UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE, HEALTH & LIFE SCIENCES

School of Medicine

Function of chicken tapasin in MHC class I antigen presentation

Andrew James van Hateren

Thesis for the degree of Doctor of Philosophy

June 2006

UNIVERSITY OF SOUTHAMPTON
ABSTRACT
FACULTY OF MEDICINE, HEALTH & LIFE SCIENCES
SCHOOL OF MEDICINE
Doctor of Philosophy

FUNCTION OF CHICKEN TAPASIN IN MHC CLASS I ANTIGEN PRESENTATION
Andrew James van Hateren

MHC class I molecules are vital for the detection and elimination of pathogen-infected or malignant cells by cytotoxic T cells and Natural Killer cells. Within the ER the peptides that are required for cell surface expression of class I molecules, and are recognised by T cells, are loaded by the "peptide-loading complex", of which tapasin is the keystone.

In the chicken MHC the tapasin, class I and TAP peptide-transporter genes are tightly linked and may have co-evolved. This might mean that these proteins can only interact when encoded by alleles of the same haplotype. To investigate this possibility, the chicken tapasin gene was sequenced from seven MHC haplotypes. This revealed chicken tapasin to have a high level of allelic polymorphism and moderate sequence diversity, which contrasts with typical mammals, such as humans, where the tapasin, class I and TAP genes are not tightly linked, and where there are far fewer and less diverse alleles of tapasin.

Analysis of the nature and distribution of polymorphic amino acids in chicken MHC-encoded proteins revealed that there are few polymorphic amino acids in the domains of the chicken tapasin and TAP proteins that are likely to interact, which suggests that these proteins might associate irrespective of haplotype. However, with the majority of the polymorphic amino acids of chicken tapasin being located within the ER luminal domains, tapasin may participate in a "haplotype-specific" interaction with the BF2 class I protein, a possibility that is supported by a phylogenetic analysis of these proteins.

A series of transfectants were produced in order to look for such an interaction, guided by the presumed sequence of events that led to the generation and evolution of a natural recombinant haplotype. This approach identified naturally polymorphic residues of class I molecules which appeared to influence the interaction with tapasin, and suggested that, following the recombinational event which generated the recombinant haplotype, tapasin and class I alleles have co-evolved in order to achieve optimal peptide-loading. When tapasin and class I alleles were mismatched, peptide-loading was shown to occur less efficiently.

Table of contents

List of figures

List of tables

Acknowledgements

There are many people that I would like to thank for helping me over the last few years. However, there are five people that I am most deeply indebted to.

The first is my supervisor at the IAH, Dr Jim Kaufman. Without his belief, vision, enthusiasm, constructive criticism and tolerance this project would not have been possible. I would also like to thank Jim for allowing me the opportunity to pursue this project, and the freedom to shape the course of the research.

The second is my supervisor at the University of Southampton, Professor Tim Elliott. Tim's insight has been a constant source of inspiration to me, and his belief, enthusiasm, advice and tolerance have always been greatly appreciated. Visits to the Elliott laboratory were always a highlight of the last few years.

I would also like to thank both Jim and Tim for their willingness to read draft versions of this thesis; your advice and comments were immensely useful.

I would also particularly like to thank my parents and my brother Nick, who have always provided valued advice, emotional support and friendship. Without their support this project would not have been possible, and I would not be the person I am today, and aspire to be. I am particularly indebted to Nick for invaluable computer and technical advice, and to my parents who helped me out during the final stages of writing, by sharing their home with me and allowing me to finally escape Compton!

I would also like to thank all the people who have helped me scientifically. At the IAH I would like to thank Larry Hunt, Denise Marston, Sally Rogers and Mick Gill. At Southampton, I would like to thank Chris Howe, Antony Antoniou, Nasia Kontouli, and especially Tony Williams (whose interest, vision and enthusiasm was greatly appreciated). I would also like to thank all those other people who have helped out in so many ways.

Abbreviations

ABC	ATP-binding cassette
AEBSF	4-(2-aminoethyl) benzenesulfonyl fluoride
ATP	adenosine triphosphate
β_2m	β_2-microglobulin
BFA	brefeldin A
β-ME	beta-mercaptoethanol
bp	base pair
BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
cM	centimorgans
CTL	$CD8^+$ cytotoxic T lymphocyte
Da	Dalton
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
ECL	enhanced chemoluminescence
ER	endoplasmic reticulum
EDTA	ethylene diamine tetra-acetic acid
Endo H	endoglycosidase H
FCS	foetal calf serum
GCG	genetics computer group
Glc	Glucose
GlcNAc	N-acetylglucosamine
HC	MHC class I heavy chain
HLA	human leukocyte antigen
HRP	horseradish peroxidase
hsp	heat shock protein
IFNγ	interferon gamma
Ig	immunoglobulin
ITIM	immuno-regulatory tyrosine-based inhibitory motif
KGB	potassium glutamate buffer
LB	Luria-Bertani medium
LMP	low molecular weight protein
LSB	Laemmli sample buffer
mAb	monoclonal antibody
Man	Mannose
MDV	Marek's disease virus
MFI	mean fluorescence intensity
MHC	major histocompatibility complex
MLR	mixed lymphocyte reaction
NK	natural killer
NKC	Natural Killer complex
NOR	nucleolar organiser region
nt	nucleotide
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
PLC	peptide-loading complex
pMHC	peptide-loaded MHC class I molecule
R10	RPMI cell culture media supplemented with 10% foetal calf serum, two mM glutamine, and antibiotics
R-BSA	RPMI cell culture media supplemented with 2 mg/ml bovine serum albumin, two mM glutamine, and antibiotics
RBC	red blood cells
Rfp-Y	restriction fragment pattern-Y
RPMI	Roswell Park Memorial Institute media
RSV	Rous sarcoma virus
RT	room temperature
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulphate
St dev	standard deviation
TAE	tris-acetate/EDTA
TAP	transporter associated with antigen processing
TCR	T cell receptor
TE	tris-EDTA
TMD	transmembrane domains
Tpn	tapasin
UTR	untranslated region
WBC	white blood cells

1. Introduction

1.1 General introduction to immunology

The immune system can be described as the ability of the host to defend itself from pathogen infection. In vertebrates, the immune response is mediated by cells and proteins that can be separated into two broad, but interacting systems; those that belong to the innate immune system, and those that belong to the adaptive immune system. In the following sections aspects of these systems are briefly described, using humans and mice as reference species.

1.1.1 The innate immune system

The innate immune response is mediated by cells and proteins that are capable of rapidly recognising and constraining or eliminating certain pathogen infections. In mammals, the innate immune response is largely mediated by non-lymphoid cells which express several DNA germline-encoded receptors that recognise common pathogen constituents. Consequently, the innate immune response may lack sufficient diversity in the total receptor repertoire to detect all pathogens, as certain pathogens have evolved mechanisms to avoid recognition by these receptors. In these cases, an adaptive immune response detects such pathogens and usually clears the infection, often relying upon the innate immune system to remove specifically targeted pathogens.

1.1.2 The adaptive immune system

The key features of the adaptive immune response are increased specificity and diversity, and immunological memory. The adaptive immune response is mediated by bone marrow-derived cells, called lymphocytes, which can be broadly separated into two types according to the location in which the cells mature, the type of antigen receptor they encode, the type of the pathogen that they recognise, and by the nature that they eliminate the pathogen. Each lymphocyte expresses one clonotypic antigen receptor that is a unique variant of a prototype receptor generated by essentially random processes including "somatic gene rearrangement" which occur during lymphocyte development. As each lymphocyte encodes a distinct antigen receptor, there is an immensely broad diversity of receptors that are collectively capable of mounting an immune response to virtually any foreign substance. Crucially, any antigen receptors that are capable of recognising the host are either removed during maturation, or otherwise controlled.

1.1.3 B lymphocytes

Lymphocytes can be divided into B or T cells. B cells mature within the bone marrow and, when activated differentiate into antibody-producing plasma cells. Antibodies protect the host from extracellular pathogens and their toxic products by neutralising their ability to enter cells; by specifically targeting or "opsonizing" pathogens for destruction by phagocytic cells of the innate immune system; and by activating a series of plasma proteins that are known as complement, which opsonize pathogens and can also directly kill pathogens.

1.1.4 T lymphocytes

T cells mature within the thymus, and most T cells protect the host from pathogens that infect cells. A major T cell subset, called αβ T cells (T cells hereafter) do not recognise pathogens directly; instead they recognise fragments of pathogen proteins, called peptides, which are expressed at the surface of host cells, bound within a glycoprotein known as a Major Histocompatibility Complex (MHC) protein. There are two classes of MHC protein which both present peptides to T cells, but differ in the cellular origin of their bound peptide. MHC class I proteins generally present peptides from proteins that are partially degraded within the cytoplasm of the host cell; whereas, MHC class II proteins generally present peptides from proteins that are partially degraded within intracellular vesicles.

There are two broad classes of T cells that specifically recognise one class of MHC protein, and are distinguished by expression of either the CD4 or CD8 co-receptor, which recognises an invariant portion of the MHC protein that the T cell receptor (TCR) recognises. $CD4^+$ T cells recognise MHC class II proteins, and can be further sub-divided into two broad classes of cells (T_{H1} or T_{H2}) that differ in their effector functions. $CD8^+$ T cells, also known as cytotoxic T lymphocytes (CTL), recognise peptide-loaded MHC class I complexes (pMHC), and can lead to the rapid induction of apoptosis in the infected host cell, in addition, CTLs can release pro-inflammatory effector molecules.

1.1.5 Major Histocompatibility Complex (MHC) molecules

The two classes of MHC proteins differ in a number of ways but have a similar structure (which is described in more detail in following sections). The most notable structural feature is the peptide-binding cleft that is formed by two membrane distal domains. It is a combination of the bound peptide and MHC protein that is recognised by T cells. For either MHC class, the peptide is bound in an elongated conformation, and is an integral part of the final protein structure, without which the proteins are unstable.

MHC proteins are encoded within the Major Histocompatibility Complex. The MHC genes are the most polymorphic genes known; often with many hundreds of variants, which are known as alleles. The proteins encoded by the MHC alleles can differ by 10-20% of their amino acids, with most alleles being relatively frequent in the population. In contrast, most non-MHC genes are not very polymorphic and if they are polymorphic, the products of these alleles often differ by only one or a few amino acids, alternatively there may be one common allele with other alleles being rare in the population. The polymorphic residues in MHC proteins are predominantly located within the peptide-binding groove and in areas of contact with TCRs. This enables the products of MHC alleles to bind distinct sets of peptides and differentially interact with TCRs. In addition to the high degree of polymorphism, the MHC is polygenic, and MHC genes are expressed from both maternally and paternally inherited chromosomes, so called co-dominant expression. This

combination of polygeny, co-dominant expression and polymorphism ensures that when a pathogen protein is degraded in the cytosol of a cell, there is the highest chance of pathogen-derived peptides being bound by MHC proteins and presented to CTLs, so initiating the immune response.

1.2 Structure and function of MHC class I molecules

MHC class I molecules are a heterodimer of proteins that are expressed at the surface of virtually all nucleated cells. The heterodimer consists of a ~44 kDa MHC-encoded heavy chain (HC) which non-covalently associates with a ~12 kDa non-MHC encoded β_2-microglobulin (β_2m) protein (figure 1.1).

The heavy chain has an evolutionarily conserved gene structure: exon one encodes the 5' untranslated region (UTR) and signal sequence; exons two through to four encode the extracellular α1, α2 and α3 domains; exons five through to eight encode the transmembrane and cytoplasmic domains and 3' UTR. Alternative splicing of some exons has been demonstrated in many species including chicken, where exon seven can be spliced out of the mRNA (Moller et al., 1991; Kaufman et al., 1992); an exon that has been demonstrated to have a role in constitutive endocytosis of class I molecules (Vega and Strominger, 1989). The exon/intron boundaries are characteristic of immunoglobulin gene super-family members, being split between the first and second nucleotides.

The extracellular protein structure is characterised by two membrane proximal Ig domains (α3 and β_2m), which support the α1 and α2 domains that combine to form the membrane distal peptide-binding domain, which consists of eight anti-parallel β strands flanked by two α helices (Bjorkman et al., 1987; Madden et al., 1991; Fremont et al., 1992). In contrast to the continuous α helix of the α1 domain, the α helix of the α2 domain is divided into three short segments: α2-1, α2-2 and α2-3 (ranging from the end of the class I molecule that binds the C-terminus of the peptide (α2-1), to the end which binds the N-terminus of the peptide (α2-3), figure 1.1c) (Wright et al., 2004). An alignment of chicken, human and murine class I mature protein sequences is provided in figure 1.2.

The class I heterodimer has three disulphide bonds: one pair within the β_2m protein (Cys25-Cys80 in human β_2m); one pair within the α3 domain (Cys203-Cys259 in HLA-A*0201); and one pair within the peptide-binding domain (Cys101-Cys164 in HLA-A*0201). Within the cytoplasmic domain, mammalian class I molecules are modified by phosphorylation (Tyr320 and Ser335 in HLA-A*0201), which may have a role in protein trafficking or recycling (Guild et al., 1983; Guild and Strominger, 1984; Lizee et al., 2003).

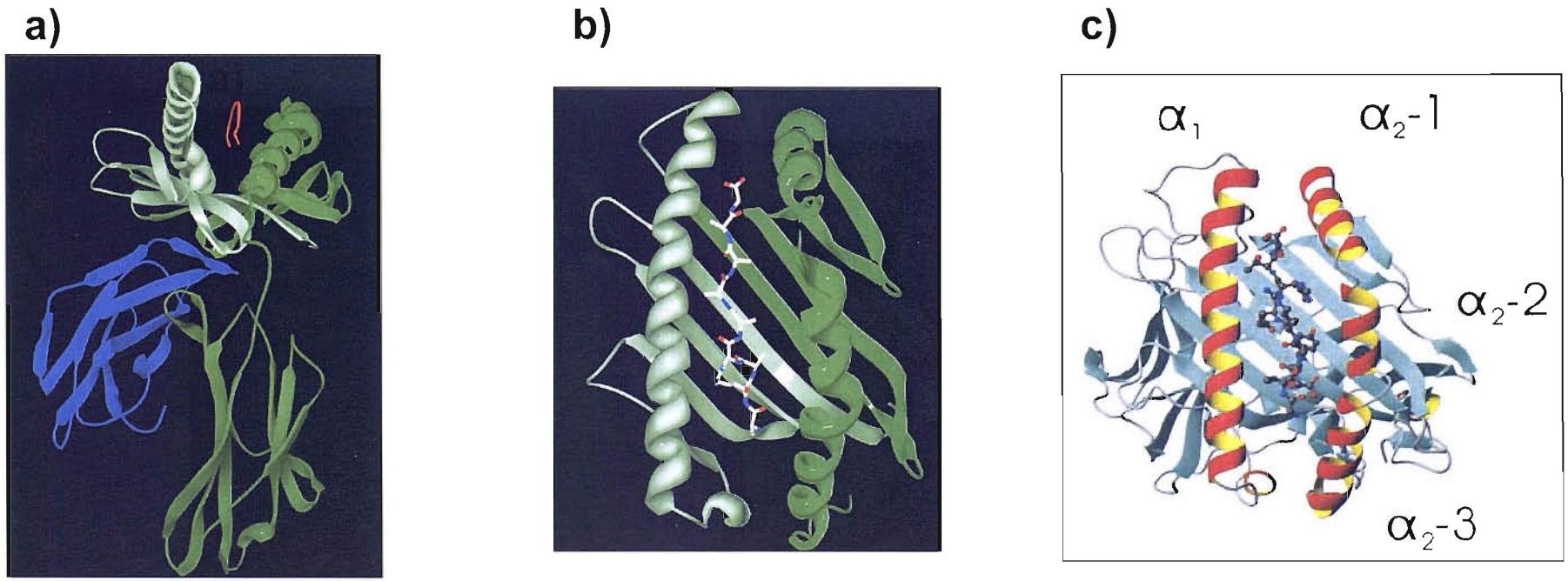


Figure 1.1 The structure of MHC class I molecules

a) The extracellular domains of a class I molecule are show in a ribbon diagram. The α1 domain is shown in light green, α2 domain in green, and the α3 domain in dark green. The $\beta_2$microglobulin protein is shown in blue. The bound peptide is shown in red.

b) The peptide binding domain is viewed from above.

c) The alpha helix of the alpha 2 domain can be divided into three portions: α2-1 (at the end of the peptide-binding groove that binds the C-terminal amino acid of the peptide); α2-2; α2-3 (at the end of the peptide-binding groove that binds the N-terminal amino acid of the peptide). Figure taken from Wright *et al* 2004.

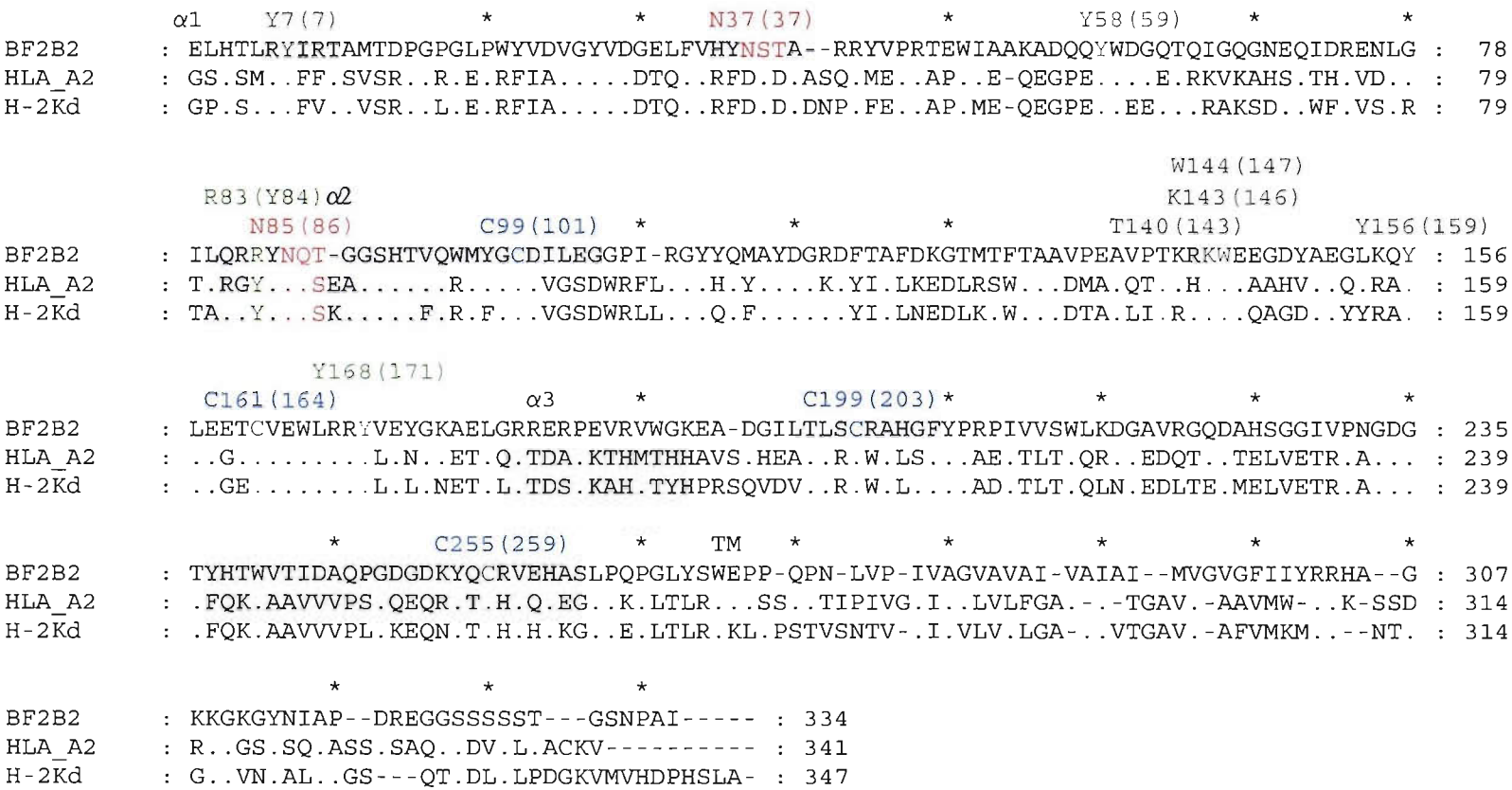


Figure 1.2 Alignment of the amino acid sequences of chicken, human and murine class I molecules

The mature proteins of chicken BF2*2, HLA-A*0201 (K02883) and H-2K^{d} (X01815) were aligned (Kaufman et al., 1992). Dots (.) indicate identities, dashes (-) indicate gaps introduced to maximize alignments. Numbers to the right show the amino acid number of the right hand amino acid. Domains are indicated above the alignment by α1, α2, α3 representing exons 2, 3 and 4, and TM representing the transmembrane and cytoplasmic domains encoded by exons 5-8 (boundaries are aligned with the second character). Residues are numbered according to their location in the chicken protein, with the number in parenthesis referring to the equivalent residue number in the human protein. Sites of N-linked glycosylation are shown in red text, cysteines involved in forming disulphide bridges are shown in blue text, and the conserved residues that form the hydrogen bonds with the N- and C-termini of the peptide are shown in green text.

