation in extracts correlated with their rate of class I presentation in vivo. This provided further evidence for a critical role of ubiquitin and the 26S proteasome in generating MHC class I peptides.

Yang et al., by analysing in vivo assembly of the proteasomal complexes, showed that 20S proteasome complexes are irreversibly assembled via 15S intermediates containing unprocessed β-type subunits. The 20S proteasome further associates reversibly with proteasome activators PA28 or ATPase complexes to form the 26S proteasome complex. The 26S complexes were predominant in the cytoplasm and a significant portion of 20S was associated with the ER. This suggests that depending upon their associated regulatory components, 26S and 20S-PA28 proteasomal complexes serve different housekeeping functions within the cell.

It seems that the 26S complexes could be responsible for antigenic peptides generated from ubiquitinated antigens, whereas the 20S or 20S-PA28 complexes could be responsible for digesting partially unfolded antigens or newly synthesized abnormal polypeptides. Antigen degradation could also occur in successive steps where large proteins are degraded by the 26S and degradation intermediates further processed by the 20S or 20S-PA28 complexes. The association of the 20S complexes with the ER membrane indicates that a final trimming of peptides to the size required by class I presentation is performed by the 20S-PA28 complexes closely located to the TAP proteins in the ER.

In conclusion the proteasome is the major protease involved in the generation of peptides for antigen presentation. During the evolution of the immune
system the ubiquitin system, the proteasome, and its activators appear to have adapted to enhance the ability to generate peptides for class I presentation. It seems that organisms are able to change subunit composition and activators to tune precisely this machinery in different tissues to meet physiological needs, as illustrated by up-regulation of LMPs and PA28 subunits during viral infection by γ-IFN stimulation.

1.6 Aims of this thesis

The aim of this work was to isolate new genes that encode proteasome subunits related to the MHC-encoded LMP2 and LMP7. We hypothesized that certain unknown proteins could replace the existing LMP2 and LMP7 proteasome subunits in the .174 mutant cell line, which has a deletion in the MHC region comprising the LMP genes. Evidence showed that the proteasome was the most likely candidate for the degradation of antigens for class I presentation. The cloning of highly homologous LMP subunits would explain the normal levels of class I expression in the .174-TAP transfected mutant cells. It would also explain, in degradation experiments, the differential sites of peptide cleavage between .174 and wild type cells including γ-interferon induced cells.

Chapter 3 of this thesis describes the cloning of MB1 and Delta subunits. The translated proteins are highly homologous to LMP7 and LMP2 respectively, indicating that these subunits are the best candidates for replacing the absent LMPs in the mutant .174 cell line. Chapter 4 describes the chromosome localization of MB1 and Delta outside of the MHC, and the possible explanation for the incorporation of the LMPs in this region. Chapters 5 describes the characterization of MB1 and Delta proteasome components. Chapter 6 analyses the quaternary structure and subunit topography of the 721 and .174 cell lines determined by immunoelectronmicroscopy of MB1, Delta, LMP2 and LMP7 polyclonal antibodies. Chapter 7 analysis MB1, Delta, LMP2 and LMP7 gene products and their differential expression, according to cell type, intracellular location and cell cycle stage.
Chapter 2

Materials and Methods

2.1 Cell lines, bacterial strains and DNA libraries.

2.1.1 Cell lines

The 53 human cell lines described were used to analyse the expression of various proteasome subunits.

1) **LCL 721**- Human B cell Epstein Barr virus transformed cell line called 721. Gift from Dr. Robert DeMars

2) **LCL 721.174**- Human mutant B cell line derived from LCL 721 called .174. Produced by γ-irradiation of LCL 721 cell line (DeMars et al. 1984). The mutant cell, .174, has a deletion spanning the entire HLA region in one copy of chromosome 6 and a large portion of the class II region missing on the other copy of chromosome 6. Gift from Dr. Robert DeMars.

3) **Raji**- Burkitt's lymphoma cell line.

4) **WJR076**- B cell line donated by Dr. Helene Teisserenc (Dr. Vincenzo Cerundolo laboratory, Molecular Immunology Group, Institute of molecular Medicine, Oxford). This cell is part of the International Histocompatibility B cell bank. This cell line does not express the LMP2 proteasome subunit due to a shorter LMP2 transcrip without the third exon.

The following B cell lines were kindly donated by the Tissue Antigen Laboratory at ICRF, provided as frozen aliquots containing approximately $10^6$ B cells. B cell were previously Epstein-Barr transformed and B cell lines were established from each donor.

5) **AWE1225-RH**- B cells from patient with Hodgkin's disease.

6) **JCH1228-RH**- B cells from patient with Hodgkin's disease.

7) **DUC1232-RH**- B cells from patient with Hodgkin's disease.

8) **DAF1435-RH**- B cells from patient with Hodgkin's disease.

9) **GIM1215-RH**- B cells from patient with Hodgkin's disease.

10) **CTH1214-RH**- B cells from patient with Hodgkin's disease.

11) **JOA1226-RH**- B cells from patient with Hodgkin's disease.
12) PHD1242-RH- B cells from patient with Hodgkins disease.
13) IVH2155-CF- B cells from patient with breast cancer
14) MPA-CF- B cells from patient with Myeloma
15) RAH1112-CF- B cells from patient with Myeloma and NHL
16) ERW-CF- B cells from patient with Hairy cell leukemia
17) MS1091- B cells from patient with Melanoma
18) RL1154-CF- B cells from patient with Waldenstrom's disease (macroglobulinemia)
19) VEH1609-CF- B cells from patient with Waldenstrom's disease (macroglobulinemia)
20) LED-NL- B cells from patient with Narcolepsy
21) LOH-CF- B cells from patient with Hodgkins disease

The following human cell lines were provided by the Cell Culture Laboratory at ICRF.

24) ZR75- Human breast carcinoma. From Dr. Malcolm Parker's laboratory at ICRF. Reference; Cancer Research (1978) 38: 8864, 4327-4339
27) 1847- Ovarian tumour. From epithelial cells of ovarian tumour. From Dr. Bridget Hill's laboratory at ICRF. Originally from Dr. R. T. Hamilton at NIH Bethesda Maryland.
28) EJ28- Bladder carcinoma. From a biopsy of primary bladder carcinoma. From Dr. Frank's laboratory at ICRF. Originally from Dr. G. R. Prout at Massachusetts Hospital Boston. Reference; Cancer Research (1986) 46; 3630-3636
29) 59M- Ovarian adenocarcinoma. From Dr. John Trowsdale's laboratory at ICRF. Reference Br.. J. Cancer (1989) 89; 527-534
30) **180D-** Human ovarian adenocarcinoma. From Dr. John Trowsdale's laboratory at ICRF. Reference Br.. J. Cancer (1989) 89; 527-534

31) **GER-** Human pancreatic adenocarcinoma. From Dr. Nick Lemones' laboratory at ICRF.

32) **HELF-** Human embryonic lung. From Dr. Ellen Solomon’s laboratory at ICRF.

33) **A818.7-** Pancreatic cancer. From Dr. Nick Lemones' laboratory at ICRF.

34) **GM7730A-** Biochemically mutant fibroblasts. From patient with fragile X-linked mental retardation Martin Bell syndrome- placental fibroblast culture. From cell Bank NIGM8 human genetic mutant cell repository.

35) **GM05401-** Biochemically mutant fibroblasts. From patient with Di George Syndrome. From NIGM8 human genetic mutant cell repository

36) **GM04286-** Biochemically mutant human lymphoblasts from 30 years old patient. Huntington’s disease Venezuelan family. From cell Bank NIGM8 human genetic mutant cell repository.

37) **GM02498-** Biochemically mutant human lymphoblasts from a 35 years old patient with Hodkin’s disease. From NIGM8 human genetic mutant cell repository.

38) **GM1416B-** Chromosomal aberrant lymphoblasts. From NIGM8 human genetic mutant cell repository.

39) **GM5131-** Chromosomal aberrant lymphoblasts. From NIGM8 human genetic mutant cell repository.

40) **GM4025B-** Chromosomal aberrant lymphoblasts. From NIGM8 human genetic mutant cell repository.

41) **ME180-** Epidermoid carcinoma of the cervix. From 26 years old caucasian female. The tumor was a highly invasive squamous cell carcinoma with irregular cell clusters and no significant keratinezation. From Dr. W. Bodmer's laboratory at ICRF. Reference J. Nat. Cancer Inst.(1980) 45: 107-122

42) **JATP-** Ovarian cancer. From Dr. Bridget Hill's laboratory at ICRF.

43) **HaPaCl-** Pancreatic cancer from Germany. From Dr. Nick Lemone's laboratory at ICRF.

44) **JAR-** Placenta choriocarcinoma- established from a from a trophoblastic tumour of the placenta of a 24 years old caucasian female. Given to cell production in 1979 by Dr. W. Bodmer. Reference In vitro (1971) 6:398

45) **5637-** Bladder carcinoma. Given to cell production by Dr. Harry White from ICRF in 1987.

46) **WI38-** Fetal lung- from normal 3 months gestation lung tissue of a caucasian female. Came to cell production from ATCC bank. Reference; Exp. Cell Research (1961) 25; 885

47) **BUTLER-** Fibroblasts from teratocarcinoma.

48) **A431-** Squamous epidermoid carcinoma.
49) HepG2- Epithelial hepatoma.
50) MCF7- Epithelial breast adenocarcinoma.
51) HT1080- Sarcoma.
52) Lovo- Epithelial colon adenocarcinoma.
53) J-6- T cell chronic lymphotic leukemia.
54) HeLa- Epitheloid carcinoma of the cervix.

2.1.1b Antibodies


2) anti-LMP2 antibody - polyclonal antibody raised against the C-terminal region of proteasome LMP2 subunit. From Dr. John Trowsdale laboratory at ICRF.

2.1.2 Bacterial strains and libraries

DH5α' *E. coli* strain (Gibco BR-Life technologies) was used for all the bacterial cell cultures.

The cDNA libraries referred in this study were constructed in the plasmid vector CDM8 (or its derivatives) according to the method of Seed (Seed, 1987). The JY B cell cDNA library was a gift from Dr. David Simmonds (ICRF, Oxford). The cDNA inserts were cloned into the CDM8 vector and electroporated into MC1061/P3 bacteria.

Human cosmid library, ICRF-DH1 was kindly provided from the Genome Analysis Laboratory at ICRF.

2.2 Bacterial preparation and manipulation

2.2.1 Bacterial cell culture

To ensure that each cell population descends from a single bacterium, colonies were isolated by streaking a L-agar plate and incubating them at 37°C until colonies become visible. Small liquid bacteria cultures were grown by inoculating 3 ml L-B medium with a single bacterial colony and culturing it overnight in a shaker at 37°C to saturation (density of 1 to 2 x10⁹ cells/ml). Large cultures were grown by diluting an overnight culture 100 fold and left overnight at 37°C under vigorous agitation to ensure proper aeration. Appropriate antibiotics were added to the cell culture.

L-B medium consists of 10g Bacto-tryptone, 5g Bacto-yeast extract, 10g NaCl, per litre, sterilised by autoclaving. L-agar plates consists of10g Bacto-tryptone, 5g Bacto-yeast extract, 10g NaCl, 15g agar, per litre, sterilised by autoclaving. Solid
medium was melted by microwaving and cooled to 50°C before antibiotics were added.

Stock solutions of Ampicillin and Kanamycin were prepared at 1000x stock solution using double distilled water, sterilised by filtration and used at final concentrations of 50µg/ml and 30µg/ml respectively. Tetracycline was dissolved in 70% ethanol for a 10x stock and used at final concentration of 12µg/ml. Stock solutions were stored at -20°C.

2.2.2 Preparation of competent E. coli cells for transformation

An overnight E. coli culture from a single colony was diluted 200 fold in L-B medium. The culture was grown at 37°C under vigorous agitation to an OD600nm of ~ 0.5 to 0.6. Cells were chilled in on ice-water bath for 15 min, transferred to a prechilled centrifuge bottle and spun at 4000 rpm in a Sorval GS3 rotor for 20 minutes at 4°C. The pellet was washed twice in the starting volume of ice-cold water and once in 1/10 of the initial volume of ice cold 10% glycerol. The pellet was then resuspended in an equal volume (to the pellet) of ice cold 10% glycerol and aliquoted into cold 0.5 ml tubes, frozen into dry ice and stored at -70°C.

2.2.3 Introduction of DNA into cells

Approximately 5-50 pg of plasmid DNA were mixed with 50µl of thawed competent cells and transferred into an ice cold electroporation cuvette. The mix of cells and DNA was pulsed at 2.5kV, 25µF, 200ohms. 1ml of L-B medium was added to the cuvette immediately after pulse, the sample transferred to a sterile culture tube and incubated at 37°C for 30 to 60 minutes. The bacteria were plated out on L-agar plates containing antibiotics.

2.3 DNA preparation and manipulation

2.3.1 Small scale preparation of plasmid or cosmid DNA

A single bacterial colony was inoculated in 5ml of L-B medium containing the appropriate antibiotic and grown by overnight shaking at 37°C. 1.5ml aliquot of each culture was transferred to an Eppendorf tube and pelleted for 1 minute by microcentrifugation. The supernatant was removed by aspiration with a drawn-out Pasteur pipette. The cell pellet was resuspended by vortexing in 100µl of cold GTE containing 4µg/ml lysozyme added just before use and left at room temperature for 5
minutes. The cells were lysed by adding 200µl of freshly prepared 0.2M NaOH, 1% SDS which was mixed by inverting the tube 5 to 6 times and leaving them for 5 minutes at 4°C. The chromosomal DNA was precipitated by adding 150ml of 5M KAc pH 4.8, mixed by inverting the tubes 5 to 6 times and left on ice for a further 5 minutes. Following a 15 minutes microcentrifugation at room temperature the clear supernatant containing the plasmid DNA was transferred to a new tube. An equal volume of phenol/chloroform/isoamylalcohol (PCIA) was added and mixed by vortexing. The tubes were microcentrifugated during 5 minutes and the upper phase was transferred to a fresh tube. The DNA was precipitated by adding two volumes of absolute ethanol and leaving for 1 hour at -20°C or 10 minutes on dry ice. The DNA was pelleted by 15 minutes microcentrifugation and the supernatant was discarded. The pellet was washed with 1ml of cold 70% ethanol, vortexed briefly and then microcentrifuged for a further 3 minutes at 4°C. The supernatant was removed by aspiration and the pellet desiccated under vacuum for 10 minutes. The pellet was resuspended in 50µl of TE pH8 containing 20mg/ml of RNase.

GTE

50mM glucose
10mM EDTA
25mM Tris/HCl pH8

KAc pH 4.8

60ml 5M KAc
11.5 ml glacial acetic acid
28.5 ml double distilled water (ddH2O)

PCIA

Phenol (melted at 65°C) was mixed with chloroform and iso-amyl alcohol in the ratio 25:24:1 and buffered by equilibrating once with 1 volume 50mM Tris base, twice with 1 volume 50mM Tris/HCl pH 8.0 and once with 1 volume TE pH8.0.

TE pH 8.0

10mM Tris/HCl pH8.0
1mM EDTA pH8.0

RNase

Pancreatic RNase A was dissolved in 15mM NaCl, 10 mM Tris/HCl pH7.5 at 10mg/ml, heated to 100°C for 15 minutes and cooled slowly to room temperature before storing at -20°C.
2.3.2 Large scale preparation of plasmid or cosmid DNA

A volume of 400 ml of L-B medium containing the required antibiotic was inoculated with 0.5 ml of an overnight culture from a single colony of *E. coli* carrying the desired plasmid or cosmid. Following overnight culture at 37°C the bacterial cells were harvested by spinning at 4,000 rpm in a Sorval GS3 rotor for 15 minutes in a 500ml centrifuge bottle at 4°C. Bacteria were resuspended in 10mls of 10mM EDTA or GTE containing 5mg/ml lysozyme and left at room temperature for 5 minutes. A volume of 20mls of 0.2mM NaOH, 0.1%SDS was added, mixed gently, and incubated for 5 minutes on ice, then 10ml of 5M KAc pH4.8 were added, mixed by inverting the tubes a few times, left on ice for a further 15 minutes and centrifuged for 15 minutes at 4,000 rpm in a Sorval GS3 rotor to precipitate cell DNA and bacterial debris. The supernatant was then filtered twice through gauze into a 50ml cylinder and the DNA was precipitated by adding 0.6 volumes of isopropanol at room temperature for 15 minutes. The DNA was recovered by 15 minutes of 5,000 rpm centrifugation at 4°C in a Sorval GS3 rotor. The pellet was resuspended in 2.5ml of TE pH8.0 buffer.

To purify the plasmid DNA, 4.4g of cesium chloride (CsCl) was dissolved in the solution of DNA and 0.4 ml of ethidium bromide (stock 10mg/ml) added. The density of the solution was checked to be around 1.55g/ml ± 0.05 g and centrifuged at 8000 rpm in a Sorval SS34 rotor for 5 minutes to precipitate bacterial proteins. Supernatant was then sealed in a Beckman Ti70.1 rotor and centrifugated overnight at 60,000 rpm. The plasmid band was collected by inserting a needle in the top of the tube and removing 0.5 to 0.7 ml of solution by suction into a 5ml syringe. An equal volume of butanol was added to extract the ethidium bromide in the solution. The contents were mixed and the phases allowed to separate. The top organic part containing the ethidium bromide was removed. Continuous extraction was performed in this way until the upper phase became colorless. The plasmid DNA was precipitated by addition of 2 volumes or TE buffer and 6 volumes of 100% ethanol followed by 15 minutes centrifugation at 10,000g at room temperature. The pellet was resuspended in 400µl of 70% ethanol, microcentrifuged at room temperature for 3 minutes, dried and resuspended in 400µl of TE. The concentration of DNA was determinated by absorbance at 260nm.
2.3.3 Restriction endonuclease digestion of DNA

Restriction endonuclease digests were performed according to standard protocols (Sambrook, Fritsch and Maniatis, 1989)

For plasmid or cosmid mapping, 100-500ng of DNA were digested in a reaction volume of 20-50µl. When specific fragments were required for subcloning or probe preparation, digests were scaled up accordingly. For preparation of genomic Southern blots, approximately 10µg of DNA were digested in a volume of 50-100µl with a 5-fold excess of enzyme.

2.3.4 Electrophoresis of DNA

Agarose gel electrophoresis was used for routine resolution, isolation or purification of plasmid DNA or cosmids. The gels were prepared with an agarose concentration appropriate for the size of DNA fragments to be separated in 1x TAE solution. The DNA samples were mixed with a 0.2 volumes of the appropriate loading buffer, loaded into the sample wells and the gel was run at a voltage and time period for optimal separation. To allow DNA visualisation, the gel was prepared with an ethidium bromide solution to a final concentration of 0.5µg/ml. For some experiments involving the isolation of DNA fragments, electrophoresis was performed using low melting point agarose.

10x TAE 48.4g/l Tris base
11.4ml/l glacial acetic acid
20mM EDTA

5x Loading buffer 50% glycerol
60mM EDTA
0.25% bromophenol blue
5x TAE or TBE

2.3.5 Labelling of DNA fragments using random hexamer priming

20-50 ng of purified DNA were diluted to a volume of 33µl of ddH20, boiled for 5 minutes and then cooled on ice. 10µl of OLB, 2µl of BSA, 3µl of [α-32p]dCTP (2-10 µCi) and 2µl of Klenow fragment enzyme (5units/µl, NLE) were added and the reaction incubated at 37°C for 60 minutes. Labelled DNA was separated from unincorporated label by passing the reaction mixture through a TE
equilibrated Sepharose G75 column poured in a 1ml syringe. Labelled probe was boiled for 5 minutes and then cooled on ice before being added to the hybridisation solution.

**OLB**

Mix solution A:B:C in a ratio 100:250:150

**Solution O**

1.25M Tris/HCl  
0.125M MgCl₂  
PpH 8.0

**Solution A**

1ml of solution O  
18μl β-mercaptoethanol  
5μl each 0.1M dATP, dTTP, dGTP (Pharmacia)

**Solution B**

2M HEPES adjusted to pH 6.6 with 4M NaOH

**Solution C**

Hexadeoxyribonucleotides (Pharmacia)  
resuspended at 90 OD₂₆₀ units/ml in 10mM Tris/HCl, 1mM EDTA.

### 2.3.6 Screening of cDNA libraries

The titre of the cDNA library was first determined by plating out several 100 fold serial dilutions (for example 10⁻², 10⁻³, 10⁻⁴, etc.) of the frozen stock library on L-agar plates containing ampicillin and tetracycline. After the overnight incubation at 37°C the number of colonies obtained in each plate was used to calculate the amount of library stock required to yield 250 000 colonies, the optimal number for hybridisation to a 20x20 cm membrane. Four 20x20 cm Hybond N⁺ membranes (Amersham) were lowered onto the surface of four L-agar 245x245 mm plates (Nunc). The appropriate amount of library stock diluted in a volume of 1ml was spotted onto the surface of each filter and spread evenly using a flamed glass spreader. These master plates were then incubated overnight at 37°C. The following day eight additional 245x245 mm L-agar plates (the replica plates) were poured and overlaid with 20x20 cm Hybond N⁺ membranes. More duplicates were prepared if necessary. To prepare duplicates of the first master plates, the membrane was lifted off the master with forceps and laid colony side up on three pieces of Whatman 3MM paper. The wetted membrane from the first replica plate was then overlaid on the master membrane, covered with three layers of Whatman paper between two perspex
2.3.6a Cosmid cloning

Genomic fragments of MB1 and Delta were obtained by screening the full human ICRF-DH1 cosmid library (kind gift from Genome Analysis Laboratory). Full length cDNA clone for MB1 and Delta were hybridized at high stringency (65°C and washed with 0.2xSSC). Three cosmids were isolated using the MB1 cDNA (called MB2-1, MB4-2 and MB5-2) as a hybridisation probe and three cosmids were isolated using the Delta probe (called C1, C2 and C3).

To confirm that cosmids containing MB1 and Delta gene were overlapping cosmids the two sets of cosmids were analysed by restriction enzyme digestion. Sets of cosmids showing similar banding patterns on EcoRI digestion, would indicate that they overlap each other.

Cosmids were also PCR tested for the presence of intronic and exonic DNA sequences. Four different sets of MB1 oligos and four different sets of Delta oligos, located 150 to 200 base pairs apart were used to confirmed the presence of introns and exons in both cosmids.
sheets. Colonies were transferred from master to replica membrane by applying pressure to the perspex sheets. An identical pattern of spots or pin marks was applied to both filters before they were pulled apart and the replica returned to its L-agar plate (colony side-up). The transfer process was then repeated for a second replica and the master membrane were returned to their L-agar plates.

Once duplicate replicas had been made from all four master plates, all twelve plates were incubated until the colonies had regrown. The master plates were then sealed with plastic film and stored at 4°C. The eight replica membranes were removed from the L-agar plates and processed by placing them (colony side up) sequentially on pads of Whatman paper soaked in denaturing solution for 10 minutes, neutralising solution (2x for 3 minutes). The membranes were rinsed by immersion in 2xSSC and the colonies were wiped off with wetted tissue paper. The membranes were dried and stored at room temperature.

<table>
<thead>
<tr>
<th>Denaturing solution</th>
<th>1.5M NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5M NaOH</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neutralising solution</th>
<th>1.5M NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5M Tris/HCl pH 7.2</td>
</tr>
<tr>
<td></td>
<td>0.001M EDTA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20xSSC</th>
<th>3M NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.3 M sodium citrate pH 7.0</td>
</tr>
</tbody>
</table>

2.3.7 Hybridisation of filters

Filters were placed in sandwich boxes containing hybridising solution pre-heated at 65°C and incubated at the same temperature for 1-6 hours. The appropriated volume of labelled probe was used to a final concentration of $5 \times 10^5$ cpm/ml of hybridisation solution. The labelled probe was boiled for 5 minutes and immediately added to fresh hybridisation mix and incubated with the filters overnight at 65°C for high stringency hybridisation or 55°C if low stringency was required. Filters were washed twice with 400ml of 6x SSC/1% SDS solution for 10 minutes at 65°C, followed by two 10 minutes washes with 0.2xSSC/1% SDS solution at 65°C for high stringency hybridisation. For low stringency hybridization the filters were washed twice with 6x SSC/1%SDS solution at 55°C for ten minutes each followed by two washes with 2xSSC/1%SDS at the same temperature for 10 minutes.
The filters were placed against Kodak XAR-5 film backed with an intensifying screen and left overnight at -70°C.