Class I molecules have between one and three N-linked glycans, the number and distribution of glycans varies between species and between alleles in a species. All class I molecules have an evolutionarily conserved N-linked glycan at one end of the α helix in the α1 domain (Asn86 in HLA-A*0201), murine class I molecules are additionally glycosylated in the α2 domain (Asn176) and some alleles (such as H-2L^{d} and H-2D^{b}) are also glycosylated in the α3 domain (Asn256). Rat class I molecules have a different distribution of glycosylation, being glycosylated in the α1 and α3 domains (Asn86 and Asn256). Chicken class I molecules are glycosylated at two locations within the α1 domain, at Asn85 (equivalent to Asn86 of HLA-A*0201) and at Asn37 (Kroemer et al., 1990; Kaufman et al., 1992). Asparagine 37 is located at the end of the β3 strand beneath the α helix of the α1 domain. The side chain of Asn37 points downwards from the peptide-binding domain and is exposed to the solvent between Ala40 and Thr10. No other classical class I molecules are glycosylated at this location; however, the non-classical murine Thymus Leukaemia T3^{b} molecule is glycosylated at this location (Pontarotti et al., 1986). T3^{b} molecules are recognised by T cells (Teitell et al., 1994), however the nature of the ligand that is recognised has not been determined, and is unlikely to be peptide, as T3^{b} molecules are efficiently expressed at the cell surface, and are recognised by T cells, in TAP-deficient transfectants (and are recognised by γδ T cells, which are not pMHC restricted) (Rodgers et al., 1995; Tsujimura et al., 2000). Therefore it is possible that the glycan attached to Asn37 may stabilise the peptide-binding domain in the absence of peptide. If so, chicken class I molecules may also be stabilised to some extent by glycosylation at this location.

Class I molecules bind peptides that are predominantly, but not exclusively, generated following the proteolysis of intracellular proteins that are usually derived from host proteins but can be derived from pathogens that replicate or reside within the cytoplasm. The peptide is usually eight to eleven amino acids long and binds in an elongated conformation within the peptide-binding groove. The chemical nature of the peptides that class I molecules bind (the peptide motif) is determined by the nature of six specificity determining pockets (A-F). The N- and C-termini of the peptide are bound via hydrogen bonds to eight conserved residues surrounding the A and F pockets. The identity of seven of these residues are conserved between humans and chickens: Y7(7), Y58(59), Y156(159), Y168(171), T140(143), K143(146), W144(147), where the residues are numbered according to their location in the mature chicken protein, with the number in parenthesis referring to the equivalent residue number in the human protein, whilst Y84, an invariant residue in mammals, is substituted for R83 in all chicken class I molecules (Kaufman et al., 1992). Polymorphic residues of the class I molecule define the size, shape and charge of the B-E pockets, resulting in the products of different class I alleles binding different sets of peptides; with between one and three pockets exerting strong specificity for peptide side chain: the so called anchor residues of the peptide.

In the periphery of the immune system, pMHC complexes are "scanned" by CTLs, whose T cell receptors (TCR) bind in a conserved orientation, and straddle the top of the peptide-binding groove in a diagonal orientation, contacting both class I molecule and peptide (Garboczi et al., 1996). TCRs are therefore restricted not just by the nature of the peptide they recognise, but also by the MHC molecule that the peptide is bound to (MHC-restriction; (Zinkernagel and Doherty, 1974)). In addition, the surface class I expression level is monitored by specific receptors expressed on Natural Killer (NK) cells (that bind to class I molecules in a distinct fashion to TCRs), and which inhibit NK activation. In numerous viral infections, the expression level of surface class I molecules is decreased (to prevent CTL recognition), however the absence of the class I molecule at the cell surface that is recognised by the inhibitory NK receptor leads to the activation of the cytotoxic functions of the NK cell (the "missing self" hypothesis) (Karre et al., 1986).

1.3 MHC class I antigen processing and presentation

The enzymes that cleave proteins into the peptides that are eventually bound by MHC class I molecules are not designated solely for this purpose; instead the MHC class I antigen processing pathway uses some of the products of the evolutionarily ancient ubiquitin-proteasome pathway (figure 1.3) which performs numerous roles, including maintaining homeostasis by recycling proteins into reusable amino acids (Rock et al., 1994).

1.3.1 The proteasome and cytoplasmic antigen processing

The first step a protein takes along the path to degradation is usually the tightly regulated attachment of polyubiquitin moieties, which is achieved by the sequential actions of ubiquitin activating, conjugating and ligating enzymes. Polyubiquitinylated proteins are cleaved by the cytosolic multicatalytic 26s proteasome, which consists of a hollow barrel shaped cylinder called the 20s core structure, inside of which proteolysis occurs, and two 19s regulatory complexes (or "caps") that attach to both ends of the 20s core structure. The 19s caps recognise polyubiquitinylated proteins, and facilitate substrate entry into the proteolytic chamber in an ATP-dependent fashion. *In vitro* studies suggest the "constitutive" proteasome displays broad target specificity, with most substrates being cleaved, albeit at different efficiencies, with the products of proteolysis being between three to 22 amino acids in length, with an average length of eight to nine amino acids; however an *in vivo* study suggests the majority of proteasomal products are 15 amino acids or greater in length (Reits et al., 2004).

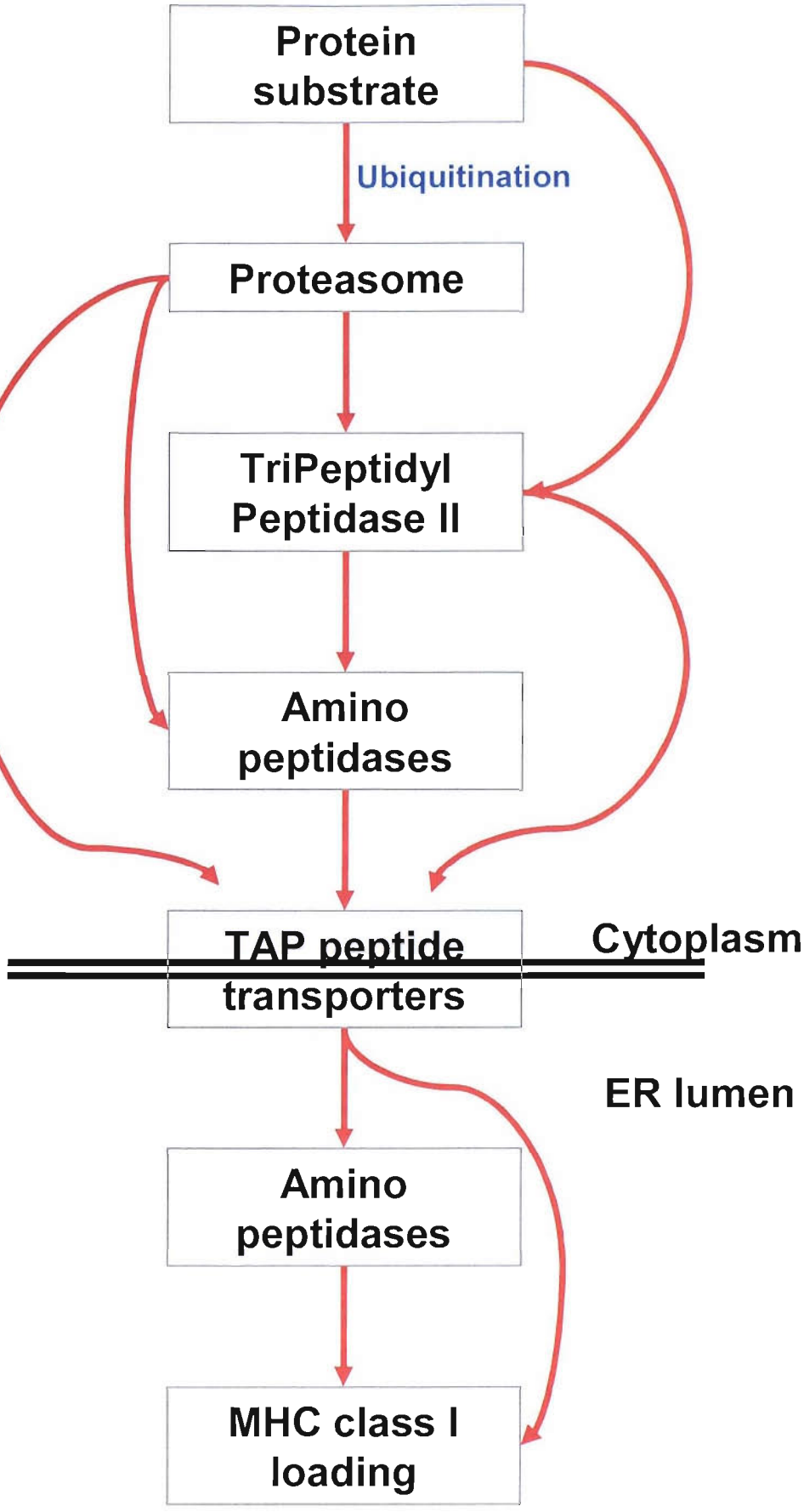


Figure 1. 3 Schematic of the major events that are involved in the MHC class I antigen processing pathway.

Historically the first indication that the proteasome is involved in class I antigen processing came following the identification of two MHC-encoded (and one non MHC-encoded) IFNγ-inducible subunits (Monaco and McDevitt, 1982), that replace the constitutive subunits which confer proteolytic specificity, forming the "immunoproteasome". Exposure to IFNγ also causes the replacement of one or both 19s caps with PA28 activator complexes. The proteolytic specificity and activity of immunoproteasomes is altered to some extent in comparison to constitutive proteasomes; with decreased cleavage after acidic residues, and increased cleavage after basic residues (Dick et al., 1996; Cardozo and Kohanski, 1998; Gileadi et al., 1999; Toes et al., 2001). It has been suggested that constitutive proteasomes efficiently degrade proteins into individual amino acids, therefore maintaining protein homeostasis; whilst immunoproteasomes produce longer proteasomal products, that are more suitable as class I ligands (Cascio et al., 2001; Strehl et al., 2005).

The role of the proteasome in class I antigen processing has since been confirmed by the use of chemical inhibitors that prohibit proteasomal degradation, leading to most class I molecules being starved of peptides, and a subsequent decrease in surface expression levels (Mo et al., 1999). These studies also demonstrated that the proteasome is largely responsible for generating the correct C-terminus of the peptide, as class I antigen presentation decreased when C-terminally extended (but not N-terminally extended) peptides were subjected to proteolysis in the presence of proteasome inhibitors.

It has been suggested that the majority of proteins that are degraded by the proteasome are not proteins that have come to the end of their useful lives, but are instead defective ribosomal products (so called DRiPs) (Yewdell et al., 1996). DRiPs are loosely defined as defective forms of newly synthesised proteins that never attain native conformation, which may be co-translationally degraded following errors in transcription, splicing, translation, folding or assembly. One study suggested, depending upon cell type, between 30-80% of newly synthesised proteins were polyubiquitinylated and degraded within a few minutes of assembly (Schubert et al., 2000), whilst another study suggested that DRiPs account for ~10% of the total protein synthesis (Bulik et al., 2005). However, the significant contribution that DRiPs make as class I ligands was revealed by two studies which showed decreasing DRiP production through the inhibition of protein synthesis led to a rapid decrease in the import of peptides into the ER for class I assembly, and to a decrease in class I export from the ER (Reits et al., 2000b; Schubert et al., 2000). The degradation of a significant proportion of proteins following their synthesis allows the rapid presentation of peptides that are derived from proteins that may be metabolically stable, or distributed to cellular compartments inaccessible to the proteasome.

The products of proteasomal degradation can be further digested by a variety of cytosolic peptidases; however one study suggested the majority of proteasomal products are subjected to

proteolysis by TriPeptidyl Peptidase II (TPPII) (Reits et al., 2004). TPPII displays exopeptidase and endopeptidase activities, and acts specifically upon peptides that are at least 15 amino acids in length. A significant role for TPPII in class I antigen processing was demonstrated by inhibition of TPPII, which produced similar effects to proteasomal inhibition (Reits et al., 2004). Other cytosolic peptidases display aminopeptidase activity with different cleavage specificities, but are likely to provide a level of redundancy (Towne et al., 2005). The combined action of cytosolic peptidases is the very efficient degradation of proteins into amino acids; with cytosolic peptides likely to be degraded within seconds (Reits et al., 2003).

Class I antigen processing is remarkably inefficient, with many proteins being degraded to generate few peptide-loaded class I complexes (pMHC). This may be a consequence of the cytosolic proteases displaying no minimal size specificity, and therefore destroying class I ligands by degrading proteins below the optimum size for class I-binding. Attempts have been made to quantify the efficiency of antigen processing and presentation, which collectively suggest one pMHC is produced from approximately a few thousand proteins digested (Princiotta et al., 2003; Yewdell et al., 2003), however efficiencies are likely to vary between cells and upon cytokine exposure.

There have been suggestions that some proteolytic products can be protected from further proteolysis by association with cytosolic peptide-binding chaperones, including heat shock proteins hsp70, hsp90 and gp96, which may shuttle the peptide directly to the ER membrane. However, the strongest evidence for the involvement of a specific chaperone in this process concerns the tailless complex polypeptide-I (TCP1) ring complex (TRiC) (Kunisawa and Shastri, 2003), which binds a subset of N-terminally extended peptides. The depletion of TRiC led to the loss of this peptide subset and to a decrease in MHC class I expression levels. However, how peptides associate with such chaperones has yet to be determined.

Whilst most peptides are degraded completely within the cytoplasm, a small proportion are transported into the lumen of the ER, where class I molecules are loaded with peptides. The peptides are transported by a heterodimer of MHC-encoded proteins called the transporters associated with antigen processing (TAP).

1.3.2 The transporters associated with antigen processing

The TAP heterodimer consists of two proteins, TAP1 and TAP2, which are members of the large and ancient ATP-binding cassette (ABC) family, whose members collectively transport a variety of molecules across membranes in an ATP-dependent fashion, and have a characteristic structure. The TAP heterodimer can be loosely divided into two domains; a transmembrane domain, which crosses the ER membrane multiple times and forms a pore through which peptides are translocated, and is responsible for determining which peptides are transported; and two cytoplasmic nucleotide-

binding domains, which alternately hydrolyse ATP to open and then close the pore that provides access to the ER lumen (Reits et al., 2000a).

The TAP heterodimer typically transports peptides that are between eight and 16 amino acids in length (van Endert et al., 1994), but longer peptides can be transported at lower efficiencies (Momburg et al., 1994a). The nature of the peptides that are transported is determined by the first three (N-terminal) amino acids and most significantly the last (C-terminal) amino acid of the peptide (Neefjes et al., 1995; van Endert et al., 1995). The human TAP heterodimer has broad specificity, and transports most peptides with the exceptions of those peptides with proline near to the N-terminus, and those peptides with acidic C-terminal amino acids (Momburg et al., 1994b). The murine TAP is more restrictive, and preferentially transports peptides with hydrophobic or aromatic residues at the C-terminus (Schumacher et al., 1994a). In the rat, two allelic families of TAP2 exist; one family is as permissive as the human TAP molecules, whilst the second family transports is more restrictive, similar to the murine TAP molecules (Livingstone et al., 1989; Livingstone et al., 1991; Powis et al., 1991; Joly et al., 1998).

1.3.3 ER aminopeptidase activity

Aminopeptidase activity within the ER was suspected long before it was demonstrated; a consequence of the low efficiency at which peptides with proline near to the N-terminus are transported by the TAP molecules, and the relatively high frequency at which proline occupies such positions in class I ligands. This suggested N-terminally extended peptides were transported into the ER, where they were subsequently trimmed (Neisig et al., 1995). Some or all of the enzymes that are responsible have now been identified. In mice the IFNγ-inducible ER aminopeptidase associated with antigen processing (ERAAP) is likely to provide the sole aminopeptidase activity (Serwold et al., 2002; Hammer et al., 2006). Whilst in humans, IFNγ-inducible ERAP1 and ERAP2 have been identified, which have subtle alterations in specificity, and can form heterodimers that act co-operatively to trim N-terminally extended peptides (Saric et al., 2002; York et al., 2002; Saveanu et al., 2005).

ERAP1 differs from cytosolic peptidases in that peptides are not trimmed completely, with products trimmed no shorter than eight amino acids in length, a size that is suitable for most class I alleles (Saric et al., 2002; York et al., 2002). The contribution that ER aminopeptidases make to class I antigen presentation has been determined by depletion by RNA interference (RNAi), or by generating knock-out mice. In most cases, depletion of ER aminopeptidases led to a decrease in class I surface expression; a result that is likely to be due to the loading of class I molecules with sub-optimal peptides that subsequently dissociate rapidly at the cell surface (Hammer et al., 2006). Depletion of ER aminopeptidases also caused a decrease in the presentation of some epitopes (York et al., 2002; Saveanu et al., 2005). However, in other cases ER aminopeptidase depletion led to an increase in epitope presentation, which suggests the usual trimming activity destroys the

epitope (York et al., 2002; Saveanu et al., 2005; Hammer et al., 2006). The significance of ER aminopeptidase depletion was however modest, unless the TAP transporters were inhibited; which suggests that a substantial proportion of N-terminally extended peptides are retrotranslocated to the cytosol to undergo further trimming, before being returned to the ER as appropriate class I ligands (Saveanu et al., 2005).

1.3.4 The early stages of class I biosynthesis

Class I heavy chains (HC) co-translationally enter the ER, where preformed oligosaccharides ($GlcNAc_2$, Man_9, Glc_3) are attached to suitable asparagine residues. Subsequently, the oligosaccharides are trimmed, generating mono-glucosylated intermediates that are bound by the lectin-binding chaperone calnexin (Diedrich et al., 2001). The association of calnexin is likely to prevent the HC aberrantly folding and consequently being rapidly degraded; however calnexin deficiency does not adversely affect class I surface expression and antigen presentation (Scott and Dawson, 1995), suggesting other chaperones provide a level of redundancy.

Calnexin also recruits the soluble thioloxidoreductase ERp57 to the newly-made HC (Lindquist et al., 1998; Morrice and Powis, 1998), which at this stage have yet to form disulphide bonds (Farmery et al., 2000). It is likely that ERp57 facilitates correct HC oxidation; however the absence of ERp57 does not prevent appropriate HC oxidation (Garbi et al., 2006), suggesting either that there is redundancy among ER thioloxidoreductases, or that non-enzymatic oxidation proceeds sufficiently to allow complete protein oxidation.

Independent of the HC, calnexin can form a "pre-loading" complex with TAP and ERp57 via the chaperone tapasin (which is discussed later) (Diedrich et al., 2001). This pre-loading complex has been proposed to act as a scaffold to which class I molecules are subsequently recruited (Diedrich et al., 2001).

1.3.5 Introduction to the MHC class I peptide-loading process

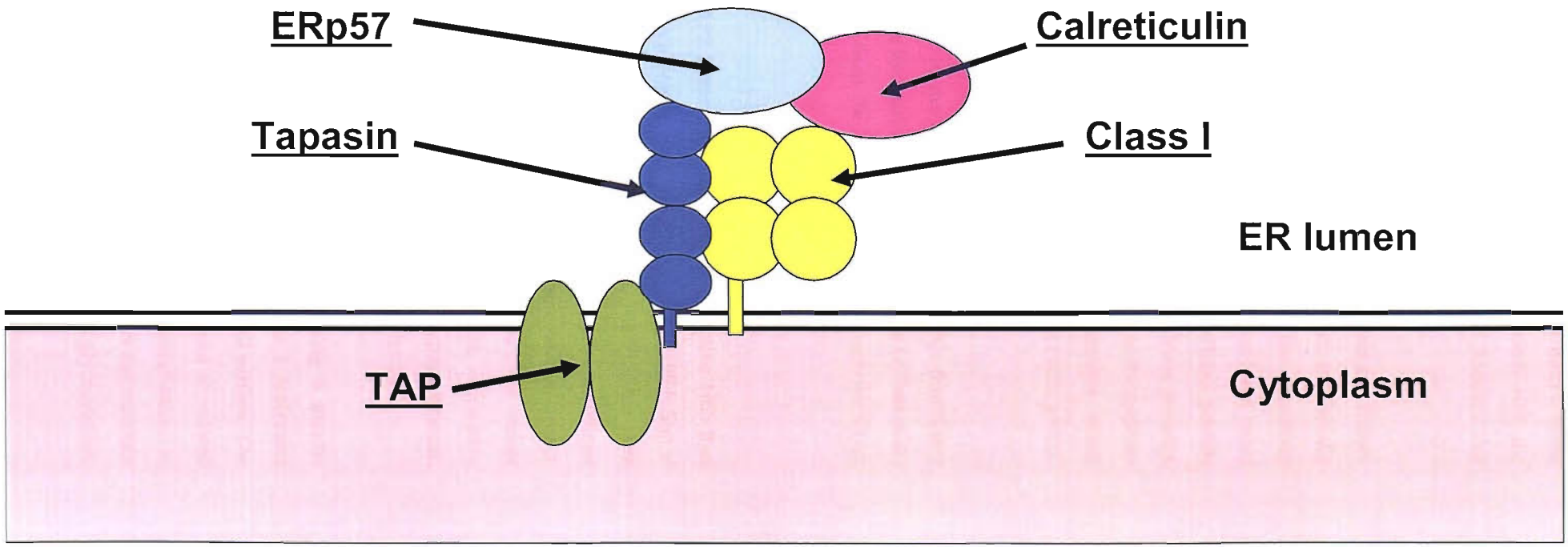
Once oxidised, the HC associates with β_2m to form a peptide-receptive dimer, at which stage the association between HC and calnexin is replaced by the HC: β_2m dimer associating with the soluble lectin-binding chaperone calreticulin (Sadasivan et al., 1996). The specificity of the interaction between calreticulin and HC: β_2m dimers is mediated solely by the monoglucosylated moiety of the HC; as HC folding status or peptide occupancy does not influence calreticulin association (Harris et al., 1998; Wearsch et al., 2004).

Once formed, most peptide-receptive HC: β_2m dimers are incorporated into "peptide-loading complexes" (PLC, figure 1.4), which transiently stabilise the interaction between HC and β_2m, and facilitate the loading of high affinity peptides into class I molecules.

Figure 1.4 The mammalian Peptide Loading Complex

Cartoon depicting the likely organisation of the mammalian PLC. Proteins are coloured and indicated for clarity. Note as many as four tapasin: class I dimers: calreticulin: ERp57 molecules may be present per TAP heterodimer.

The interactions between proteins of the PLC are likely to be collectively strengthened. The MHC heavy chain is tethered via tapasin to the TAP heterodimer, with the ER luminal domains of tapasin interacting with the class I molecule, and the transmembrane and/or cytoplasmic domains interacting with the N-terminal domains of the TAP heterodimer. ERp57 is likely to be maintained in the PLC by a transient disulphide bond with tapasin, and a low affinity interaction with calreticulin. Calreticulin may be maintained in the PLC by associating with the monoglucosylated glycan of the HC, and by the interaction with ERp57.



Peptides within the ER are available for binding to class I molecules, but will only form a stable pMHC complex if the peptide is of an appropriate length to form hydrogen bonds with the conserved residues at either end of the peptide-binding groove, and has suitable residues that complement the specificity determining pockets of the class I molecule.

It is likely that the production of stable pMHC complexes is a two-step process (Lewis and Elliott, 1998). In the first step newly-assembled peptide-empty class I: β_2m heterodimers are initially stabilised by binding a low affinity peptide, which may not fully complement the specificity determining pockets of the class I molecule, but which prevents the heterodimer from rapid denaturation and degradation. This pre-peptide cargo is then replaced by a higher affinity peptide through a peptide-optimisation process, which occurs through the actions of numerous proteins, some of which are dedicated solely for this purpose, that are collectively referred to as the peptide-loading complex. These proteins act in a co-ordinated and complementary fashion which collectively ensures the loading of a broad range of high affinity peptides, producing stable pMHC complexes.

High affinity peptide occupancy induces a change in the conformation of the HC: β_2m dimer (Smith et al., 1992), and increases the strength of the interaction between HC and β_2m (Elliott et al., 1991). Peptide-loaded class I molecules dissociate from the PLC and progress through the Golgi apparatus *en route* to the cell surface.

1.3.6 The peptide-loading complex (PLC)

The human PLC consists of the TAP1 and TAP2 heterodimer, the peptide-receptive HC: β_2m dimer, calreticulin, ERp57 and the class I-specific chaperone tapasin (Ortmann et al., 1994; Suh et al., 1994; Sadasivan et al., 1996; Hughes and Cresswell, 1998; Li et al., 1999). The PLC has been estimated to constitute four HC: β_2m dimers and four tapasin molecules per TAP heterodimer (Ortmann et al., 1997), with calreticulin and ERp57 likely to be present at the same ratio as tapasin. The interactions between components of the PLC have yet to be precisely determined, but mutant cell lines or knock-out mice have been individually generated for all PLC components, which have aided the partial characterisation of the role of each protein in the peptide-loading process. However, the co-operative nature of the protein interactions has complicated this analysis.

The role of calreticulin in class I antigen presentation has been assessed using a murine calreticulin-deficient fibroblast cell line (Gao et al., 2002). In these cells, class I molecules associated with otherwise normal PLCs, but trafficked at an increased rate to the cell surface. The class I molecules presented some antigens to CTL at a reduced level, were expressed at reduced steady state levels, and were more receptive to exogenous peptide-induced stabilisation in comparison to wild-type cells. However, it must be noted that murine HC: β_2m dimers can maintain an association with calnexin, a likely consequence of the additional site of N-linked glycosylation,

which may suggest calreticulin-deficient human cell lines may exhibit a more severe phenotype. The defects in class I expression that are apparent in the absence of calreticulin are likely to be the result of sub-optimal peptide-loading. As the components of the PLC associate despite the absence of calreticulin, it is possible that calreticulin provides a post-PLC stage of quality control to the peptide-loading process (Gao et al., 2002; Wright et al., 2004). In such a process, it is likely that pMHC complexes that have exited the ER loaded with sub-optimal peptides will rapidly dissociate, returning the peptide-binding groove to a molten and flexible conformation. Such a conformation may be recognised by the enzyme UDP glucosyl glycoprotein transferase (UGGT), which reglucosylates the N-linked glycan of the class I molecule. Once the HC is monoglucosylated, calreticulin is able to bind and return such molecules to the ER via KDEL-mediated vesicular recyclying. Within the ER, peptide-loading may proceed as before and when pMHC complexes are released from the PLC, their monoglucosylated glycan is deglucosylated, removing the calreticulin recognition site, and allowing pMHC complexes to egress from the ER.

1.3.7 An introduction to the role of tapasin in the PLC

Tapasin was identified as a component of the PLC after it was reported that peptide-empty class I molecules associated with the TAP heterodimers (Ortmann et al., 1994; Suh et al., 1994), and that a previously unidentified ~48 kDa protein also co-immunoprecipitated with these molecules (Ortmann et al., 1994). Subsequently it was found that the defect in a mutant human B cell line (721.220), which exhibited deficiencies in class I surface expression (Greenwood et al., 1994), and in which class I molecules did not associate with the TAP heterodimers (Grandea et al., 1995), was due to the absence of an uncharacterised MHC-encoded protein (Grandea et al., 1995) that was proposed to bridge the class I and TAP proteins, which was named the "TAP associated protein" or tapasin (Sadasivan et al., 1996). Transfection of a cDNA encoding tapasin restored the interaction between class I and TAP molecules, and increased the level of surface expressed class I molecules (Ortmann et al., 1997), demonstrating that tapasin performs a critical role in the PLC.

The availability of mammalian cell lines that are deficient for tapasin expression, and of tapasin knock-out mice has allowed the partial characterisation of the role that tapasin plays in class I antigen presentation. Tapasin knock-out mice exhibit similar defects in class I expression to the human tapasin-deficient 721.220 cell line (Greenwood et al., 1994; Grandea et al., 1995; Garbi et al., 2000; Grandea et al., 2000). In tapasin knock-out mice, class I surface expression levels are reduced ~10 fold, with the few class I molecules that reach the cell surface exhibiting decreased stability and an impaired ability to present some antigens to CTL. As a likely consequence of the low surface expression levels of class I molecules in tapasin knock-out mice, there is a significant reduction in the numbers of $CD8^+$ T cells, which show deficiencies in the positive or negative selection of CTL specific for some antigens (Grandea et al., 2000). However, these deficiencies in class I expression are not as severe as is evident in TAP-deficient or β_2m-deficient cells and mice,

which suggests that tapasin optimises the peptide-loading process, which can occur to some extent in the absence of tapasin.

Three inter-related roles have been ascribed to tapasin: firstly, acting as a molecular bridge linking peptide-receptive HC: β_2m dimers to the N-terminal domains of the TAP heterodimer (Lehner et al., 1998; Bangia et al., 1999; Tan et al., 2002; Koch et al., 2004); secondly, stabilising the TAP heterodimer in a poorly defined way, thereby increasing TAP protein expression levels and consequently increasing the amount (but not the rate) of peptide import (Bangia et al., 1999; Garbi et al., 2003)}; and thirdly, facilitating the peptide-loading process, leading to the production of stable pMHC complexes that are loaded with high affinity peptides that exist at the cell surface for elongated periods of time (Williams et al., 2002; Howarth et al., 2004).

Based upon two lines of evidence it has also been proposed that tapasin retains peptide-empty or sub-optimally loaded class I molecules within the cell. Firstly, tapasin prevented empty H-2K^b molecules from being expressed at the cell surface of insect cells, unless a source of appropriate peptide was provided (Schoenhals et al., 1999). Secondly, the rate of maturation of H-2K^b molecules expressed in human tapasin-deficient cells was slowed by tapasin expression (Barnden et al., 2000). However this ER retention function was questioned by two findings: firstly, class I molecules trafficked to the cell surface at the same rate in tapasin knock-out mice as in wild-type cells, but were loaded with sub-optimal peptides that reduced their stability and expression levels (Grandea et al., 2000). Secondly, there are other reports of mammalian class I molecules that traffic at the same rate in tapasin-deficient cells as in wild-type cells, and whose overall quality of peptide-loading is substantially improved by tapasin-mediated peptide-loading (Grandea et al., 1997; Lewis et al., 1998). Collectively these studies suggest that tapasin does not act exclusively to retain peptide-empty or sub-optimally loaded class I molecules, as the rate of maturation of class I molecules expressed in the absence of tapasin would be expected to be faster than is apparent when peptide-loading occurs in the presence of tapasin. Instead, it would appear tapasin improves the peptide-loading process, increasing the production of stable pMHC complexes.

Whilst all murine class I alleles exhibit defects in their expression when peptide-loading occurs in the absence of tapasin, human class I alleles exhibit variable "dependence" upon tapasin to be expressed at the cell surface and present antigen to CTL (Greenwood et al., 1994; Peh et al., 1998). This finding led to HLA alleles initially being categorised as either "tapasin-dependent", for example HLA-B*4402, or "tapasin-independent", for example HLA-B*2705, or varying between these extremes. However, closer analysis has revealed that whilst for some class I alleles the total level of surface expressed class I molecules or the ability to present certain antigens to CTL may be unaffected by the absence of tapasin, it is clear that those class I molecules expressed in the absence of tapasin are loaded with an allele-specific proportion of sub-optimal peptides, and that the stability of tapasin-independent alleles is improved by tapasin when it is present (Williams et

al., 2002). This demonstrates that peptide-loading of all class I alleles is improved by the action of tapasin. It has been suggested that class I alleles that are less dependent upon tapasin to achieve peptide-loading reflect an adaptation of the host to pathogen subversion of the action of the PLC (Purcell et al., 2001; Zernich et al., 2004); which is targeted by a variety of viral proteins (including tapasin (Bennett et al., 1999; Lybarger et al., 2003; Park et al., 2004)).

1.3.8 Tapasin gene and protein structures

Translation of the human tapasin cDNA sequence suggests a 428 amino acid type I transmembrane protein (Ortmann et al., 1997). The human tapasin gene consists of eight exons (Herberg et al., 1998), with the signal sequence being encoded by the first exon and part of exon two, which also encodes the first 49 amino acids of the ER luminal portion of the mature protein. The structure of the N-terminal domains of the ER luminal portion of tapasin remains poorly defined, although it was noted from homology searching that exon three displays weak homology to an Ig domain (Herberg et al., 1998). However, two Ig domains are predicted within the ER luminal portion; an Ig-V domain is likely to be encoded by exon four; and a membrane-proximal Ig-C domain by exon five. A connecting peptide and part of the transmembrane domain are encoded by exon six, with the remainder of the transmembrane domain and part of the cytoplasmic tail being encoded by exon seven. Exon eight encodes the remainder of the tail and the 3' UTR (Li et al., 1997; Ortmann et al., 1997; Herberg et al., 1998). A schematic of the predicted structure of the human tapasin protein is provided in figure 1.5. A dilysine motif (KKxx) is found at the C-terminus of the cytoplasmic tail, which is recognised by subunits of the COP1 coatomer, which return proteins from the Golgi back to the ER (Paulsson et al., 2002).

To date there is no x-ray crystal structure of tapasin, but structural predictions and biophysical studies have provided tantalising clues. Biophysical studies suggest tapasin assumes an elongated structure in solution, and controlled proteolysis suggests the ER luminal portion is broadly defined into two domains (potentially divided into further sub-domains) separated by a flexible linker ~90 amino acids from the N-terminus (Chen et al., 2002). Tapasin has been proposed to adopt a structure similar to that of class I molecules (Mayer and Klein, 2001). This proposal was based upon the similarity between the membrane-proximal Ig-C domain of tapasin (exon 5) and the class I α3 domain; whilst other domains of tapasin which exhibit less homology with Ig domains (exons 3 and 4) may resemble a modified peptide-binding groove; with the extreme N-terminus of tapasin (~100 amino acids) remaining structurally undefined.

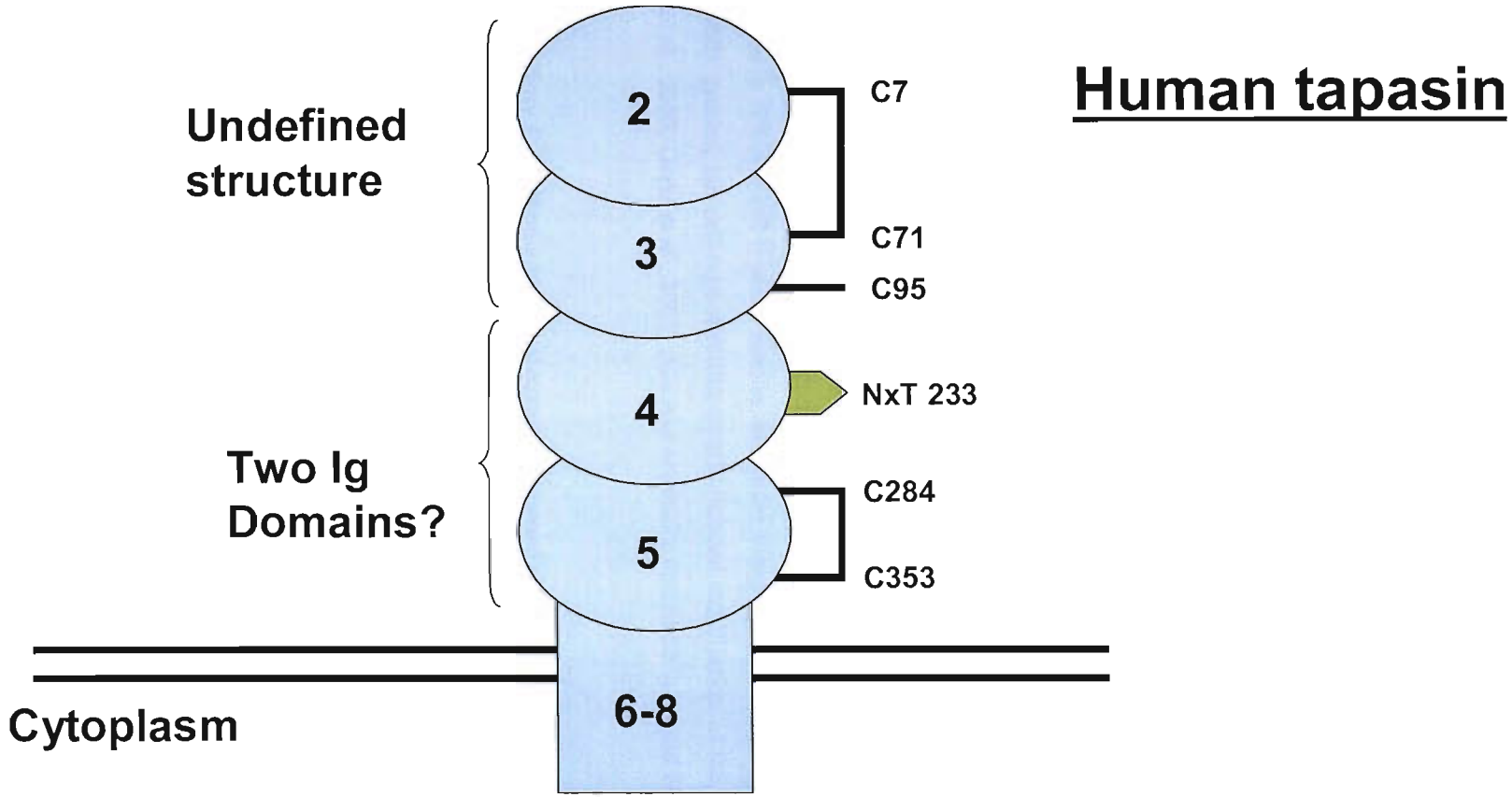


Figure 1.5 Cartoon depicting the likely structural organisation of the human tapasin protein
Ovals or rectangles depict potential domains encoded by one or more exons, which are numbered within the shape. The site of N-linked glycosylation is shown in green. Potential disulphide bonds are shown as black lines, together with the numbering of cysteine residues. The first residue of the predicted mature protein is numbered +1, assuming a 20 amino acid signal sequence.

1.3.9 Partial characterisation of structural elements of tapasin, and the identification of specific residues of tapasin that interact with other PLC constituents

By truncating tapasin at the C-terminus, a soluble tapasin molecule which comprised only the ER luminal domains was shown to interact with class I molecules, and to facilitate an increase in their surface expression levels and presentation of antigen to CTL (Lehner et al., 1998). However, this soluble tapasin molecule did not associate with the TAP molecules, a finding which initially questioned the relevance of the interaction between class I and TAP molecules. Subsequently, studies have demonstrated that soluble tapasin is less able to facilitate optimal peptide-loading in comparison to full-length tapasin (Tan et al., 2002; Williams et al., 2002).

Progressive N-terminal truncations of tapasin demonstrated that the transmembrane and cytoplasmic domains mediate association of tapasin with the TAP heterodimer (Bangia et al., 1999). The association of N-terminally truncated tapasin molecules with the TAP heterodimer led to an increase in TAP protein expression levels, and consequently the import of peptides into the ER. Despite these beneficial effects on the TAP molecules, class I molecules were not recruited to the PLC, and the deficiencies in class I expression remained; suggesting the major function of tapasin was to promote optimal peptide-loading, rather than acting to increase the availability of peptides in the ER. However, the interaction between tapasin and TAP molecules may vary in significance in different cell types or species, as studies using tapasin knock-out mice have shown significantly greater, but cell-type specific, tapasin-induced stabilisation of the TAP heterodimer (Garbi et al., 2000; Garbi et al., 2003) than was originally described in human cells (Bangia et al., 1999).

Beyond the N- or C-terminally truncated mutants discussed above there have only been limited and piecemeal attempts to determine the influence of structural elements in tapasin upon its function. However, two studies have revealed some of the residues in tapasin that are involved in the association with the TAP heterodimer. It was predicted that the interaction between tapasin and TAP molecules may be mediated by a lysine residue within the transmembrane of tapasin (K408 in human tapasin) (Ortmann et al., 1997), which is conserved in most species, but is replaced by arginine in chicken tapasin. Accordingly, mutation of this residue in human tapasin (K408A or K408W) decreased both the strength of the interaction with TAP and the tapasin-induced stabilisation of TAP, and adversely affected the rate of maturation and surface expression levels of class I molecules (Petersen et al., 2005), which collectively suggests that mutation of this residue adversely affected the ability of the class I molecules to become loaded with peptides. In a separate study, it was found that a point mutation (L410F) within a predicted leucine zipper motif (L396, L403, L410 and L417) in the transmembrane domain of human tapasin failed to stabilise TAP protein expression levels, and the HLA-B*4402 molecules displayed sub-optimal expression at the cell surface (Tan et al., 2002). This mutation also partially released tapasin from ER retention,

suggesting that L410 within the possible leucine zipper motif contributes to ER retention in addition to the dilysine motif.

The region of the TAP heterodimer with which tapasin interacts has been determined to be the N-terminal four transmembrane domains of TAP1 (amino acids 2-166) and the N-terminal three transmembrane domains of TAP2 (amino acids 2-122) (Koch et al., 2004; Leonhardt et al., 2005). Deletion mutants lacking these "N-domains" form functional peptide-binding and peptide-transporting heterodimers in the ER membrane (forming a core transmembrane segment that is predicted for most ABC transporters), but are unable to bind tapasin (Koch et al., 2004; Procko et al., 2005). Interestingly, deletion of the N-terminal tapasin-binding region of TAP2, but not TAP1, resulted in a failure to form stable intracellular pMHC complexes, suggesting a more significant role for TAP2 than TAP1 in the PLC (Leonhardt et al., 2005). For both TAP proteins, these N-terminal tapasin-binding domains are the least conserved regions of the TAP proteins, when multiple species are compared (Walker et al., 2005). Whilst the exact details of the association between tapasin and the TAP heterodimer remains to be solved, collectively these studies provide further evidence that the bridging of class I molecules to the TAP heterodimer is required for complete optimisation of the peptide repertoire.

Two other mutagenesis studies have been conducted to characterise whether any of the cysteine residues in mammalian tapasin (figure 1.5) are involved in internal or intermolecular disulphide bonds. By mutation, it was determined that C7 and C71 were likely to form a disulphide bond at the N-terminus of the protein, whilst C95 mediated a transient disulphide bond with ERp57 (Dick et al., 2002); with virtually all tapasin molecules being disulphide bonded to ERp57 (Peaper et al., 2005). Importantly, mutation of C95 in human tapasin resulted in the failure to incorporate ERp57 into the PLC (Dick et al., 2002). However, the expression of tapasin C95 mutants in human cells produced a slightly different phenotype to that of mice which have a B cell specific deletion of ERp57 (referred to as ERp57 knock-out mice hereafter) (Garbi et al., 2006). In both systems, slight reductions in class I surface expression levels were noted, which may vary in an allele or species-specific fashion. In ERp57 knock-out mice, few HC (or calreticulin) molecules associate stably with the PLC, with the class I molecules trafficking at a faster rate to the cell surface, where they are fully oxidised, but presumably sub-optimally peptide-loaded. In comparison, if ERp57 is present but excluded from the PLC by mutation of C95 in tapasin, class I molecules are recruited to the PLC in usual amounts, traffic at the same rate as they do in wild-type tapasin expressing cells, but are incompletely oxidised.

In both situations, class I alleles exhibit variable degrees of deficiencies in their expression, which suggests a repertoire of slightly less optimal peptides were loaded into the class I molecules. Three potential roles have been suggested to explain how ERp57 may contribute to optimising the peptide repertoire (Dick et al., 2002; Wright et al., 2004; Garbi et al., 2006). Firstly, the interaction between

ERp57 and tapasin may induce a conformational change at the N-terminus of tapasin, which may be necessary for tapasin to function fully. Secondly, ERp57 may play a direct role in iterative cycles of class I molecules binding and then releasing peptides that do not form sufficient stabilising contacts, which may involve sequential reduction and re-oxidation of the disulphide bond within the peptide-binding domain. Lastly, ERp57 may periodically reduce the intermolecular disulphide bond with tapasin, which may be necessary to allow the escape of pMHC complexes from the PLC (Wright et al., 2004).

The differences that exist between the two systems may be reconciled in a number of ways. The variable strength of HC association with the PLC may be explained because ERp57 is involved in aspects of class I assembly prior to PLC incorporation, in which tapasin is not involved. It has been suggested that the complete absence of ERp57 at an early (pre-PLC) stage of assembly may impair the subsequent recruitment of HC to the PLC (Elliott, 2006), which can proceed normally in the presence of tapasin C95 mutants. The differences in the rates of maturation may be a reflection of the incorporation of less calreticulin to the PLC in ERp57 knock-out mice; a phenotype consistent with the absence of calreticulin (Gao et al., 2002), or decreased HC association with calreticulin (Lewis and Elliott, 1998). The reason for the difference in HC oxidation status remains unclear, but it has been proposed that ERp57 may be unusually reactive if it is present but prevented from associating with tapasin (perhaps through the second cysteine residue of the pair that is present in the thioredoxin motif that disulphide bonds with tapasin) (Elliott, 2006).