Colonies that were positive in both auto radiographs of the filters were used to identify the location of the positive signal on the master plate. Bacterial cells were removed from a region of about 1mm² around the signal and transferred to 1ml of L-Broth. Several dilutions of this stock were made and plated out on 140mm diameter L-agar plates and incubated overnight at 37°C. Plates containing well separated colonies were selected for the preparation of secondary colony lifts. Appropriately sized Hybond N+ membranes were overlaid on the colonies for approximately 1 minute. During this time the position of the filter was recorded by marking a pattern of matching ink dots on the membrane and underside of the plate. The membrane was removed with forceps and placed colony side up on sequential pads of Whatman paper soaked in denaturing solution and neutralising solution. The secondary plate was incubated at 37°C until the colonies had regrown and then stored at 4°C. The secondary membranes were hybridised with the original probe of interest to identify positive secondary colonies which were then individually picked into L-broth. Minipreps were prepared as described in section 2.2.1, digested with Xho I to excise the inserts and the digests were separated by agarose gel electrophoresis and transferred to membranes. The filters were then probed with the original probe.

Southern hybridisation solution

6x SSC
5x Denhardt's solution
0.5% w/v SDS
10% v/v dextran sulphate
40µg/ml denatured salmon sperm DNA

50x Denhardt's
5g polyvinilpyrroridone
5g BSA
5g Ficoll 400
to 500ml in ddH₂O

To prepare denatured salmon sperm DNA: 500mg salmon sperm DNA were cut into small pieces with scissors. The DNA was dissolved in 50 ml of ddH₂O by stirring overnight. The solution was sheared by three passes through a 19G needle and boiled for 10 minutes.
2.3.8 Southern blotting

Restriction enzyme digestion were carried out with a 5 fold excess of enzyme in the recommended buffer. DNA fragments were separated on 0.6-2% agarose gels in 1x TBE buffer. The gel was soaked in 0.2M NaOH, 0.6M NaCl for 30 minutes to denature the DNA. The gel was washed in destilled water and then neutralised with two incubations of 20 minutes in 0.5M Tris pH 7.6, 0.5M NaCl. The DNA was capillary blotted onto nitrocellulose filters (Hybond N+) in 20x SSC. The filters were briefly rinsed in 2xSSC before pre hybridisation. Hybridisation was as described in section 2.2.7.

2.3.9 Sequencing of double stranded plasmid DNA

Approximately 2μg of high quality plasmid DNA was mixed with 2μl 20mM EDTA, 2μl 2M NaOH and ddH2O to a final volume of 20 μl. The mixture was incubated for 5 minutes at room temperature and then neutralised with 3μl 3M Na Ac pH 5.2 and 7μl ddH2O. 75 μl of absolute ethanol was added and the denatured DNA precipitated at -70°C for 15 minutes. The DNA was pelleted by microcentrifugation at 4°C for 10 minutes, washed in 100μl 70% ethanol, repelleted and the supernatant removed with a drawn-out Pasteur pipette. The pellet was briefly dried under vacuum and resuspended in 7 μl of ddH2O, 2μl of Sequenase reaction buffer and 1μl (approximately 0.5pmol) of sequencing oligonucleotide primer. Sequencing was performed using the Sequenase kit from USB based on the chain termination method of Sanger (Sanger, Nicklen et al, 1977). The mixture was heated at 65°C for 2 minutes and then slowly cooled over 30-60 minutes to 30-35°C. The primer was extended by adding to this template/primer mix: 1μl 0.1 M DTT, 2μl 1x labelling mix, 0.5μl [α-35S] dATP and 2μl Sequenase enzyme diluted 1:7 in TE pH 8.0. After the reaction was incubated for 2-5 minutes at room temperature, 3.5μl was transferred to each four pre-warmed tubes containing 2.5 μl of the four dideoxy nucleotide mixes. These reactions were incubated at 37°C for 5 minutes and then stopped by addition of 4 μl stop solution.

<table>
<thead>
<tr>
<th>5x reaction buffer</th>
<th>200mM tris/HCl pH 7.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100mM MgCl2</td>
</tr>
<tr>
<td></td>
<td>250mM NaCl</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5x labelling mix</th>
<th>7.5 μM each of dGTP, dCTP and dTTP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>diluted to 1x with ddH2O</td>
</tr>
</tbody>
</table>
Dideoxy (dd) termination mixes:

- 80μl of dGTP, dATP, dCTP and dTTP
- 8μl of either ddGTP, ddATP, ddCTP or ddTTP
- 50 μM NaCl

Stop solution

- 95% formamide
- 20mM EDTA
- 0.05% bromophenol blue
- 0.05% xilene cyanol

### 2.3.10 Denaturing polyacrilamide gel electrophoresis

Sequencing gels were run using vertical electrophoresis apparatus. The glass plates were washed with scouring powder rinsed in dH2O, and washed again with ethanol. The large plate was siliconized, allowed to dry and rinsed again with water. Spacers of 0.4mm were placed between plates and the sides and the bottom were sealed with electrical tape. A volume of 40 ml polyacrylamide mix (Sequagel) was prepared and the gel poured using a 60ml plastic syringe. The flat side of the comb was inserted and the gel was left to polymerise. The gel was assembled in the tank and pre-warmed by electrophoresing in 1x TBE for 30 minutes. Excess urea was removed from the loading wells and 2.5μl of each sample loaded. Electrophoresis was performed at 40mA constant current and was terminated after 60-100 minutes depending on the sequence to be resolved.

Following the electrophoresis, the plates were separated and the gel fixed in 10% glacial acetic acid, 10% methanol for 15 minutes. The gel was transferred to Whatman 3MM paper, covered in Saran wrap and dried at 80°C for 30 minutes on Bio-Rad slab drier. Autoradiography was performed overnight using Kodak XAR-5 film at room temperature.

Sequagel mix

- 9.6mls concentrate
- 26.4mls diluent
- 4mls buffer
- 90μl 10% ammonium persulphate
- 70μl TEMED

10x TBE

- Tris base 108g/litre
- Boric acid 55g/litre
- EDTA 9.3g/litre
2.4 RNA preparation and manipulation

2.4.1 RNA preparation

Total RNA was isolated from approximately $10^8$ cells using TRIZOL reagent (Gibco life technologies). Cells were homogenised by adding 1ml of TRIZOL per $10^7$ cells in suspension, and per 10cm$^2$ monolayer cells. The cell lysate was passed several times through a pipette and incubated 5 minutes at room temperature to dissociate nucleoprotein complexes. 0.2ml of chloroform was added per 1ml of TRIZOL the tubes were incubated at room temperature 3 minutes and shaken vigorously by hand for 15 seconds. The samples were centrifuged at 12000 g for 15 minutes at 4°C and the aqueous phase containing the RNA was collected. The RNA was precipitated by adding 0.5ml of isopropyl alcohol per 1ml of TRIZOL used in the initial homogenisation. Samples were incubated for 10 minutes at room temperature, centrifugated at 4°C at 12000 g. The RNA pellet was washed once with 75% ethanol, adding 1ml of ethanol per 1 ml of TRIZOL. The pellet was vortexed and centrifuged at 7500 g for 5 minutes at 4°C. The pellet was air-dry and redissolved in RNase-free water by passing the solution a few times through a pipette tip, and incubating for 10 minutes at 55°C. The RNA concentration was assessed by absorbance: $OD_{260}=1$ equivalent to approximately 40μg RNA/ml.

2.4.3 Electrophoresis of RNA

The gels were prepared depending on the size of the mRNA, usually varying between 1 to 1.4% of agarose concentration for RNA smaller than 2Kb. For 100 ml of 1%gel, 1g of agar was microwave melted in ddH$_2$O-depc water, and cooled down to 60°C in a water-bath, 18ml of 37% formaldehyde and 10ml of 10x MOPS were added. The solution was immediately poured into the gel mold. A volume of 20μg of total RNA was mixed to a final volume of 20μl with 3.5μl of 37% formaldehyde, 2μl of 5x MOPS and 10μl of deionised formamide. Samples were mixed, briefly spun, and denatured by heating at 65°C for 15 minutes and chilled on ice, and 2μl of loading buffer was added prior loading the gel. The gel was run overnight at 20V in 1x MOPS buffer using a recycling buffer system in vertical apparatus.

The gel was rinsed in water, stained with 1μg/ml of ethidium bromide for approximately 30 minutes and destained in ddH$_2$O until the 28S and 18S ribosomal RNA bands were visible.
MOPS 10x
800ml depc treated ddH2O
41.8g MOPS pH 7 with NaOH
16.6ml 3M NaAc pH 7
20ml 0.5M EDTA pH 8

Loading Buffer
1mM EDTA pH8
0.25% Bromophenol blue
0.25% Xylene cyanol
50% glycerol

Deionised formamide:
Formamide is stirred for 1 hour with AG 501-X8 (Biorad), and then filtered twice through Whatmann N°1 paper.

2.4.2 Northern blotting

The gel was rinsed several times in ddH2O-depc (diethyl pyrocarbonate) and soaked for 45 minutes in 20x SSC. The RNA was capillary blotted overnight in 20x SSC onto a Hybond N nitrocellulose membrane. The wells and positions of ribosomal RNA were marked onto filters, followed by a rinsing in water and fixation by baking at 80°C for 2 hours under vacuum. Filters were washed with 5 x SSC prior to hybridisation.

2.4.4 Hybridisation of blots

Filters were prehybridised at 42°C in northern hybridisation mix for 1 hour. The labelled probe was boiled with 10μl of 10mg/ml salmon sperm DNA per ml of hybridisation solution for 5 minutes, and then added to the hybridisation mix. Hybridisation was performed at 42°C overnight and the filters were washed to 0.2x SSC, 0.1% SDS at room temperature. Autoradiography was at -70°C with an intensifying screen. Relative intensities of RNA bands were quantified by LKB ultrascan.

Northern hybridisation solution
25ml formamide
10ml 50% dextran sulphate
10 ml 5M NaCl
2.5 ml 20% SDS
2.5 ml ddH2O
2.5 Preparation of antisera, immunoprecipitations and western blotting

2.5.1 Production of polyclonal rabbit antisera

Peptides used for production of polyclonal rabbit antisera were synthesised by the Peptide Synthesis laboratory, ICRF. 15 amino acid sequences corresponding to the C-terminal predicted open reading frames of MB1 and Delta genes were used for synthesis. 6.25mg of peptide was conjugated to an equal weight of keyhole limpet haemocyanin. The peptide and the carrier protein were dissolved in 0.1M NaHCO3 at a concentration of 2 carrier mg/ml. A fresh vial of 25% ultra pure grade glutaraldehyde was thawed and added to the solution to a final concentration of 0.05%. The mixture was stirred at room temperature overnight. Glycine ethyl ester pH8, was added to 0.1M followed by a 30 minute room temperature incubation. The coupled peptide was precipitated by addition of 4-5 volumes ice cold acetone and incubation at -70°C for 30 minutes. The coupled protein was pelleted at 10000g for 10 minutes. The pellet was resuspended in 0.9% NaCl at 1mg/ml.

Rabbit immunisation schedules were carried out by staff of the ICRF animal unit, Clare Hall. Briefly, following a pre-bleed 500mg of peptide/carrier sample was injected subcutaneously in complete Freund's adjuvant. 21 days later a further 500mg of sample was injected in incomplete Freund's adjuvant. This was repeated every 28 days until bleedout (usually following the seventh immunisation). Test bleeds were taken 1 week after each injection.

2.5.2 Immunoprecipitation

1-2 x 10^7 cells were washed in PBS and preincubated at 37°C for 1 to 2 1/2 hours in methionine free RPMI medium, 10% FCS, glutamine. 100-200 µCi 35S methionine were added and incubated at 37°C for 1 hour. When 12 hour incubations were necessary they were performed by adding 10% of complete RPMI medium, 10% FCS and glutamine. The cells were washed twice in PBS and lysed by adding 1ml of lysis buffer followed by a 20 minutes incubation on ice. The nuclei were then pelleted and discarded by centrifugation at 12,000 rpm at 4°C for 10 minutes. In order to pre-clear immunoglobulin, 0.2ml of 10% washed Staphylococcus A was added to the supernatant and agitated overnight at 4°C. Staphylococcus A was removed by centrifugation and the appropriate antiserum or antibody added to the cleared lysate. In some cases, SDS was added to 2% and the lysate was heated to
95°C for 5 minutes to dissociate the proteasome prior to precipitation. Following incubation for 90 minutes at 4°C, BSA was added to a final concentration of 1% together with 100μl of 5% (w/v) protein A-sepharose in lysis buffer. Tubes were rotated at 4°C for 45 minutes and then washed four times in 1ml of washing buffer. Prior to the final wash, tubes were changed in order to eliminate any protein which had non-specifically bound to the plastic tube. The lysis buffer was removed and the bead-bound protein was eluted in 50μl of sample buffer. Tubes were then heated for 4 minutes at 95°C, centrifuged for 30 seconds and the supernatant loaded onto 10-15% denaturing SDS-polyacrylamide gels for electrophoresis.

The gel was fixed in 7% acetic acid, 40% methanol, treated for 15 minutes with Amplify (Amersham) and dried on 3MM paper. The gel was exposed to Kodak XAR-5 film for an appropriated length of time at either room temperature or -70°C with an intensifying screen.

**Lysis buffer**
- 0.5% v/v Nonidet-P40
- 0.5% Mega 9
- 150mM NaCl
- 5mM EDTA
- 50mM Tris pH 7.5
- 2mM PMSF
- 5mM iodoacetamide

**Wash buffer**
- 0.5% v/v Nonidet-P40
- 150mM NaCl
- 5mM EDTA
- 50mM Tris pH 7.5
- 2mM PMSF
- 5mM iodoacetamide

**Sample buffer**
- 4ml ddH2O
- 1ml 0.5M Tris/HCl pH 6.8
- 0.8ml glycerol
- 0.8ml 20% SDS
- 0.4ml 2-mercaptoethanol
- 0.4ml 0.05% bromophenol blue

80
2.5.3 Pulse chase analysis

5x 10^7 cells were washed in PBS and incubated in 2ml methionine free RPMI, 10% FCS for 1 hour. 0.5μCi of ^{35}S methionine was added and incubated for a further 3 hours. The cells were washed in PBS and resuspended in 150 ml RPMI, 10% FCS. Aliquots of 30 ml of cells were removed after 1 hr, 5 hr and 22 hr of chase and lysed (as above). All of these lysates were heated to 95°C for 5 minutes in 2% SDS prior precipitation. At 0 hr chase an additional 30 ml aliquot of cells were lysed by precipitation with the same antisera without SDS. The samples were run on denaturing 15% acrylamide gel, fixed, treated with Amplify and dried before exposure to film.

2.5.4 Reprecipitation

1x 10^7 cells were washed in PBS and incubated in 5ml of methionine free RPMI medium, 10% FCS for 1 hour. 1μCi of ^{35}S methionine was added and incubated for 2 hours. An extra 20ml of methionine free RPMI medium and 3 ml of complete RPMI medium were added and the cells were incubated for an extra 12 hours. The cells were washed in PBS and lysed in 10 ml of lysis buffer, the nuclei removed and the supernatant precleared overnight with Staph A (as in section 2.4.2).

The total lysate was split in two parts, 80μl of MCP21 monoclonal and 3μl of rabbit anti-mouse was added to one half, and 200μl of anti-LMP2 was added to the other half, and incubated for 1 1/2 hour. BSA and protein A-sepharose were added and incubated for a further 1 hour (as in section 2.2.4). The bead-bound proteins were washed 4 times in washing buffer. Each half was resuspended in 250μl of lysis buffer containing 2% SDS, and heated to 95°C for 5 minutes. The beads were pelleted and the supernatant collected. A further aliquot of lysis buffer containing 2% SDS was added to the beads, heated and collected and the two aliquots were combined. 1/6 of each half was collected and stored. The remaining 5/6 of each half was diluted to 5ml in lysis buffer without SDS and precleared overnight with Staph A to remove any remaining antibody. After pelleting the Staph A in each half, the supernatant was collected and divided in four 1ml aliquots that were reprecipitated with one of the following antisera, αLMP7, αLMP2, αDelta, αMB1 or a mock precipitation without any second antibody. The precipitations were carried out as described and the samples were run on a 15% acrylamide gel.
2.5.5 Denaturing polyacrylamide protein gel electrophoresis

A solution of 15% separating gel was prepared as below and polymerization initiated by addition of fresh 10% APS and TEMED. The gel was allowed to set for 1 hour, then the stacking gel, containing APS and TEMED was poured. Small gels were run at 20-25mA and maximum of 150V, for 1-2 hours until the dye reaches the bottom of the gel. Large gels were run overnight at 70V.

15% separating polyacrylamide gel (40ml)

- 10ml 1.5M Tris/HCl pH 8.8
- 0.2ml 20% SDS
- 20ml acrylamide solution: bis (30% T, 2.6% C)
- 9.6ml ddH2O
- 0.2ml 10% APS
- 0.02 ml TEMED

4% stacking polyacrylamide gel (10ml)

- 2.5 ml 1.5M Tris/HCl pH 6.8
- 0.05 ml 20% SDS
- 1.3ml acrylamide solution (30% T, 2.6% C)
- 6.1ml ddH2O
- 0.05ml 10% APS
- 0.01 ml TEMED

2.5.6 Western blotting

Lysates of approximately 1x 10^7 cells were prepared as previously described (section 2.4.1). The total lysate was collected and the protein levels were quantified using the BioRad protein assay kit. Equal amounts of protein were loaded per lane, per gel. 10μg to 40μg of protein were loaded depending on the size of the combs and protein gel apparatus used. The gel was run at constant 25mA at room temperature until the required separation was achieved, as judged of a set of molecular weight 'rainbow' markers.

The transfer apparatus was assembled in a shallow tray containing blotting buffer in the following order, on the top of each other in the transfer apparatus; sponge, three pieces of 3MM paper, Hybond-C membrane (pre-treated as manufactures indications), gel without the stacking part, three pieces of 3MM paper and another sponge. Each of them were previously soaked in blotting buffer and
2.5.6a Two dimensional polyacrylamide gel electrophoresis

Immunoprecipitations: Monoclonal antibody MCP21 was immobilized on CNBr-activated Sepharose CL 4B to a concentration of 2-3 mg of immunoglobulin/ml of gel. Tissues were homogenized in 5 vol. of 50 mM Tris/HCl/17% glycerol, pH 7.4 in a Potter-Elvehjem homogenizer. Homogenates were cleared by centrifugation at 12000 g for 5 minutes, and 400 μl of supernatant was incubated on a rocking platform for 1 hour with 20-60 μl of Sepharose with immobilized antibody. The Sepharose beads were washed in 3x15 ml of 20 mM Tris-HCl/20 mM NaCl/0.1 mM Na₂EDTA/1 mM MgCl₂/0.5% Nonidet P40 (NP40)/0.1% SDS/17% glycerol, pH 7.5, with intervening centrifugations at 40 g for 5 minutes and then in 15 ml of 50 mM ammonium acetate/acetic acid, pH 7.0. The final pellet was resuspended into 3 ml of distilled water, freeze-dried, and suspended into 100 μl of sample buffer for isoelectric focusing.

Electrophoresis: Immunoprecipitates were dissolved in a sample buffer with 9 M urea, 2% NP40, 5% mercaptoethanol, 1.5% 2D Pharmalyte pH 3-10, 1% Ampholine 5-8 and 0.5% Ampholine 9-11 (Pharmacia, Uppsala, Sweden). The subunits, in 10 μl of sample buffer with 20 μg of proteasomes, were separated in a Hoefer (San Francisco, CA, USA). The first dimension was non-equilibrium pH-gradient electrophoresis towards the cathode for 3 h at 400 V. The second dimensions runs were in Laemmli SDS/PAGE gels with 12.5% acrylamide. Proteins were blotted on to BA83 nitrocellulose sheets (Shleicher und Shull, Dassel, Germany) in semi-dry blotter with 48 mM Tris/39 mM glycine/0.01% SDS in 20% (v/v) methanol. The blots were reversibly stained with Ponceau S, blocked with 5% fat-free dried milk and incubated overnight with MB1 antisera in 50 mM Tris-HCl/150 mM NaCl/5 mM Na₃/0.01% Tween-20, pH 7.4. Antibodies bound to blots were detected with peroxidase-coupled rabbit antibody to mouse immunoglobulins (Dako) and tetramethylbenzidine/H₂O₂ as substrate.
bubbles were removed by rolling with a Pasteur pipette. The gel sandwich was transferred to the chamber and run at 100V at 4°C for 1 hour or 30V at room temperature overnight.

The membrane was removed from the sandwich, washed in ddH₂O and stained with Ponceau S for 5 minutes. The membrane was washed again with ddH₂O to visualise the protein transfer, pattern and concentration. The stain was removed with several washes in PBS and the membrane was incubated for 1 hr with gentle rocking at room temperature in 10-20 ml of blocking solution, prepared as follow: 5% w/vol nonfat dry milk, 0.2% Tween-20 (polyoxyethylene 20) and 0.02% sodium azide in PBS. Alternatively the membrane could be stored at 4°C after being wetted in PBS 10mM sodium azide. The appropriate dilution of antibody was prepared in 10 ml of blocking solution and the membrane was incubated for a further 1 hr at room temperature with gentle rocking. Membranes were washed in 0.2% Tween-20 PBS twice briefly and three times for 10 minutes each. Washes were shaken vigorously to achieve low background staining. A 1:1000 to 1:2000 dilution of the second layer antibody (horseradish peroxidase-conjugated anti-rabbit or anti-mouse immunoglobulin) was made in 10 ml of blocking solution and incubated at room temperature with gentle rocking for a further 45-60 minutes. The membrane was washed in Tween-20 PBS as before.

The washed membrane was transferred to a glass plate, or plastic paper, with the protein side up. ECL was performed using the Amersham ECL detection kit by mixing 1ml of each provided solution and poured immediately onto the membrane. After 1 minute the membrane was dried from behind using paper tissues. The membrane was sealed in plastic bags attached to X-ray cassettes and taken to dark room for exposure to film. A range of exposure times (from 10 seconds to 20 minutes, as required) were used to obtain the ideal exposure. Relative intensities of protein bands were quantified by LKB ultrascan.

2.5.7 Nuclear extracts

Nuclear extracts were prepared by a variation of the method of Schreiber (Schreiber, Matthias et al, 1989). 1x 10⁷ cells were washed in PBS and pelleted by centrifugation at 1300 rpm for 3 minutes in a bench top centrifuge. The pellet was resuspended in 1ml of PBS, transferred to an Eppendorf tube and pelleted again by spinning for 15 seconds in a microfuge. PBS was removed by aspiration and the cell pellet was resuspended in 400μl of cold buffer A by gentle pipetting in a yellow tip. The cells were allowed to swell on ice for 5 minutes, after which 5μl of a 10%
solution of Nonidet NP-40 was added and the tube vigorously vortexed for 10
seconds. The homogenate was centrifuged for 30 seconds in a microfuge and the
supernatant containing cytoplasm and RNA was transferred to a fresh tube. Glycerol
was added to the cytoplasmic lysate to a 10% final concentration and stored at -70°C.
The nuclear pellet was resuspended in 50-100μl ice-cold buffer C and the tube was
vigorously rocked at 4°C for 1 hour on a shaking platform. The nuclear extract was
centrifuged for 10 minutes in a microfuge at 4°C and the supernatant was stored at
-70°C.

Buffer A
10mM Hepes pH 7.9
1.5mM MgCl₂
10mM KCl
0.5mM DTT (freshly added)
protease inhibitors (freshly added)

Buffer C
20mM Hepes (pH7.9 at 4°C)
25% w/v glycerol
0.42M NaCl
1.5mM MgCl₂
0.2mM EDTA
0.5mM DTT (freshly added)
protease inhibitors (freshly added)

Protease inhibitors
0.5mM PMSF
5μg/ml Aprotinin
5μg/ml Pepstatin
30μg/ml Leupeptin

2.5.8 Cell culture synchronization and γ-interferon treatment

To arrest cells in G₁ phase, cultures were prepared using the method of
Tobey et al (Tobey et al, 1990). Exponentially growing suspension cells were
cultured in RPMI medium supplemented with 10% FCS and glutamine until the FCS
was completely depleted from the medium. An initial culture of 2x 10⁵ cells per ml
of medium was used for each experiment, and samples were collected after 0hr, 74hr
and 120 hr of culture. The cells were then washed in PBS and nuclear and
cytoplasmic extracts were prepared as before (section 2.5.7). Protein concentration
was measured for each lysate. 40μg of protein were run in a 15% protein gel, and
analysed for protein expression with the desired antisera.
For γ-interferon treatment, cells were induced with 300 units of recombinant human interferon-γ per ml of medium for 3 days. Cells were lysed (as in section 2.4.7), protein concentration was measured, and samples run in a 15% protein gel.