The two cysteine residues within the membrane proximal domain (exon 5) allowed the prediction of this domain as an Ig-C domain (Ortmann et al., 1997). Subsequently, it has been shown that these cysteines are involved in a disulphide bond; the presence of which stabilised tapasin itself, and induced stabilisation of TAP, and are required to associate with H-2L^d (Turnquist et al., 2004). Other mutagenesis studies have examined the role of this Ig domain; for example human or murine tapasin deletion mutants that are missing nine amino acids (334-342) from this Ig domain are unable to associate with H-2L^d molecules, but are apparently stable proteins that associate with and increase TAP expression levels (Turnquist et al., 2001; Turnquist et al., 2004). Perhaps more significantly two amino acid point mutations (H334F & H335Y) within this domain apparently act as "hypermorphs", in that they decrease the proportion of sub-optimally loaded H-2L^d molecules expressed at the cell surface (Turnquist et al., 2001). Whilst much remains to be determined about the molecular explanations for the observed phenotypes of these mutants, these findings collectively suggest the membrane proximal Ig domain mediates an important role in the ability of tapasin to facilitate efficient antigen presentation.

1.3.10 Regions of the class I molecule with which tapasin interacts

By analysing the surface expression and PLC incorporation characteristics of mutant class I molecules it has been determined that tapasin interacts with two regions of the class I molecule

(figure 1.6). One of these regions comprises a solvent exposed loop in the α2 domain formed by residues 125-133 in chicken class I sequences (128-136 in mammalian class I molecules), that connects the short α2-1 helix to the β sheets surrounding the F pocket, as most non-conservative mutations within this region prevent, or decrease the ability to associate with the PLC, and result in sub-optimal peptide-loading (Lewis et al., 1996; Peace-Brewer et al., 1996; Yu et al., 1999). Interestingly two mutations within this loop (H-2L^d K131D and HLA-A*0201 S132C) demonstrate increased PLC association and improvements in the quality of their peptide-loading (Lewis et al., 1996; Yu et al., 1999).

Several lines of evidence suggest the loop formed by residues 125-133 is relatively flexible: the short α2-1 helix that lies adjacent to the 125-133 loop, and which forms a side of the peptide-binding groove surrounding the F pocket, has been observed in slightly altered conformations in crystal structures of several class I alleles, and even in crystal structures of the same class I allele bound by different peptides, and both the 125-133 loop and especially the short α2-1 helix are predicted to adopt different conformations in models of peptide-empty and peptide-occupied MHC molecules (Elliott, 1997; Wright et al., 2004; Zacharias and Springer, 2004; Hansen et al., 2005) (figure 1.7).

In addition, tapasin is likely to associate with a solvent exposed β strand in the α3 domain formed by residues 218-223 (or 218-225) in chicken class I molecules (222-227 or 222-229 in mammalian class I molecules) (Kulig et al., 1998; Suh et al., 1999; Harris et al., 2001), which also constitutes the binding site for the CD8 co-receptor. Mutations within this region decrease the ability of the class I mutant to interact with the PLC, and lead to deficiencies in surface expression.

Other residues of the class I molecule have been shown to influence the association with tapasin indirectly. From relatively early on it was noted that particular residues within the peptide-binding domain appeared to determine the extent to which class I alleles are found in association with the PLC (Neisig et al., 1996). However, as there has been only one report of a peptide derived from tapasin being eluted from a class I molecule, it would seem unlikely that tapasin directly enters the peptide-binding groove to contact these residues (Turnquist et al., 2001).

Despite this, attention has focused on two specific residues of mammalian class I molecules, 114 and 116 (111 and 113 in chicken class I molecules), which are respectively located next to or within the F pocket of the peptide-binding groove. In one study it was proposed that mammalian class I alleles with acidic residues at position 114 are strictly dependent upon tapasin to achieve optimal peptide-loading, whilst those mammalian class I alleles with basic residues at position 114 are independent of tapasin to achieve optimal peptide-loading, and that mammalian class I alleles with neutral residues at positions 114 fall between the two extremes (Park et al., 2003).

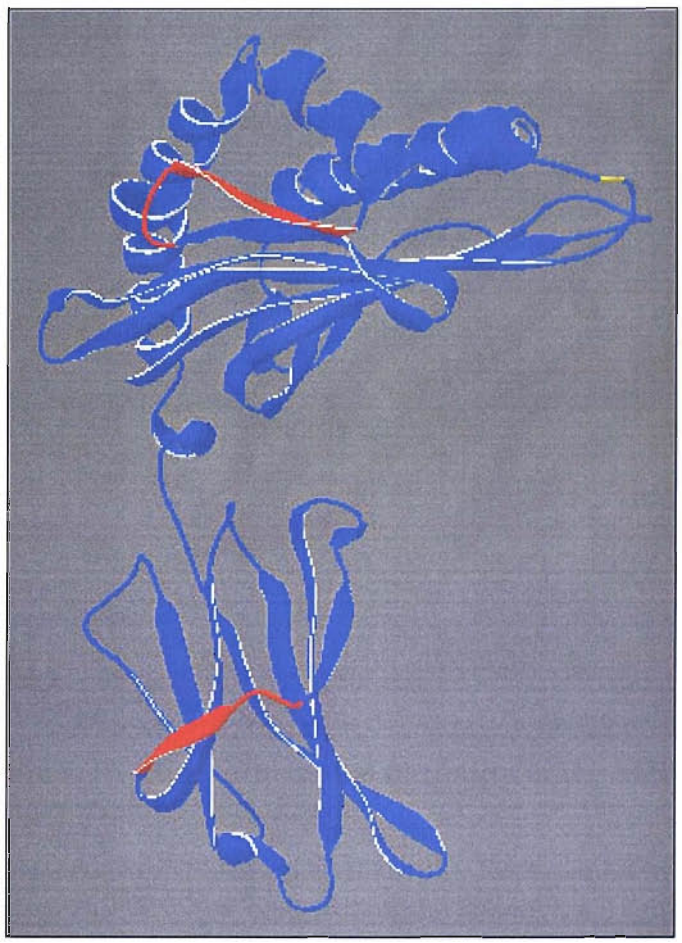


Figure 1.6 3D structures of a mouse MHC class I protein

A model of the H-2K^b class I molecule, shown in ribbon format. β_2m omitted for clarity.

Asn86 is shown in yellow, the site of N-linked glucosylation, where calreticulin can bind to monoglucosylated glycans.

α2 residues 128-136, and α3 residues 222-227, to which tapasin binds are shown in red.

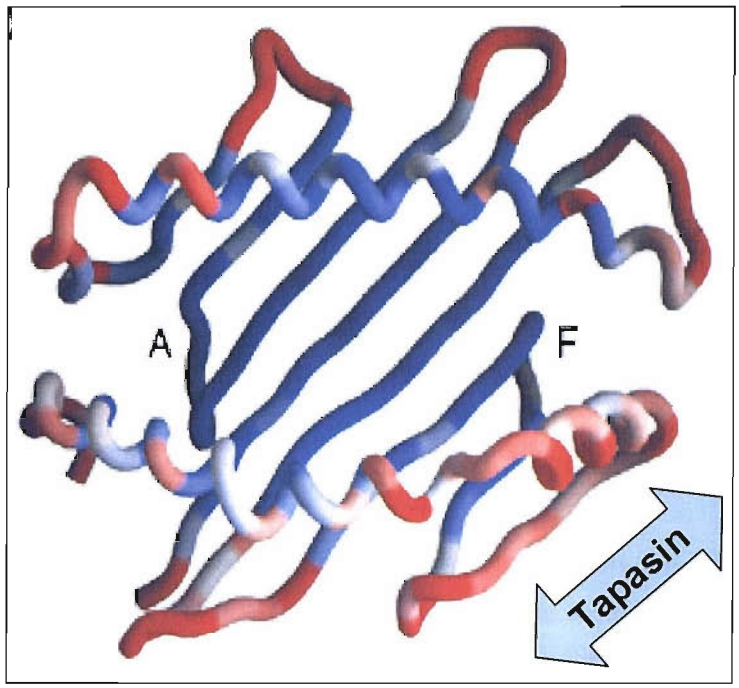


Figure 1.7 Conformational flexibility of the MHC class I peptide binding domain

Figure take from "Hansen *et al* (2005). Recognition of open conformers of classical MHC by chaperones and monoclonal antibodies. Immunol Rev *207*, 100-111."

The peptide binding domain of H-2K^{b} is shown in ribbon format. The regions of the class I molecule that exhibit the greatest variability in crystal structures (i.e. the regions of greatest flexibility) are coloured in red, with the least variability (the least flexible) regions shown in blue. The loop with which tapasin associates is indicated. The ends of the class I molecule at which the A and F specificity determining pockets are located are indicated.

By producing point mutants, it was demonstrated that the strictly tapasin-dependent HLA-B*4402 allele could be made to resemble a tapasin-independent allele, by the D114H mutation. This D114H mutant associated to a lesser extent with tapasin, and was able to load peptide and transit to the cell surface in the absence of tapasin. Reciprocally, the relatively tapasin-independent HLA-B*2705 allele could be made to be strictly dependent upon tapasin to achieve optimal peptide-loading by the H114D mutation.

However it is likely that the overall structural nature of the portion of the peptide-binding groove that surrounds the F pocket determines the level of dependence on tapasin that a class I allele demonstrates, rather than a specific, single residue. For example HLA-B*4402 and HLA-B*4405, which may dictate the extremes of tapasin dependence and PLC association (Williams et al., 2002; Zernich et al., 2004), both have acidic residues at 114 (D114), and differ only at residue 116 (D116 or Y116 respectively). This substitution has minimal effect on peptide repertoire, but drastically alters the route through which peptide-loading occurs, with HLA-B*4405 molecules failing to be detected in the PLC even when TAP peptide transport is inhibited (Zernich et al., 2004).

1.3.11 Evidence that the peptide repertoire is optimised, and that tapasin is involved in this process

The first indication that class I molecules optimise their peptide repertoires before they reach the cell surface came from a study of murine cells infected with the bacteria *Listeria monocytogenes* (Sijts and Pamer, 1997). In this study two peptides derived from the p60 antigen bound to H-2K^b with comparable affinities, but differed in their rates of dissociation. By preventing export to the cell surface via the drug brefeldin A (BFA), it was found that the peptide with the faster dissociation rate was released from intracellularly-retained class I complexes at a rate that was significantly increased in comparison to non-BFA treated cells, whilst the peptide with the slower dissociation rate was virtually unaffected by BFA treatment. In addition, the class I molecules were able to bind new peptide ligands during their intracellular retention.

Further evidence for a peptide-optimisation process was provided by the study of an HLA-A*0201 T134K mutant, which fails to interact with components of the PLC, and traffics rapidly to the cell surface, where the molecules are unstable and are rapidly lost (Lewis et al., 1996; Peace-Brewer et al., 1996; Lewis and Elliott, 1998). However, T134K molecules are not expressed at the surface of TAP-deficient cells, suggesting that sub-optimal peptides can partially stabilise the class I molecules, and permit their exit from the ER, and that the T134K mutation causes a failure to optimise this sub-optimal peptide repertoire.

Similar evidence supporting a peptide-optimisation step was provided by two studies examining the non-classical class I H-2M3 molecule (Chun et al., 2001; Lybarger et al., 2001). H-2M3 molecules bind a highly restricted repertoire of peptides, which are only produced from mitochondrial and

bacterial proteins. Therefore the majority of H-2M3 molecules are usually retained within the ER, pending the provision of appropriate peptides. However, when expressed in the absence of TAP, H-2M3 molecules are rapidly degraded (Chun et al., 2001). This suggests that TAP provided a repertoire of peptides which stabilised the pMHC molecule, but which did not allow export. In the presence of appropriate peptides, tapasin mediates the loading and optimisation of the peptide repertoire, releasing H-2M3 molecules from retention.

Further evidence that tapasin participates in a peptide-optimisation step came from a comparison of the peptide repertoires eluted from HLA-A*0201 molecules that were expressed in tapasin-expressing or tapasin-deficient cells (Barber et al., 2001). Substantially fewer peptides were recovered from tapasin-deficient cells despite both cell types expressing similar amounts of HLA-A*0201 molecules at the cell surface, which is likely to reflect an increased dissociation of sub-optimal peptides during the isolation of the class I molecules that were loaded in the absence of tapasin. A similar decrease in the recovery of peptides eluted from HLA-B*2705 molecules expressed in tapasin-deficient cells has also been observed (Purcell et al., 2001). This suggests that tapasin improved the repertoire of peptides that were loaded into the class I molecules.

By monitoring the expression levels of specific pMHC complexes that were loaded in the presence or absence of tapasin, it was noted that tapasin promoted the expression of pMHC complexes that contained peptides that conferred long half-lives to a greater extent than pMHC complexes that contained peptides that conferred short half-lives (Howarth et al., 2004). In the absence of tapasin this hierarchy of pMHC expression levels that is based upon pMHC half-life collapsed. In cells lacking calreticulin or PLC-associated ERp57, the hierarchy was maintained; suggesting tapasin is the key component in optimising the peptide repertoire to produce more stable pMHC complexes.

Making use of the correlation between the thermal-stability of a radiolabelled cohort of pMHC complexes and the affinity (or quality) of their bound peptide cargo, it was elegantly shown that for all class I alleles, tapasin improved both the rate and the extent to which the peptide repertoire was optimised over increasing periods of time (Williams et al., 2002). The enhanced thermal-stability of pMHC complexes was interpreted as the replacement of low affinity peptides with a repertoire of higher affinity peptides ("peptide editing"). Whilst tapasin-dependent alleles, such as HLA-B*4402, strictly relied upon tapasin to optimise their peptide cargoes, alleles such as HLA-B*2705 or HLA-B*4405, that are classified as tapasin-independent alleles, were capable of self-optimising their peptide repertoire in the absence of tapasin, albeit less efficiently than when tapasin was present. This suggests that the peptide-optimisation process is an inherent characteristic of class I molecules, which occurs to various extents in different class I alleles, and that this peptide-optimisation process is enhanced by the action of tapasin.

However, this interpretation was questioned by another report where the results of two assays were combined to suggest that tapasin did not aid the production of pMHC complexes containing high affinity peptides. The first assay showed that the stability of cell surface pMHC complexes that were loaded in tapasin-deficient cells was substantially decreased in comparison to tapasin-expressing cells. Secondly, the peptides that were eluted from HLA-B*0801 or HLA-A*0201 molecules expressed in tapasin-deficient cells subsequently bound to the respective class I molecules in an *in vitro* binding assay with equal, or greater affinity than those peptides that were recovered from tapasin-expressing cells (Zarling et al., 2003). This result was interpreted as a demonstration that tapasin did not edit the peptide repertoire towards higher affinity peptides, and instead facilitated peptide-loading through two actions. Firstly, tapasin was proposed to stabilise peptide-receptive class I molecules, allowing a broader repertoire of peptides to be loaded, a proportion of which constituted lower affinity peptides than were loaded in the absence of tapasin. Secondly, tapasin mediated a conformational change in peptide-loaded class I molecules, which stabilised the structure of the pMHC complexes irrespective of affinity of the bound peptide, therefore conferring increased stability to pMHC complexes loaded in the presence of tapasin. This interpretation relies upon the assumption that no low affinity peptides were lost during the isolation of the class I molecules that were loaded in the absence of tapasin, which would increase the average binding affinity of the remaining peptides. To address this concern, the authors argue that no low affinity peptides were lost during the purification process, as the amount of peptides that were recovered was in accordance with the level of surface expression.

However, this proposal for the action of tapasin is inconsistent with two further pieces of evidence. Firstly, it was noted in the study in which the cell surface expression of specific pMHC complexes was monitored by flow cytometry (Howarth et al., 2004), that whilst the expression level of specific pMHC complexes loaded in the presence or absence of tapasin varied, the rate at which the specific pMHC complexes decayed from the cell surface of BFA treated cells was indistinguishable. This suggests that tapasin did not mediate a change in conformation that "locked" the peptide into the class I molecule. Secondly, it was noted that the peptides whose pMHC half-life determined the hierarchy of cell surface expression levels of specific pMHC complexes, all displayed comparable binding affinities to peptide-empty class I molecules (Howarth et al., 2004). This suggests that the two assays (*in vitro* peptide-binding assays and cell surface stability assays) do not measure the same attributes of class I peptide-loading. *In vitro* peptide-binding assays may measure the ability of a peptide to bind to peptide-empty class I molecules at neutral pH; whilst the cell surface expression level and stability of class I molecules may be influenced by many other factors occurring in living cells, potentially including constitutive cycles of endocytosis and recycling of class I molecules back to the cell surface, which may act to promote the dissociation of sub-optimal pMHC complexes. Therefore, it appears that the cell surface expression level and stability of pMHC complexes is determined principally by the rate of dissociation of pMHC complexes.

1.3.12 How might tapasin optimise the peptide repertoire?

How tapasin facilitates the production of stable pMHC complexes remains to be determined. Only one report has been published in which tapasin was found associated with a radiolabelled and cross-linked peptide (together with the TAP and class I proteins) (Li et al., 1997), the absence of other similar observations suggests that tapasin does not physically transfer peptides into the peptide-binding groove.

However, as discussed, the regions of the class I molecule with which tapasin interacts have been determined. It is possible that tapasin may contact the α3 domain of the class I molecule to stabilise the interaction between proteins, whilst facilitating peptide-loading via contacting the flexible loop in the α2 domain of the class I molecule (125-133 in chicken class I molecules, 128-136 in mammalian class I molecules) (Wright et al., 2004; Elliott and Williams, 2005). It is likely that the flexibility of this α2 loop is influenced by residues within the peptide-binding groove or the adjacent short α2-1 helix, with the degree of flexibility of this loop being specifically monitored by tapasin (Wright et al., 2004; Zacharias and Springer, 2004; Elliott and Williams, 2005). The extent of the flexibility of the α2-1 helix in a peptide-receptive class I molecule may determine the ability of a class I allele to self-optimise its peptide repertoire in the absence of tapasin (Elliott and Williams, 2005). For example, class I alleles which can self-optimise their peptide repertoire in the absence of tapasin may exhibit less flexibility in the α2-1 helix, and therefore may never adopt a closed and peptide-unreceptive conformation, and always maintain a degree of peptide-receptivity; alternatively these alleles may possess the ability to spontaneously recover from peptide-unreceptive conformations. In contrast, class I alleles that are strictly dependent upon tapasin to facilitate peptide-loading may have an increased amount of flexibility in this α2-1 helix, which may allow the peptide-binding groove to adopt such collapsed and peptide-unreceptive conformations in the absence of tapasin, from which they can not spontaneously recover.

It has been shown that tapasin remains in association with class I molecules for a short while after they are loaded with peptides (Sadasivan et al., 1996; Li et al., 1999). During this fleeting association, the peptide-optimisation process is likely to occur. Such a possibility is consistent with the tapasin-mediated optimisation of the peptide repertoires of tapasin-independent alleles, such as HLA-B*2705 and HLA-B*4405, which suggests tapasin acts upon sub-optimally peptide-loaded MHC molecules. In addition, the strict dependence upon tapasin for peptide-loading that other class I alleles display, suggests that tapasin facilitates the peptide-loading of peptide-empty class I molecules. Therefore tapasin may act upon peptide-empty and sub-optimally peptide-loaded class I molecules.

Three mechanisms have been suggested to explain how tapasin may optimise the peptide repertoire towards higher affinity peptides that produce stable pMHC complexes with long half lives (Elliott and Williams, 2005).

Tapasin may specifically interact with an open and peptide-receptive conformation of class I molecules. The maintenance of class I molecules in such open conformations may raise the energy threshold that is needed for peptides to trigger the change in conformation that secures a peptide within the peptide-binding groove. In addition, an open conformation may promote the release of sub-optimal peptides by disrupting any hydrogen bonds formed with the class I molecules, a process only high affinity peptides may be able to withstand (Elliott, 1997; Wright et al., 2004).

Alternatively, tapasin may interact specifically with closed and peptide-unreceptive conformations of class I molecules. Such conformations may be the result of a collapsed and empty peptide-binding groove, or may be formed by the binding of a sub-optimal peptide which does not form sufficiently stabilising contacts to prevent the tapasin-mediated re-opening of the peptide-binding groove (Wright et al., 2004).

It is also possible that tapasin catalyses the change in conformation that occurs when peptide-binding grooves close around peptides; a function that may not be mutually exclusive to the above possible mechanisms of action.

1.3.13 Post peptide-loading events

Once peptide-loading has occurred and class I molecules have adopted a closed conformation, pMHC complexes dissociate from the PLC and progress to the cell surface, a process that had been assumed to occur through non-selective vesicular traffic. However, it has been shown that the rate limiting step of the entire antigen presentation process is not the PLC-mediated peptide-loading process, but is instead the recognition of pMHC complexes by cargo receptors, such as Bap29/31, that mediate vesicular transport (Spiliotis et al., 2000; Paquet et al., 2004).

Once pMHC complexes have left the ER, two potential processes (or quality control check-points) have been suggested to exist that allow further peptide-optimisation to raise the stability of initially sub-optimally peptide-loaded class I molecules. The first check-point concerns the ER retrieval cycle of UGGT and calreticulin that was described earlier. A second check-point has been suggested to operate beyond the Golgi apparatus of human cell lines, which concerns "proprotein convertases", which are proteolytic enzymes such as furin that operate in low pH vesicles (Leonhardt et al., 2005). In such low pH vesicles, sub-optimally peptide-loaded class I molecules may rapidly lose their peptide cargo, and furin as well as other proteases could trim vesicular polypeptides into peptides that may then bind to peptide-receptive class I molecules that are then returned to the cell surface.