2.5.9 Large scale cell cultures

Proteasomes from 721 and .174 cell lines were isolated and purified from large scale cell cultures in order to analyze and compare their quaternary structure.

Volumes of 2 to 10 liters of 721 and .174 cell culture, corresponding to $10^9$ to $10^{10}$ cells, were used to precipitate 20S proteasomes. Cell pellets were homogenized for 3 x 30 seconds in a Waring blender with 2 volumes of 5mM Tris-HCl, pH 7.4, 1mm EDTA, 5mM mercaptoethanol and centrifuged at 20,000g for 45 min (crude homogenate). A 5-14% (w/v) PEG precipitate was prepared and dissolved in 20mM Tris-HCl, pH7.4, 5mM mercaptobetanol, 1mM EDTA and 20% glycerol. Extracts were then clarified by centrifugation (27,000g for 45 min) and used for immunoaffinity. Proteasomes from 721 and .174 cell lines were isolated with a column of MCP21 monoclonal antibody, which recognizes an α-type proteasome subunit.

2.5.10 Immunoaffinity column

A column of monoclonal antibody MCP21 immobilized on Sepharose CL-4B (1.6 x 6.7 cm; 2.5 mg IgG/ml) was prepared and equilibrated with 25mM Tris-HCl, pH7.5, before it was loaded with the PEG fraction. The enzyme activity was measured at various steps of the flow-through. The column was washed with 25mM Tris-HCl buffers: pH 7.5 (280ml); pH8.0 with 50% ethylene glycol 60ml); pH8.0 with 50ml NaCl (60ml) and, finally, pH8.0 with 0.2M NaCl (180ml). A total of 0.4 mg of proteasomes were eluted with 2M NaCl in 25mM Tris-HCL, pH 8.0, and desalted by gel filtration. The affinity column could be used several times without detectable loss of binding capacity.

2.5.11 Ammonium sulphate fractionation method

Proteasomes were purified from packed 721 or .174 cells diluted with an equal volume of 20mMTris buffer, pH7.4, containing 20mM NaCl, 1mM EDTA and 1mM DTT (buffer A) and the cells were lysed by five passes of the pestle through a
Potter homogeniser. The lysate was centrifuged at 80,000g for 60min (50.2Ti rotor; Beckman). The supernatant was then removed and ammonium sulphate fractionation was carried out at 35% saturation by adding 205 mg ammonium sulphate per ml lysate, with continuous stirring for a further 30-45 min. After centrifugation at 40,000 for 40 min to pellet the precipitate, the supernatant was decanted into a beaker and a further 0.205 gm/ml of ammonium sulphate was added to bring the final ammonium sulphate concentration up to 60% saturation. After stirring for 60min the mixture was centrifuged at 40,000g for 40 min. Pelleted material was resuspended to a final volume of 50ml with 20% volume lysis buffer, and dialysed overnight against 20mM Tris buffer pH7.4, 1mm MgCl2, 1mM EDTA, 1mM DTT and 20% glycerol. The dialysed sample was then loaded onto DEAE Sepharose column followed by application of a linear gradient of increasing NaCl. This resulted in the elution of about 1mg of proteasomes at 0.20M NaCl as determined by Bradford protein assay. The active fractions obtained from the DEAE-Sepharose column were made to 5mM potassium phosphate (pH7.6) plus 20% glycerol and loaded onto a hydroxyapatite column (3cm x 10cm). The protein was eluted by a 400ml gradients (5-200mM potassium phosphate). The most active fractions were collected, dialysed against buffer A and loaded onto a Mono-Q column. The protein was eluted using a salt gradient (20-500 mM NaCl). The protein was dialysed against 20mM NaCl in 20mM Tris-HCl buffer, pH 7.5 and concentrated using Centricon microconcentrates. Final protein concentrations were determined using the Bradford protein assays.

**Electron microscopy of 20S proteasome-antibody complexes:**

Proteasomes from wild type and mutant cell lines and antisera at a molar ratio of 1:50 were mixed in a volume of 0.1ml 20mM Tris buffer, pH7.5 plus 20mM NaCl and incubated for 1h at 20°C. Aliquots of proteasomes, from both cell lines, and proteasome-antibody complexes were collected on air-glow-discharged carbon coated 400 mesh copper grids and negatively stained by floating on a drop of 1% (w/v) uranyl acetate for 20 seconds. Excess stain was blotted and the grid was dried in air prior to examination in a Zeiss 10CR transmission electron microscope at 100kV. See page 139 for results and analysis.
Chapter 3

Cloning of MB1 and Delta Proteasome Subunits

3.1 Introduction

A number of mutant cell lines have been described that are deficient in antigen processing (Salter et al. 1985, Ljunggreen et al. 1990, DeMars et al. 1985, Cerundolo et al. 1990). These cells had normal class I gene sequences but were deficient in class I expression. The class I products became sequestered in the ER, suggesting a defect in peptide supply. In the cell line .174, this phenotype was mapped to a region in the MHC stretching from DPA2 to the telomeric end of class II. This suggested that genes evolved in antigen processing lay within this deletion.

Two genes with homology to the transporters were discovered in this region, the TAP1 and TAP2 genes (Trowsdale et al. 1990, Deverson et al. 1990, Spies et al. 1990). Their involvement with antigen processing came with the realisation that mutations in either of the two genes led to the phenotype described above and class I expression could be restored by transfection of the appropriate wild type TAP gene (Kelly et al. 1991, Spies et al. 1991).

Circumstantial evidence for the existence of further MHC genes coding for proteins involved in the degradatory step of antigen processing came with the description of the H2-linked LMP complex in the mouse (LMP's for low molecular weight polypeptide) (Monaco et al. 1986). This complex consisted of ~16 components which have a 2D gel pattern similar to that of biochemically purified human proteasomes. Subsequently, the genes of two proteasome subunits, LMP2 and LMP7, were found to be intimately associated with TAP genes. The MHC encoded LMP genes were obvious candidates for involvement in producing peptides for transport into the ER.

However, two reports suggested that the LMP genes were not essential for processing of antigen for class I presentation. Mutant cells, .174 and T2, lacking the LMP genes, could express class I glycoproteins at the cell surface at normal levels and could present intracellular antigens after transfection with TAP genes (Arnold et
Further experiments suggested that the MHC-encoded proteasome subunits could have a specialized function, producing peptides better suited for class I presentation (Driscoll et al. 1993, Gaczynska et al. 1993). LMP containing proteasomes could change the cleavage specificity, producing a pool of peptides with carboxyl-termini with higher affinity for class I binding.

We therefore reasoned that the lack of any obvious effect on class I presentation in the LMP negative mutant cell line was due to expression of two other highly-related proteasome subunits. We cloned the genes encoding a new proteasome subunit, MB1, which is closely related to LMP7, and that encoding a second subunit, Delta, which is closely related to LMP2.

### 3.2 Cloning and sequence analysis of MB1 gene

In order to obtain an LMP7-related sequence encoding a proteasome subunit, we screened cDNA libraries with LMP7 cDNA probe under non-stringent hybridization conditions.

The LMP7 cDNA had been previously cloned from a CEM T cell cDNA library in the plasmid vector CDM8 (Glynne et al., 1991). The LMP7 cDNA was removed from the vector and an insert of approximately 1.3kb was obtained and purified. The LMP7 probe was then labelled and used to screen a JY B cell cDNA library (described in section 2.3.6). A complete JY cDNA library was obtained by plating out 4 large 20x20 plates, and for each master plate 4 duplicates were prepared. Two duplicates were hybridized with the LMP7 probe under high stringent conditions (65°C and 0.1xSSC), and the other two duplicates were washed at low stringency (55°C and 6xSSC). The patterns of spots obtained in the autoradiographs under low and high stringent conditions were compared. Spots from the low stringency hybridization that were not found at high stringency were selected, subcloned and analysed by sequencing.

A total of 12 clones were obtained and rescreened at low stringency with the LMP7 probe, resulting in 5 positive clones. The 5 clones were then sequenced and checked against EMBL and Genbank nucleic acid databanks. Only one clone, called MB1, was identified as a new β-type proteasome subunit. The 1.0 kb MB1 insert was oligo-sequenced in both directions by primers designed for previously sequenced regions. One end of the MB1 gene sequence was very similar to that of the LMP7 β-proteasome subunit. The other end had a poly A tail. The cDNA sequence showed an open reading frame of around 639 bp, encoding a protein of 208
amino acids. The sequence of the MB1 clone, together with the predicted protein sequence is shown in figure 3.1.

During assembly of Archaebacterium and eukaryotic proteasomes the propeptide of the β-subunits is posttranslationally removed (Glynne et al. 1993, Fruh et al. 1992, Frentzel et al. 1994, Zwickl et al. 1994). This was confirmed by the crystal structure of the Archaebacteria proteasome showing that β-subunits have the amino terminal region cleaved off in order to be incorporated in the 20S complex (Löwe et al.1995). The mature form of the MB1 protein could therefore be deduced as 204 amino-acids long, with its first amino acid at the Thr^ residue (equivalent to position 60 in the MB1 protein sequence). The processed form of MB1 encodes a protein with Mr 22.5 and pl 9.0 similar to that observed on western blots and 2D-gels. The unprocessed forms of β-subunits' proteins have molecular weights varying from 25-30KDa each, and they encompass a pro-peptide leader sequence of 30-50 amino acids. However, the MB1 cDNA clone did not have the initiating methionine and according to western blot data an extra 50-60 amino-acids sequence encoding an unprocessed MB1 form of around 30-35 kDa is predicted (figure 5.2).

Sequencing of the MB1 gene from a cosmid clone containing the complete 5' region of the MB1 gene revealed a potential leader sequence of 59 amino acids (Abdulla et al. in preparation). Western blot data of protein extracts from different cell lines showed two MB1 protein bands separately of about 8-12kDa (figure 5.2), consistent with cleavage of the leader sequence of 59 amino acids. The MB1 gene has also been referred to as X or epsilon (Akiyama et al. 1993, Lee et al. 1990).

The protein sequence from the MB1 open reading frame was compared with other protein sequences to test its homology to previously published proteasome subunits. The MB1 protein sequence was checked against the following data banks: PDB, SwissProt, SPupdate, PIR, GenPep and GPupdate, containing at a total of 161,547 sequences. Matches were found with around 98 proteasome subunits. The highest scores were; subunit RAX (rat), C1 (chicken), LMP7 (rat), LMP7 (mouse), LMP7 (human), PRE2 yeast and PTS1 (yeast).

The high similarity between MB1 and LMP7 was obvious when comparing the amino acid sequence of the two proteins (figure 3.2). The degree of identity between both subunits at the protein level is 68.1%, calculated by the molecular biology program GAP (GAP creates alignment with the largest number of matched bases and the fewest number of gaps). Further analysis of the MB1 protein sequence was performed by comparing the MB1 protein sequence with other human
Figure 3.1: Sequence of MB1 cDNA clone. The MB1 amino-acid sequence is derived from the longest open reading frame, and is shown underneath by using the single amino-acid letter code. The region corresponding to the start of the mature form of the protein is indicated by an asterisk. Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point has its NH2-terminus freed by cleavage of the pro-peptide, and it corresponds to the active-site nucleophile of the protein. The second conserved residue is Lys at position 92, and it corresponds to the putative proton acceptor-donor residue. The two glycine-rich sequences are in direct proximity to the active-site residues according to the proteasome crystal structure.
Figure 3.2: Comparison of amino-acid sequence of MB1 and LMP7b human proteasome β-subunits. The region corresponding to the start of the mature form of the proteins is indicated by an asterisk. Identical and conserved residues are shown in red and blue, respectively. Numbers listed on the left side are residue numbers. Conserved amino-acids are defined: A, S, T, P and G; N, D, E and Q; R, K and H; I, V, L and M; F, Y and W. The reported sequence of human LMP7 is shown (Glynne et al. 1991). Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point corresponds to the active-site nucleophile of the protein, position 60 for MB1 and position 74 for LMP7b. The second conserved residue is Lys at positions 92 for MB1 and position 106 for LMP7b, corresponding to the putative proton acceptor-donor residue. The two glycine-rich sequences are at positions 133-135 and 174-176 for MB1, and 197-199 and 238-240 for LMP7b. Both positions are in direct proximity to the active-site residues.

Conserved amino acids are defined according to the nature of their side-groups. Short branched-chain amino acid group-A S T P G; Long branched-chain amino acid group-I V L M; C=O branched-chain amino acid group-N D E Q; HN branched-chain amino acid group- R K H; Aromatic branched-chain amino acid group-F Y W
β-subunits recently published. MB1 was compared to the following subunits; Delta, LMP2, Z, MECL1, HC5, HC7, HC10 and HN3. The sequences were analysed in pairs using the GAP program. Figure 3.3 shows that MB1 is highly homologous to LMP7 and that the degree of relatedness is far higher than that observed with the other β-subunits, varying from 20.2% for HC7 to 32.4% for Delta. Given the high similarity between the MB1 and LMP7 subunits, we suggest that the MB1 gene would be the most probable candidate to replace the LMP7 gene in the .174 mutant cell.

**Active site residues within the MB1 sequence**

The 20S proteasome forms the proteolytic core of the 26S complex. However, analysis of 20S proteasome subunits has revealed no obvious homology with other known proteases (DeMartino et al. 1991). Early studies using proteasome inhibitors have indicated that the 20S proteasome may be an unusual type of cysteine or serine protease (Rock et al. 1994). Further information on the nature of the active site was obtained by mutating serine and histidine residues in the archaebacterial β-subunits, and it was concluded that the proteasome was not a cysteine or serine protease (Seemüller et al. 1995a). The catalytic mechanism of the 20S proteasome from *Thermoplasma acidophilum* was discovered by deleting the amino-terminal threonine or mutating this to alanine, resulting in inactivation of the enzyme (Seemüller et al. 1995b). This data, in combination with the crystal structure of a proteasome-inhibitor complex showed that the nucleophilic attack was mediated by the amino-terminal threonine of processed β-subunits (Löwe et al. 1995). It was realized that the proteasome is a new type of protease: a threonine protease.

Proteolysis also requires a second conserved residue that carries an unprotonated nucleophilic group (Seemüller et al. 1995b). Its function is to strip the proton from the active-site nucleophile, thus initiating the attack, and then to donate the proton back to the leaving NH2-terminal group of the cleaved substrate. Studies performed in *Thermoplasma acidophilum* proposed two possible proton acceptors -donors that were highly conserved in the potentially active β subunits: Lys33 and the amino-group of Thr1 (Seemüller et al. 1995b). Mutation of Lys33 showed that the proteasome became inactive but could fold and assemble correctly, leading to the conclusion that Lys33 was essential for functional reasons. However, the results did not show whether Lys33 is directly involved in proton transfer or whether it serves to polarise the Thr1 amino-group. On the basis of the crystal structure both roles appear to be possible, but the Thr1 amino group was favored as the primary acceptor-donor.
Figure 3.3: Similarities of primary structures of various human proteasome β subunits. The sources of sequence data are as follows: Delta and Z (DeMartino et al. 1991), HC5 (Tamura et al. 1994), HC10 and HN3 (Nothwang et al. 1994), MECL1 (Larsen et al. 1993), LMP7 (Glynne et al. 1991), LMP2 (Kelly et al. 1991), MB1 (Belich et al. 1994). Percentages of overall amino acid sequence identity between proteasome β-subunits were obtained by a using the GAF molecular biology program. Sequences were compared in pairs. The percentage identity was deduced by comparing the amino acid sequences of the mature form of each β-subunit, which are MB1 (60-263), LMP7b (73-276), Delta (35-239), LMP2 (21-219), Z (44-277), MECL1 (40-273), HC5 (38-241), HC7 (2-201), HC10 (9-205) and HN3 (46-237) (for the amino acid residues' location in each sequence see figure 3.8).
Two other conserved regions in the amino-acid sequence have been proposed to be involved in the active sites of the β-subunits, the pattern Gly^{128}\text{SerGly} and Ser^{169}\text{GlyGly}, are in direct proximity to the active-sites' residues and are conserved in active subunits (Seemüller et al. 1995b).

Inhibition studies in mammalian cells using lactacystin, a Streptomyces metabolite that inhibits cell cycle progression, identified the 20S proteasome as its specific cellular target (Fenteany et al. 1995). It was shown that lactacystin covalently modifies the highly conserved amino-terminal threonine (Thr\textsuperscript{1}) of the proteasome MB\textsubscript{1} subunit. The conclusion from this is that the threonine residue has a catalytic role, and subunit MB\textsubscript{1} may be a core component of an amino-terminal-threonine protease activity of the proteasome. Three distinct peptidase activities of this enzyme complex (trypsin-like, chymotrypsin-like, and peptidylglutamyl-peptide hydrolyzing activities) were inhibited by lactacystin, the first two irreversibly and all at different rates.

Analysis of the MB\textsubscript{1} sequence (figure 3.1) shows that it contains the two active site nucleophiles, the Thr\textsubscript{1} (residue 60 in the MB\textsubscript{1} sequence), and the putative second conserved residue Lys\textsuperscript{33} (residue 92). It also contains the Gly^{128}\text{SerGly} (residues 188-190) and Ser^{169}\text{GlyGly} (residues 229-231), demonstrating that the MB\textsubscript{1} is an active β-type proteasome subunit. Interestingly, when the same positions are analysed in the LMP\textsubscript{7} protein sequence they contain exactly the same amino-acids as in MB\textsubscript{1} (figure 3.2). However, the flanking regions of the active residues of MB\textsubscript{1} and LMP\textsubscript{7} are different, and this could account for a possible differential structural conformation affecting the activity of the complex. In addition, cells treated with lactacystin were shown to bind and inhibit only the MB\textsubscript{1} subunit, and not LMP\textsubscript{7} (Fenteany et al. 1995).

It has been shown that proteasomes containing LMP\textsubscript{2}/LMP\textsubscript{7} subunits have increased trypsin-like and chymotrypsin-like activities and those with MB\textsubscript{1} and Delta have increased peptidylglutamyl-peptide hydrolysing and decreased trypsin-like and chymotrypsin-like activity. However, inhibitor studies with lactacystin showed that the MB\textsubscript{1} subunit has trypsin-like and chymotrypsin-like activities. Thus, the differential activity properties observed in the proteasomes with Delta and MB\textsubscript{1} are probably due to the Delta subunit increasing the peptidylglutamyl-peptide hydrolysing activity. The other hypothesis is that the MB\textsubscript{1} activity could be modified by incorporation of a different subunit in the complex.
Thus, protein data analysis shows that both genes, MB1 and LMP7, are active β-subunits with similar active residues, but it is not yet known if they have different activity properties.

3.3 Cloning and sequence analysis of Delta gene

In order to obtain a cDNA encoding the human Delta proteasome subunit, we screened a human B-cell cDNA library with a Delta polymerase chain reaction (PCR) probe under high-stringency hybridization conditions.

The human sequence encoding the proteasome subunit Delta was previously published, and it showed similarity to the other β-proteasome subunits having the highest homology to LMP2 (DeMartino et al. 1991). A 131-nucleotide PCR probe corresponding to the 5' region of the published Delta sequence was obtained with the two oligos 5'ATCGCCAATCGAGTGACTGAC A A3' and 3'GTCGAGCCAAAGGTGTCGTAACT5'. The human B-cell c-DNA library JY was screened at high stringency (65°C and washed with 0.1 x SSC) re-screened, and one positive cDNA clone was obtained. The 0.8 kb insert was oligo-sequenced in both directions by designing primers according to the previously published Delta sequence. The cDNA sequence showed an open reading frame of around 700bp, encoding a protein of 239 amino acids. The sequence of the Delta clone, together with the predicted protein sequence is shown in figure 3.4.

The sequence obtained from the Delta open reading frame protein was identical to but longer than the previously published sequence. It contained the Thr residue and a further 13 amino acids. However, the Delta sequence did not contain the initiating methionine AUG. The predicted Mr for the mature form of the Delta protein, which is 205 amino acids long, was 21.8 kDa and the isoelectric point 4.6, both of which are consistent with the observed values for Delta protein on western blots and 2-D gels. Recently, a longer Delta sequence has been identified which contains the probable pro-peptide region of 34 amino-acids, consistent with the unprocessed form of Mr around 26 kDa (EMBL D29012 accession number).

The high similarity between Delta and LMP2 was shown by comparing the amino acid sequences of both proteins (figure 3.5). The degree of identity between both subunits at the protein level is 60.8% as calculated by the molecular biology program, GAP. Comparison of the Delta protein sequence to other human β-subunits (LMP7, MB1, Z, MECL1, HC5, HC7, HC10 and HN3) shows that the degree of relatedness to LMP2 is far higher than observed with the other β-subunits,
Delta cDNA clone. The Delta amino-acid sequence is derived from the longest open reading frame, and is shown underneath by using the single amino-acid letter code. The region corresponding to the start of the mature form of the protein is indicated by an asterisk. Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point has its NH2-terminus freed by cleavage of the pro-peptide, and it corresponds to the active-site nucleophile of the protein. The second conserved residue is Lys at position 67, and it corresponds to the putative proton acceptor-donor residue. The two glycine-rich sequences are in direct proximity to the active-site residues, positions 163-165 and 204-206 respectively.
Figure 3.5: Comparison of amino-acid sequence of the Delta and LMP2 human proteasome β-subunits. The region corresponding to the start of the mature form of the proteins is indicated by an asterisk. Identical and conserved residues are shown in red and blue, respectively. Numbers listed on the left side are residue numbers. Conserved amino-acids are defined: A, S, T, P and G; N, D, E and Q; R, K and H; I, V, L and M; F, Y and W. Reported sequence of human LMP2 is shown (Kelly et al. 1991). Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point corresponds to the active-site nucleophile of the protein, positions 35 for Delta and 21 for LMP2. The second conserved residue is Lys at positions 67 for Delta and 53 for LMP2, corresponding to the putative proton acceptor-donor residue. The two glycine-rich sequences are at positions 163-165 and 204-206 for Delta, and 148-150 and 189-191 for LMP2 is in direct proximity to the active-site residues.

Conserved amino acids are defined according to the nature of their side-groups. Short branched-chain amino acid group- A S T P G; Long branched-chain amino acid group-I V L M; C=O branched-chain amino acid group-N D E Q; HN branched-chain amino acid group- R K H; Aromatic branched-chain amino acid group-F Y W
which varies from 16.8% for HC7 to 32.4% for LMP7. As observed for MB1 and LMP7, the high similarity between Delta and LMP2 subunits suggests that the Delta gene is the most probable candidate to replace the LMP2 gene in the .174 mutant cell.

**Active site residues within the Delta sequence**

Sequence analysis of the Delta protein (figure 3.4) shows that it contains the two active site nucleophiles, the Thr\(^1\) (residue 35 in the Delta sequence), and the second conserved residue Lys\(^{33}\) (residue 67). It also contains the Gly\(^{128}\)SerGly (residues 163-165) and Ser\(^{169}\)GlyGly regions (residues 204-206). As for MB1, Delta is also an active \(\beta\)-type proteasome subunit.

When the equivalent putative active residues were compared in the LMP2 protein it was shown that both subunits have exactly the same composition (figure 3.5). However, in contrast to the MB1 and LMP7 analysis, the flanking regions of the active residues of Delta and LMP2 are identical, which could relate to the lack of peptidyl glutamyl hydrolase activity associated with both subunits (Gaczinska et al. 1994). In addition, proteasomes containing high levels of LMP2, or over-expressing the LMP2 gene by transfection, have an increased hydrolysis after basic residues. However, it is not known if this property is associated uniquely with the LMP2 subunit, or if the incorporation of LMP2 in the complex is accompanied by the incorporation of other subunits, that in association would change the specificity of the proteasome.

**3.4 mRNA expression of MB1 and Delta**

Northern analysis of both LMP genes has previously shown that their mRNA is constitutively expressed in B and T lymphoblastoid cell lines and up-regulated in \(\gamma\)-interferon treated cells (Kelly et al. 1991, Glynne et al. 1991). Similar results were also observed for the class I and TAPs genes. In order to study the mechanism of the apparent switch between MB1/Delta and LMP2/LMP7 gene expression, we compared mRNA levels for the four genes in both wild-type and LMP-deficient B cells.

Both MB1 and Delta mRNAs were increased 3-4 fold in the mutant cell .174 compared to the wild type 721 cells (figure 3.6). Total mRNA was isolated as described in section 2.3, and equal amounts of mRNA were used from each sample. The relative intensities of the mRNA bands, and MB1 and Delta protein levels from western blot data, were quantified by LKB ultrascan. MB1 and Delta protein levels
Figure 3.6: Analysis of mRNA expression in wild-type and LMP-deficient cells.

(A) Northern hybridization analysis of Delta and MB1 expression; LMP7 is also shown for comparison. Total RNA was taken from lymphoblastoid cell lines 721 (wild-type; lanes 3, 4, 6, 8, 10 and 12) and .174 (mutant; lanes 1, 3, 5, 7, 9, 11). Blots were then probed with LMP7 (lanes 1-4), Delta(5-8) and MB1 (9-12) cDNA inserts labelled by random priming. The upper bands are most probably the result of 28S ribosomal RNA cross-hybridization.

(B) Ethidium bromide staining of the blotted gel.
were increased 10-20 fold in the mutant .174 cell (data shown in chapter 5). From this result it appears that expression of the LMPs leads to a slight reduction in the mRNA levels for MB1 and Delta, and most probably post-transcriptional control is responsible for most of the decrease seen at the protein level.

A possible explanation for post-transcriptional control would be that the degradation of MB1 and Delta subunits follows their lack of incorporation into the proteasome complex in the presence of the LMP subunits. The possibility that MB1 and Delta are exchanged respectively by LMP7 and LMP2 in such a way that MB1 and Delta are released from the complex and then replaced by the LMPs can be ruled out because, so far, free processed form of either LMP2, LMP7, MB1 or Delta have not been observed, whereas their unprocessed forms have been detected by antisera raised against C-termini region of the protein common to both mature and immature forms (data shown in chapter 5). Thus, it seems that precursor processing most likely occurs during the assembly process and that once the formation of the proteasomal complex is completed, the subunits cannot be exchanged.

Preferential incorporation of LMP7 and LMP2 over MB1 and Delta could be due to higher affinities for other proteasome subunits or by higher concentration of the LMP precursor.

3.5 Analysis of the proteasome β-subunits

As discussed in the Introduction, the 20S eukaryotic proteasome has been shown to have up to five different proteolytic components, but the lack of similarity to any other proteases and the initial inconclusive studies with inhibitors were inconsistent with any of the known protease families. Various studies with Thermoplasma proteasomes showed that they have only one activity, the chymotrypsin-like one, although the complex could hydrolyse any peptide bond (Wenzel and Baumeister, 1993, 1995). The assembly of the complex was shown to be directed by the α subunits, which were essential for the β subunit assembly and pro-peptide cleavage. When the pro-peptide was removed by genetic deletion, the β subunits could form aggregates and regain some proteolytic activity, supporting the idea that the β subunits are the proteolytic core of the 20S complex (Zwickl et al. 1995).

Figure 3.7 shows an analysis of all currently known human β subunits according to their molecular weight, isoelectric point and pro-peptide size. An alignment of the human β-type sequences (figure 3.8) shows that there are
<table>
<thead>
<tr>
<th>Subunit size</th>
<th>Mature</th>
<th>Immature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mr</td>
<td>Size</td>
</tr>
<tr>
<td>MB1</td>
<td>22.46</td>
<td>204</td>
</tr>
<tr>
<td>LMP7</td>
<td>22.66</td>
<td>204</td>
</tr>
<tr>
<td>Delta</td>
<td>21.90</td>
<td>205</td>
</tr>
<tr>
<td>LMP2</td>
<td>21.23</td>
<td>199</td>
</tr>
<tr>
<td>Z</td>
<td>25.30</td>
<td>234</td>
</tr>
<tr>
<td>MECL1</td>
<td>24.65</td>
<td>234</td>
</tr>
<tr>
<td>HC5</td>
<td>22.48</td>
<td>204</td>
</tr>
<tr>
<td>HC10</td>
<td>22.05</td>
<td>197</td>
</tr>
<tr>
<td>HN3</td>
<td>24.38</td>
<td>218</td>
</tr>
<tr>
<td>HC7</td>
<td>22.48</td>
<td>200</td>
</tr>
</tbody>
</table>

Figure 3.7: Analysis of the protein sequences of various human proteasome β-subunits.

The molecular weight and isoelectric point of the mature and immature forms of the β-subunits were obtained by using the molecular biology software for protein analysis from GCG. The Mr of the mature form of the proteins was calculated according to the predicted cleavage point at the Thr1 positions. The same starting point was used to calculate the Ip of the mature form of the proteins. Putative starting sites for HC5, HC7 and HC10 were calculated at equivalent positions to the Thr1. As an exception, subunit HN3 has its starting point eight residues prior to Thr1.

The amino acid residues for the mature form of each β-subunit are MB1 (60-263), LMP7 (69-272 LMP7a and 73-276 for LMP7b), Delta (35-239), LMP2 (21-219), Z (44-277), MECL1 (40-273), HC5 (38-241), HC7 (2-201), HC10 (9-205) and HN3 (46-237) (for the amino acid residues located in each sequence, see figure 3.8). The immature forms of β-subunits were calculated according to published protein sequences that contained the Met residues as putative starting point upstream to the Thr1 residue. Published protein sequences for HC10 and HC7 subunits were not complete so their Mr were not calculated. The sizes of putative pro­peptides were obtained by subtracting the protein sizes of the mature and immature forms.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>lmp7a</td>
<td>KVESTDVSDLLHQREANQ ..</td>
</tr>
<tr>
<td>lmp7b</td>
<td>KVESTDVSDLLHQREANQ ..</td>
</tr>
<tr>
<td>mbl</td>
<td>RVSNDVADLHEKYSGSTP ..</td>
</tr>
<tr>
<td>delta</td>
<td>VLGDPQPKFAVATLPPA ..</td>
</tr>
<tr>
<td>lmp2</td>
<td>VILGNELPKFYDE ..</td>
</tr>
<tr>
<td>z</td>
<td>AVLTEKIPILEEVLEETVQMTDS</td>
</tr>
<tr>
<td>mcl1</td>
<td>AVLQTVKPIELVEEVQAMEVE</td>
</tr>
<tr>
<td>hc5</td>
<td>..</td>
</tr>
<tr>
<td>hc10</td>
<td>..</td>
</tr>
<tr>
<td>hc7</td>
<td>..</td>
</tr>
<tr>
<td>hn3</td>
<td>..</td>
</tr>
</tbody>
</table>

Figure 3.8: Comparison of amino-acid sequences of the published human proteasome β-subunits. The regions corresponding to the start of the mature form of the proteins are indicated by asterisks at the top of the sequence alignment. Subunit HN3 has its starting point located at eight residues prior the starting site residues of the other β-subunits. Identical and conserved residues are shown in red and blue, respectively. Numbers listed on the left side are residue numbers. Conserved amino acids are defined as: A, S, T, P and G; N, D, E and Q; R, K and H; I, V, L and M; F, Y and W. Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point corresponds to the active-site nucleophile of the protein, and it is not present in all β-subunits. Subunits HC5, HC7, HC10 and HN3 are inactive subunits. The second conserved residue is Lys^{33} (boxed) and it corresponds to the putative proton acceptor-donor residue for each β-subunit. The two glycine-rich sequences (boxed) are in direct proximity to the active-site residue, they are also not present in all subunits. Additional residues that are conserved in all β-subunits are marked by an asterisk below the sequence alignment.

Conserved amino acids are defined according to the nature of their side-groups.

- Short branched-chain amino acid group- A S T P G
- Long branched-chain amino acid group- I V L M
- C=O branched-chain amino acid group- N D E Q
- HN branched-chain amino acid group- R K H
- Aromatic branched-chain amino acid group- F Y W
few conserved residues, and apart from glycines the only completely conserved 
residues are two Asp residues at positions 17 and 51 respectively, numbering from 
Thr^1 (note that each published sequence has a different length). Comparison of 
Thermoplasma and eukaryotic α and β subunits shows that Asp^51 is the only 
conserved residue, and when mutated to asparagine it showed a surprisingly increased 
activity of the mutant by three times (Seemüller et al. 1995a). Further mutation or 
deletion of the first two threonines in the Thermoplasma β subunits generated 
proteasomes that were inactive but correctly folded, implicating the Thr^1 as the active 
site nucleophile (Seemüller et al. 1995b). Sequence alignment from human β 
subunits (figure 3.8) shows that HC5, HC7 and HC10 type β subunits are unlikely to 
carry proteolytic activity, whereas all other β subunits are potentially active. Further 
experiments showed that Lys^33 is the possible proton-acceptor donor and mutants for 
the Lys^33 residue folded and assembled correctly but remained inactive. Sequence 
analysis of subunit HN3 shows that it lacks the Lys^33 residue and it is most likely to 
be proteolytically inactive.

The two glycine-rich sequences, Gly^{128}Ser Gly and Ser^{169}GlyGly, in 
direct proximity to the active-site residues, are conserved in active subunits but not in 
HN3, HC5, HC7 and HC10. The other β subunits, LMP2, LMP7, MB1, Delta, 
MECL1 and Z belong to three different groups, and each group contains one subunit 
that is induced by γ-interferon; LMP2 is the inducible homologue of Delta, LMP7 of 
MB1 and MECL1 of Z.

3.6 Conclusion

This chapter has described the cloning of a new proteasome β subunit, 
called MB1, highly related to the LMP7 subunit. Subsequently, a second subunit 
homologue to LMP2, called Delta, was identified and cloned from previously 
published sequences. Sequence analysis of all β subunits showed that MB1 and Delta 
are the most probable candidates for replacing LMP7 and LMP2 subunits, 
respectively, in the mutant cell .174. The .174 mutant has shown no major defect in 
class I expression at the cell surface, suggesting that the subtle phenotype of the 
LMP-deficient cell line results from the compensatory expression of the MB1 and 
Delta subunits. The mRNA levels of MB1 and Delta appear to be up-regulated in the 
mutant cell and only slightly reduced in cells containing the LMP genes, indicating 
that post-transcriptional control could be responsible for most of the reduction in 
expression of MB1 and Delta subunits. Sequence analysis of all human β subunits 
showed that six of them (MB1, Delta, Z, LMP7, LMP2 and MECL1) are probably 
active.
Recently, experiments with LMP2 and LMP7 mutant mice have shown that these β subunits influence antigen presentation (Van Kaer et al. 1994, Fehling et al. 1994). The peptidase activity of cells from LMP2 negative mice resembled those observed in the .174 mutant cell, showing significantly lower levels of degradation of hydrophobic and basic substrates and almost two-fold higher levels of degradation of acidic substrate when compared with wild type mice (Van Kaer et al. 1994). Similarly to our data, the loss of LMP2 subunit in the mutant mice was associated with up regulation of Delta subunit.

The data described in this chapter has also been confirmed by two other groups. Akiyama et al showed that interferon-γ treated cells increased expression of the MHC-encoded LMP2 and LMP7 subunits and decreased expression of subunit X (identical to Delta) and subunit Y (identical to MB1) (Akiyama et al. 1994). Yang et al. showed similar results with interferon-γ treated cells. However, the subunits with decreased levels of expression were not identified by molecular cloning (Yang et al. 1994).
Chapter 4

Mapping of MB1 and Delta Genes

4.1 Introduction

As discussed in the introduction, the *LMP2* and *LMP7* loci are located close to each other on chromosome 6p21 within the MHC region, which also contains the *TAPI* and *TAP2* genes (Beck et al. 1992). The TAPs and LMPs form a cluster of tightly linked genes that have been proposed to function during the immune response. A number of genes involved in immune response are up-regulated by γ-IFN, including TAPs and LMPs. Evolutionary relationship analysis of these four genes suggests that TAP1 and TAP2 genes most probably arose by a duplication event. LMP2 and LMP7 are also related and could have been derived from a common primordial gene. It was previously proposed that an initial TAP and LMP gene pair would duplicate, forming two groups of genes, followed by an inversion of one LMP gene (Beck et al. 1992).

In order to understand if there was any relationship between MB1 and Delta loci, given that the LMP2 and LMP7 are in close proximity to each other in the MHC, both genes were examined for their chromosomal localization. MB1 and Delta genomic clones were isolated from cosmid libraries by hybridization, and the cosmids were then used to map the genes by *in situ* hybridization (FISH).

MB1 and Delta genes were localized on different chromosomes, mapping to 14q11.2 and 17p13, respectively, raising interesting questions about the evolution of the four β subunits' genes.

4.2 Chromosome localization of MB1 gene

Genomic fragments of MB1 were obtained by screening the full human ICRF-DH1 cosmid library (kind gift from Genome Analysis Laboratory). The full length MB1 cDNA clone was hybridized at high stringency (65°C and washed with 0.2xSSC). Three positives clones were obtained, called MB2-1, MB4-2 and MB5-2. The cosmids were analysed by restriction enzyme digestion and they showed similar banding patterns on EcoRI digestion, indicating that they overlapped each other. The
longer cosmid, MB4, contained a ~40kb insert, and the other cosmid insert varied between 34-37 kb in size. The cosmids were also PCR tested for the presence of intronic and exonic DNA sequences in the MB1 gene. Four different sets of MB1 oligos, located 150 to 200 base pairs apart, confirmed the presence of introns and exons in all MB1 cosmids. For cosmid cloning see materials and methods 2.3.6a.

The comids containing MB1 (MB2-1, MB4-2, and MB5-2) were labelled by nick translation with biotin-dATP and hybridized to chromosome spreads. Chromosome spreads were obtained from phytoemagglutin-stimulated blood T-lymphocytes of a healthy donor. Metaphase chromosomes were prepared from normal human lymphocytes after incorporation of bromodeoxyuridine (BrdU) during the second half of S phase of the cell cycle. The probe detection was performed with avidin-Texas Red (figure 4.1) and a replication of G-banding pattern was obtained using an anti-BrdU monoclonal antibody conjugated to fluorescein isothiocyanate (FITC).

The three MB1 cosmids independently hybridized at the same region on chromosome 14q11.2-q12. These results were based on observation of more than 20 prometaphase chromosomes per cosmid analysed, and the band assignments were also confirmed by G-banding. Fluorescence signals were not consistently observed on other chromosomes. Digitized images from hybridization signals and banding patterns were taken with a cooled CCD (charge-coupled device) camera (Photometrics) and merged using computer software which was developed by T. Rand and D. C. Ward. The FISH method described above was performed in collaboration with Dr. G. Senger at the Human Cytogenetics Laboratory at ICRF.

4.3 Chromosome localization of Delta gene

Genomic fragments containing the Delta gene were obtained as for MB1. The ICRF-DH1 cosmid library (kind gift from Genome Analysis Laboratory) was screened with a full length Delta cDNA clone and hybridized in high stringency conditions as described above. Three positive clones were obtained, called C1, C2 and C3. The cosmids were also analysed by restriction enzyme digestion and they showed similar banding patterns as a result of EcoRI digestion, indicating that they overlapped each other. The cosmids varied in size between 34-40 kb, with the C1 cosmid being the largest.
Figure 4.1: Chromosome localization of MB1 gene by fluorescence in situ hybridization (FISH). Hybridization signals of MB1 (yellow) are visible on chromosome 14 in the region 14q11.2-14q12 (arrows). Ideogram of chromosome 14 shows the localization of the MB1 gene.
The FISH method was performed the same way as described for the MB1 mapping. The three Delta cosmids (C1, C2 and C3) mapped independently to chromosome 17p13.2-13.3 (figure 4.2). The Q-band pattern in figure 4.2 was obtained by counter-staining the chromosomes with DAPI. Interestingly, analysis of some prometaphases of C1, C2 and C3 cosmids showed a visible signal on chromosome 6p21.31 which corresponds to the location of the MHC region. However, this signal on chromosome 6p21.31 was not reproducible in other experiments. It is not known if the 6p21.31 signal was due to a cross-reaction of the cosmids containing the Delta gene with LMP2 gene on the MHC. The FISH method used to map the Delta gene was also performed in collaboration with Dr. G. Senger at the Human Cytogenetics Laboratory at ICRF.

Following the finding that Delta was localized on chromosome 17, we decided to confirm this result by screening a flow-sorted chromosome 17-specific cosmid library with a Delta cDNA probe. Three cosmids positive for the Delta gene were obtained by using the hybridization conditions described above. The cosmid clones were called X17-1, X17-4, and X17-5, and they showed a similar pattern as a result from EcoRI restriction digestion.

The six cosmids obtained with Delta cDNA hybridization (C1, C2, C3, X17-1, X17-4 and X17-5) were then retested for the presence of the Delta gene by PCR. The PCR was performed by using 3 sets of Delta oligonucleotide probes (each set at 200-300 base pairs apart) that comprised the 5', 3' and middle regions of the Delta cDNA sequence. The size of the PCR from the middle and 3' regions was consistent with the presence of a short intronic sequence with a similar size in all the cosmids (data not shown).

4.4 Models for the genetic origins of the MB1, Delta and LMP genes

Amino acid sequence comparisons of proteasome β subunit genes (figure 3.3) reveals that MB1 and Delta differ significantly from each other (only 31.2% identity) as do LMP2 and LMP7 (only 31.6% identity), suggesting that the two pairs are not recent duplicates of each other. Considering the much stronger homology of MB1 and LMP7, and Delta and MB1, respectively, it seems likely that MB1 and LMP7 originated from one common ancestral gene, and Delta and LMP2 from another.
Figure 4.2: Chromosome localization of Delta gene by fluorescence in situ hybridization (FISH). Hybridization signals of Delta (red) are visible on chromosome 17 in the region 17p13 (arrows). Ideogram of chromosome 17 shows the localization of the Delta gene.
Further sequence comparison of proteasome subunits from different species provided candidates for the primordial loci for each pair in other species; the yeast gene PRE2 is most similar to MB1, and PRE3 is most similar to Delta (Figure 4.3) (Heinemeyer et al. 1993, Enekel et al. 1994). Sequence analysis shows that the similarity of the yeast PRE2 to the MB1 gene (64.9% identity) was higher than the similarity to the LMP7 gene (59.3% identity). Similar conclusions were obtained by comparing the yeast PRE3 sequence with Delta (54.4% identity) and LMP2 (equal to 44.8%). These data suggest that MB1 and Delta are the more ancient of the four human genes.

Molecular genetic studies on yeast have shown that most of the proteasome subunits are essential for cell proliferation, showing that deletions of the PRE2 and PRE3 genes were lethal (Heinemeyer et al. 1993, Enekel et al. 1994). The essential role of PRE2 and PRE3 for cell proliferation suggested that their human homologues, MB1 and Delta, have a similar constitutive housekeeping role in protein degradation. Early experiments have shown that LMP2 and LMP7 subunits are not essential for degradation, since human cells lacking the LMP genes grow normally, indicating that these genes are dispensable (Arnold et al. 1992, Momburg et al. 1992). Furthermore, it has been observed that the complexity of the proteasome subunit composition during evolution correlates with an increasing diversity in proteolytic specificities. Five different proteolytic activities have been found in vertebrates, proteasome as compared to only one in the archaeabacterial proteasome.

It is believed that all subunits of eukaryotic proteasomes have evolved from an ancient pair, one α, one β, as found in Thermoplasma α and β subunits. The fact that LMP7 and MB1 protein sequences share 68.1% identity indicates that these two loci are derived from a single branch of the proteasome subunits' evolutionary tree. A similar argument holds for the LMP2 and Delta pair which are 60.8% identical (Belich et al. 1994). The LMP7 and MB1 genes were presumably derived from a common primordial sequence more recently than they separated from other proteasome β subunits' sequences. Since species such as S. cerevisae have an MB1 counterpart (PRE2) which is slightly more distantly related to LMP7, it has been assumed that LMP7 was more recently derived from MB1 by duplication and divergence.
Conserved amino acids are defined according to the nature of their side-groups.

Short branched-chain amino acid group- A S T P G
Long branched-chain amino acid group- I V L M
C=O branched-chain amino acid group- N D E Q
HN branched-chain amino acid group- R K H
Aromatic branched-chain amino acid group- F Y W

Figure 4.3: Similarities of primary structures of yeast and human proteasome subunits. The source of sequences data are as follows: MB1 (Belich et al. 1994) Delta (DeMartino et al. 1991), LMP7 (Glynne et al. 1991), LMP2 (Kelly et al. 1991), PRE2 (Heinemeyer et al. 1993), PRE3 (Enekel et al. 1994).

A) Percentage of overall amino acid sequence identity between proteasome β-subunits were obtained by using the GAP molecular biology program. Sequences were compared in pairs. The percentage of identity was deduced by comparing the amino acid sequences of the mature form of each β-subunit, which are MB1 (60-263), LMP7.b (73-276), Delta (35-239), LMP2 (21-219), PRE2 (76-287) and PRE3 (1-193) (for the amino acid residues' location in each sequence, see below).

B) Comparison of amino-acid sequence of MB1 and LMP7 human proteasome β-subunits with PRE2 yeast subunit. The region corresponding to the start of the mature form of the proteins is indicated by an asterisk. Identical and conserved residues are shown in red and blue, respectively. Numbers listed on the left side are residue numbers. Conserved amino-acids are defined: A, S, T, P and G; N, D, E and Q; R, K and H; I, V, L and M; F, Y and W. Residues involved in the proteolytic activity of the protein are boxed. The first threonine at the cleavage point corresponds to the active-site nucleophile of the protein. The second conserved residue is Lys at positions 92 for MB1, position 106 for LMP7.b, and position 108 for PRE2, corresponding to the putative proton acceptor-donor residue. The two glycine-rich sequences are at positions 133-135 and 174-176 for MB1, 197-199 and 238-240 for LMP7b, and 205-207 and 246-248 for PRE2. Both positions are in direct proximity to the active-site residues.

C) Comparison of amino-acid sequences of Delta and LMP2 human proteasome β-subunits with PRE3 yeast subunit. For legend, see as above. The first threonine at the cleavage point corresponds to the active-site nucleophile of the protein, positions 35 for Delta, 21 for LMP2, and it is absent in PRE3 due to the short cDNA from published sequence. The second conserved residue is Lys at position 67 for Delta, 53 for LMP2, and 30 for PRE3, corresponding to the putative proton acceptor-donor residue. The two glycine-rich sequences are at positions 163-165 and 148-150 and 189-191 for LMP2, 25-27 and 166-168 for PRE3, and they are in direct proximity to the active-site residues.
Two models have been proposed to explain the genetic origin of the *LMP* genes (figure 4.4) (Belich et al. 1994): 1) The primordial genes, ancestral proteasome subunits, were initially linked and, after duplication and mutations, remained linked on 6p21 (the *LMPs*) but *Delta* and *MB1* split onto different chromosomes. 2) The primordial genes for *Delta* and *MB1* duplicated and mutated in different locations. Copies (the *LMPs*) were brought together to form a cluster of genes, with the *TAPs*, within the MHC. The *LMPs* share less sequence homology (amino-acid identity) with their respective yeast homologues, described above, than do *MB1* and *Delta*, and *MB1* and *Delta* encode constitutively expressed subunits, found in most tissues, whereas *LMP2* and *LMP7* are up-regulated in cells of the immune system. It seems reasonable to propose that the LMP loci arose later in the evolution of vertebrates than the *Delta* and *MB1* genes.

Would be interesting to ascertain whether this theory could be extended to the TAP genes. As the TAP genes are closely linked to the LMPs, they may have been a part of this duplicated region and could have homologues near to MB1 and Delta. On the other hand, it would be interesting to know if the cluster of TAP/LMP genes has been linked together in the MHC due to an evolutionary advantage by having the antigen processing and presentation genes linked.

Support for the notion that TAPs and LMPs had evolved in linkage to class I genes to provide a functional advantage is provided by analysis of the *cim* phenotype in the rat where different RT1.A alleles seem to be better provided with peptides from particular linked TAP2 alleles (Joly et al. 1994, Howard 1993). However, the TAP loci are far more polymorphic in the rat, compared to the human, and the advantage conferred by the different TAP alleles in the cim system may not apply to humans. It may be unnecessary to have alternative TAP alleles in view of the large number of class I genes in humans.

### 4.5 Conclusion

The eukaryotic proteasome has up to 15 different subunits that have been classified as either α or β subunits. After finding that two β subunits, LMP2 and LMP7, mapped in the MHC, several groups have been investigating where the other proteasome subunits are located. Three α-type subunits genes have been localized. HC2, PROS-27, and HC3 map to human chromosome 11q15.1, 14q13, and 6q27 respectively (Bey et al. 1993, Okumura et al. 1995). In addition to the LMPs, three other β-subunits have been localized. Chromosome mapping of MB1 and Delta were performed as described above, and they were localized respectively to chromosomes...
Figure 4.4: Models for the genetic origin of the LMP genes. Comparisons of similarities between proteasome subunits genes (Fig. 3.3) reveal that LMP2 and LMP7 differ significantly from each other. Similarly, MB1 and Delta are not recent duplicates of each other. It is suggested that MB1 and LMP7 have developed from one common ancestral gene, and Delta and LMP2 from another. Comparison of proteasome sequences from different species provides candidates for these primordial genes loci: the yeast PRE2 is more related to MB1, and PRE3 is more related to Delta, than they are to LMP7 and LMP2 respectively. Two models have been proposed to explain the observed linkage of the LMP genes within the MHC.