1.3.14 Summary of the MHC class I antigen processing and presentation processes

The production of the peptides that are bound by class I molecules is an extremely inefficient process; with proteolysis of proteins proceeding rapidly and completely within the cytoplasm of a

cell. Only a minor proportion of the peptides that are produced escape complete degradation by being transported by the TAP heterodimer into the ER lumen. In this compartment further peptide trimming occurs, however such processes appear to generate more class I ligands than are destroyed; suggesting that ER aminopeptidases and class I molecules may have co-evolved in their functions.

In contrast to the inefficient generation of potential class I peptide ligands, the actual peptide-loading process is highly orchestrated, with several quality control check-points ensuring only appropriate pMHC complexes populate the cell surface. Within the ER, the PLC is formed from multiple proteins which act in a co-ordinated and co-operative fashion to produce pMHC complexes loaded with high-affinity peptides. Once pMHC complexes have progressed beyond the ER, two further quality control check-points may ensure the cell surface expression of stable, long-lived pMHC complexes that may be needed to ensure recognition by rare antigen-specific CTL.

1.4 The genomic organisation and evolution of the Major Histocompatibility Complex

The MHC was discovered during investigations into tissue transplant survival in inbred strains of mice, which suggested a multi-allelic locus determined transplant rejection (Snell and Jackson, 1958). The MHC has since been the focus of intensive research, and has been found to encode the class I and class II molecules, and to determine: serological alloantigens; allograft survival or rejection; strong mixed lymphocyte reactions (MLR); and cellular co-operation in the immune system.

The MHC is a gene dense region that encodes a high percentage of proteins with functions in the immune response. The clustering of genes with related function may be advantageous in an evolutionary context, as it may allow related gene expression, co-evolution or exchange of DNA by non-homologous gene conversion. There are many autoimmune diseases and responses to pathogen infections that are associated with the MHC, but the gene loci or alleles that are responsible have not always been determined.

The MHC is present in all jawed vertebrates, which are thought to have separated from jawless fish ~555 million years ago (MYA) (Flajnik and Kasahara, 2001). Despite variation in the organisation and number of genes that are found in jawed vertebrates, there is strict conservation of the structure and function of the class I and class II molecules.

The MHC is thought to have undergone two rounds of genome-wide duplication events before the appearance of jawed vertebrates giving rise to four paralogous regions (Kasahara et al., 1996;

Katsanis et al., 1996). Subsequently these paralogous regions have undergone differential gene inversion, translocation, expansion or deletion.

In the following sections the genomic organisation of the MHC is discussed, starting with mammalian species and proceeding backwards in evolutionary time to the earliest jawed vertebrates: cartilaginous fish. A schematic of the likely timescale of divergence of species from a common ancestor is provided in figure 1.8. The genomic organisation of the avian MHC is omitted from this initial discussion, and is discussed in detail later.

1.4.1 The genomic organisation and evolution of mammalian MHC regions

The MHC found in all mammals is a large and complex region that can be divided into three regions containing genes with broadly related functions (originally the regions were defined by recombination events). All mammals share similar genomic organisation of their MHC, with the class III region separating the class I and class II regions. Some of the genes within the MHC are members of multigene families, and can be highly polymorphic. The large size of the MHC of mammals is a consequence of the large number of genes, large intergenic distances, large intron sizes, and of the many repetitive elements and pseudogenes being present within the region.

In general, the class III region is extremely well conserved in all mammals, with the class II region also being well conserved; however there are differences between mammals in which isotype has prevailed, and the extent of gene expansion. The class I region is the least conserved amongst mammals, although there are conserved "framework" genes amongst variable numbers of class I and class I-related genes. A schematic representing the genomic organisation of the MHC of humans, rodents and chickens is provided in figure 1.9.

1.4.2 The MHC of humans

The human MHC is called the Human Leukocyte Antigen (HLA) system, and spans ~3.6 Mb region on chromosome 6, and encodes more than 200 genes, over 120 of which are thought to be expressed (1999).

The class I region spans ~1.8 Mb and contains 118 genes, of which 30 are expressed, 65 are pseudogenes, and 23 are potential coding sequences. This region contains genes that encode: HLA-A, -B and –C, the classical class I molecules (that are expressed on all nucleated cells, are highly polymorphic and present peptides to CTL); non-classical class I molecules called HLA-E, -F and -G and class I-related molecules including MICA and MICB (which are less polymorphic, have a restricted tissue expression and have divergent functions); as well as proteins that perform other functions.

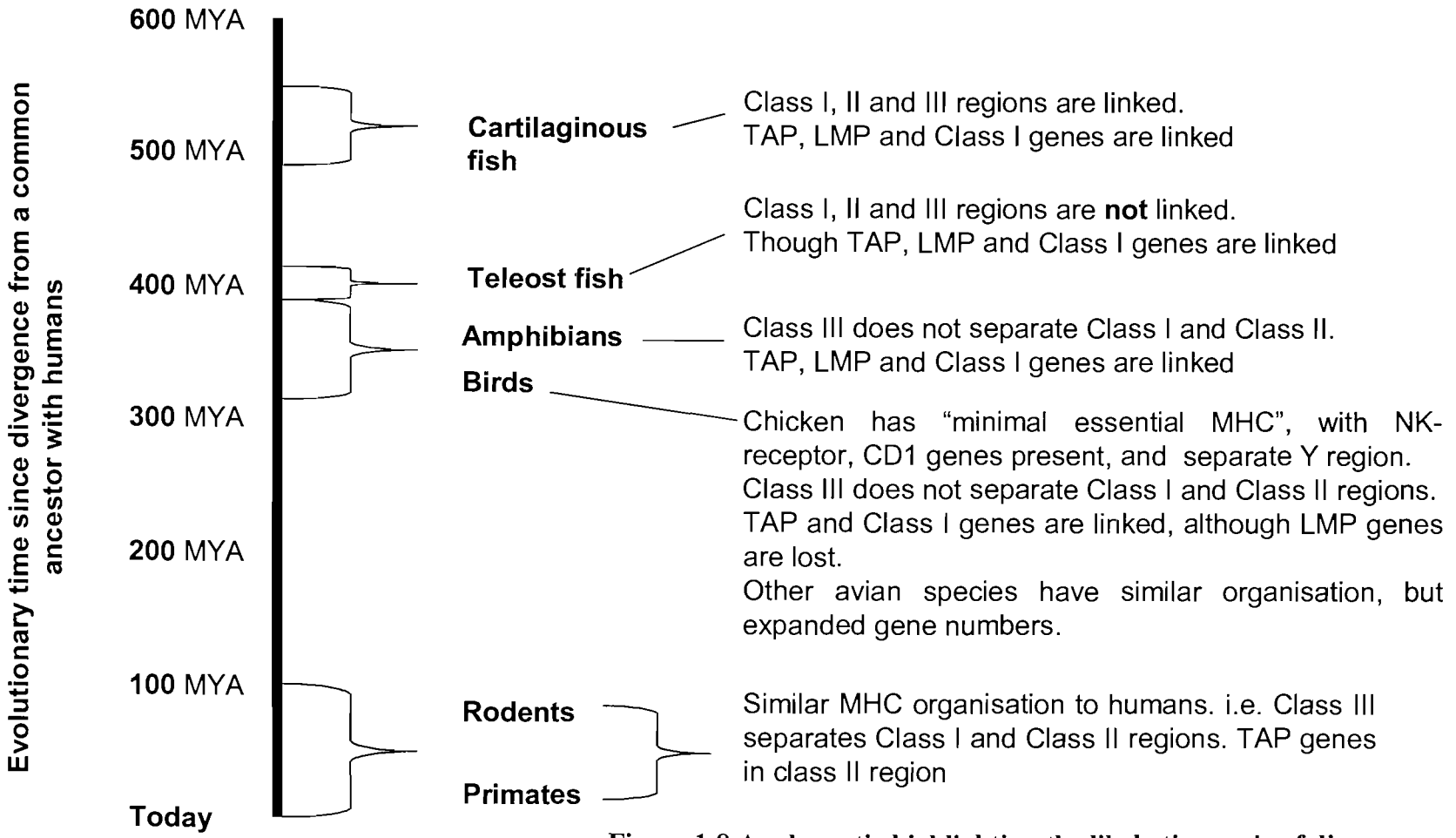


Figure 1.8 A schematic highlighting the likely timescale of divergence of species, and notes on the genomic organisation of their MHC regions.

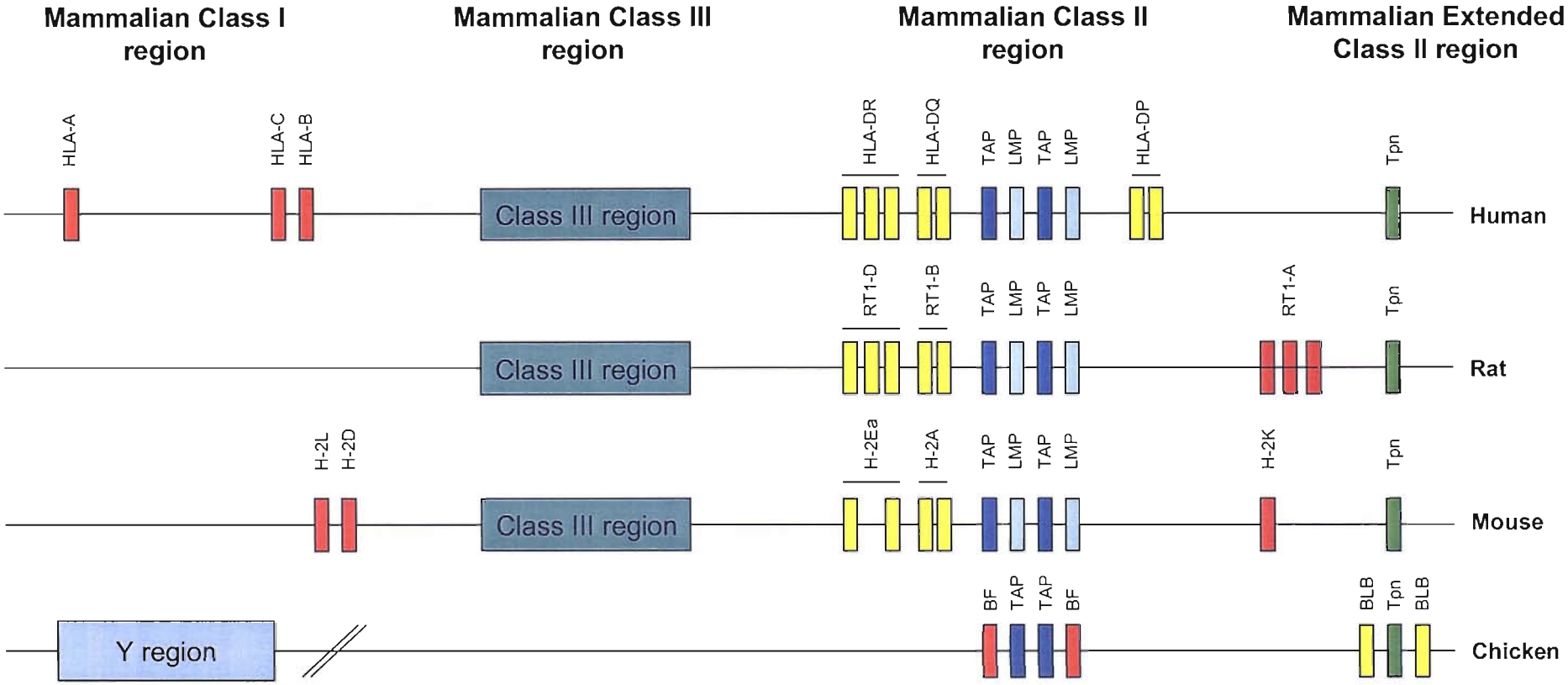


Figure 1.9 Comparisons of the genomic organisation of the MHC of humans, rodents and chickens

The schematic is not to scale, but compares the number and location of particular MHC encoded genes. For all species, classical class I genes are red, classical class II genes are yellow, TAP genes are deep blue, LMP genes are light blue, and the tapasin gene is green. The genes are aligned according to location, with the TAP and tapasin genes in a conserved location throughout.

The class III region spans ~1 Mb and is the most gene dense region of the human genome, and contains no pseudogenes. Some of the proteins that are encoded by genes in this region have immune-related functions, for example various complement components and cytokines. However, most of the proteins that are encoded by genes in this region perform non-immune related functions.

The class II region is at the centromeric end of the MHC, and spans ~1.2 Mb. The exact number of genes varies according to MHC haplotype, but in the composite of haplotypes that were sequenced there were 33 expressed genes and 26 pseudogenes. This region encodes a large number of proteins that perform immune-related functions, including the classical class II molecules that present peptides to $CD4^+$ helper T cells. Other genes in this region encode proteins that facilitate peptide-loading of class I molecules, for example the TAP peptide transporters, the tapasin protein that orchestrates optimal peptide-loading, and two inducible proteasome components (called Low Molecular weight Polypeptides, LMP) that preferentially generate class I binding peptides.

1.4.3 The HLA and disease association

There are many diseases that are associated with the HLA, which may be a consequence of the large number of proteins that are involved in the immune response being encoded by genes within the region. However, there are no strong associations between the HLA and resistance or susceptibility to infectious diseases; the few detectable associations with infectious pathogens are weak. In the clearest example, that of HLA-B*5301 and malaria, MHC haplotypes differ only slightly in their level of resistance or susceptibility (Hill et al., 1991). In fact the only strong disease associations with the HLA are with biochemical disorders or autoimmune diseases such as ankylosing spondylitis, where 95% of patients have the HLA-B*27 allele (Edwards et al., 2000). However, the association of a specific locus with a disease or autoimmune disorder is not always clear due to linkage disequilibria within regions of the HLA.

The high level of polymorphism in the MHC is thought to be a response to the selective pressure of pathogen infections (Kaufman et al., 1995). For example, new MHC alleles may appear as a response to novel pathogens (for example, the Black Death is thought to have led to the selection of HLA-A*0201), or be a result of population structure; whereby isolated populations of a species counter the infection of a common pathogen via a small number of MHC alleles that are unique to each population, so that collectively there is a large number of alleles within the species. In addition, even if a new MHC allele appears that has been selected to counter a novel pathogen infection, but is otherwise disadvantageous to the species, it is unlikely to be quickly lost from the population due to the high level of polymorphism in the MHC which means that there is only a small chance of being homozygous for the deleterious MHC allele. In addition, there is an advantage to being heterozygous for MHC alleles (overdominant selection).

1.4.4 The MHC of rodents

Rodents diverged from a common ancestor of humans ~100 MYA, and share similar, but not identical MHC organisations to humans. They also differ in the presence or absence of some genes, for example rodents lack MIC-related genes within their MHC.

The murine MHC (H-2) is located on chromosome 17 and differs to the human MHC in that one classical class I locus, H-2K, has translocated from the class I region to reside centromeric to the class II region. There are also numerous non-classical class I gene clusters.

The rat MHC (RT1) spans 3.8 Mb on chromosome 20, and differs to the human MHC in that there are eight clusters of class I genes. The classical class I proteins are encoded by the RT1-A cluster, which is located centromeric to the class II region, similar to the murine H-2K locus, however the rat has no orthologues of the H-2D and H-2L loci. This places all of the genes encoding the classical class I proteins and the proteins that facilitate their peptide-loading in relatively close proximity; which allows a level of co-evolution (and is discussed in more detail later in this chapter).

1.4.5 The genomic organisation and evolution of non-mammalian MHC regions

In non-mammalian vertebrates the genomic organisation of the MHC is different to that of mammals, in that the class III region does not separate the class I and class II regions.

1.4.6 The MHC of amphibians

The frog *Xenopus*, which diverged from an ancestor common to man ~350 MYA, has only one classical class I locus, which is in linkage disequilibria with TAP1, TAP2 and LMP loci; a situation evident in other species and which is discussed below. There are two ancient allelic families of linked TAP, LMP and class I genes that appear to have evolved as unit. *Xenopus* has other non-classical class I genes that are linked with each other, and are located on the same chromosome as the MHC, but they are not linked with the MHC. *Xenopus* is also of interest because the genome is polyploid (*Xenopus* species range from diploid (2n) to dodecaploid (12n)), however the MHC is generally diploid, with the duplicated MHC genes being silenced or deleted.

1.4.7 The MHC of teleost fish

Bony fish are thought to have diverged from an ancestor common to humans ~400 MYA. The MHC of many species of bony fish has been found to differ in organisation from all other jawed vertebrates. In teleost fish, the class I and class II genes are unlinked, and are located on separate chromosomes. In addition, the genes that are usually found in the class III region are scattered across the genome. Similar to *Xenopus*, there is linkage of class I, TAP, LMP and RING3 (an evolutionarily conserved nuclear serine/threonine kinase of unknown function).

1.4.8 The genomic organisation of the MHC of cartilaginous fish

Cartilaginous fish are the earliest jawed vertebrate to diverge from an ancestor common to humans, an event thought to have occurred ~500 MYA. In contrast to teleost fish, there is linkage of some class I, II and III genes, suggesting that the translocation of class I and class II genes to separate chromosomes was unique to the lineage leading to teleost fish. As with amphibians and teleost fish, there is linkage of class I, TAP and LMP genes in what has been described as the "true class I region", suggesting that these genes, whose products perform related functions in class I antigen processing and presentation, may have been present in this arrangement in the primordial MHC. This suggests the translocation of TAP, LMP (and tapasin) genes from the class I genes is unique to mammals.

1.5 The chicken MHC

The chicken MHC was originally described not as a histocompatibility complex, but as a polymorphic serological blood group (Briles et al., 1950), the so called B complex. This complex controls all of the attributes of an MHC (including allograft rejection (Schierman and Nordskog, 1961; Schierman and Nordskog, 1963), MLR (Miggiano et al., 1974), cellular co-operation of the immune response (Vainio and Ratcliffe, 1984; Vainio et al., 1988) and disease resistance (Bacon et al., 1981; Kaufman, 2000)). Historically, the chicken MHC (the BF/BL region) was the second MHC to be described.

The chicken MHC has been the subject of intensive research for five reasons. Firstly, the poultry industry is of enormous economic value as more than 30 billion chickens are bred each year, primarily for meat or egg production. Secondly, the chicken is subject to a large variety of natural pathogens that are studied both in the laboratory and in the field; some of which have become more virulent following the introduction of increasingly intensive rearing practices and new vaccination strategies. Thirdly, a number of inbred and congenic chicken lines are kept, which facilitates the study of infectious pathogens. Fourthly, the gene loci that control immune responses tend to be smaller and simpler, and therefore easier to comprehend at the molecular level. Lastly and perhaps the most striking difference to the MHC of typical mammals, is the fact that the chicken MHC haplotype can determine resistance or susceptibility to a variety of pathogens.

On the basis of two lines of evidence the chicken MHC has always been thought to be smaller and simpler than the MHC of typical mammals: firstly, the failure to demonstrate the expression of more than one class I (BF) and one class II (BL) antigen; secondly, of the inability to detect recombination between the loci encoding the BF and BL antigens (Ziegler and Pink, 1976). However, the major breakthrough in the characterisation of the chicken MHC came following an analysis at the genomic level of cosmid clusters encoding class I and class IIB genes that were

found in close proximity (Guillemot et al., 1988). This analysis showed that a large class III region did not separate the class I and class IIB genes, in common with other non-mammalian vertebrates.

The chicken MHC differs in other ways to that of any mammalian MHC. Firstly, there is a unique class of highly polymorphic molecules, called BG molecules, which are expressed at the cell surface of erythrocytes and other cell types. Secondly, there is another genomic region (originally named Rfp-Y, but since renamed the Y region) containing class I and class IIB genes that segregates independently from the BF/BL region (Briles et al., 1993; Miller et al., 1994). This region encodes non-classical MHC genes.

1.5.1 Location of the B locus

The chicken genome has now been sequenced in its entirety (Hillier et al., 2004) and is about one third the size of the human genome, constituting approximately one billion base pairs. Fifty percent of the genome is distributed on five pairs of macrochromosomes, the remaining 50% is divided between five pairs of chromosomes of intermediate size and 29 pairs of microchromosomes.

The B complex was mapped to microchromosome 16 together with the Nucleolar Organisation Region (which consists of many tandem repeats of rRNA genes) via an analysis of aneuploid chickens (Bloom and Bacon, 1985). Unfortunately, in the recent sequencing of the entire chicken genome, microchromosome 16 is one of the poorest chromosomes to be assembled and annotated (Wallis et al., 2004), and is of limited use in the analysis of the chicken MHC.

The initial molecular characterisation of the chicken MHC came following the cross-hybridisation of a human class IIB (HLA-DQB) probe with 17 cosmid clones of a genomic DNA library (Guillemot et al., 1988). These cosmids were arranged into four clusters, with cluster 1 encoding the BF/BL region, which has since been fully sequenced (Kaufman et al., 1999b) to reveal the presence of two classical class I and two classical class IIB genes. The other three clusters map to the non-classical Y region.

1.5.2 Description of the BF/BL region

The BF/BL region has all of the hallmarks of an MHC: it contains the classical class I and class IIB genes; it determines serological alloantigens, rapid allograft rejection, strong MLR and cellular co-operation in the immune response, attributes that are not shared by the Y region.

Analysis of the BF/BL sequence confirmed the chicken MHC is smaller and simpler than the MHC of typical mammals (figure 1.10).