A) The primordial genes were initially linked and, after duplication and mutations, remained linked on 6p21 (the LMPs), but MB1 and Delta split onto different chromosomes.

B) The primordial genes for MB1 and Delta duplicated and mutated in different locations. Copies (the LMPs) were brought together to form a cluster of genes, with the TAPs, within the MHC. The LMPs share less sequence homology (amino-acid-identity) with their respective yeast homologues, described above, than do MB1 and Delta, and MB1 and Delta encode constitutively expressed subunits, found in most tissues, whereas LMP2 and LMP7 are up-regulated in cells of the immune system. It is hypothesized that the LMP loci arose later in the evolution of vertebrates than the MB1 and Delta genes.
14q11.2-12 and 17p13 (Belich et al. 1994). In addition, the β-subunit MECL-1 was localized on chromosome 16q22.1 (Larsen et al. 1993).

Thus, these preliminary results suggest that the genes for proteasome subunits are located on different chromosomes, and the different locations of these genes suggest the existence of some mechanism regulating their coordinated expression. However, a specific consensus region common to the upstream region of all the genes for proteasome subunits of various eukaryotes, such as yeast, Drosophila, and human, has yet to be identified. Therefore, further studies are necessary to clarify the mechanisms regulating the coordinated expression of the genes of the multiple subunit proteasome.

The MECL1 β subunit locus was discovered by analysing a cluster of genes on chromosome 16q22.1 that had apparently unrelated functions (Larsen et al. 1993). The expression of the MECL1 gene has been shown to be induced by γ-INF, which suggests a similar function to that of the LMPs. Recently, a new β subunit called Z has been cloned and sequence analysis showed high structural similarity to MECL1 (57.7% identical). γ-INF induction showed reciprocal expression of Z and MECL1, suggesting that along with the replacement of MB1 and Delta by LMP7 and LMP2, this event produces proteasomes that are more appropriate for antigen processing through the class I pathway (Tanaka et al., personal communication). Figure 4.5 shows a model of the 20S proteasome with proposed β subunit arrangement.

The finding that the proteasome recruits different subunits indicates that it participates not only in processes common to all organisms, but also in certain species-specific functions. It appears that the 20S complex acquires functional diversity by changing subunit composition in response to environmental stimuli, such as γ-INF. This increase in the number of β-subunits by duplication results in an increased functional diversity.

Sequence analysis showed that the similarities between MB1 and PRE2, and Delta and PRE3 are greater than those between MB1 and LMP7, and Delta and LMP2, leading to the conclusion that the closest homologues of the yeast genes are MB1 and Delta. The evolution of the LMP genes is interesting, considering that they are nonessential genes but with an apparent function in modulating the immune response. The LMPs presumably evolved from MB1- and Delta-like ancestral genes, most probably by gene duplication and mutation. Thus, the immune system may have
Figure 4.5: Model of the 20S proteasome. The particle is a cylindrical structure composed of four stacked rings; each ring has seven subunits in a pseudo-helical arrangement. The proteasome is composed of two types of subunits, designated α and β. The outer rings contain the α subunits and the inner rings the β ones. Each subunit can occur twice in the complex, leading to a structure comprising a complex dimer. The model proposes that MB1, Delta and Z subunits can be exchanged for LMP7, LMP2 and MECL1, respectively. Although the spatial relationships of the various subunits are not yet known, they were drawn adjacent to each other. Preliminary evidence suggests that mixed 20S particles can exist in cells expressing the six known β subunits. It is proposed that subunit exchange provides a mechanism for altering single components of a complex structure, to modulate its catalytic activity. Different subunit exchanges could be used in other tissues.
used preexisting proteasome genes to acquire a specific antigen processing function. Similarly, MB1 and Delta probably evolved from ancestral MB1-like and Delta-like genes, but it is still not known if the ancestral genes were in the same chromosome or not. Further studies, searching for new genes located close to MB1 and Delta, and the Z subunit, will be used to clarify questions about the evolution and function of the proteasome complex.
Chapter 5

Analysis of MB1 and Delta Proteasome Components

5.1 Summary

This chapter describes the characterization of MB1 and Delta proteasome components. The experiments were performed by using antisera raised against the C-termini of each protein. MB1 and Delta subunits are part of the 20S proteasome complex, and are processed from an immature to a mature form due to a cleavage at the N-terminus which results in loss of ~3kDa from Delta, and ~8-12 kDa from MB1. The molecular weight and isoelectric point of the mature forms of MB1 and Delta are consistent with those predicted by analysis of their protein sequences considering the Thr^1 residue as the cleavage point for the mature forms. Cleavage of the precursor form probably occurs during the assembly process and once the proteasome complex is formed the subunits cannot be exchanged. Treatment of γ-IFN decreases the amount of detectable MB1 and Delta proteins in western blots, at the same time that it increases LMP2 and LMP7 expression, suggesting a reciprocal expression of these subunits. Analysis of the mutant cell, .174, showed that the MB1 and Delta protein levels are up-regulated most probably by compensating the absence of the LMP2 and LMP7 genes. The levels of expression of MB1 and Delta proteins were higher than expected from the northern data, suggesting that their protein expression are posttranscriptional regulated. These data show that MB1 and Delta subunits are compensating the absence of LMP7 and LMP2 in the mutant cell, and are the housekeeping proteasome subunits that fail to bind to proteasome and are degraded when LMPs are present.

5.2 Introduction

It has been previously discussed that LMP2 and LMP7 proteasome subunits are encoded adjacent to the TAP1 and TAP2 genes, and they are all up-regulated by γ-IFN, a major immunomodulatory cytokine (Yang et al. 1992). Thus, the proteasome composition changes upon induction with γ-IFN, leading to the appearance of some subunits and disappearance of others. The expression of LMP2 and LMP7 is strongly induced by γ-IFN, resulting in the increased production of their
precursor forms which are processed upon incorporation into the proteasome (Glynne et al. 1993, Früh et al. 1992, Frentzel et al. 1993). However, it is not known which subunits were replaced by the LMPs and how the additional subunits disappeared from the complex.

Recently, it was also shown that γ-IFN changed the specificity of the proteasome for peptide degradation, increasing their activities for cleavage of peptide bonds on the carboxyl side of basic and hydrophobic amino-acid residues (Gaczynska et al. 1993, Driscoll et al. 1993, Boes et al. 1994). Thus, proteasomes containing the LMP subunits favour production of peptides with carboxyl termini suitable for the appropriate pocket in the HLA class I peptide binding groove. In addition, cells of the mutant human B-cell line 174 expressed class I molecules at normal levels at the cell surface, but showed different peptidase activity from wild-type proteasomes, favouring cleavage after acidic residues (Driscoll et al. 1993, Gaczynska et al. 1993).

Chapter 3 describes the cloning of MB1 and Delta genes encoding new proteasome subunits, which were shown to be highly related to LMP7 and LMP2, respectively. MB1 and Delta genes were the most probable candidates to compensate for the absence of the LMPs in the mutant cell.

In order to investigate MB1 and Delta protein expression levels, and understand the relationship of these proteins to the proteasome complex and to the LMP subunits, we raised antisera against the C termini of the predicted protein sequences. We chose the C terminus region, of MB1 and Delta proteins, due to their low homology compared to other β type subunits, and we would expect little chance of a cross reactivity with other subunits. The same procedure was previously used to obtain antiserum against LMP2 and LMP7 subunits, and they were successful on immunoprecipitations and western blots (Glynne et al. 1991).

Peptides corresponding to the C-terminal 14 amino acids of the predicted MB1 protein sequence, and to the C-terminal 15 amino acids of Delta, were synthesised chemically. The MB1 peptide corresponds to the DNVADLHEKYSGST sequence, and Delta to the GDQIPKFAVTLPAPA sequence. In order to find if any other protein could cross-react with MB1 or Delta, peptide sequences were checked against the following data banks: PDB, SwissProt, SPupdate, PIR, GenPep and GPupdate, and no matches were observed.

The production of the antisera is described in section 2.4.1. Rabbit pre bleeds and subsequent test bleeds were tested on lysates, at various dilutions, by
western blots and immunoprecipitations from different cell types. Western blots of various cytoplasmic cell lysates with anti-MB1 polyclonal antiserum gave bands of ~22.5 kDa and ~30-35kDa that correspond to the mature and immature forms of the MB1 protein, respectively. The anti-Delta polyclonal antiserum gave bands of ~25.3 kDa and ~21.9 kDa for the mature and immature forms, respectively. The above experiments are described in detail in the following sections.

5.3 MB1 is a subunit of the 20S proteasome complex

Chapter 3 describes the isolation of the MB1 cDNA clone by using low stringency hybridization with the LMP7 probe. Computer-assisted homology analysis showed that the primary structure of the subunit MB1 had significant similarities with those of other proteasome subunits, which themselves have high inter-subunit similarity, that can be classified into two subfamilies with high similarity to an archaeabacterial proteasome α- and β-subunit, respectively. As shown in figures 3.2 and 3.3 the MB1 subunit belongs to the β-type family and shows remarkable similarity to subunit LMP7 (68.1% identical). Thus, the sequence data itself suggests that MB1 protein is most probably part of the proteasome complex.

In order to further characterize the MB1 subunit as part of the proteasome complex, the anti-MB1 antiserum was used for immunoblot analysis in two-dimensional polyacrylamide gel electrophoresis (PAGE), and this showed that MB1 is a component of the proteasome extracted from placenta (figure 5.1a, b). The antiserum bound specifically to a single protein spot that has the isoelectric point (Ip) and relative molecular mass (Mr) predicted for the protein encoded by the MB1 cDNA, such as 8.76 Ip and 22.5 kDa respectively. For 2D-gel method see page 82a.

Immunoprecipitations from total cell lysates, from B-cells and non B-cells, were analysed on non-denaturing gels in order to test the ability of the anti-MB1 antiserum to recognize the native form of the MB1 protein. The metabolically labelled cell extracts were immunoprecipitated with anti-MB1 antisera and no specific bands were visualized on autoradiography of the gel. We concluded that our antisera raised against the C-terminal MB1 sequence does not recognize a native form of this subunit.

To assay the specificity of the antiserum, lysates were made from different cell lines, such as fibroblast, breast carcinoma, choriocarcinoma placenta, ovarian tumour, cervix carcinoma and B cell. Western blots of these lysates with anti-MB1 serum gave a band of ~22.5 kDa and a second band of ~30-35kDa (figure 5.2). The
Figure 5.1: Two-dimensional PAGE analysis of proteasomes purified from human placenta and immunoblot of MB1 subunit. 
(A) Schematic representation of the 20S proteasome subunit pattern after staining with Ponceau-S. The numbers correspond to the following subunits: 1-Zeta, 2-Delta, 3-HC8, 4-HN3, 5-Z, 6-HC10, 7-HC9, 8-Iota, 9-HC7, 10-HC2, 11-HC3, 12-XAPC7, 13-HC5 and MB1. (B) Immunoblot analysis using anti-MB1 polyclonal antibody. Proteasomes were purified from human placenta by affinity chromatography on immobilized antibody MCP21. The two-dimensional gel electrophoresis results from non-equilibrium pH gradient electrophoresis (NEPHGE) followed by SDS-polyacrylamide gel. Proteins were transferred to membranes and stained with Ponceau-S before incubation with anti-MB1 antiserum. Detection was performed as described in section 2.5.6.
Figure 5.2: Assay of the specificity of the anti-MB1 antiserum. Western blots of lysates from the following cell lines were probed with anti-MB1 at a concentration of 1:500; W138 (normal fibroblast), MCF7 (breast carcinoma), JAR (choriocarcinoma placenta), 1847 (ovarian tumour), ME180 (cervix carcinoma), and WJR076 (B cell). 30µg of cytoplasmic cell lysates were evenly loaded per lane. The anti-MB1 antiserum recognized a band of ~22.5 kDa, and band of ~30-35kDa. The two protein bands detected by western blot with the anti-MB1 correspond to the mature and immature forms of this subunit.
two protein bands detected by western blot with the anti-MB1 correspond to the mature and immature forms of this subunit, respectively.

The MB1 subunit was further characterized as part of the proteasome complex by immunoelectron microscopy of proteasome-antibody conjugates, and it was localized in the two juxtaposed inner rings of the human proteasome complex. Detailed results of MB1 immunoelectron microscopy localization are described in chapter 6.

Recently, other groups have characterized the MB1 subunit as part of the 20S proteasome. The MB1 subunit, referred to as X, was identified from 2-D gels of 20S proteasome (Akiyama et al. 1994). X subunit was isolated and proteolytically fragmented, followed by subsequent purification by high-performance liquid chromatography (HPLC), and amino-acid sequence of the X fragments. A full length complementary DNA encoding X (MB1) was isolated from a human cDNA library, and it was shown to be identical to MB1, thus re-confirming that MB1 is part of the proteasome complex. The MB1 subunit has also been referred to as the epsilon chain, which protein sequence was previously obtained by similar proteolytical fragmentation and HPLC analysis (Lee et al. 1990, Tamura et al. 1995). A short peptide sequence, corresponding to the 23 N-terminal amino acids of the epsilon chain, showed 21 residues, out of 23, similar to the MB1 sequence, suggesting that MB1 is similar but not identical to the epsilon chain. Further analysis is necessary to clarify if MB1 and epsilon are the same protein, but it is likely that the differences observed are due to sequencing artifacts or to polymorphism.

5.4 Delta is a subunit of the 20S proteasome complex

Delta was originally identified from purified human erythrocyte proteasomes (DeMartino et al. 1991). The subunit was isolated by 2D-PAGE, transferred to filters and stained. Peptides were then purified and subunit Delta yielded sequence information by Edman degradation. Degenerate PCR oligos were employed as templates to synthesize products that were further used to screen cDNA libraries. The Delta protein sequence was obtained from a cDNA clone, and analysed for Mr and Ip, which corresponded to its position on the 2D-PAGE. These data demonstrated that Delta is part of the 20s proteasome complex.

As discussed in chapter 3, subunit Delta was the most obvious candidate to replace LMP2 in the mutant cell .174. In order to text this hypothesis we cloned subunit Delta using PCR templates and raised an antiserum against the C-terminal
region of the protein. Immunoprecipitations using anti-Delta antiserum were carried out on labelled cells, and these showed precipitation of a stack of proteins between 20 and 30 kDa, as expected from precipitation of the 20S proteasome complex (Figure 5.3 a, b).

A direct comparison between anti-Delta and anti-proteasome precipitations was made. 20S proteasomes, from .174 and 721 cell lines, were precipitated using MCP21 antiserum (which recognizes a proteasome α-subunit, kindly provided by Dr. K. Hendil), and anti-LMP7 polyclonal previously described (Glynne et al. 1994). The above immunoprecipitations were performed using lysates from native conditions and they showed similar migration on SDS-PAGE (Figure 5.3 a).

In order to determine which of the bands in the proteasome stack was Delta, anti-Delta precipitates from metabolically labelled cell lines were tested on denatured lysates prepared by adding SDS to 1% and heat at 95°C for 5 minutes. The anti-Delta antiserum was tested on .174 mutant cell and 721 wild-type, and separated on SDS PAGE. Anti-Delta recognized a protein of ~22.5 kDa, which corresponds to the mature form of the protein, and a ~25kDa band that most probably corresponded to the unprocessed form of Delta (figure 5.3 b). The data showed that Delta is expressed at higher levels in the mutant cell compared to wild type, as expected from the absence of LMP2 in .174. The simplest interpretation of the above results is that anti-Delta recognizes a ~22.5 kDa protein (this band was very faint in the original picture and due to photographic limitations it is not able to be visualised) that is part of a large complex, and a longer form of the same protein that corresponds to the unprocessed Delta. As shown before, the Delta antiserum was raised against the C-terminus of the protein. Thus, as we expected, it recognized both forms of the same protein since the cleavage point for the mature forms is located in the N-terminal region.

The specificity of the Delta antiserum was further characterized on western blots, and lysates were made from different cell lines, such as fibroblast, breast carcinoma, choriocarcinoma, ovarian tumour, cervical carcinoma and B cell. Western blots of these lysates with anti-Delta gave a band of ~22kDa and a second band of ~25kDa (figure 5.4). The two protein bands detected by western blot with the anti-Delta antiserum correspond to the mature and immature forms of this subunit. Further analysis of the Delta protein are described in chapter 7.
Figure 5.3: Immunoprecipitations of proteins by the anti-Delta antiserum.
(A) Anti-Delta precipitates a stack of proteins between 20 to 30 kDa, similar to those precipitated with anti-LMP7 antibody, and with anti-α subunit MCP21 antibody. Lanes 1, 2, and 3; precipitations were carried out as described in Material and Methods. Cells were incubated with [35S]-methionine for 1 hour. .174 cell lysates were precipitated with either MCP21 (anti-proteasome monoclonal, 1/100 dilution) lane 2, or anti-Delta antiserum (1/100 dilution) lane 1. Lane 3 shows cell lysate of 721 immunoprecipitated with anti-LMP7 antiserum (1/100 dilution).
(B) Precipitation of Delta after dissociation of the proteasome. .174 (lanes 1 and 2) and 721 (lane 3) were precipitated with anti-Delta antiserum, both with (lanes 2 and 3) and without (lane 1) prior dissociation of the proteasome by heating to 95°C for 5 minutes in 2% SDS. Anti-Delta was used at a concentration of 1:100. Two bands of ~22 and ~25 kDa are specifically precipitated by anti-Delta after SDS treatment. The 22kDa band comigrates with the Delta band in the proteasome stack.
Figure 5.4: Assay of the specificity of the anti-Delta antiserum. Western blots of lysates from the following cell lines were probed with anti-Delta at a concentration of 1:1000; W138 (normal fibroblast), MCF7 (breast carcinoma), JAR (choriocarcinoma placenta), 1847 (ovarian tumour), ME180 (cervix carcinoma), and WJR076 (B cell). 30µg of cytoplasmic cell lysates were loaded per lane. The anti-Delta antiserum recognized a band of ~22kDa, and a band of ~25kDa in some tracks. The two protein bands detected by western blot with anti-Delta correspond to the mature and immature forms of this subunit.
Additional evidence that the Delta subunit is part of the proteasome complex is shown by immunoelectron microscopy of proteasome-antibody conjugates. Delta is localized in the inner rings of the human proteasome complex. Detailed results of the immunoelectron microscopy localization are described in chapter 6.

5.5 Delta is synthesized as a proprotein and is post-translationally processed

It has been previously shown, in archaebacteria and eukaryotes, that propeptides are post-translationally removed from β subunits. LMP2 and LMP7 subunits are synthesized as proproteins and post-translationally processed (Glynne et al. 1993, Frue et al. 1992, Frentzel et al. 1993). Similar results were observed for T. acidophilum and E. coli β subunits showing that 8 amino acid propeptides are removed by the post-translational process. Other proteinases, such as the E coli ATP-dependent proteinase Clp, α-lytic proteinase, subtilisin, have been shown to cleave off proregions autocatalytically during activation. These propeptides can vary in size, and are essential for folding of an active mature proteinase and have inhibitory activity against the mature proteinase (Baker et al. 1993).

Recent studies with the mouse LMP2 and LMP7 subunits showed that proteasome assembly takes place via distinct precursor particles, called 11S and 13S. These precursor particles contain unprocessed protein subunits and subunit processing and maturation appear to be concomitant events (Frentzel et al. 1994). The processing, and possibly also modification of subunits, takes place in these precursor complexes. Once processing is complete, two proproteasome molecules, each representing one half of the proproteasome are combined to form a 20S enzyme complex. The 13S complexes contain many α-type subunits, supporting the observation that processing of the β subunits from archaebacteria proteasome requires the presence of α subunits (Zwickl et al. 1992).

Immunoprecipitations and western blot analysis from various cell lysates showed that two proteins were recognized by the anti-Delta antiserum, and we predicted that they would correspond to the processed and unprocessed forms of Delta (figures 5.3 and 5.4). However, we could not rule out the possibility that a different translation of the Delta could start at an internal methionine. Figure 5.3b shows that the two proteins precipitated by anti-Delta were observed from the lysates which had been boiled in SDS to disrupt non-covalent protein:protein interactions. The 22kDa
To find out whether there was a precursor/product relationship between these two proteins, a pulse-chase experiment was performed. This showed that the ~25kDa protein was processed into the ~22kDa protein with a half-life of 5 hours (figure 5.5). The 22kDa protein was stable after 48 hours chase time. Thus, as for the LMPs, Delta is also processed from a propeptide form by cleavage of a 3kDa peptide chain. For pulse-chase experiments see section 2.5.3 and figure 5.5 legend.

Similar pulse-chase experiments could not be performed for the MB1 subunit, since the antiserum did not recognize any protein by immunoprecipitation. Two-dimensional gel electrophoresis (figure 5.1) showed MB1 as part of proteasome complex, and it is very likely that the two proteins recognized by the anti-MB1 reagent on western blots, figure 5.2, correspond to the unprocessed and processed forms of this protein.

5.6 MB1 and Delta are up-regulated in LMP-deficient cells

As described in previous chapters, the role of proteasome in antigen presentation, until recently, had not been generally accepted because of the finding that antigen presentation took place in the absence of LMP2 and LMP7. This conclusion came from experiments using the mutant cell .174, deficient in class I assembly and presentation of peptides. Two groups showed that transfection of human TAP1 and TAP2 genes into .174 cells restores class I assembly and peptide binding to almost similar levels to those observed in the wild type parent 721 (Arnold et al. 1992, Momburg et al. 1992). Thus, it indicated that LMP2 and LMP7 were not essential for the production of antigenic peptides. However, these studies were widely interpreted as indicating that the LMP subunits (and by association the proteasome) are not important for class I antigen presentation.

Momburg et al., in order to study the effect of the LMPs on individual epitopes transfected rat TAP cDNAs into the LMP negative T/B cell fusion mutant line, T2 (Momburg et al. 1992). Again, class I levels were increased to near wild type levels. Presentation of a minor histocompatibility antigen, HA-2, to a CTL clone was restored. However, the ability of the transfected cell line to present viral peptides was less than that of the parental line T1 (50% presentation for influenza infection, 30% for influenza protein expressed from a vaccinia recombinant construct). It was
Figure 5.5: Delta is processed from a ~25.3kDa precursor. The first two lanes show precipitates from native lysates, and the last three lanes show precipitates that were dissociated before precipitation by using 2% SDS and heating at 95°C for 5 minutes. 174 cells were labelled for 30 minutes, and lysates were made at different time points. Anti-Delta was used at a concentration of 1:100. Chase times after labelling are shown above each lane in hours. Lane 1 shows precipitate from MCP21 monoclonal, and lanes 2 to 5 from anti-Delta antisera.

Mature form of the protein is indicated by '*'.
argued that this decrease in presentation could be due to the absence of the LMPs in these cells.

We reasoned that the subtle phenotype of the TAP-transfected .174 mutant cells might reflect the expression in these cells of the alternative proteasome subunits. The most obvious candidates were MB1 and Delta. In order to test this hypothesis we looked at the expression of MB1 and Delta in the mutant .174, as compared to expression in the parental line, 721. Both proteins were highly expressed in the mutant .174 but were barely detectable in the wild-type line (figure 5.6). Thus, we concluded that the lack of a gene, or genes mapping within the deletion found in the .174 cell was associated with the increase in expression of MB1 and Delta. We then suggested that the expression of the LMPs in wild-type B cells, causes the diminution of the two otherwise constitutive proteasome subunits, MB1 and Delta. These data reinforced our hypothesis that MB1 and Delta were replacing LMP7 and LMP2 in the mutant cell, and in conjunction with the western blot analysis shown in the previous sections, MB1 and Delta were most probably the housekeeping proteasome subunits.

5.7 Levels of MB1 and Delta protein decrease when LMP7 and LMP2 are up-regulated by interferon-γ

γ-interferon is one of the cytokines secreted by activated T lymphocytes and natural killer cells, and it functions as an immunomodulatory and antiviral cytokine in several cell types, such as macrophages, helper T cell, natural killer and B cells (Farrar and Schreiber 1993). As previously discussed, this cytokine has the ability to increase the cellular content of LMP2 and LMP7 and thus to change the subunit composition of the 20S and 26S complexes (Yang et al. 1992).