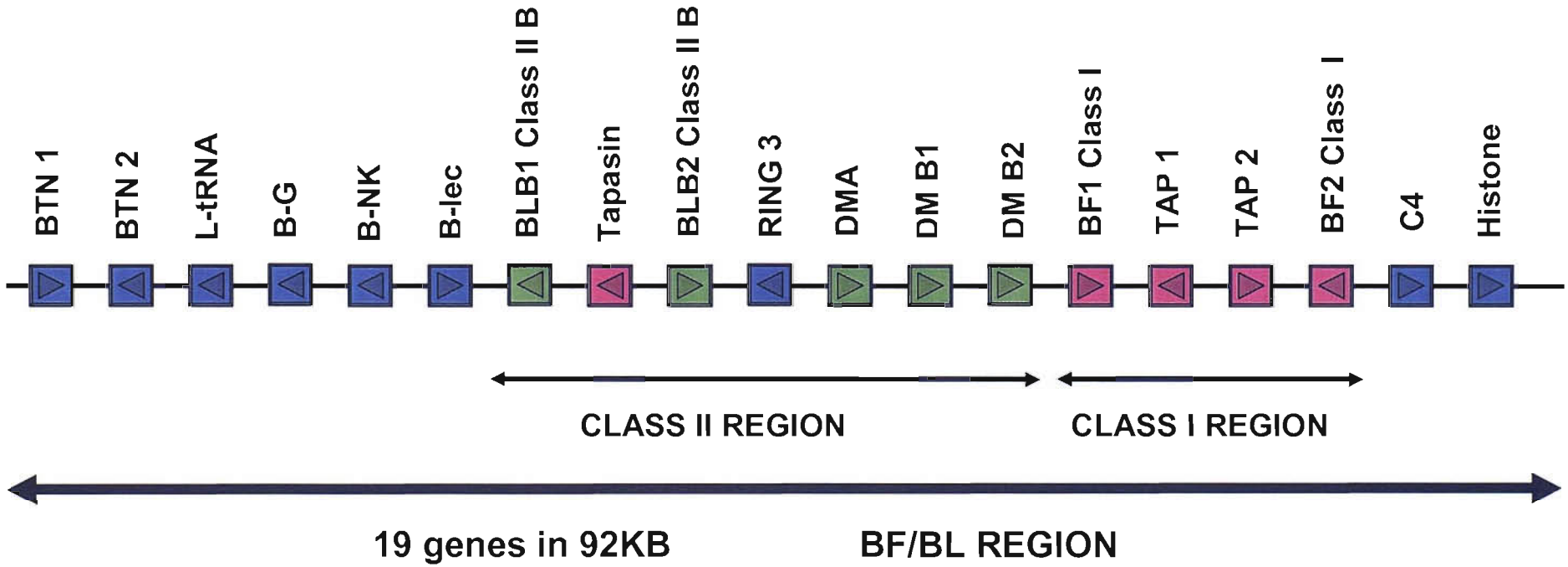


Figure 1.10 The chicken MHC (BF/BL region) from the B12 haplotype.

Arrows indicate transcriptional orientation. Genes are colour coded according to related protein function: purple for class I, green for class II, proteins with other functions are coloured in blue. Not drawn to scale.

The BF/BL region contains 19 genes in only 92 kb. The central portion of the BF/BL region (containing the class IIB and class I genes) has 11 genes in only 44 kb without any multigene families, repetitive elements, or pseudogenes, which prompted the description of the chicken MHC as a "minimal essential MHC", constituting only the core genes required for an MHC region (Kaufman et al., 1995; Kaufman et al., 1999b). The genes in the BF/BL region are much smaller than their human equivalents, which is mainly a consequence of reduction in intron size and intergenic distances, which are general features of the chicken genome.

Most of the genes in the chicken MHC have orthologues found in the MHC of mammals, however many of the expected genes are absent. For example, the class II α chain gene is not encoded in the chicken MHC, where it is in mammals, but is a non-polymorphic, single copy gene located ~5.6 cM from the BF/BL region (Salomonsen et al., 2003). In addition, whilst the class I and TAP genes are present, the LMP genes are not (and have not yet been found in the genome).

The absence of the LMP genes from the chicken genome may explain why some chicken MHC haplotypes (such as B4) have class I peptide-binding motifs with acidic C-terminal amino acids, which is not found for any mammalian class I peptide-binding motif; as peptides with these residues at their C-termini are not preferentially produced by the immunoproteasomes of mammals. It is possible that certain chicken class I alleles have evolved to use peptides with these C-terminal residues, whilst in other species that express LMP proteins, class I molecules have evolved to match the low abundance of these peptides in the repertoire of peptides produced following IFNγ exposure.

Many of the genes that are present in the BF/BL region are not arranged in the same order as mammals. For example, the TAP and tapasin genes are not separated by a large genetic distance from the class I genes, with the TAP genes being situated between the two class I genes and the tapasin gene between the two class IIB genes. The C4 gene represents one end of the class III region (Maruoka et al., 2005), which does not separate the class I and II regions, which is in agreement with the genomic organisation of other non-mammalian vertebrates (the class I region is defined as the class I and TAP genes, and the class II region is defined as the class IIB, RING3 and DM genes).

There are genes present in the BF/BL region that may not be expected to be, based upon the MHC of typical mammals, for example there are two genes with characteristics of C-type lectins, which are closely related to mammalian NK receptors. One of these genes, called B-NK, encodes the first non-mammalian NK receptor to be described. B-NK is expressed on NK cells and contains a functional immuno-regulatory tyrosine-based inhibitory motif (ITIM); the second gene, called B-lec, encodes an early activation antigen most similar to the human lectin-like transcript 1 (Kaufman et al., 1999b; Rogers et al., 2005). The mammalian NK receptors to which these genes are most

similar are located in the Natural Killer Complex (NKC) outside of the MHC; the chromosome on which they are located is likely to be another region that is paralogous to the MHC. In addition, there are two CD1 genes that are located ~50 kb from the class I genes in the BF/BL region (Maruoka et al., 2005; Miller et al., 2005; Salomonsen et al., 2005). As with the C-type lectins, CD1 genes are not found in the MHC of any mammal. The simplest explanation for the presence of these genes within (or in close proximity to) the BF/BL region is that these genes were present in the primordial MHC. It could be envisaged that following the genome-wide duplication events, differential gene deletion led to the CD1 and NK receptor genes being lost from the MHC, with the MHC and either the CD1 or NK receptor genes being lost from the separate paralogous regions in which the CD1 or NK receptors genes are found in mammals.

One of the most notable features of the chicken MHC is the virtual absence of recombination events occurring within the BF/BL region. In over 30,000 informative typings, not a single recombination event was observed to separate serologically detected BF and BL antigens, which allowed the calculation of the map distance between BF and BL loci to be <0.01 cM (Skjodt et al., 1985; Plachy et al., 1992). However, there is evidence for one recombinant haplotype: the B19v1 haplotype, which shares a serologically indistinguishable BL antigen with the B12 haplotype, but has a BF antigen that is serologically indistinguishable from the B15 haplotype (Simonsen et al., 1982); suggesting a recombination event occurred between the BF and BL loci in the B12 and B15 haplotypes, which is discussed in detail in a later chapter. The low rate of recombination means that the entire BF/BL region is likely to be transmitted as a unit through many generations, which increases the probability that genes within this region may evolve to interact with the alleles of other genes that are specifically encoded by the same haplotype; in addition, the longer the linked genes can co-evolve, the more specific the interaction is likely to be (Kaufman, 1999; Kaufman et al., 1999a).

1.5.3 BF: the class I molecules of the chicken

The BF/BL region contains two classical class I genes, BF1 and BF2, that are located either side of the TAP genes. The chicken class I genes and molecules show all of the characteristics expected of mammalian classical class I genes (Kroemer et al., 1990).

MHC homozygous chickens have only a single dominantly-expressed class I molecule (Wallny et al., 2006), an observation based upon three pieces of evidence: firstly, it was evident at the RNA level (through counting the number of cDNA clones amplified by primers specific for both BF1 and BF2 genes); secondly, there was a dominantly expressed class I protein at the protein level (revealed by N-terminal protein sequencing, and from two-dimensional electrophoresis); thirdly, at the peptide level (which was historically the first observation to be made; whereby a single clear peptide motif was eluted from MHC homozygous chicken blood cells, which is not the case in mammals, where multiple class I alleles of differing peptide-binding specificity produce a complex

pool of peptides). The dominantly expressed gene corresponds to the BF2 locus, and is often referred to as the "major" gene, and the BF1 locus has been referred to as the "minor" gene.

A molecular explanation for the BF1 locus to be expressed at a lower level than the BF2 locus has been revealed at the level of transcription (Kaufman et al., 1999a). The enhancer A, which is the most important element for high level transcription of mammalian class I genes (Kimura et al., 1986; Gobin et al., 1998), is absent from the promoter region of the BF1 locus in two haplotypes (B12 and B19v1), and is diverged in three other haplotypes (B2, B4 and B21), whilst the BF1 gene is disrupted (most likely by a large insertion) in two other haplotypes (B14 and B15). In contrast, the promoter of the BF2 locus does not vary significantly between the seven haplotypes sequenced.

The expression of a single dominantly expressed class I molecule is reminiscent of the situation in some non-mammalian jawed vertebrates (Kaufman, 1999), and it appears the chicken has gone to great lengths to down-regulate expression from the BF1 locus. The reason the chicken pursues such a strategy is not immediately apparent; it would seem advantageous to express multiple class I alleles each of different peptide-binding specificities (like mammals), which would ensure a greater diversity of peptides derived from infectious pathogens or abnormal self-proteins are presented to CTL. In contrast, the predominant expression of a single class I molecule means the chicken may be susceptible to pathogen infection or malignancy simply because no pathogen/tumour-derived peptides can bind to the expressed class I molecule, a situation which has been confirmed in response to infection from Rous Sarcoma virus (Wallny et al., 2006). The most likely explanation for the evolution of a single dominantly expressed class I molecule is thought to be a consequence of the small size of the chicken genome in general, which may have been driven by an unknown pressure (Kaufman et al., 1995). This may have resulted in multigene families being much smaller in size than in mammals (including the MHC, TCR and antibody gene segments).

Another notable feature of the class I molecules of the chicken is their relative expression levels at the cell surface (Kaufman et al., 1995). A hierarchy of surface expression levels has been documented: whereby the B21 haplotype has the lowest expression of class I molecules; the B2, B6 and B14 haplotypes express a higher level of class I molecules; the B4, B12 and B15 haplotypes express an even higher amount of class I molecules; and the B19v1 haplotype expresses the most class I molecules at the cell surface. In addition, the level of surface expression varies according to cell type: the difference between the B19v1 and B21 haplotypes is 7-10 fold on erythrocytes, 5-7 fold on thrombocytes (that perform functions equivalent to platelets), and 2-5 fold on white blood cells (Kaufman et al., 1995). Such hierarchies of expression are not immediately apparent in mammals, where the expression of multiple class I alleles "averages out" any differences in total surface expression levels. The molecular reason for such differences in surface expression levels has not been fully determined, but is unlikely to be a consequence of differences in either: levels of

transcription or translation; mRNA or protein stability; β_2m association, but may be explained by the rate or extent to which class I molecules traffic to the cell surface (Kaufman et al., 1999a).

1.5.4 Other genes in the BF/BL region

There are two classical class IIB genes, BLB1 and BLB2, which flank the tapasin gene. Similar to the BF2 genes, the BLB2 gene is expressed at a higher level than the BLB1 gene, assessed by cDNA clones amplified by primers specific for both BLB genes (Pharr et al., 1998; Jacob et al., 2000). However, the BLB gene promoter regions do not account for this differential transcript expression (J. Jacob and J. Kaufman, unpublished).

There are two copies of DMB genes and one DMA gene whose products combine to form the non-classical class II-related DM heterodimer that chaperones and facilitates the peptide-loading of class II molecules. However, there are no genes present in the BF/BL region that encode the DO protein which modulates the action of DM.

1.5.5 The BG region

The chicken has a third class of highly polymorphic molecule, called BG molecules, that are expressed primarily as non-glycosylated dimers at the surface of many cell types (originally thought to be expressed exclusively on erythrocytes, where they are expressed at a greater level than BF molecules; a wider distribution of expression has subsequently been demonstrated) (Kaufman et al., 1991). The BG molecules are members of a large multigene family with no known mammalian equivalent, although BG molecules have been described in other avian species (Shiina et al., 2004). There is one BG gene in the BF/BL region, and more than 12 BG genes tentatively mapped to the BG region, that is ~100 kb from the BF/BL region. Recombination events have been observed to separate the BG antigens from the BF/BL antigens; however the recombination rate remains low, with the calculated map distance between the BG and the BF/BL regions being 0.037 cM (Plachy et al., 1992).

The function of BG molecules has not been determined, but three immunological phenomena have been ascribed: firstly, the adjuvant effect, where the presence of BG molecules in the same cell or liposome boosts the immunogenicity of an otherwise poorly immunogenic molecule; secondly, the preferential response is characterised by a primary humoral response following immunisation with BG molecules being unusually fast and higher titre than usual primary responses; thirdly, the presence of natural antibodies in naive animals specific to BG molecules (Kaufman et al., 1991; Salomonsen et al., 1991; Kaufman and Salomonsen, 1992). The highly polymorphic and immunogenic BG molecules were presumably the antigens detected by alloantisera in the initial description of the B complex.

1.5.6 The Y region

Following the detection of inconsistencies in restriction fragment length polymorphism analyses of serologically defined B haplotypes, a second genomic region containing genes that cross-hybridised with BF and BL probes was identified, which segregates independently of the BF/BL region (Briles et al., 1993). This region was originally called the Rfp-Y region (restriction fragment length polymorphism-Y), but has subsequently been renamed the Y region (Miller et al., 2004).

The Y region was assigned to microchromosome 16 via an analysis of aneuploid chickens (Miller et al., 1996), however it is not linked to the BF/BL region, due to being separated by the repetitive and recombinogenic Nucleolar Organisation Region (NOR), which consists of many tandem repeats of rRNA genes in a 6 Mb region (Bloom and Bacon, 1985).

The exact number of class I-like genes (YF) and class II-like genes (YL) that are found in this region varies with Y haplotype, but two YF have been analysed from one Y haplotype (Afanassieff et al., 2001). The YF1 gene is polymorphic, and encodes a product that is widely expressed and associates with β_2m. However, substitutions within the peptide-binding domain suggest the YF1 protein binds ligand in a novel fashion, and is likely to serve a distinct role to that of the classical class I proteins. The YF2 gene is disrupted by a large repeat and lacks a polyadenylation sequence, suggesting no gene product is produced. Three YL genes have been identified within the same Y haplotype, but all lack significant polymorphism and have disrupted promoter regions, suggesting these genes are pseudogenes. In addition, two C-type lectin genes are found in the Y region (Bernot et al., 1994; Rogers et al., 2003). As with the BF/BL region, recombination events rarely separate the YF and YL loci, with a map distance calculated to be <0.25 cM (Pharr et al., 1997).

The Y region determines at best only a poor MLR, and a significantly slower rate of transplant rejection of Y-incompatible allografts (17-21 days) in comparison to B-incompatible allografts (<8 days), with Y-compatible grafts being rejected after 37-43 days (Pharr et al., 1996; Juul-Madsen et al., 1997; Thoraval et al., 2003). The rejection of Y incompatible allografts is likely to be mediated by Y region linked genes, or minor histocompatibility antigens encoded in this region (epitopes that are generated from polymorphic genes of the Y region).

1.5.7 Disease associations with the BF/BL region

In contrast to the lack of association of disease resistance or susceptibility to the HLA system, are the striking examples where the MHC haplotype of a chicken determines life or death in response to both small and large pathogens (Kaufman, 2000). For example, the B5 haplotype confers better resistance to *Eimeria tenella* infection than the B2 haplotype (Clare et al., 1985). In addition, the B1 haplotype was more resistant than the B19 haplotype to bacterial *Pasteurella multicidia* infection, which causes fowl cholera (Lamont et al., 1987). In another study the B15 haplotype conferred a higher level of resistance to chicks inappropriately vaccinated with a less attenuated

Infectious Bronchitis virus (IBV) vaccine than the B13 (B4) or B21 haplotypes (Bacon et al., 2004). However, the best examples concern the small retrovirus Rous Sarcoma virus (RSV), and the large α-herpes virus Marek's disease virus (Plachy et al., 1992; Kaufman et al., 1995).

RSV is a small retrovirus consisting of four genes flanked by long terminal repeats. Three genes (*gag, pol* and *env*) are involved in viral replication, whilst the fourth gene, *v-src*, is a mutated copy of the cellular proto-oncogene *c-src* (Schwartz et al., 1983). The *c-src* encodes a kinase that is regulated via the binding of a phosphorylated tyrosine residue encoded within the 3' tail of the gene. The most significant differences between *v-src* and *c-src* concern the 3' tail, where the *v-src* lacks the tyrosine residue that is usually phosphorylated in order to regulate the kinase activity (Kaufman et al., 1995; Kaufman, 2000). Consequently, RSV infection leads to immediate viral transformation of the cell, and rapid tumour growth. As was best demonstrated by the infection of MHC congenic CC (B4) and CB (B12) lines with the RSV-Prague C virus (Plachy et al., 1992), tumours either progress leading to death (for example the B4 or B15 haplotypes), or regress leading to survival (characterised by the B12 haplotype), depending upon B haplotype. The mechanism of MHC haplotype-determined resistance to RSV infection has been shown to be simply the result of whether the peptide-binding motif of the dominantly expressed class I molecule allows the binding of peptides derived from RSV (Wallny et al., 2006). In the B12 haplotype, a peptide derived from the *v*-src 3' tail can bind to the class I molecule and is presented to CTL, which leads to the subsequent clearance of infected cells. However, in the B4 and B15 haplotypes, no suitable peptides can bind to the class I molecules, leading to tumour progression and death (Kaufman et al., 1995; Wallny et al., 2006).

The strongest association of any MHC and disease resistance or susceptibility concerns Marek's disease virus (MDV). MDV is a large virus encoding nearly 100 proteins, so the mechanism of MHC-determined survival to MDV infection is unlikely to be the result of whether a peptide binds to the dominantly expressed class I molecule. MDV infections have become increasingly virulent, a result that is likely to be caused by an increase in intensive rearing practices, and in viral "break-through" following the introduction of increasingly less attenuated MDV vaccines.

The B19 haplotype confers susceptibility to MDV, whilst the B21 haplotype confers resistance, with other haplotypes correlating between these extremes; which suggest MDV resistance is not simply the result of the presence or absence of a specific MHC allele. In fact, the hierarchy of MDV resistance correlates inversely with class I surface expression levels; with the most resistant B21 haplotype expressing the lowest number of class I molecules at the cell surface, and the most susceptible B19 haplotype expressing the most, with the other haplotypes ranked between. The mechanism of disease resistance has not been determined, but may be related to NK activity, as NK activity is highest in the resistant B21 haplotype and lowest in the susceptible B19 haplotype (Sharma and Okazaki, 1981; Kaufman and Salomonsen, 1997; Kaufman, 2000).

1.5.8 The genomic organisation and evolution of the MHC of other avian species

The MHC of the quail

The quail (*Coturnix japonica, Coja*) is thought to have diverged from a common ancestor of the chicken 40-80 MYA (van Tuinen and Hedges, 2001; Shiina et al., 2004). As with other avian species, the organisation of the quail MHC is similar to that of the chicken, but the quail MHC has undergone two segmental duplication events (of the class I genes alone, and of the class IIB, BG and lectin genes) (Shiina et al., 2004). The quail MHC encodes 41 gene loci in 180 kb. Fifteen of these genes are expressed, including six class IIB and four class I genes, and the tapasin, TAP1, TAP2, RING3 and C4 genes. A further 12 genes may be expressed (including four class IIB, three BG and five NK-receptor genes), whilst 14 genes are likely to be pseudogenes.

The quail MHC has a total of seven class I genes in a class I region twice as long as the equivalent region of the chicken MHC. Three of the class I genes are pseudogenes (Coja-F, -G and -H), whilst two of the four expressed class I genes are likely to be non-classical (Coja-D1 and -D2), based upon low and tissue-specific expression patterns, leaving only two classical class I genes (Coja-B1 and -E), with Coja-B1 being located next to TAP2. The quail class I genes are most similar to each other suggesting that they arose through duplication after the separation of quail and chickens. Similar to the class I genes, there are only two classical class IIB genes, with the other class IIB genes being non-classical. The tapasin gene occupies an equivalent position to that found in the chicken, and is 90% homologous to chicken tapasin and only 33-34% homologous to human and murine tapasin.

The MHC of the duck

Fowl (Galliformes) are thought to have diverged from ducks (Anseriformes) ~90 MYA (van Tuinen and Hedges, 2001). A portion of the MHC of the mallard duck (*Anas platyrhynchos*) has been characterised (Moon et al., 2005). A total of five classical class I genes were found, which were arranged sequentially in the 37 kb genomic fragment that was sequenced, that also contained the TAP1 and TAP2 genes. The class I gene (UAA) which was located next to the TAP2 gene was found to be pre-dominantly expressed (Mesa et al., 2004), similar to the chicken and quail MHC regions. Three of the remaining four class I genes (UBA, UCA, and UEA) are disrupted (either the promoter was disrupted, the polyadenylation sequence was absent, or a stop codon was found in frame), with the UDA gene being expressed at a significantly lower transcript level than UAA, which appeared to be due to a disrupted promoter; suggesting UDA represents the equivalent of the chicken BF1 locus. Phylogenetic analysis of the class I sequences suggested they are most similar to each other, and arose by duplication event after speciation. No evidence for other class I genes was found by Southern blot analysis, suggesting there was no equivalent of the Y region.

The MHC of other avian species

Little progress has been made in characterising the MHC of other birds. The analysis that has been conducted to date has shown multiple MHC genes, but has failed to distinguish non-classical from classical regions, and to address gene expression levels.

1.5.9 Conclusions concerning the comparisons of the MHC found in jawed vertebrates

The genomic organisation of the MHC of typical mammals appears to be the exception, rather than the rule. Analysis of lower jawed vertebrates suggests that the order of genes in mammals is the product of translocation events unique to mammals, followed by multiple events that collectively increased the diversity of the mammalian MHC. For example, the translocation of the class I genes away from the "true class I region" (that contains the TAP and LMP genes and which is likely to represent the primordial gene arrangement) may have freed the class I genes of typical mammals from functional constraints imposed by the tightly linked class I antigen processing and presenting genes (which may have co-evolved with the class I genes), allowing the class I genes to duplicate and diverge (Kaufman et al., 1995; Nonaka et al., 1997).

The analysis of some lower jawed vertebrates has revealed multiple class I genes, the majority of which have been silenced, leading to the predominant expression of a single class I molecule. The variety of mechanisms that are used to silence the class I genes suggests there is a strong selective pressure for this strategy. The advantage this brings is not immediately apparent, in fact it would appear to severely restrict the diversity of peptides that could be presented. However, the expression of a single class I molecule may have been driven by a limited size of TCR and MHC multigene families in lower vertebrates (Kaufman et al., 1995). If so, if more MHC molecules were to be expressed without increasing the TCR repertoire, then deletion of self reactive TCR may severely limit the number of the TCR in the periphery. Despite this, the presence of multiple (but silenced) MHC genes may be advantageous at times of infection from novel pathogens; as recombination processes including gene conversion and allelic exchange may create new MHC alleles that may protect the host from infection.

1.6 The discovery of chicken tapasin

As part of the initial characterisation of the genomic structure of the chicken MHC, a gene, referred to as "gene 8.4", was identified (Guillemot et al., 1988). The initial description of gene 8.4 and its potential product was limited to the genomic location (between BLB1 and BLB2, figure 1.11a), transcriptional activity (which was detected in numerous tissues), and to encoding a non-MHC class I or non-MHC class II protein (despite showing 25-30% similarities to immunoglobulin, MHC class I and class II amino acid sequences).

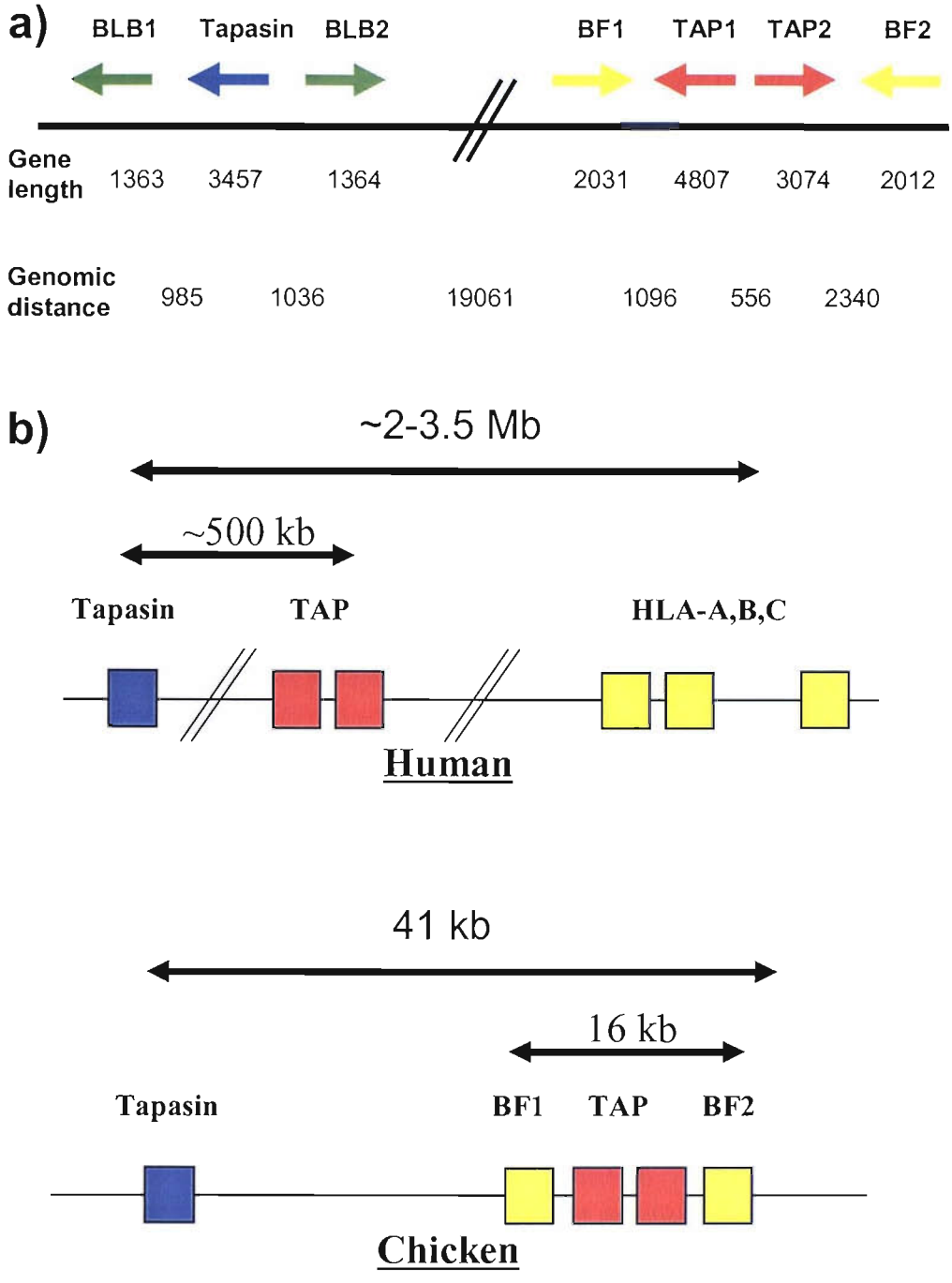


Figure 1.11 Schematic representing gene locations within the chicken MHC

a) Schematic representing the locations of tapasin, class I major (BF2) and minor (BF1), TAP1 and TAP2, class IIB major (BLB2) and minor (BLB1). Not drawn to scale. Gene orientation, length and genomic distances between start and stop codons (in bp) are detailed (taken from the B12 reference sequence AL023516, Kaufman 1999).
b) Cartoon depicting the genomic organisation of the chicken and human MHC regions. MHC class I genes are shown as yellow boxes, TAP genes shown as red boxes, and tapasin shown as a blue box. Arrows show approximate genomic distances between loci.

Nearly ten years later, two laboratories independently identified gene 8.4 as the chicken orthologue of tapasin by homology searching, and sequenced the gene in its entirety (Frangoulis et al., 1999; Kaufman et al., 1999b). Both of these sequences were from the B12 haplotype, and were given the accession codes AJ004999 (Frangoulis et al., 1999) and AL023516 (Kaufman et al., 1999b). The AJ004999 sequence derives from the cβ8 cosmid, which was made using non-replicative methods at the same time as the c4.5, cβ12 and cBF23 cosmids that were sequenced to form the AL023516 sequence, using genomic DNA isolated from the liver of a B12 haplotype chicken of the CB line (Guillemot et al., 1988; Kaufman et al., 1999b).

Analysis of the chicken tapasin nucleotide sequence shows that chicken tapasin shares a similar gene structure to its mammalian counterparts. However, the chicken tapasin locus is smaller than found in typical mammals; largely a consequence of a reduction in the size of two introns.

The chicken tapasin gene encodes a protein of 430 amino acids, 18 amino acids smaller than the human tapasin protein, with which it shares 42-46% amino acid similarity (an alignment of chicken, human and murine tapasin protein sequences is provided in figure 1.12).

Both human and chicken tapasin proteins share some interesting structural similarities, but also differ in other aspects (figure 1.13). Whilst the predicted mature protein of chicken tapasin has six cysteine residues, there are only five cysteine residues in the predicted mature human tapasin protein (as with other mammals). In figures 1.12 and 1.13 the cysteine residues are sequentially numbered according to location in either chicken or human tapasin protein sequences.

In both chicken and human tapasin proteins there is conservation of the potential disulphide bonds within the membrane proximal Ig domain (involving cysteines 6 and 7 within the domain encoded by exon 5) and at the N-terminus (involving cysteines 1 and 2 within the domain(s) encoded by exons 2 and 3). Chicken tapasin has an additional pair of cysteine residues (involving cysteines 4 and 5) within the domain encoded by exon 4, which facilitated the identification of this domain as an Ig domain. However, chicken tapasin lacks an equivalent of cysteine 95 (cysteine 3) in the mature human tapasin protein that forms a transient disulphide bond with ERp57 (Dick et al., 2002). Another difference between chicken and human tapasin proteins concerns the location of the single N-linked glycan. In mammals, tapasin is N-glycosylated within the domain encoded by exon 4, whilst in chicken tapasin the glycan is towards the end of the domain encoded by exon 2.


1
2
Cys1
3
Human : M-KSLSLLLAVALGLATAVS---AGPAVIECWFVEDASGK---------GL---AKRPGAL---------LLR-QGPGEPPPRPDLDPELYLSVHDPAGA : 74
Murine : M-KPLLLLVAVALGLATFVSVVSAGPEAIECWFVEDAGGG---------GL---SKKPATL---------LLR-HGPRGPPPRPDLDPKLYFKVDDPAGM : 77
Chicken : MAAGLRLLLA---GL------------CWSQFRVEDAASPPPPPAPVRCALLEGVGRGGGLPGGGNARPALLRFGGDAETPPEPGPEPEVTFNVSDPWGT : 85
Cys2
Cys3
4
Human : LQAAFRRYPRGAPAPHCEMSRFVPLPASAKWASGLTPAQNCPRALDGAWLMVSISSPVLSLSSLLRPQPEPQQEPVLITMATVVLTVLTHTPAPRVRLGQ : 174
Murine : LLAAFRRYPAGASAPHCEMSRFIPFPASAKWARSLSPEQNCPRALDGDWLLVSVSSTLFSLSSLLRPQPEPLREPVVITMATVVLTVLTHNPAPRVQLGK : 177
Chicken : LTPLGV--PPRTP-PSCELNPTNPQTGSDPWSRPLHPDARSPPTAGGQWWVAAVGTPQYGVTALL--QGGMGTEGT-ITAA-VALAVLTHTPTLRARVGS : 178
----A----
Cys4
Cys5
Human : DALLDLSFAYMPPTSEAASSLAPGPPPFGLEWRRQHLGKGHLLLAATPGLNGQMPAAQEGAVAFAAWDDDEPWGPWTGNGTFWLPRVQPFQEGTYLATIH : 274
Murine : DAVLDLRFAYAPSALEGSPSLDAGPPPFGLEWRRQHRGKGHLLLAATPGLAGRMPPAQEKATAFAAWDDDEPWGPWTGNGTFWLPAVKPSQEGVYLGTVH : 277
Chicken : PIHLHCAFAA-PPSS------------FVLEWRHQNRGAGRVLLAYDSS-TARAPRATPGAELLLGTRD----GDGVTAVTLRLARPSPGDEGTYICSVF : 260
---B---- ---C--- --C'-- --D--- ---E---- ---F----
5
Cys6
Human : LPYLQGQVTLELAVYKPPKVSLMPATLARAAPGEAPPELLCLVSHFYPSGGLEVEWELRGGPGGRSQKAEG---QRWLSALRHHSDGSVSLSGHLQPPPV : 371
Murine : LPYLQGQVSLELTVHKGPRVSLTPAPVVWAAPGEAPPELLCLASHFFPAEGLEVKWELRGGPGGSSRKVEG---KTWLSTIRHHSDGSVSQSGHLQLPPV : 374
Chicken : LPHGHTQTVLQLHVFEPPKVTLSPKNLV-VAPGTSA-ELRCHVSGFYPL-DVTVTWQRRAGGSGTSQSPRDTVMDSWTSGHRQAADGTYSRTAAARLIPA : 357
-----G----- -----A----- -----B----- ---C---- ----D---- ----E-----
Cys7
6
7
8
Human : TTEQHGARYACRIHHPSLPASGRSAEVTLEVAGLSGPSLEDSVGLFLSAFLLLGLFKALGW-AAVYLSTCKDSKKKAE*---------- : 448
Murine : TAKQHGVHYVCRVYHSSLPASGRSADVTLEVAGFSGPSIEDGIGLFLSAFLLLGLLKVLGWLAAYWTIPEVSKEKATAASLTIPRNSKKSQ : 465
Chicken : RPQHHGDIYSCVVTHTAL-AKPMRVSVRLLLAGTEGPHLEDITGLFLVAFVLCGLIR---W-LYPKAARPKEETKKSQ*---------- : 430
---F--- -----G-----

Figure 1.12 Alignments of human, murine and chicken tapasin protein sequences.

The alignment was produced in two steps: firstly, human and murine sequences were aligned according to Grandea et al., 1998; secondly, chicken was aligned according to (Frangoulis et al., 1999). Gaps are introduced to facilitate alignment, however note that other alignments are possible. The start of a new exon is indicated by numbers above the text. Dilysine motifs are shown in italics; sites of N-linked glycosylation are shown in blue. Predicted signal sequences are shown in grey, human (Ortmann et al., 1997), murine (Grandea et al., 1998), chicken (this study, although a 15 amino acid signal sequence was predicted by (Frangoulis et al., 1999)). Predicted transmembrane TM domains are underlined, human (Ortmann et al., 1997), murine (Grandea et al., 1998), chicken (Frangoulis et al., 1999). Cysteine residues within the predicted ER luminal domains are shown in red, and are sequentially numbered. The position of immunoglobulin β strands is shown beneath the alignment, and was aligned using the chicken tapasin sequence (Teng et al., 2002).

A polymorphism in human tapasin is described at position 260 (R/T) (Copeman et al., 1998), and polymorphisms exist between two published murine tapasin sequences at positions 17 (F/V), 274 (G/A), 294 (G/A), and 326 (A/S) (Grandea et al., 1998; Li et al., 1999).

Human: AF009510, murine: AF043943 (Grandea et al., 1998) is shown, which differs to AF106278 (Li et al., 1999) at four positions, chicken: AL023516 (Kaufman et al., 1999).

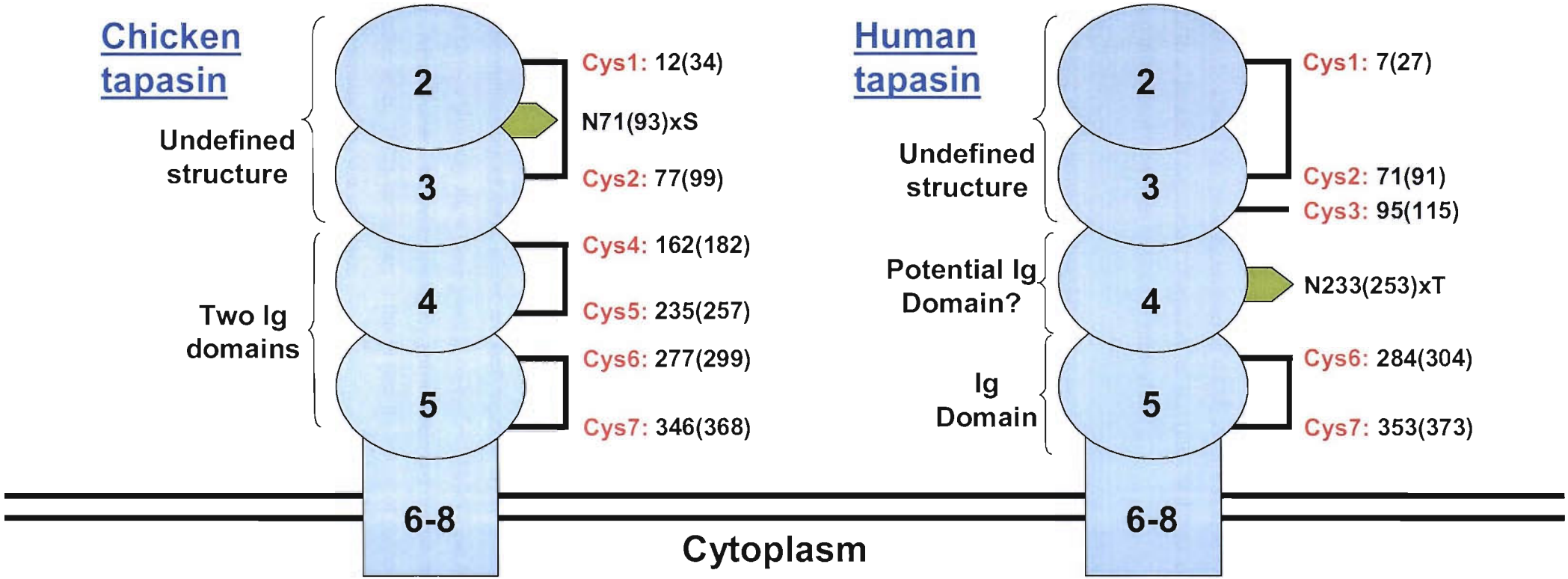


Figure 1.13 Cartoons depicting the likely structural organisations of chicken and human tapasin proteins
Ovals or rectangles depict potential domains encoded by one or more exons, which are numbered within the shape. Potential disulphide bonds are shown as black lines and are sequentially numbered (Cys1-7) according to location in the mature tapasin protein sequences that are found in different species, and are also numbered according to likely location in the mature chicken and human protein sequences (assuming a 22 or 20 amino acid signal sequence for chicken or human proteins), with the number in parentheses referring to the location of the cysteine residue in the primary sequence. The site of N-linked glycosylation is shown in green, and is numbered in a similar fashion.

In humans, the tapasin locus is separated by a relatively large recombinational distance from the class I (~2 Mb from HLA-B and –C and ~3.5 Mb from HLA-A) and TAP loci (~500 kb) (figure 1.11b). Extensive sequencing of the human tapasin gene has revealed only one coding change at position 240 of the mature protein (arginine or threonine) (Copeman et al., 1998; Furukawa et al., 1998; Herberg et al., 1998). There is a report of an additional polymorphic residue at residue 412 (asparagine instead of aspartic acid) (Li et al., 1997). However, this polymorphism was not found in either of the other studies (Furukawa et al., 1998; Williams et al., 2000), and may be specific to the B lymphoblastoid cell line sequenced; perhaps it is a spontaneous mutation. The frequency of the two tapasin alleles has been determined as 0.47/0.53 in a sample of the UK population and 0.37/0.63 in a Japanese population (Furukawa et al., 1998; Williams et al., 2000). To date no functional differences have been ascribed to either allele.

The relative lack of sequence polymorphism, and the large recombinational distance between the human tapasin and class I genes, means it is likely that human tapasin has evolved to function with whichever class I molecules are expressed, although this has not been fully investigated. The observed variations in tapasin-dependence of class I molecules are therefore likely to be a property determined by the class I molecule.

In contrast to the large recombinational distances separating human tapasin, TAP and class I genes, the chicken tapasin gene is separated by ~41 kb from the TAP and the class I (BF) genes (figure 1.11). The extent of polymorphism in the chicken tapasin gene was unknown at the onset of the project.

1.7 Initial aims of the project

The close proximity between tapasin, TAP and BF genes (Kaufman et al., 1999b), and the low level of recombination within the BF/BL region (Skjodt et al., 1985) means there is an extremely small chance of recombination separating combinations of alleles of these genes. Therefore it is likely that combinations of these alleles are kept together in the same haplotype for extended periods of time, which may allow these genes to co-evolve. Whilst it has been documented that the chicken TAP and class I are genes are highly polymorphic and diverse in sequence (Miller et al., 2004; Walker et al., 2005), at the start of this project the extent of polymorphism and diversity in the chicken tapasin gene was unknown.

Aim one: Is chicken tapasin polymorphic?

The initial aim of this project was to determine the extent of allelic polymorphism and sequence diversity in the chicken tapasin gene among seven MHC haplotypes that are kept at IAH Compton. If chicken tapasin were polymorphic and diverse in sequence, this would contrast to the situation in

typical mammals, where the tapasin gene is nearly monomorphic and located distant to the class I genes.

Aim two: Is tapasin involved in a haplotype-specific interaction with the class I and TAP proteins?

In the event that chicken tapasin is polymorphic, the limited opportunity to separate alleles of class I, TAP and tapasin may mean that the products of the class I, TAP and tapasin alleles that are found on the same haplotype have evolved to specifically interact with each other, in other words these proteins are involved in a "haplotype-specific" interaction. If tapasin is involved in a haplotype-specific interaction and the combination of alleles were to be disrupted, perhaps by a recombination event, then the products of a "mismatched" combination of alleles may be less able to work together, or may be completely incapable of working together.

If tapasin is involved in a haplotype-specific interaction, polymorphic amino acids in tapasin may partly contribute to the specificity of the interaction (through interacting with polymorphic residues of the class I and TAP proteins) or they may completely determine the specificity of the interaction (through interacting with invariant residues of the class I and TAP proteins).

In a haplotype-specific interaction, it is possible that tapasin proteins are only able to associate with the TAP and class I molecules that are encoded by the same haplotype. This would bring only those class I and TAP molecules that have complementary peptide-binding and peptide-translocating requirements into close proximity, and would prevent class I molecules with different peptide-binding motifs from associating with inappropriate TAP peptide transporters. Such a possibility would provide a molecular explanation for the predominant expression of a single dominant class I molecule, as additional class I molecules may be unable to interact productively with tapasin and may be supplied with inappropriate peptides by the TAP molecules.

Therefore the second aim of this project is to determine whether chicken tapasin is involved in a haplotype-specific interaction with the class I and TAP proteins.

Aim three: Characterise other influences that polymorphic residues in chicken tapasin alleles may have on MHC class I peptide-loading

In the event that chicken tapasin is polymorphic it is possible that the products of one (or more) tapasin allele(s) may be capable of facilitating optimal peptide-loading to a greater or lesser extent than other tapasin alleles, irrespective of whether tapasin is involved in a haplotype-specific interaction. Such variance in function may have been selected at times of pathogen pressure experienced by particular haplotypes.

If tapasin were polymorphic and involved in a haplotype-specific interaction, it is possible the products of tapasin alleles vary in function, but can only interact with the class I and TAP molecules encoded by alleles of the same haplotype. This may result in the extent of optimal peptide-loading varying between haplotypes.

Alternatively, if chicken tapasin were polymorphic but not involved in a haplotype-specific interaction, it is possible that the products of tapasin alleles vary in function, and are capable of interacting with TAP and class I alleles of any haplotype. Whilst it would be expected that a "superior" tapasin would be selected to fix throughout the population, this may have yet to occur, perhaps because of the limited opportunity for recombination to allow the superior tapasin allele to replace other alleles.