In order to investigate whether γ-IFN alters the expression of the various proteasome components, we used western immunoblotting to analyse a panel of cell lines, both with and without addition of γ-IFN (figure 5.7). To determine how LMP2, LMP7, MB1 and Delta were affected by this cytokine the cells were treated for 3 days with 300 units of γ-IFN. Western blots from the following lysates were analysed: 721, .174, HT1080 (fibrosarcoma), Lovo (colon carcinoma), Raji (Burkitt's lymphoma), HeLa (cervical carcinoma), and J-6 (T cell). As a result, up-regulation of expression of the LMP proteins was seen in response to γ-IFN. The induction of the LMPs in the B cell line 721 was small, due to the high constitutive levels in these cell lines. As we expected, MB1 and Delta were expressed at high levels in cells that had not been treated with cytokine, but were present only in small amounts in cells treated with γ-IFN. On the basis of these findings, we propose that MB1 and Delta are
Figure 5.6: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in 721 and 174 cell lines. Extracts of 721 (lanes 1, 3, 5, and 7) and 174 (lanes 2, 4, 6, and 8) were used. The asterisk on the right of lane 8 indicates the position of the mature form of the protein, and the upper band corresponds to the precursor form of Delta. 20μg of cytoplasmic lysate were loaded per lane.

Mature form of the protein is indicated by ' * '.

-30 kD

-21.5 kD
Figure 5.7: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in cell lines cultured with and without addition of γ-interferon. Protein expression after γ-interferon treatment is indicated by "+", and cells not treated are indicated by "-". Cells were treated with 300 units of γ-interferon for 3 days prior to lysis. Cell lines: 721 (B cell), .174 (mutant B cell), HT1080 (fibrosarcoma), Lovo (colon carcinoma), Raji (Burkitt's lymphoma), HeLa (cervical carcinoma), and J-6 (T cell). The asterisk on the right shows the positions of the processed proteasome subunits. Analysis of protein pattern from anti-Delta and anti-MB1 antisera shows that under γ-interferon induction these proteins have decreased expression. Analysis of the protein pattern from anti-LMP2 and anti-LMP7 antisera shows that most cell lines under γ-interferon induction increased expression of these proteins.
constitutive proteasome subunits that in the presence of γ-IFN are substituted by LMP7 and LMP2.

In order to understand the mechanism of the apparent switch between MB1/Delta and LMP2/LMP7 gene expression, we compared mRNA levels for the four genes in both the 721 wild-type and .174 LMP-deficient B cell (figure 3.6). Both MB1 and Delta showed a 3 to 4 fold increase in mRNA levels as compared to 721 cells. Quantification of western blots (figure 5.6) showed that MB1 and Delta protein levels were increased by 10 to 20 fold in the mutant cell. It appears, from this analysis, that the expression of the LMPs leads to a slight reduction in the mRNA levels of MB1 and Delta, and post-transcriptional control may be responsible for most of the decrease seen at the protein level. One possible mechanism would be that the degradation of MB1 and Delta subunits follows their displacement from, or lack of incorporation into, the proteasome complex in the presence of the LMP subunits.

Supporting these findings, another group showed that X (identical to MB1) and Y (identical to Delta) protein levels were decreased when LMP2 and LMP7 were up-regulated by γ-IFN treatment (Akiyama et al. 1994). Similarly altered composition of proteasome was noticed in HeLa cells treated with γ-IFN as determined by 2Dgel electrophoresis. Two of the spots that increased in intensity were identified as the LMPs (Yang et al. 1992).

5.8 Conclusion

This chapter described the characterization of MB1 and Delta proteins. This characterization was possible primarily by producing antisera against both proteins. Using the MB1 antiserum it was possible to show that this protein is incorporated into proteasome complexes isolated from human placenta analysed on 2D-PAGE. Using Delta antiserum in conjunction with other anti-proteasome antibodies it was possible to show that they immunoprecipitate similar 20-30 kDa complexes.

Sequence analysis from different proteasome β-subunits, as discussed in chapter 3, indicated that the N-termini of the proteins are cleaved off from a propeptide. As expected, the Delta subunit was shown to be synthesized with an N-terminal extension which is cleaved off (with a half life of 5 hours), and is not present in the mature form of the protein.
Analysis of MB1 and Delta expression in .174 mutant cells strongly suggests that they are the candidates to replace LMP2 and LMP7. Moreover, coincident reciprocal expression of LMPs, and MB1 and Delta in γ-IFN treated cells reinforces the hypothesis that incorporation of the LMPs replaces the MB1 and Delta subunits. Thus, the incorporation of the LMPs into the proteasome must afford some advantage to the organism in terms of antigen processing and cell lines that lack the LMPs express MB1 and Delta subunits, allowing sufficient peptides to be generated, so that class I expression and antigen presentation are not significantly impaired. Furthermore, the coincident expression of the LMPs, TAPs and surface class I expression after IFN-γ treatment (Yang et al. 1992), suggest that all three classes of genes have related functions. It appears that incorporation of the LMPs into the proteasome after IFN-γ treatment may bias proteasome activity towards degradation products better suited to class I presentation (Gaczynska et al. 1993, Driscoll et al 1993, Boes et al. 1994).
Chapter 6

Localization of LMP2, LMP7, MB1, and Delta Subunits in Proteasomes.

6.1 Summary

This chapter analyses the quaternary structure and subunit topography of the 20S proteasomes from 721 and .174 B-cell lines determined by immuno-electronmicroscopy using LMP2, LMP7, MB1, and Delta polyclonal antibodies. The experiments described in this chapter were performed in collaboration with Dr. R. Newman at the Protein Structure Laboratory at ICRF.

The data indicate that proteasomes from 721 and .174 cell lines have a similar quaternary structure shown by negative stain electronmicroscopy. Antibodies to LMP2, LMP7, and Delta proteins bound to inner ring, β-type 20S proteasome subunits, from the 721 cell line. Thus, proteasome complexes from 721 cells can have the mature forms of either LMP2 or Delta, or both subunits. In addition, immunoprecipitation experiments using Raji B-cell line showed that LMP2 and Delta subunits could exist as part of the same 20S complex. Furthermore, antibodies to MB1 and Delta proteins bound to the 20S proteasome inner ring in .174 cells.

The data outlined in this chapter are consistent with the idea that MB1 and Delta/ LMP2 and LMP7 occupy similar positions on the inner rings of the proteasome.

6.2 Introduction

Several studies using electron microscopy have shown that the quaternary structure of 20S proteasome complexes is conserved from Thermoplasma to humans (Puehler et. al. 1994). The 20S proteasome appears as a ring with a diameter of 11nm, or in side-on projection as a rectangle of 11nm x 15nm (Kopp et al 1986). In addition, image analysis has provided the location of several centres of density, later confirmed by crystal structure as the proteasome subunits (Baumeister et al. 1988, Lowe et al 1995). Immunoelectronmicroscopy of the prokaryotic proteasome revealed that α subunits were located in the two end disks, whereas the two inner
rings were formed by β subunits (Grziwa et al. 1991). However, the exact stoichiometry of the subunits and their location in the complex structure of the eukaryotic proteasome particle was not determined.

The crystal structure resolution of the 20S proteasome confirmed the previous analysis, showing that it is a cylindrical particle composed of four stacked rings or disks (Löwe et al. 1995). The two inner rings consisted of seven β subunits each and the two outer rings consisted of seven α subunits each, with stoichiometry α7, β7, β7, α7. The 20S proteasome has a narrow channel that controls access to the inner compartment.

Two kinds of α-type subunits have been localized in the end disks of human proteasome (Koop et al. 1993). Each kind of subunit occurred twice in a proteasomal complex, once in each of the two end disks, suggesting that the proteasome is a complex dimer with its general architecture similar to that of the archebacterial complex in that α-type subunits build the outer disks whereas the β-type subunits form the inner rings. This hypothesis, that proteasomes are made of two identical halves, has been recently confirmed by observing that the β type subunit HsN3 was present in two copies, one in each ring (Kopp et al. 1995). Using electron microscopy, the binding sites of the monoclonal antibody with specificity for the subunit HsN3 have been located in the two juxtaposed inner rings of the 20S complex. The experiments performed so far have not yet defined the absolute configuration of the arrangement for each subunit in the proteasome.

6.3 721 and .174 cells have similar 20S proteasome structure.

Proteasomes from 721 and .174 cell lines were isolated and affinity purified from large scale cell cultures, described in section 2.5.9, in order to analyze and compare the quaternary structure of their 20S proteasomes.

The three-dimensional structures of proteasomes isolated from 721 and .174 cell lines were analysed and compared by electron microscopy. On electron micrographs of negatively stained speciemens from 721 and .174 cell lines the cylindrical 20S proteasome showed two projections (figure. 6.1 a,b). Some 20S proteasomes are in a rectangular side-view orientation, with four characteristic striations, while other proteasome are in an end-on, annular view, orientation. The quaternary structure of the 20S proteasomes from 721 and .174 were similar at this level. Immunoaffinity procedure was performed as described in section 2.5.10.
Figure 6.1: Electron micrographs of negatively stained 20S proteasomes from (A)721 and (B).174 cell lines. Proteasomes are seen in projection, in either a 'side-on' rectangular view, showing four rings or an 'end-on' annular view, showing a central cavity. Scale bar equivalent to 15nm. Aliquots of proteasomes, from both cell lines, were collected on air-glow-discharged carbon coated 400 mesh copper grids and negatively stained by floating on a drop of 1% (w/v) uranyl acetate for 20s. Excess stain was blotted and the grid was dried in air prior to examination in a Zeiss 10CR transmission electron microscope at 100kV.
6.4  721 cells have LMP2, LMP7 and Delta proteasome subunits localized in the inner rings.

To further characterize the proteasomes isolated from 721 cell lines, the structural distribution of the LMP2, LMP7, MB1 and Delta subunits was determined by immuno-electronmicroscopy of whole proteasomes, using antibodies specific for the four subunits. The knowledge of the arrangement of the various subunits in the complex helps to elucidate the molecular basis of the catalytic functions performed by the proteasome. The 20S proteasomes used for subunit localization studies were purified by using an ammonium sulphate fractionation method (section 2.5.11).

Figure 6.2 shows electronmicrographs of negatively stained 721 proteasomes incubated mixed in equimolar proportion with affinity purified polyclonal antibodies for LMP2, LMP7, MB1 and Delta subunits (for method see page 86). Antibodies to LMP2, LMP7, and Delta bound to inner rings of 20S proteasomes from the 721 cell line. No labeling of 20S proteasomes from 721 cells by anti-MB1 antibodies was observed.

Interpretation of images of antibody labeled molecules requires some knowledge of the three dimensional structure of the molecule. Proteasomes joined together by a single antibody molecule have higher degree of flexibility, giving a variable arrangement of the two proteasomes on the grid. The majority of the 20S proteasome-antibody complexes show the proteasome side-by-side with the antibody molecule attached to the two juxtaposed inner rings (figure. 6.2 a, b and d). Occasionally antibodies are seen joined to one proteasome in the side-on-orientation and one in the end-on orientation (figure 6.2 b). Here, too, the antibody appears to bind to the central part of the proteasomes seen side-on (figure 6.2 b). Only, occasionally, the antibody appears bound to two end-on view proteasomes (figure 6.2 e). The rectangular 'side-on' or annular 'end-on' 20S proteasomes have an immunoglobulin molecule linked to the central region. Antibodies can occur as links between similar subunits on neighboring 20S proteasomes.

Western blot analysis of the 721 cell line, previously shown, indicates that the Delta subunit is expressed at low levels. However, it is not known if proteasomes can have LMP2 and Delta subunits in the same complex. It could be that there are three subsets of proteasomes, one containing LMP2 and LMP7, other containing LMP7 and Delta, and other containing LMP2, LMP7 and Delta subunits. The later one could probably be formed by the assembly of two different proteasome halves.
Figure 6.2: Electron micrographs of negatively stained IgG labelled 20S proteasomes from 721 cells. 20s proteasomes from 721 cells labelled using specific antiserum for the β-type subunits (A, B) LMP7, (C) LMP2, (D) Delta, and (E) Delta. IgG molecules can be seen linking proteasomes together. (B) Two proteasomes from 721 cell with 'end-on' view are linked together and are in turn linked centrally to two 20S proteasomes with 'side-on' view. One of these is further linked to another 20s proteasome. In (D) and (E), anti-Delta antibodies bound to 20S proteasomes from 721, the latter in an 'end-on' view, are joined via the inner rings of the proteasomes. Proteasomes from 721 cells and antiserum at a molar ratio of 1:50 were mixed in a volume of 0.1 ml 20 mM Tris-HCl buffer, pH 7.5 plus 20 mM NaCl and incubated for 1 h at 22°C. Electron micrographs of negatively stained material were performed as described in figure 6.1. Scale bar equivalent to 15nm.
To further investigate if LMP2 and Delta subunits are part of the same proteasome complex, metabolically labeled Raji B-cells were used to visualize if the mature forms of both proteins could be associated in the same complex (figure 6.3 a, b). Labeled cell lysates were first precipitated with monoclonal antibody, MCP21 (anti-α subunit), or LMP2, and then re-precipitated with one of the following antiserum, LMP7, LMP2, Delta, or MB1, as described in section 2.5.4 One aliquot was mock precipitated without the addition of a second antibody, used as a control to check for the presence or absence of the first antiserum used, MCP21 or LMP2.

Figure 6.3 a shows that Raji proteasomes precipitated with MCP21 contained LMP2, LMP7 and Delta subunits. The presence of MB1 subunits could not be detected since the MB1 antiserum did not work on immunoprecipitations. These data showed that LMP2 and Delta were associated with proteasome α-subunits precipitated by MCP21. However, when anti-LMP2 was used in the first precipitation, Delta was co-precipitated (Fig. 6.3 b). Thus, in Raji cells LMP2 and Delta were associated in the same 20S complex. It is not known if LMP2 and Delta can be associated in the 20S complex in cells other than Raji.

6.5 .174 cells contain MB1 and Delta proteasomes subunits localized in the inner rings

Proteasomes isolated from .174 cell lines where characterized by the structural distribution of the MB1 and Delta subunits. Subunit distribution was determined by immuno-electronmicroscopy of whole proteasomes, using anti-MB1 and anti-Delta antiserum. Antibodies to MB1 and Delta subunits bound to 20S proteasome inner rings isolated from .174 cells, confirming that these subunits are part of the complex. The proteasomes were isolated by the ammonium sulphate fractionation method described in section 2.5.11

Figure 6.4, a to d, are examples of electronmicrographs showing negatively stained .174 proteasomes mixed in equimolar proportions with Delta or MB1 polyclonal antibodies. Figures 6.4 c and d show proteasome side-by-side with an anti-MB1 antibody molecule attached to the two juxtaposed inner rings. In figure 6.4 b the Delta antibody appears bound to two end-on view proteasomes
Fig. 6.3: LMP2 and Delta subunits are part of the same 20S proteasome complex in Raji cells. Raji B-cells were preincubated in methionine-free RPMI 10% FCS for 1 hour. [35S]-Methionine. Labelled cell lysates were first precipitated with (A) monoclonal antibody MCP21 (anti-α subunit), or (B) LMP2. The proteins from (A) and (B) precipitation were eluted from protein A sepharose beads by heating to 95°C in 2% SDS and the eluate divided into six aliquotes (lane 1 to 12). One aliquote from (A, lane 1) MCP21 precipitation, and one from (B, lane 7) LMP2 precipitation, were mixed directly with sample buffer, showing similar 20S subunit patterns from analysed proteasomes. The other five aliquotes (A lanes 2 to 6; and B lanes 8 to 12) were precleared with Staph A to remove any undenatured MCP21 or anti-LMP2. Four aliquotes for MCP21 and four for LMP2 precipitates, were then re-precipitated with one of the following antiserum, LMP7 (lanes 2 and 8), LMP2 (lanes 3 and 9), Delta (lanes 4 an 10), or MB1 (lanes 5 and 11). The fifth aliquotes (lanes 6 and 12) were mock precipitates without the addition of a second antibody, used as a control to check for the presence or absence of the first antiserum used, MCP21 or LMP2.
Figure 6.4: Electron micrographs of negatively stained IgG labelled 20S proteasomes from .174 cells. 20S proteasomes from .174 cells labelled using specific antisera for the β-type subunits (A) and (B) Delta, and (C) and (D) MB1, showing binding of antibody to the proteasome inner rings. Scale bar equivalent to 15nm. Schematic representation is shown in the left of figure (A) and in the bottom of panels (B) (C) and (D). Proteasomes from .174 cells and antisera at a molar ratio of 1:50 were mixed in a volume of 0.1 ml 20 mM Tris-HCl buffer, pH 7.5 plus 20 mM NaCl and incubated for 1 h at 22°C. Electron micrographs of negatively stained samples were performed as described in figure 6.1.
6.6 Conclusion

The results presented here show that proteasomes from 721 and .174 cell lines have similar quaternary structure but do not reveal the precise arrangement of LMP2, LMP7, MB1 and Delta subunits. However, it is clear that all four subunits occupy inner sites, corresponding to catalytic β-type subunits. Proteasomes from the 721 cell line have LMP7, LMP2 and Delta localized in the inner rings. By analysing another B-cell line, Raji, it was possible to show that LMP2 and Delta subunits localize in the same 20S complex. Finally, analysis of .174 shows that anti-MB1 and anti-Delta physically bind the 20S complex, supporting the proposed role of these subunits in directly substituting the LMP ones.
Chapter 7

Expression of MB1, Delta, LMP2 and LMP7 Gene Products

7.1 Summary

Proteasomes are present in cells from eukaryotes ranging from humans to yeast (Rivett 1993a). Moreover, proteasomes have been found in both the cytoplasm and the nuclei of a variety of mammalian cells (Arrigo et al. 1988, Tanaka et al. 1989, Kumatori et al. 1990, Kanayama et al. 1991) and also various lower eukaryotes (Beyette et al. 1992, Haass et al. 1989, Akhayat et al. 1987, Pal et al. 1988), suggesting that their diverse roles depend on their differential localization and differential subunit expression in the cell. However, many of the proteasome functions, and specific subunit functions, are still unknown. One way to obtain information on their physiological roles is to study their expression in different cell types and cells in abnormal states. This chapter analyses the expression of the LMP2, LMP7, Delta and MB1 proteasome subunits in 46 different cell types, including normal and abnormal cell lines.

This analysis was initially applied to identify cells that were natural mutants for any of the subunits. One mutant cell, an LMP2 mutant B-cell called WJR076, identified by H. Teisserenc, was studied. The data showed an enormous variability in the expression levels of the 4 subunits according to cell type. Not all subunits were expressed, and when they were, their levels of expression varied widely. This variability of subunit expression might be a reflection of their differential roles in proteasome function.

In order to clarify this extreme variety of subunit expression, 13 cell lines were analysed for cellular and nuclear expression, and interestingly, it was observed that some cell lines expressed a high molecular weight form of MB1 or Delta proteins, the functions of which are still unclear. To investigate this observation, we analysed expression of the LMP, MB1 and Delta subunits in the nuclei and the cytoplasm of 5 different cell lines under γ-interferon induction.
Since proteasomes are involved in cyclin breakdown, we wondered how the cell cycle may be altered with different proteasome composition. To obtain some evidence for a possible function of any of the subunits during the cell cycle, the levels of expression of these subunits were analysed in synchronized cells. The synchronization method applied here was to arrest the cells at G1 by medium depletion. Preliminary experiments suggested that the LMPs might also have some specific function during G1 stage and/or during cell starvation.

7.2 Introduction

Proteasomes have been found in all eukaryotic cells examined. However, the level of proteasome expression and cellular localization have shown to vary depending on cell type, stage of cell development and stage of cell cycle. The differential intracellular localization of proteasomes and differential levels of expression are most probably a reflection of the multiple functions of the complex. As discussed in earlier chapters, the proteasome has been recently involved in the regulation of diverse cellular functions, such as DNA repair, cell cycle progression, signal transduction, transcription and antigen presentation.

In order to elucidate the intracellular location of proteasomes, many studies using immunohistochemical localization in a variety of tissues showed that proteasomes localized in the nucleus and cytoplasm, and some cases associated with the endoplasmic reticulum or around the nucleolus (Kanamura et al. 1988, Tanaka et al. 1989, Haass et al. 1989, Akhayat et al. 1987, Pal et al. 1988, Rivett et al. 1992).

The relative amounts of proteasome in the nucleus and cytoplasm analysed showed clear differences depending on cell type. In some cases, proteasomes are almost entirely cytoplasmic and apparently absent from the nucleus, such as crustacean muscle, and in other cases are predominantly localized in the nuclei as seen in human primary cancer cells of the kidney (Beyette et al. 1992, Kumatori et al. 1990, Kanayama et al. 1991).

The localization of proteasomes was also shown to vary within a single cell type, as observed when cells were induced to differentiate (Shimbara et al. 1992). Changes in proteasome cellular location were also observed according to cell cycle stages showing that proteasomes can be rapidly transported from the cytoplasm to the nucleus (Kahawara et al. 1992a, Amsterdam et al. 1993). Subsequently, the implication of proteasome in the cell cycle has been confirmed by its role in the degradation of cyclin and cyclin-dependent kinase inhibitor p27 (Glotzer et al. 1991,
Changes in proteasome localization have also been noticed according to stages of embryonic development (Haass and Kloetzel 1989, Klein et al. 1990, Pal et al. 1994).

Many studies of proteasome function and its expression in different cell types have been performed by analysing cells in abnormal states (Kumatori et al. 1990). Analysis of leukaemic cells showed considerably increased concentrations of proteasomes predominantly localized in the nuclei. In addition, immunohistochemical staining showed that serum proteasome levels were high in patients with hematological malignancies such as acute leukemia, chronic myelogenous leukemia, non-Hodgkin's lymphoma, and myeloma (Wada et al. 1993). The proteasomes in these patients were strongly expressed especially in the nuclei of tumour cells. These data suggested that the elevated levels of serum proteasomes in these patients were derived from tumour cells, and the increased production of proteasomes by malignant blood cells may be involved in transformation and proliferation of these cells.

Studies of intracellular localization of proteasome in relation to cell growth conditions showed that when cells are grown under favorable conditions the proteasomes were mainly detected in the nuclei. However, when the medium's nutrients become depleted the pattern changes showing a much less pronounced detection of proteasomes in the nuclei. It was noticed that proteasomes from cells grown at low density are localized mainly in the nuclei with only weak staining in the cytoplasm (McHiels et al. 1995). However, when cells were grown at high density the nuclear localization was hardly observed and the cytoplasmic localization remained the same. It is known that starvation arrests cells at the G1 stage of the cell cycle, and according to the above experiments there is a reduced level of proteasome localization in the nuclei during this stage.

When synchronized cells were used to analyse proteasome intracellular location it was noticed that they localised in the nucleus and cytoplasm at all phases of the cycle. However, different immunofluorescence patterns were observed (Palmer et al. 1994). In the nucleus, a low intensity staining was observed in early S phase which changed to a more intense labelling during S to G(1) phase. In the cytoplasm, proteasomes were concentrated around the nuclear region at G(1) and at the start of S phase and gradually moved towards the periphery of the cell as the cell cycle progressed to G(2). No cell cycle-dependent changes were detected in the rate of synthesis or level of proteasomes.
Not much is known about the subunit composition of the complex and any functional implications during cell starvation and cell cycle stage. It is reasonable to believe that different subunits might have different functions, specific to cell cycle stages, nutritional conditions, and intracellular localization. So far, the only example of exchangeable proteasome subunits are the LMP2, LMP7 and MECL1 subunits, expression of which is stimulated by γ-interferon.

In order to analyse the proteasome composition in different tissues, during different physiological conditions, several cell samples were analysed in a variety of conditions. This chapter shows a preliminary characterization of some proteasome subunits using specific antisera in the following situations.

1) B cell lines
2) Non-B cells
3) Nucleus and cytoplasm
4) γ-interferon treatment
5) Cell cycle arrest.

7.3 Analysis of LMP2, LMP7, MB1 and Delta expression in B cell lines

In a first experiment we surveyed a range of B cell lines to get an idea of the range of variation of expressing proteasome subunits. MB1, Delta, LMP2 and LMP7 were screened, by western blot, in a number of different human B cell lines. We were also hoping to identify cell lines that were defective in the expression of any subunit. The B cell lines analysed were kindly provided by the Tissue Antigen Laboratory at ICRF, and they are from patients or individuals coming from families that carry identified disease. The 17 B cell lines analysed included members of Hodgkin's or Waldestron's macro disease families, patients carrying breast cancer, myeloma, leukemia, melanoma, or Narcolepsy. The diseases were incidental and we did not expect to find any association of proteasome differences in the B cell lines with these diseases.

Figure 7.1 summarizes the results obtained by western blots of the 17 cell lines and their respective proteasome subunit composition. This figure contains the B cell lines' names and their origins, and it also specifies for each cell line the processed and unprocessed forms of LMP2, LMP7, MB1 and Delta proteins. The corresponding western blots are shown in figures 7.2 and 7.3.
<table>
<thead>
<tr>
<th>Cell line</th>
<th>LMP2</th>
<th>LMP7</th>
<th>MB1</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U</td>
<td>P</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>1) AWE1225-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) JCH1228-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) DUC1232-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) DAF1435-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5) GIM1215-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6) CTH1214-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) JOA1226-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8) PHD1242-RH</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9) IVH2155-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Breast cancer B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10) MPA-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Myeloma B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11) RAH1112-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Myeloma+NHL B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12) ERW-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hairy cell leuk. B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13) MS1091-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Melanoma B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14) RL1154-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Waldenstrom's B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15) VEH1609-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Waldenstrom's B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16) LED-NL</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Narcolepsy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17) LOH-CF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hodgkins B cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.1: List of the 17 human B cell lines from patients or individuals from families that carry some abnormality. The name of the cell lines are shown on the first column and their respective origins are described below the cell name. The presence of each of the following subunits, LMP2, LMP7, MB1 and Delta, are marked by "+". The processed and unprocessed forms of each subunit are designated respectively as "P" and "U", and the observed ~70kDa high molecular weight bands for MB1 and delta are designated as "70".
Figure 7.2: Analysis of LMP2, LMP7, MB1 and Delta protein expression in human B cell lines from patients with Hodgkin's disease; cell lines AWE1225-RH, JCH1228-RH, DUC1232-RH, DAF1437-RH, GIM1215-RH, CTH1214-RH, JOA1226-RH and PHD1242-RH. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 721 and .174 cell lines were included as controls. 10μg of total cytoplasmic lysates were loaded per lane per cell. Western blots were probed with the following antiserum A) anti-LMP2; B) anti-LMP7; C) anti-Delta; D) anti-MB1. Equivalent levels of protein were loaded in each track.
Figure 7.3: Analysis of LMP2, LMP7, MB1 and Delta protein expression in human B cell lines from patients carrying the following diseases: cell lines IVH2155-CF-breast cancer, MPA-CF-myeloma, RAH1112-CF-myeloma/NHL, ERW-CF-hairy cell leukaemia, MS1091-CF-melanoma, RL1154-CF-Waldenstrom's macro, VEH1609-CF-Waldenstrom's macro, LED-NL-narcolepsy, and LOH-CF Hodgkins. 721 and .174 cell lines were included as controls. 721 was used on LMP2 and LMP7 westerns, and .174 for MB1 and Delta ones. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 10μg of total cytoplasmic lysates were loaded per lane/per cell. Western blots were probed with the following antisera A) anti-LMP2; B) anti-LMP7; C) anti-Delta; D) anti-MB1.
Western blot analysis of the 17 B cell lines showed a pattern of subunit expression very similar to that obtained from normal B cells. As expected from previous B cell analysis (chapter 5), the mature forms of LMP2 and LMP7 subunits were highly expressed, the mature Delta subunit was expressed at a lower level compared to non-B cells or to the .174 mutant. The data also shows that only the unprocessed form of MB1 subunit was expressed. This observation led us to postulate that MB1 subunit is not processed in the presence of LMP7 protein.

All B cells lines studied here expressed the mature form of LMP2, LMP7 and Delta subunits with little variation in levels of protein expression, and most cell lines expressed only the unprocessed form MB1 protein. According to this, it appears that MB1 and Delta proteins behave differently; MB1 remains unprocessed in the presence of the LMP subunits, but Delta is present and processed.

7.4 Analysis of LMP2, LMP7, MB1 and Delta subunit expression in non-B cells, tumour cells and mutant cells

Subsequently, we analysed a further 26 non-B cell human cell lines. These cell lines include a variety of carcinomas, such as bladder carcinoma, fibroblast teratocarcinoma, ovarian adenocarcinoma, pancreatic carcinoma, breast carcinoma, epidermoid cervix carcinoma, choriocarcinoma, hepatocellular carcinoma and mutant fibroblasts. Normal cell lines from early developmental stages, such as embryonic lung and foetal lung were also included.

Figure 7.4 summarizes the results obtained on western blots of the 26 different cell lines and their respective proteasome subunit composition. This figure contains the name of each cell line and its origin, and it specifies for each cell line the presence of processed and/or unprocessed forms of LMP2, LMP7, MB1 and Delta subunits. In addition to the unprocessed and processed forms of MB1 and Delta, high molecular weight protein bands of ~70kDa were recognized with the MB1 and Delta antisera. These bands were present on 5 cell lines. Western blots for expression of MB1, Delta, LMP2 and LMP7 subunits on the 26 cell lines are shown in figures 7.5 to 7.8 (figure 7.5 B lane 2 and D lane 4; figure 7.6 B lane 5; figure 7.7 D lane 6; figure 7.8 B lane 2). These high molecular weight bands could be longer unprocessed forms of these subunits that could have novel functions or could be cross-reacting proteins.
<table>
<thead>
<tr>
<th>Cell line</th>
<th>LMP2</th>
<th>LMP7</th>
<th>MB1</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U P</td>
<td>U P</td>
<td>70 U P</td>
<td>70 U P</td>
</tr>
<tr>
<td>1) EJ28</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bladder carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2) BUTLER</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fibroblast.teratoc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3) S9M</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ovarian adenocarc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) GM04286</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hungtinton disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5) 180D</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ovarian carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6) GER</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pancreatic adenocarc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) HELF</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Embryonic lung</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8) A818.7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9) GM05401</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Biochem. mutant fibrob.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10) GM7730A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Biochem. mutant fibrob.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11) SKOV3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ovarian cancer cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12) OVCA433</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ovarian cancer cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13) ZR75</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Breast carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14) SKBR3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Adenocarcinoma breast</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15) L132</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Epithelial embryonic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16) GM5131</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Chromos. aberrant fibrob.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17) 1847</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ovarian tumour</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18) ME180</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Epidermoid cervix carc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19) JATP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Choriocarc. placenta</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20) A431</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Epidermoid carc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21) HaPaC1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22) JAR</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Choriocarc. placenta</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23) 5637</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Bladder carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24) HepG2</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hepatocellular carc. liver</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25) MCF7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Breast adenocarc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26) WI38</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Foetal lung</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.4: List of the 26 human cell lines, which include non-B cells, tumour cells, and mutant cells. The name of the cell lines are shown on the first column and their respective origins are described below the cell name. The presence of each of the following subunits, LMP2, LMP7, MB1 and Delta, are marked by "+". The processed and unprocessed forms of each subunit are designated respectively as "P" and "U", and the observed ~70kDa high molecular weight bands for MB1 and Delta are designated as "70".
Figure 7.5: Analysis of LMP2, LMP7, MB1 and Delta protein expression in various human cell lines including non B cells, tumour cells, and mutant cells. Cell lines: EJ28-bladder carcinoma, BUTLER-fibroblast teratocarcinoma, 59M-ovarian adenocarcinoma, GM04286-Huntington disease, 180D-ovarian carcinoma, GER-pancreatic adenocarcinoma, and HELF-embryonic lung. 721 and .174 cell lines were included as controls. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 10mg of total cytoplasmic lysates were loaded per lane/per cell. Western blots were probed with the following antisera A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.6: Analysis of LMP2, LMP7, MB1 and Delta protein expression in various human cell lines including non B cells, tumour cells and mutant cells. Cell lines: A818.7-pancreatic cancer, GM05401-biochemically mutant fibroblast, GM7730A-biochemically mutant fibroblast, SKOV3-ovarian cancer cell, OVCA433-ovarian cancer cell, ZR75-breast cancer cell, SKBR3-adenocarcinoma of the breast, and L132-embryonic epithelial. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 40μg of total cytoplasmic lysates were loaded per lane per cell. Western blots were probed with the following antisera: A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.7: Analysis of LMP2, LMP7, MB1 and Delta protein expression in various human cell lines including non B cells, tumour cells, and mutant cells. Cell lines; GM5131-chromosome aberrant fibroblast, 1847-ovarian tumour, ME180-epidermoid cervix carcinoma, JATP-choriocarcinoma placenta, A431-epidermoid carcinoma, HaPaCl-pancreatic cancer, and JAR-choriocarcinoma placenta. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 40 µg of total cytoplasmic lysates were loaded per lane/per cell. Western blots were probed with the following antisera; A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.8: Analysis of LMP2, LMP7, MB1 and Delta protein expression in various human cell lines including non B cells, tumour cells, and mutant cells. Cell lines: 5637-bladder carcinoma, HepG2-hepatocellular carcinoma liver, MCF7-breast adenocarcinoma, and WI38-foetal lung. 721 cell line was included as control. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000 and MB1 at 1:500. 40μg of total cytoplasmic lysates were loaded per lane/per cell. Western blots were probed with the following antisera: A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
By comparing figure 7.1 (summary of subunit expression on B cells) with figure 7.4 (summary of proteasome expression on non-B cells and mutant cells) it is clear that there is a large diversity in proteasome subunit expression in the latter group.

The cell lines were provided by the Cell Line Bank at the ICRF. The first 8 cell lines analysed here, shown in figure 7.5, were analysed by using a total of 5μg of cytoplasmic lysate for each cell. This amount of cell lysate was later increased to 30-40μg in order to obtain better western blot detection, especially when using the anti-MB1 antiserum. The remaining 19 cell lines are shown in figures 7.6, 7.7 and 7.8.

The main conclusions drawn from the analysis of the 26 cell lines for MB1, Delta, LMP2 and LMP7 subunits' expression are;

1) LMP2 and Delta, MB1 and LMP7 may not have a simple 1:1 relationship as previously proposed.

2) Many cell lines expressed all four proteasome subunits, suggesting the presence of a mixed set of proteasomes. This indicates that proteasomes may be present with a complex arrangement of subunits.

3) Most non B cells preferentially express MB1 and Delta subunits.

4) High levels of processed Delta subunits were found in all non B cells analysed.

5) High molecular weight forms of MB1 and Delta (cell lines BUTLER, GM04286 and OVCA433) subunits were observed by western blot analysis. However, the nature of these high molecular weight protein bands are still unknown. The possibility of cross reaction with a non-proteasome protein has not been ruled out.

7.5 Analysis of LMP2, LMP7, MB1 and Delta expression in the nucleus and cytoplasm of various cell lines

The above section showed that MB1, Delta, LMP2 and LMP7 subunits exhibit a wide range of expression at different levels and combinations. In addition, high molecular bands for MB1 and Delta subunits were detected on 5 different cell lines. In order to further clarify some of these observations, 10 cell lines were analysed for their subunit expression in the nucleus and cytoplasm. Nuclear and
cellular extracts for the following cells were obtained and analysed for MB1, Delta, LMP2 and LMP7 expression by western blots: B-cells 721 and .174, HeLa (cervix carcinoma), JAR (choriocarcinoma placenta), MCF7 (breast adenocarcinoma), W138 (foetal lung), 5637 (bladder carcinoma), HepG2 (hepatocellular carcinoma liver) and GM513 (chromosome aberrant fibroblast). In addition to the above cell lines, the WJR076 B-cell line was included here (kindly provided by Dr. H. Tesserenc) and analysed for the lack of LMP2 expression.

As previously discussed in the introduction of this chapter, it has been extensively shown that proteasome subunits can be expressed in the nucleus and cytoplasm of cells. However, it was not known how MB1, Delta, LMP2 and LMP7 were distributed in the two compartments.

Figure 7.9 shows the first analysis of nuclear and cytoplasmic expression of MB1, Delta, LMP2 and LMP7 subunits. The monoclonal antibody MCP21, raised against an α type subunit, was included as a positive control. Cytoplasmic and nuclear extracts were obtained according to the protocol in section 2.5.7. B-cell lines 721, .174 and HeLa cells were included. In order to obtain stronger signals on western blots, 40 μg of each nuclear or cytoplasmic protein extract were used per lane, per cell line.

Analysis of this experiment showed that LMP2, LMP7 and Delta subunits are expressed in the nucleus and cytoplasm of cells (figure 7.9 a, b and c, lanes 1 to 6). However, the MB1 subunit showed an inconsistent pattern of expression. The previously detected ~70kDa cytoplasmic band (observed from cytoplasmic lysates from GM04286 and HaPaC1 cell lines) was found in nuclear lysates of 721 and .174 cell lines. In addition to the ~70kDa band, another band of ~46kDa was observed on HeLa cells. Analysis of the α-type subunit using the MCP21 monoclonal showed that the equivalent subunit was not expressed in nuclear lysates (figure 7.10 e, lanes 2, 4, and 6).

In order to investigate the specificity of the MB1 antiserum for possible cross-reaction with another protein, western blots containing the same samples as described above were incubated with 100μg/ml of C-terminus MB1 peptide used in the production of the MB1-antiserum (figure 7.10 c). As seen in figure 7.10 c, the bands observed in the control western blot in figure 7.10 A had disappeared due to the competition of the peptide with the MB1 antiserum. Only a faint band on HeLa cytoplasmic lysates remained, and this could indicate that this 46kDa form is not the
Figure 7.9: Analysis of LMP2, LMP7, MB1 and Delta, protein expression in the cytoplasm and nuclei of the following cell lines; 721, .174 and HeLa-cervix carcinoma. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively. The anti-α subunit antisera MCP21 was also included in this analysis. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, MB1 at 1:500, and MCP21 at 1:2000. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera A) anti-LMP2; B) anti-Delta; C) anti-LMP7; D) anti-MB1; E) MCP21.
Figure 7.10: Analysis of the specificity of the anti-MB1 antiserum. MB1 antiserum was tested in the cytoplasm and nuclear lysates of 721, .174 and HeLa-cervix carcinoma. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively. Anti-MB1 antiserum was used at a concentration of 1:500. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were treated as follows, prior to adding the anti-MB1 reagent A) cell lysates were heated at 95°C for 5 minutes; B) cell lysates were heated at 95°C for 20 minutes; C) MB1 peptide was added at a concentration of 100μg/ml; D) 1A1.3B peptide was added at a concentration of 100μg/ml (Hosking et al. 1995).
MB1 subunit. So far, searches of data banks have not shown any other protein with sequences similar to the MB1 C-terminus.

Figure 7.10 d shows a similar peptide competition experiment, using a different peptide (raised against an internal region of the 1A1.3B breast cancer gene) to incubated western blots filters prior to the MB1 antisera incubation. In this case, all the different MB1 bands were still present.

The high molecular weight MB1 band observed in the nuclei could be an aggregate of two or more proteins. To investigate this hypothesis, samples from the same cell lysates as previously used, were heated at 95°C for a longer period of time, up to 20 minutes, in an attempt to dissociate any aggregates. However, as shown in figure 7.10 c, after submitting the samples to 20 minutes' incubation at high temperature, the high molecular MB1 band was still present. This observation appears to rule out the hypothesis that MB1 could be associated with another protein. However, it is not impossible that a more stable association would still exist by post lysis protein:protein interaction.

Strikingly, during the time that the ~70kDa MB1 form was observed, it was discovered that the MB1 subunit could associate in vitro with the enzyme topoisomerase II (personal communication by Dr. R. Charkraverty at Dr. I. Hickson's laboratory, Oxford). This association was observed by using the two-hybrid system when studying proteins that would associate with topoisomerase II. Topoisomerase II is an enzyme involved in the DNA topological conversion, and DNA relaxation and decatenation reactions. These data reinforced the idea that the high molecular weight MB1 found in nuclei might have a specific function. Topoisomerase II inhibitors alter the rate of cell progression through G2 phase by accumulating cells at this stage. It is not known how the MB1 and topoisomerase II are associated. Further experiments to clarify the MB1 and topoisomerase II association are in progress.

By using the same two-hybrid system for fishing proteins that interact, it was recently observed that another proteasome subunit, in this case the α-type Hu 28kDa subunit interacts with the HBx non structural protein of hepatitis B virus (HBV) (Fisher et al. 1995). These data support the model of HBx acting indirectly on transcriptional processes; by binding to a specific proteasome α-subunit, HBx might interfere with the degradative processes, thereby enhancing the half-life of different transcriptional factors and other regulatory proteins. The interaction with the Hu 28kDa proteasome subunit could explain the remarkably pleiotropic effects of HBx, such as its capacity to transactivate a wide range of viral and cellular promoters.
These type of interactions mask the potential nuclear targeting signal in the α-subunit, preventing a subset of proteasomes from entering the nucleus, and consequently enhancing the half-life of transcription factors and other nuclear regulatory proteins.

Figure 7.11 shows the MB1 Delta, LMP2 and LMP7 subunit expression in the nucleus and cytoplasm of JAR, MCF7 and W138 cell lines. Figure 7.12 shows similar analysis for 5637, HepG2 and GM5131 cell lines. These six cell lines were previously analysed by cytoplasmic expression of the four β subunits, and as expected none of them expressed LMP2 or LMP7 subunits, either on the cytoplasm or nuclei (figure 7.11 a and c and figure 7.12 a and c). Analysis of the Delta subunit, on the same cells, shows that the processed form of Delta is expressed at similar levels in the cytoplasm and nucleus of most of these cells (figure 7.11 b and figure 7.12 b). However, the high molecular weight Delta subunit was observed only in the cytoplasmic lysate of the 5637 cell line (figure 7.12 b lanes 1 and 2). In addition, it appears that HepG2 expresses higher levels of unprocessed form of Delta in the cytoplasm compared to the nucleus (figure 7.13 b lanes 3 and 4).

Analysis of MB1 expression on the same cell lines is shown in figures 7.11 d and 7.12 d. Cell lines JAR, MCF7 and HepG2 showed similar levels of this subunit in both nucleus and cytoplasm. However, W138, 5637 and GM5131 cell lines showed a different level of MB1 expression. W138 (foetal lung cell line) expressed the unprocessed form of MB1 subunit in the cytoplasm only, and it was completely absent in the nuclei (figure 7.11 d lanes 5 and 6). 5637 cell line (bladder carcinoma) expressed slightly higher MB1 levels in the cytoplasm, contrary to GM5131 (chromosome aberrant fibroblast) which expressed higher MB1 levels in the nuclei (figure 7.12 lanes 1, 2, 5, and 6).

The last cell line analysed for nuclear and cytoplasmic expression was the LMP2-mutant B cell line WJR07. This cell is part of the International Histocompatibility B cell bank and it was kindly provided by Dr. H. Teisserenc. Figure 7.13 shows that this cell line does not express the LMP2 subunit (figure 7.13 a lanes 1 and 2). LMP7 and Delta are normally expressed at similar levels in cytoplasm and nucleus (figure 7.13 b and c, lanes 1 and 2). The MB1 subunit has similar expression to other B-cell types. A high molecular MB1 band was observed in nuclear lysates (figure 7.13 d lanes 2 and 6).
Figure 7.11: Analysis of LMP2, LMP7, MB1 and Delta, protein expression in the cytoplasm and nuclei of the following cell lines: JAR-choriocarcinoma placenta, MCF7-breast adenocarcinoma, and WI38-foetal lung. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, MB1 at 1:500, and MCP21 at 1:2000 (these cell lines were previously tested with the positive control 721 cell line, and none of them gave any positive signal with anti-LMP2 and anti-LMP7 antiserum). 40µg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera A) anti-LMP2; B) anti-LMP7; C) anti-Delta; D) anti-MB1.
Figure 7.12: Analysis of LMP2, LMP7, MB1 and Delta, protein expression in the cytoplasm and nuclei of the following cell lines; 5637-bladder carcinoma, HepG2-hepatocellular carcinoma liver, GM5131-chromosome aberrant fibroblast. 721 cell lines was used as a control. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera; A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.13: Analysis of LMP2, LMP7, MB1 and Delta, protein expression in the cytoplasm and nuclei of the WJR076 LMP-2 mutant Bcell line. 721 and .174 cell lines were used as controls. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera; A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Interestingly, after analysing a number of different cell extracts of 721 and .174 cell lines it was noticed that sometimes the unprocessed form of MB1 could be detected in cytoplasmic lysates from these cell lines and sometimes not. For example, in the 721 cell line lysate in figure 7.9 d lane 1 shows the unprocessed form of MB1 protein, but not from lysates obtained later and used on the western blot of figure 7.14 d lane 3 and 4. Similar results were observed for .174. Figure 7.10 d lane 3 has a 30kDa MB1 band which is absent on .174 cytoplasmic lysate shown in figure 7.13 d lane 5. It is not known why there is a different expression of the same subunits according to cell extracts, but the most probable explanation is that cells' extracts were made with cultures that were growing at different growth phases. As discussed in the introduction, high density cell line cultures have lower levels of expression for some proteasome subunits, compared to low density cultures (McHiels et al. 1995).

In summary, proteasome subunits show great variation in their expression between different cell lines and in their cellular distribution. LMP2 and LMP7 subunits are expressed in the nuclei and cytoplasm of cells. Delta subunit is also expressed in the nuclei and cytoplasm, but in some cases a ~70kDa high molecular weight band was noticed in cytoplasmic lysates. The MB1 subunit showed an unusual distribution. Sometimes the unprocessed MB1 form was localized only on the cytoplasm and in other cases it showed a differential distribution on the nucleus and cytoplasm. Finally, it was observed that a high molecular weight ~70kDa MB1 band was expressed on the nucleus of a few cell lines, such as B-cells 721, .174 and WJR076.

7.6 Analysis of LMP2, LMP7, MB1 and Delta expression in cell lines treated by γ-interferon

It has been previously shown that γ-interferon greatly increases the levels of LMP2 and LMP7 mRNAs and consequently their protein levels. Furthermore, the same cytokine does not affect the expression of MB1 and Delta mRNA, suggesting that the decreased expression of these subunits is due to post-transcriptional regulation. However, it was not known if γ-interferon would have a different effect on the expression of proteasome subunits in the cytoplasm and nuclei.

In order to further investigate the cytoplasmic and nuclear expression of LMP2, LMP7, MB1 and Delta subunits, cells lines were cultured in the presence of γ-interferon. The same approach was used to obtain more information about the mature and immature forms of MB1 and Delta and their unusual high molecular weight 70kDa protein bands expressed in some cell lines.
Four cell lines were analysed for MB1, Delta, LMP2 and LMP7 expression in the cytoplasm and nucleus after 3 days of induction by γ-interferon at a concentration of 100U/ml. Cells lines included in this study were; B-cell 721 and .174, HELF (embryonic lung), and HeLa (cervix carcinoma).

The effect of γ-interferon differed from cell line to cell line. γ-interferon affected the expression of MB1 and Delta subunits, and in particular the ~70kDa forms of these proteins, suggesting that the high molecular weight bands are most probably different forms of these proteins with a specific, but unknown, function (figures 7.15 d lanes 2 and 4, and 7.17 b lanes 1 and 3).

Figure 7.14 shows the nuclear and cytoplasmic expression of the four proteasome subunits in the 721 cell line with and without γ-interferon. In general, the expression levels of the four subunits was not strongly affected by γ-interferon induction. It shows that γ-interferon slightly increased the levels of LMP2 expression in the nuclei compared to 721 control cells (figure 7.14 a lanes 3, 4). Higher levels of the unprocessed form of MB1 were found in the cytoplasm, most probably due to its substitution by the LMP7 subunit.

Investigation of .174 cell line revealed that Delta subunit, in the presence of γ-interferon, expresses slightly higher levels in the cytoplasm and slightly lower levels in the nuclei (figure 7.15 b lanes 1 to 4). Similarly, in the presence of γ-interferon, the mature MB1 form was slightly elevated in expression of the cytoplasm and it appears to disappear in the nuclei under cytokine induction. It was also noticed that the 70kDa form of MB1 band expression was slightly higher in the nuclei of induced cells.

Figure 7.16 shows the effect of γ-interferon on HeLa cells. As previously observed, LMP2 and LMP7 subunits are up-regulated by γ-interferon. In this experiment, the level of both subunits increased in the cytoplasm and nuclei (figure 7.16 a, b lanes 1 to 4). The expression of the Delta subunit decreased under γ-interferon induction in both cytoplasm and nuclei. Interestingly, cytoplasmic levels of processed MB1 totally disappeared in γ-interferon cultured cells in both cytoplasm and nuclei, but higher levels of the unprocessed form of MB1 appear only in the cytoplasmic portion of the cell, and not in the nuclei (figure 7.16 d, lanes 1 to 4).
Figure 7.14: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of 721 cell line cultured with and without addition of γ-interferon. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively, and after γ-interferon treatment are indicated by "CY" and "NY" respectively. 721 cells were treated with 300 units of γ-interferon for 3 days prior to lysis. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera; A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.15: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of .174 cell line cultured with and without addition of γ-interferon. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively, and after γ-interferon treatment are indicated by "Cy" and "Nγ" respectively. .174 cells were treated with 300 units of γ-interferon for 3 days prior lysis to LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40µg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera: A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.16: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of HeLa cell line cultured with and without addition of γ-interferon. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively, and after γ-interferon treatment are indicated by "Cy" and 'Ny' respectively. HeLa cells were treated with 300 units of γ-interferon for 3 days prior lysis to LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40μg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera: A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Interesting results obtained from γ-interferon treatment were observed in the HELF cell line. Figure 7.17 shows that the HELF cells induced with this cytokine showed higher levels of LMP2 and LMP7 expression compared to untreated cells, as observed before on other cell lines. However, an unexpected result was found when analysing subunit Delta, showing that the 70kDa form of this subunit is expressed on the cytoplasm under γ-interferon treatment (figure 7.17 b lane 3). The 70kDa cytoplasmic form of Delta subunit has been observed on other cell lines, such as BUTLER, OVCA433, and 5637. This observation is consistent with the notion that the 70kDa protein bands detected on western blots are new forms of Delta proteins, or Delta subunit associated with other proteins.

In conclusion, γ-interferon treatment increases the nuclear and cytoplasmic displacement of MB1 and Delta by LMP7 and LMP2, as seen in 721, HeLa and HELF cell lines. In addition, γ-interferon might be involved in the induction of different forms of proteasome subunits, as observed for the high molecular weight MB1 and Delta subunits.

7.7 Analysis of LMP2, LMP7, MB1 and Delta expression in synchronized cells

To obtain some evidence for a possible function of any of the four proteasome subunits during the cell cycle, we analysed the levels of expression of LMP2, LMP7, MB1 and Delta subunits during the G1 stage of cell synchronization. The synchronization method applied here was to arrest the cells at the G1 stage by medium depletion. Cells were grown for up to 5 days without adding new medium to the cell culture, consequently leading to cell starvation.

721 B-cell, Raji B-cell, and HeLa cell lines were cultured for up to 5 days without adding new medium, leading to synchronization of all the cells at the G1 cell cycle stage and cell starvation. Samples of 721 and Raji cell cultures were collected from day 1 (start), day 3 and day 5. Samples of HeLa cells were collected at day 1 (start) and day 5. Each sample was investigated for MB1, Delta, LMP2 and LMP7 expression in the cytoplasm and nucleus.

Figure 7.18 shows western blots of Delta and MB1 protein expression from nucleus and cytoplasm of 721, Raji and HeLa cells at different time points of cell culture. The processed forms of Delta and MB1 subunits were expressed at similar levels in the cytoplasm of 721, Raji and HeLa, leading to the conclusion that
Figure 7.17: Analysis of MB1, Delta, LMP2 and LMP7 protein expression in the cytoplasm and nuclei of HELF cell line cultured with and without addition of γ-interferon. Cytoplasmic and nuclear lysates are indicated by "C" and "N" respectively, and after γ-interferon treatment are indicated by "Cy" and "Ny" respectively. HELF cells were treated with 300 units of γ-interferon for 3 days prior to lysis. LMP2, LMP7 and Delta antisera were used at concentrations of 1:1000, and MB1 at 1:500. 40pg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were probed with the following antisera: A) anti-LMP2; B) anti-Delta; C) anti-LMP7; and D) anti-MB1.
Figure 7.18: Analysis of MB1 and Delta protein expression in the cytoplasm and nuclei of synchronized cell lines. 721, Raji and HeLa cells were arrested at the G1 stage of cell cycle by growing the cells for 5 days without feeding the culture. Aliquots of each cell culture were collected at the starting time, after 3 days, and after 5 days of culture, indicated on the figure as "0", "3", and "5" respectively. Delta antisera was used at concentration of 1:1000, and MB1 at 1:500. 40µg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were performed as follows; A) anti-Delta antisera probed on cytoplasmic lysates; B) anti-Delta antisera probed on nuclear lysates; C) anti-MB1 antisera probed on cytoplasmic lysates; D) anti-MB1 antisera probed on nuclear lysates.
MB1 and Delta are not affected by cell starvation (figure 7.18 a and c, lanes 1 to 8). Similar results were observed for the nuclear expression of the MB1 subunit. MB1 expression levels did not change during cell starvation or at the G1 cell cycle stage in any of the cell lines analysed here. However, Delta subunit decreased its nuclear expression when at the G1 cell cycle stage and/or under cell starvation (figure 7.18 b lanes 1 to 8). This phenomenon was easily observed on Raji and HeLa cell lines due to the higher expression of Delta in these cells compared to 721 cell line. The observation that Delta subunit expression decreased in the nucleus of cells at the G1 stage or in high density cultures has been observed with other 20S proteasome subunits, as discussed in the introduction.

Most interesting were the results obtained from the LMP2 and LMP7 investigation under the same conditions. Figure 7.19 shows that LMP2 and LMP7 subunits increased their levels of expression when submitted to cell starvation and consequently arrested at the G1 stage of cell cycle. This phenomenon was observed on all three cell lines analysed here (figure 7.19 a and c, lanes 1 to 8). It appears that cells at the G1 stage might need these subunits for a special function, most probably linked to degradation processes that occur at the end of mitosis and cell growth. The other possibility might be that during cell starvation LMP2 and LMP7 subunits are necessary to elevate the intracellular protein degradation necessary for the survival of the cell due to lack of nutrients.

The finding that LMP2 and LMP7 levels increased at the G1 stage and/or during cell starvation resembles their expression when γ-interferon is present. It might be that LMP2 and LMP7 subunits not only produce peptide more suitable for class I presentation, but also have the ability to elevate the degradation of proteins inside the cell. In this case LMP2 and LMP7 are not only up-regulated by γ-interferon, but also, an alternative system might exist to increase LMP2 and LMP7 expression.

Finally, LMP2 and LMP7 nuclear expression was analysed at different cell culture stages. LMP2 and LMP7 nuclear levels increased in the 721 cell line, most probably due to their higher expression in the 721 cell line compared to Raji and .174 cells (figure 7.19 b, d lanes 1 to 3). In Raji and HeLa cells it was not very clear if the nuclear expression of LMP2 and LMP7 subunits was increased. It appears that expression of the LMP7 subunit did not change in Raji and HeLa cells (figure 7.19 d, lanes 4 to 8) and LMP2 expression was only slightly increased in HeLa cells at the G1 cell cycle stage (figure 7.19 b, lanes 7 and 8).
Figure 7.19: Analysis of LMP2 and LMP7 protein expression in the cytoplasm and nuclei of synchronized cell lines. 721, Raji and HeLa cells were arrested at the G1 stage of cell cycle by growing the cells for 5 days without feeding the culture. Aliquots of each cell culture were collected at the starting time, after 3 days, and after 5 days of culture, indicated on the figure as "0", "3", and "5" respectively. LMP2 and LMP7 antisera were used at a concentration of 1:1000. 40µg of total cytoplasmic or nuclear lysates were loaded per lane/per cell. Western blots were performed as follows; A) anti-LMP2 antisera probed on cytoplasmic lysates; B) anti-LMP2 antisera probed on nuclear lysates; C) anti-LMP7 antisera probed on cytoplasmic lysates; D) anti-LMP7 antisera probed on nuclear lysates.
The observation that in some cell lines LMP2 and LMP7 subunit levels increased in the cells during cell starvation/or at G1 stage of cell cycle has not been observed before. Interestingly, this finding is in opposition to previous observations from cells arrested at the G1 cell cycle stage, or medium depletion, where reduced levels of expression of some proteasome subunits expressed in the nuclei were shown (McHills et al. 1995).

In summary, the data above show that LMP2 and LMP7 subunits increase their expression levels when cell cultures are arrested at the G1 cell cycle stage or when the medium is depleted. This observation raises questions about the function of LMP2 and LMP7 subunits, other than processing of antigen better suited to class I presentation.

7.8 Conclusion

This chapter analysed the expression of MB1, Delta, LMP2 and LMP7 subunits in a variety of cell types. The analysis shows that these four subunits are expressed at different levels, indicating that 20S proteasomes are formed from different combinations of the available subunits. It is possible that some proteasome subunits have specific functions according to cell type, embryonic development stage, and cell cycle stage.

It appears that γ-interferon affects the high molecular weight MB1 and Delta bands detected on western blots, increasing the possibility of the existence of new forms of these subunits. Finally, cell synchronization experiments suggested that LMP2 and LMP7 subunit might also have an additional function on the 20S proteasome, such as to speed up protein degradation. It seems likely that since the proteasome is a complex which is used for breakdown of most, if not all, soluble proteins in the cell, this organelle is subjected to subtle changes in composition under regulation. My studies have started to explore the variation under proteasome composition in different conditions but further work is needed to determine the importance of this variation.
Chapter 8

Concluding Remarks

The objective of my research has been to identify and characterize new human 20S proteasome subunits that are homologous to the human β subunits LMP2 and LMP7. I am particularly interested in the following questions: How is the mutant CTL 721.174 cell line able to process peptides for class I presentation in the absence of the LMP genes? Is the quaternary structure of the 20S proteasome from .174 cell line similar to the wild type? In which chromosome region are the putative LMP homologue genes located and what is the genetic origin of these new β subunits? How is the expression of the LMP subunits and their homologues regulated? What effect does γ-IFN have on the expression of the LMP2 and LMP7 subunit homologues? What is the level of expression and intracellular location of the LMP subunits and their homologues in different cell types and at cell cycle stages?

The studies presented in chapter 3 suggest that MB1 and Delta gene subunits are most probably the LMP7 and LMP2 subunit homologues. I discovered the MB1 gene by screening cDNA libraries under non-stringent conditions with a LMP7 probe. By DNA and protein sequencing analysis I showed that the MB1 gene is highly similar to the LMP7 subunit. Protein analysis showed that MB1 is 68.1% identical to the LMP7 sequence, the highest level of identity known among all 20S subunits. The MB1 gene is probably an active subunit in view of the presence of the highly conserved amino-termini threonine residue, called Thr^, which mediates the nucleophilic attack, and due to the presence of the putative conserved residue Lys. I identified the Delta subunit as the LMP2 homologue, by analysing previously published proteasome sequences. I showed by protein analysis that Delta has the highest level of protein identity to LMP2, 60.8%, when compared to all the other β subunits. I obtained the cDNA encoding the human Delta by screening a human B-cell cDNA library with a Delta polymerase chain reaction (PCR) probe. I showed by protein analysis that Delta contains the two active site nucleophiles, the Thr^ and Lys, which indicates that Delta is most probably an active β-type subunit.
Therefore, by sequence homology my studies suggested that MB1 and Delta genes are the best candidates to substitute for LMP7 and LMP2 in the processing of antigens for class I presentation in the mutant .174 cell line.

I analysed the mRNA levels of MB1 and Delta in the mutant .174 and showed that they were 3-4 fold higher compared to the wild type 721 cells line. By comparing protein levels, I showed that MB1 and Delta subunits increased their expression 10-20 fold in the .174 mutant cell line. According to these results, I conclude that the expression of the LMPs leads to only a slight reduction in the mRNA levels of MB1 and Delta, suggesting that post-transcriptional control is responsible for most of the decrease seen at the protein level in the 721 cell line. I propose that the post-transcriptional control would entail the degradation of MB1 and Delta by their lack of incorporation into proteasome complexes. The preferential incorporation of LMP2 and LMP7 could be due to higher affinity to the other proteasome subunits.

In order to understand if there was any genetic relationship between MB1 and Delta loci in view of the close proximity of LMP2 and LMP7 in the MHC, I examined (chapter 4) the chromosomal localization of MB1 and Delta loci. I obtained MB1 and Delta genomic clones from cosmid libraries and used them for fluorescence in situ hybridization (FISH). I localized MB1 and Delta genes to different chromosomes, 14q11.2 and 17p13 respectively. The different chromosomal location of MB1 and Delta raised interesting questions about the evolution of these four β-type subunits.

By comparing amino-acid sequence of β proteasome subunits I showed that MB1 and Delta differ significantly from each other (31.2% amino acid identity) as do LMP2 and LMP7 (36% amino acid identity), suggesting that they are not recent duplicates of each other. By considering the higher homology of MB1 to LMP7 and Delta to LMP2 I suggest that it is more likely that MB1 and LMP7 originated from one common ancestral gene and Delta and LMP2 from another. By comparing MB1, Delta, LMP7 and LMP2 protein sequence with proteasome subunits from other species I showed that MB1 and Delta have higher similarity to yeast PRE2 and PRE3 genes than LMP7 and LMP2 genes.

The above analysis indicates that MB1 and Delta are the more ancient of the four human genes, and the LMP loci arose later in the evolution of vertebrates. I propose two models for the genetic origin of the LMP genes. In one model the ancestral MB1 and Delta genes were linked and, after duplication and mutation,
remained linked on 6p21 (the LMPs), but MB1 and Delta split onto different chromosomes. In another model, the primordial MB1 and Delta genes were localized in different chromosomes, where they duplicated and mutated. Copies of each one of the ancestral genes were brought together to form a cluster on 6p21 (the LMPs) together with the TAPs within the MHC. The other duplicates of the ancestral genes (MB1 and Delta) localized on chromosomes 14 and 17 respectively.

The β proteasome subunit MECL-1 was shown to be similar to the LMPs by its ability to be stimulated by γ-IFN (Goettrup et al. 1996). Similarly, a new human subunit called Z was shown to have a reciprocal expression to that of MECL-1 (Dr. Keiji Tanaka personal communication), suggesting it has similar role as the MB1 and Delta subunits. A model for the role to these subunits was proposed in figure 4.5. Further, it would be informative to examine genes that are located close to MB1, Delta and Z loci, in order to obtain more information about the evolution of these β subunits. Recently, the mouse subunit MECL-1 was shown to be induced by IFN-γ and to replace a barely characterized β subunit designated MC14, homologue to the human Z subunit (Groettrup et al. 1996).

Using antisera raised against the C-termini of MB1 and Delta protein I was able to further characterize these β subunits. In chapter 5 I showed that MB1 and Delta proteins are part of the 20S proteasome complex. By immunoprecipitation of Delta and pulse-chase analysis, I showed that Delta is processed from an immature to a mature form due to N-terminus cleavage, resulting in ~3kDa loss. Although the MB1 antisera was unable to work in immunoprecipitation, by comparing the SDS-PAGE localization of processed MB1 in two-dimensional gels and western blot analysis I conclude that the two MB1 protein bands observed on western blots were most probably the mature and immature forms of MB1 protein. Having compared the two different MB1 molecular weight bands observed on western blot, I concluded that the mature MB1 form is a result of ~8-12 kDa loss due to cleavage of the N-terminal peptide. The molecular weight and isoelectric point of the mature and immature forms of MB1 and Delta proteins shown on western blots and immunoprecipitation were consistent with those predicted by analysis of the protein sequences considering the Thr1 residue as the cleavage point for the mature protein.

Therefore, my studies suggest that MB1 and Delta are part of the proteasome complex, and similarly to previously analysed β type subunit, they are processed from an immature to mature form, for association with active β subunits that are incorporated within the 20S proteasome complex.
Using the antisera raised against MB1 and Delta I investigated the expression of both subunits in the mutant .174 and wild type 721 cell lines. I showed that MB1 and Delta subunits exhibit a 10-20 fold increase in protein expression in the mutant cell compared to the wild type. Thus, these data show that MB1 and Delta subunits are most probably compensating for the LMP7 and LMP2 defect.

In order to obtain further information about the role of MB1 and Delta as possible LMPs homologues, I analysed the effect that γ-IFN treatment has on the levels of MB1 and Delta expression compared to its effect on the levels of LMP7 and LMP2. It has been previously shown that LMP2 and LMP7 protein levels are substantially increased when cells are stimulated by γ-IFN. By choosing a variety of cell types, including cell lines that normally express either low levels of LMPs, or none at all, I was able to show that γ-IFN treatment increases LMP2 and LMP7 expression at the same time that MB1 and Delta expression levels decrease. Therefore, I concluded that MB1 and Delta have a reciprocal expression to that of the LMPs subunits, suggesting that the proteasome recruits different subunits in response to environmental stimuli. This finding indicates the proteasome participates not only in protein turnover, but also in specialized functions, such as antigen processing for class I presentation.

The results of my analysis of MB1, Delta, LMP2 and LMP7 protein expression in the mutant .174 cell line and in cells treated with γ-IFN strongly suggest that MB1 and Delta are proteasome housekeeping subunits that are substituted by LMP7 and LMP2 subunits in order to generate products more suitable for class I presentation.

I further investigated the quaternary structure and subunit topography of the 20S proteasomes from 721 and .174 cell lines determined by immunoelectron microscopy using the LMP2, LMP7, MB1 and Delta polyclonal antibodies. I showed in chapter 6 that proteasomes isolated from the .174 cell line, lacking the LMPs, have the same quaternary structure as the wild type 721 cell line. This suggests that in the absence of the LMPs, other proteasome subunits are being expressed in order to generate stable and fully assembled 20S proteasome complex. In addition, I showed that MB1 and Delta subunits are localized in the inner rings of the .174 proteasome.

Similarly, LMP2, LMP7 and Delta are physically localized in the inner rings of 20S proteasomes isolated from wild type 721 cell lines, showing that these subunits are correctly located as expected from β-type proteasome subunits. It was also noticed that the MB1 subunit did not immunolocalize in the inner rings in the
721 cell line, indicating that the MB1 subunit is not incorporated into the fully assembled 20S complex.

By studying proteasomes extracted from 721 cell lines and analysing for immunolocalization of the four β subunits, I suggest that there are at least three different 721 proteasome species, one that expresses the LMP2 subunit only, one that expresses Delta only, and one that expresses both subunits. Therefore, if further analysis of other β-type subunits, such as MECL1 and Z are performed it is possible that an even larger number of proteasome species will exist.

In addition, by using Raji cell line extracts and performing sequential immunoprecipitations with LMP2, LMP7 and Delta antiserum I was able to show that LMP2 and Delta subunits can be integrated in the same 20S proteasome complex.

In conclusion, the results presented in chapter 6 show that proteasomes from 721 and .174 cell lines have similar quaternary structure and the immunolocalization of MB1 and Delta subunits into .174 cells supports the proposed role of these subunits in directly substituting the LMP.

The data described in chapters 3 to 6 show that MB1 and Delta are homologues of the LMP subunits. However, very little is known about the specific function of each of these subunits in intracellular protein degradation. In order to obtain further information about the function of each of these subunits, I tried to identify cell lines that were defective in the expression of any of these subunits. So far, the only known β subunit mutant cell lines are the previously described .174 and T2 cell lines, which have a large number of genes deleted, making the characterization of specific proteasome subunits difficult. In chapter 7 I analysed a large number of cell lines to search for MB1, Delta, LMP2 or LMP7 mutations. So far, only one cell line has been identified (by Dr. Teisserenc) and it is defective for the expression of the LMP2 subunit. Experiments to identify the specific function of these subunits are in progress.

By analysing B-cell and non B cell lines for LMP2, LMP7, MB1 and Delta expression I showed that these subunits are expressed at different levels and compositions, and that 20S proteasomes are formed from many different combinations of the available subunits. Therefore, if other subunits were investigated, such as MECL-1 and Z, an even larger variability of 20S proteasome will be identified.
Using a panel of various B cell lines I showed that LMP2, LMP7 and Delta subunits are processed and incorporated into 20S proteasomes. However, I show that in B cells the MB1 subunit is not processed into mature form, and therefore I conclude that in the presence of LMP7, the MB1 subunit is not processed and incorporated into 20S proteasomes. Accordingly, I postulate that MB1 and Delta subunits behave differently, MB1 remaining unprocessed in the presence of LMP7 whereas Delta is present and processed.

When analysing a variety of non B cell lines, I showed that most non B cells express MB1 and the Delta subunits, and Delta subunit is highly expressed. Many cell lines analysed express the four β subunits, therefore indicating that proteasomes may be present with a complex arrangement of subunits. In addition, I showed that some cell lines express a high molecular weight form of MB1 or Delta subunits of unknown nature. Further experiments will be necessary to clarify this observation. However, it has been recently shown that α and β proteasome subunits can associate with proteins other than proteasomes.

By studying the expression of LMP2, LMP7, Delta and MB1 subunits in the nuclei and cytoplasm I show that the first three subunits have similar patterns of distribution in both cellular compartments. However, MB1 subunit show an unusual distribution with cellular and nuclear expression of mature and immature forms varying considerably. Furthermore, I was able to show that a high molecular weight MB1 form was also expressed on the nucleus of B cell lines, such as .174, 721 and WJR076.

By further analysing the intracellular expression of MB1, Delta, LMP2 and LMP7 subunits under γ-IFN induction I showed that this cytokine helps to increase the nuclear and cytoplasmic displacement of MB1 and Delta, by LMP7 and LMP2. I also show that in the HELF cell line γ-IFN induces the cytoplasmic expression of a ~70kDa Delta form, and in .174 cell line it increases the nuclear expression of a ~70kDa MB1 form. Therefore, I suggest that these are either different forms of MB1 and Delta subunits or are associations of MB1 or Delta subunits with other proteins.

In the last part of chapter 7 I analysed the expression levels and subcellular localization of LMP2, LMP7, MB1 and Delta in synchronized cells at G1 stage by serum starvation. I used this approach to discover if any of these β subunits would have a specific function in protein degradation during cell cycle. By growing cells for up to 5 days without adding new medium to the cell culture and collecting samples at
various time points I was able to show that LMP2 and LMP7 behave differently to Delta and MB1 subunits. I showed that LMP2 and LMP7 subunits increased their cytoplasmic expression levels at the G1 cell cycle stage. In addition, it appears that LMP2 and LMP7 nuclear levels increase at the G1 stage of the cell cycle, as observed in the 721 cell line.

Similar analysis of the MB1 and Delta subunits showed that their cytoplasmic levels remained the same despite increased expression of the LMPs. Cell cycle arrest and/or serum starvation only affected the nuclear expression of Delta by decreasing it slightly, an effect not observed for the MB1 subunit.

It is possible that the LMP2 and LMP7 subunits are differently expressed at G1 compared to MB1 and Delta. A system of up-regulating the LMPs might exist, other than by γ-IFN induction. It could be that during the G1 stage of cell cycle or during cell starvation more proteasomes are required for protein breakdown and cells tend to produce more LMP2 and LMP7 containing proteasomes. The cytoplasmic expression of MB1 and Delta subunits did not decrease during increased LMPs expression. Perhaps all four subunits are essential for protein degradation. LMP2 and LMP7 subunits may be preferentially required at the G1 stage, most probably linked to degradation processes that occurs at the end of mitosis and cell growth, or during cell starvation where intracellular degradation is elevated for the survival of the cell during lack of nutrients.

Very little is known about the specific function of LMP2, LMP7, MB1 and Delta subunits. Yeast cell experiments have shown that mutation in 13 of the 14 existing 20S subunits are lethal, and with the exception of LMP2, LMP7, MECL1 and their homologues, similar lethal phenotypes are expected for all other 20S mammalian subunits. Another approach to studying proteasome subunits has been to use specific proteasome inhibitors. Until now the only specific proteasome inhibitor found has been lactacystin, specifically binding to the MB1 subunit. However, prolonged treatment with lactacystin is lethal to cells, making it impracticable for analysis of the MB1 function.

One recent approach to studying protein function has been the use of the yeast two hybrid system for protein-protein interaction. Three 20S proteasome subunits have been characterized using these approaches: 1) α subunit Hu interacts with a nonstructural protein of hepatitis B virus interfering in the degradative processes and enhancing the half life of different transcription factors and other regulatory proteins (Fischer et al. 1995). 2) Viral transactivator Tax protein encoded
by human T-cell leukaemia virus HTLV-I binds to two 20S proteasome subunits, β subunit HsN3 and α subunit HC9 (Rousset et al. 1996). Tax also binds to p105/NF-κB1 protein inducing nuclear translocation of NFκB factors. The dual interaction of Tax with p105 and HC9 favours recruitment of p105 to the proteasome, increasing the processing of p105 into p50. 3) Preliminary experiments performed studying proteins that would associate with topoisomerase II show that the MB1 subunit can associate in vitro with this enzyme (personal communication by Dr. R. Charkraverty at Dr. I. Hickson's laboratory, Oxford). Topoisomerase II is an enzyme involved in the DNA topological conversion, and DNA relaxation and decatenation reactions. Further experiments using the two-hybrid system would be useful to characterize specific subunit function.
Bibliography