Polymorphic residues in chicken tapasin may therefore provide clues as to the general mechanism of how tapasin facilitates and optimises class I peptide-loading. Additionally, polymorphic residues in chicken tapasin may also explain phenomena attributed to the chicken MHC, including differential disease resistance and cell surface class I expression levels. Therefore, the third aim of the project was to characterise the influence that any polymorphisms in chicken tapasin have on class I peptide-loading.

1.7.1 Examples of co-evolution occurring between genes encoded within the MHC

As a precedent, there are at least three examples of MHC genes co-evolving within their haplotypes. The first example concerns two pairs of mouse class II genes: the I-Aα and the I-Aβ pair; and the I-Eα and I-Eβ pair. The I-Aα and I-Aβ genes, whose products constitute the I-A heterodimer, are found in close proximity to each other and are both polymorphic (Germain et al., 1985). Molecules of I-Aα and I-Aβ can interact only when both are encoded from alleles from the same haplotype, but not when the alleles are from different haplotypes. Therefore I-Aα and I-Aβ have co-evolved, with the specificity of the interaction being mediated by polymorphic amino acid interactions that can only occur when the molecules are encoded by alleles present on the same haplotype. The I-Eα and I-Eβ genes have evolved in a separate fashion. Although the genetic distance between the I-Eα and I-Eβ genes is small, the genes are separated by a hotspot of recombination occurring within the I-Eβ locus. It appears that the I-Eα gene product has evolved to act as a monomorphic average "best-fit" molecule capable of interacting with whichever polymorphic I-Eβ chain partner is encoded within the haplotype. A similar situation exists in the human MHC, which involves a haplotype-specific interaction occurring between the polymorphic α and β chains of the HLA-DP and HLA-DQ heterodimers, whilst the monomorphic HLA-DRα protein can interact with the product of any HLA-DRβ allele.

The second example of co-evolving polymorphic genes is the rat TAP and MHC class I genes. Certain class I alleles such as RT1.A^{a}, which bind peptides with C-terminal basic amino acids, are

generally linked to the TAP2A allele, which is capable of transporting a broad repertoire of peptides; whilst other class I alleles, such as RT1.A^u, that bind peptides with C-terminal hydrophobic residues are generally linked to the TAP2B allele, which generally transports peptides with hydrophobic residues at the C-terminus (Livingstone et al., 1989; Livingstone et al., 1991; Powis et al., 1991; Joly et al., 1998). However, if class I alleles such as RT1.A^a are expressed with the TAP2B allele with which they are not usually expressed (and which transports a more restrictive repertoire of peptides), such as occurs in recombinant strains of laboratory rat, then the surface expression level of these class I alleles is dramatically reduced (the so called "cim" effect), as a result of inappropriate peptide supply. Such co-evolution between polymorphic TAP and class I genes is possible in the rat because the recombinational distance separating these genes is greatly reduced compared to the human and mouse.

A third example of genes co-evolving within the MHC comes from the chicken MHC. The dominantly expressed class I gene BF2 is highly polymorphic, with different alleles binding different repertoires of peptides. The chicken TAP genes, which are located between the two BF genes (figure 1.11) are also polymorphic both at the sequence level and at the level of function (Walker et al., 2005). The evidence that the chicken TAP genes may be co-evolving with the BF2 gene is provided by a correlation in the properties of the polymorphic amino acids in the TAP molecules and in the peptide-binding groove of the class I molecules in some haplotypes (Kaufman et al., 1999a). In the B4 haplotype the peptide motif eluted from class I molecules contains three acidic residues, at peptide positions 2, 5 and at the C terminus, which coincide with the BF2 gene having positively charged residues (R9, R80, R113 and R152) in the peptide-binding groove. The TAP1 allele found in the B4 haplotype uniquely has positively charged residues that are negatively charged in all other haplotypes at two positions (K54 and K131). Whilst K54 and K131 do not align with residues implicated in binding peptides in mammalian TAP molecules (Nijenhuis and Hammerling, 1996; Nijenhuis et al., 1996), it is possible that they may contribute to the selection of peptides for transport in chicken TAP proteins, and may select for transport only those peptides bearing opposing charges, so tailoring peptide provision for the BF2 molecule encoded by the B4 haplotype. In contrast, in the human and murine MHC regions, the TAP proteins are not very polymorphic, and the few sequence polymorphisms that are found do not affect the peptide transport specificity (Schumacher et al., 1994b; Obst et al., 1995; Daniel et al., 1997). As the human and murine TAP genes are separated from the class I genes by greater recombinational distances than in the chicken MHC (figures 1.9 and 1.11), it has been suggested that the human and murine TAP molecules have evolved as functionally monomorphic "best-fit" molecules that supply a repertoire of peptides that is appropriate for whichever class I molecules are present (Kaufman, 1999; Kaufman et al., 1999a).

1.7.2 Are all chicken class I alleles dependent upon tapasin?

Studying the influence that any polymorphic residues in chicken tapasin alleles have on class I peptide-loading will inevitably use the expression characteristics of the class I molecule (for example surface expression level) as a reflection of the ability of the tapasin allele to facilitate optimal peptide-loading. However, regardless of the extent of polymorphism in chicken tapasin, it is unknown whether the products of the chicken class I alleles are all equally dependent upon tapasin to achieve optimal peptide-loading. For example, it has been documented that whilst all human class I alleles benefit from the action of tapasin, they do so to various extents; with some class I alleles possessing an enhanced intrinsic ability to self-optimise their peptide repertoire in the absence of tapasin.

Therefore different chicken class I alleles may rely upon tapasin to different extents in order to achieve optimal peptide-loading. Some chicken class I alleles could be capable of relatively efficiently self-optimising their peptide repertoire without the assistance of tapasin, and could be referred to as tapasin-independent alleles, similar to alleles such as HLA-B*2705 and B*4405. The extent to which such tapasin-independent chicken class I alleles exist and may profit from tapasin is also unknown; they may benefit (like all human class I alleles); or they may be capable of becoming optimally peptide-loaded entirely independently of tapasin, defining a new category of "completely-tapasin-independent" class I alleles. Such alleles may be completely unaffected by pathogen subversion of PLC activity. Alternatively, the chicken class I alleles may strictly require tapasin in order to become optimally peptide-loaded (similar to HLA-B*4402), and could be referred to as tapasin-dependent alleles. Therefore this possible variation in the abilities of different class I alleles to self-optimise their peptide repertoires may confound the comparison of the expression characteristics of chicken class I alleles.

In the event that chicken tapasin is polymorphic, it is possible that the products of particular tapasin alleles are capable of facilitating optimal peptide-loading to a greater or lesser extent than other tapasin alleles. Additionally, chicken class I alleles may have evolved to function specifically with the tapasin allele that is encoded by the same haplotype; in which case the class I alleles could be referred to as "tapasin-specific" alleles. Either or both of these possibilities may confound the comparison of the same class I allele expressed with different tapasin alleles; as the tapasin alleles may vary in function, and the class I alleles may be most optimally peptide-loaded by a specific tapasin allele.

1.7.3 Possible strategies to characterise the functional influence of polymorphic residues in tapasin on class I peptide-loading

If tapasin were found to be polymorphic, the ideal method for comparing the expression characteristics of class I alleles that were expressed in the presence of different tapasin alleles would be to reconstitute a tapasin-deficient and/or class I-deficient chicken cell line with different

combinations of chicken tapasin and class I alleles. Unfortunately such cell lines do not exist; neither do TAP-deficient chicken cell lines; which could have been used to determine whether the products of chicken tapasin and TAP alleles that are encoded from a variety of haplotypes interact differentially. In addition, the study of MHC class I antigen presentation in chickens is hampered by a lack of certain serological reagents, including:

- There are no antibodies that are specific for particular conformations of class I molecules (for example, pMHC heterodimers vs. free HC). However, there are antibodies against chicken class I and β_2m proteins which work in all types of assay; including western blots, immunoprecipitations and flow cytometry.
- There are no antibodies to certain components of the chicken PLC, including tapasin or TAP, and attempts to raise such antibodies have so far been unsuccessful (K. Skjodt, A. van Hateren, B. Walker and J. Kaufman unpublished).
- However, there are commercially available antibodies against calnexin and calreticulin, which were raised against peptide sequences derived from the mammalian proteins, that cross-react with proteins of the appropriate molecular weight in chicken lysates (details in methods).

Therefore three options are to reconstitute either: 1) mammalian; 2) insect; or 3) chicken cell lines with different combinations of chicken tapasin, TAP and class I alleles.

Mammalian cells:

A number of well-studied mutant cell lines have defects in genes involved in class I antigen presentation. For example, the human tapasin-deficient 721.220 cell line could be used (Greenwood et al., 1994). The advantage of using a tapasin-deficient cell line from a mammalian species is that, although it is unknown whether the mammalian tapasin protein can interact productively with chicken class I molecules, the absence of any tapasin molecules eliminates any possibility of such an interaction occurring.

Human cells offer the advantage that the proteasome cleavage specificity is likely to produce peptides suitable for most chicken class I alleles, but possibly not for B4, as the generation of peptides with C-terminal acidic residues is not favoured by the human proteasome (class I peptide motifs for various haplotypes are listed in later chapters). The human TAP peptide transport profile has also been well characterised, and is likely to transport peptides suitable for most class I alleles. In addition there is the potential to use a number of well-defined monoclonal antibodies that are specific for human proteins that are not available for their chicken orthologues.

However there are examples documented of molecular incompatibilities occurring between class I and tapasin proteins derived from different mammalian species (Peh et al., 2000). Therefore it is possible that the chicken proteins may not be able to interact with the mammalian orthologues of their normal molecular partners.

Insect cells:

Invertebrate cells lack an MHC and therefore offer the opportunity to reconstitute the exact combinations of genes of interest. However, a prerequisite for studying the interactions between class I and tapasin molecules is that the insect cell line must first be reconstituted with functional TAP molecules in order to supply suitable peptides for class I peptide-loading. However attempts to reconstitute insect cells with functional chicken TAP genes (with or without tapasin) have so far proven unsuccessful (B. Walker, A. van Hateren, J. Kaufman unpublished). In addition, it may be necessary to express the avian orthologues of ERp57, calreticulin and calnexin in order reconstitute the exact function of a normal chicken peptide-loading complex.

Chicken cells:

Chicken cells offer the advantage that all the appropriate chaperones and other molecules are already present, therefore limiting the amount of "reconstitution" that is needed. However chicken cell lines with specific defects in class I antigen presentation are not available. Therefore all chicken cell lines will express endogenous alleles of tapasin, TAP and class I molecules which may complicate the analysis of any proteins transfected into such cells.

An additional complexity in using chicken cell lines to compare combinations of transfected class I and tapasin alleles is that the products of the polymorphic chicken TAP alleles are much more restrictive in their transport specificity of index peptides than their mammalian orthologues; with cell lines of different haplotypes exclusively transporting distinctive sets of peptides (J. Kaufman and B. Walker, personal communication; Walker et al., 2005). This may prohibit the characterisation of the effect that mismatching tapasin and class I alleles has; as class I molecules must be supplied with a suitable peptide repertoire within the ER before the effect of potential tapasin incompatibility could become evident.

1.7.4 Summary of aims

1. To determine if chicken tapasin is polymorphic at the sequence level in the seven MHC haplotypes that are kept at IAH Compton.
2. To determine whether tapasin participates in a haplotype-specific interaction with the class I and TAP molecules.
3. To determine other influences of polymorphic residues in tapasin on class I peptide-loading.

2. Methods

2.1 Descriptions of materials

2.1.1 Materials

All chemicals were purchased from Sigma unless stated. Components of propriety reagents are detailed, where known. All media and buffers were prepared using ion-exchange purified or double-distilled water. Reagents were sterilised, where necessary, by autoclaving for 20 minutes at 15 psi, or by filtration through a 0.2 μm filter.

2.1.2 Buffers and solutions

RNAse-free H_20	DNAse and RNAse-free water
PBSa	0.15 M NaCl, 2.5 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.0
TE	10 mM Tris, 1mM EDTA, pH 7.5
2x KGB	200 mM potassium glutamate, 50 mM Tris-acetate (pH 7.5), 20 mM magnesium acetate, 0.1 mg/ml BSA (fraction V), 1 mM β-mercaptoethanol
TFB1	30 mM potassium acetate, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM $RbCl_2$, 15% glycerol, pH 5.8
TFB2	10 mM MOPS (pH 6.5), 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15% glycerol, pH 6.5
TAE (50x)	2 M Tris, 1 M glacial acetic acid, 0.05 M EDTA (pH 8)
Agarose gel loading buffer	50% glycerol, 0.25% xylene cyanol, 0.25% bromophenol blue
dNTP mixture	5 mM each dATP, dCTP, dGTP and dTTP
N-P40 lysis buffer	0.5% N-P40, 150 mM NaCl, 50 mM Tris pH 8, 10 mM EDTA, 0.1% NaN_3, protease inhibitors, 10 mM iodoacetamide
RIPA lysis buffer	150 mM NaCl, 1.0% Igepal CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0, protease inhibitors, 10 mM iodoacetamide
LSB (5x)	60 mM Tris HCl (pH 6.8), 40% (v/v) glycerol, 10% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.05% (w/v) bromophenol blue
SDS-PAGE running buffer	2 M Glycine, 0.25M Tris, 1% (w/v) SDS
Western blotting buffer	2 M Glycine, 0.25M Tris
TBS Tween	50 mM Tris HCl (pH 8), 150 mM NaCl, 0.1% Tween-20
FACS wash	PBSa containing 0.05% (w/v) BSA and 0.1% NaN_3

2.1.3 Restriction enzymes

Restriction endonuclease digestions were carried out at 37°C and as detailed below.

Enzyme	Restriction site	KGB concentration	Manufacturer
*BamH*I	G'GATCC	0.5-1.0x	Invitrogen
*Cla*I	AT'CGAT	0.5-2.0x	Invitrogen
*EcoR*I	G'AATTC	0.5-2.0x	Invitrogen
*Hind*III	A'AGCTT	0.5-2.0x	Invitrogen
*Kpn*I	GGTAC'C	0.5-1.0x	Invitrogen
*Not*I	GC'GGCCGC	0.5-2.0x	Invitrogen
*Pvu*I	CGAT'CG	0.5-2.0x	New England Biolabs
*Sac*I	GAGCT'C	0.5-1.0x	New England Biolabs
*Sca*I	AGT'ACT	1.0-2.0x	Gibco BRL
*Xba*I	T'CTAGA	0.5-1.5x	Gibco BRL
*Xho*I	C'TCGAG	0.5-1.5x	Gibco BRL

2.1.4 Bacterial strains

E. coli DH5α — *sup*E44 β*lac* U169 (φ80 *lacZ*ΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *relA*

2.1.5 Growth media

Luri-Bertani (LB) and cell culture media were prepared by the Microbiological Services department, IAH Compton.

LB — 1% (w/v) Bacto tryptone, 0.5% Bacto yeast extract, 0.5% (w/v) NaCl

LB^{Amp} — LB, 150 μg/ml ampicillin

LB agar — LB, 1.5% (w/v) agar

RPMI — Roswell Park Memorial Institute (RPMI 1640) supplemented with 2 g/l sodium bicarbonate and L-glutamine

2.1.6 Synthesis and sequence of primers

Primers were purchased from MWG Biotech, TAGN or Sigma as lyophilised powder and were resuspended in TE at 100 μM, numbered and stored at -20°C. Primer sequences are provided at the end of the methods chapter.

2.1.7 Synthesis of peptides

Peptides were made by the protein chemistry laboratory at IAH Compton. Peptides were dissolved in a final concentration of 0.9% DMSO in PBS.

2.1.8 Description of vectors

pIST

For sequencing; a derivative of pBSII KS(+), modified by the insertion of two *Xcm*I sites, which creates a T-tailed vector when cut (I. Shaw, unpublished).

Eukaryotic expression vectors

All vectors contained the human CMV immediate-early promoter for constitutive expression of the DNA insert in mammalian cells

pCIpac

A derivative of pCIneo (Promega), modified by the replacement of neomycin resistance gene with the puromycin resistance gene. Kindly provided by Clive Tregaskes.

pcDNA3.1+neo

Encodes the neomycin resistance gene. (Invitrogen).

pcDNA6

Encodes the blasticidin resistance gene. (Invitrogen).

pMCFR

Encodes the puromycin resistance gene, kindly provided by Prof. T. Elliott, containing human tapasin cDNA cloned between the 5' *Kpn*I and 3' *Cla*I restriction sites (Ortmann et al., 1997).

2.1.9 Cell lines and growth media and conditions

721.220

Human HLA-A and –B negative, tapasin-deficient cell line (Greenwood et al., 1994). Maintained at 37°C in 5% CO_2 incubators, in RPMI 1640 supplemented with 10% Foetal Calf Serum (FCS), two mM glutamine, and antibiotics (penicillin and strepamycin, each at 100 units/ml) (medium referred to as RPMI 10% FCS or R10).

TG15

REV transformed lymphocytes from a homozygous B15 chicken, were provided courtesy of Dr Thomas Goebel (University of Munich). Maintained at 41°C in 5% CO_2 incubators, in R10.

2.1.10 MHC haplotypes

Line	Line hereafter referred to as	MHC haplotype	Place of development
6_1	6	B2	Regional Poultry Research Lab (East Lansing, MI)
7_2	7	B2	Regional Poultry Research Lab (East Lansing, MI)
C-B4	4	B4	Subline derived from Reaseheath line C, Northern Poultry Breeding station (Reaseheath, UK)
C-B12	12	B12	Subline derived from Reaseheath line C, Northern Poultry Breeding station (Reaseheath, UK)
WL	14	B14	Wellcome Research Lab (Beckenham, UK)
151	15	B15	Regional Poultry Research Lab (East Lansing, MI)
P	19v1	B19v1	Originated from Line P-2a, Cornell Uni (Ithaca, NY)
0	0	B21	Regional Poultry Research Lab (East Lansing, MI)
N	N	B21	Regional Poultry Research Lab (East Lansing, MI)

2.2 General nucleic acid methods

2.2.1 DNA extraction from erythrocytes

DNA was extracted from chicken erythrocytes by salting out. Briefly two ml of blood was diluted in PBSa, underlayed with two ml of Ficoll-Paque (Pharmacia) and centrifuged at 223 g. After pipetting off the Ficoll-Paque, the erythrocytes were washed twice with 12 ml PBSa and centrifuged for five minutes at 322 g. Ten µl of the resulting packed erythrocytes were suspended in five ml of 300 mM NaCl, 10 mM EDTA pH 8.0, 10 mM Tris pH 8.2 and 50 µl of 10 mg/ml proteinase K. This was mixed and incubated at 37°C for 10 minutes. 100 µl of 10% SDS was added and the cells were incubated overnight at 37°C. Following cooling to room temperature, 1.3 ml of saturated NaCl was added and the cells shaken vigorously for 15 s. The resulting liquid was centrifuged at 3,000 g for 20 minutes. The supernatant was transferred to a clean tube and DNA precipitated with four ml of isopropanol for 15 minutes. After centrifugation at 3,000 g for 15 minutes the pellet was washed twice with 80% ethanol, the ethanol drained and the pellet air dried. The pellet was resuspended in 200 µl of TE and stored at 4°C. (Performed by Jim Kaufman and previous lab members)

2.2.2 Quantification of nucleic acid

Nucleic acid concentrations were estimated by spectrometry, assuming an optical density (OD) of 1.00 corresponds to a concentration of 50 µg of double stranded DNA at a wavelength of 260 nm. Purity was estimated by calculating the $OD_{260/280}$ ratio.

2.2.3 Details of genomic DNA

Line	MHC haplotype	Lab #	Wing band number	Sex	Date of bleed	$A_{260/280}$	Yield µg/ml
6	2	2	928	F	27/5/98	1.821	509
7	2	17	1760	M	27/5/98	1.867	569
C	4	28	481	F	27/5/98	1.865	550
12	12	43	2370	F	27/5/98	2.250	360
W	14	55	26	F	11/2/98	1.695	500
15	15	69	1509	F	27/5/98	1.927	539
15	15	64*	1505*	F	27/5/98	1.827	694
15	15	66*	1507*	F	27/5/98	1.970	670
P	19v1	71	2895	M	27/5/98	1.897	550
0	21	86	2361	M	27/5/98	1.885	584
N	21	96	233	M	27/5/98	1.979	514

* Sources of genomic DNA used to confirm sequence of B15 genomic sequence

2.2.4 Isolation of RNA

Total RNA was isolated from 10^7 white blood cells of MHC homozygous chickens using the RNeasy minikit (Qiagen) following the manufacturer's instructions with the sample being homogenised by passing through a 20 g needle five times. Additional on-column DNase digestion was performed, as described in protocol.

2.2.5 Details of source of RNA

Line	MHC haplotype	Wing band number	Date
6	2	0808	21/2/03
7	2	0332	21/2/03
C	4	0409	21/2/03
12	12	0505	21/2/03
W	14	1228	21/2/03
15	15	1430	21/2/03
P	19v1	1008	21/2/03
0	21	1601	21/2/03
N	21	1102	21/2/03

2.2.6 cDNA production

cDNA was produced using Superscript III reverse transcriptase (Invitrogen) according to the manufacturer's instructions, with an additional step to potentially relieve RNA secondary structures: four µl RNA and one µl reverse primer (detailed below), were incubated at 70°C for 10 minutes, prior to the addition of the remaining components required for the first strand cDNA synthesis, and incubated at 55°C for two hours. Following denaturing of reverse transcriptase by treatment at 70°C for 15 minutes, RNA was degraded by incubation with RNase H (Invitrogen) at 37°C for 20 minutes. Resulting cDNA was stored at –20°C.

Gene(s)	Oligo used
Tapasin	C469
BF2, β_2m	Oligo dT

2.2.7 Polymerase chain reaction (PCR)

PCR amplifications were performed in a T3 Thermocycler (Biometra). Unless stated all amplifications were performed with 15 pmol of each primer, 0.2 mM total dNTPs, 1x enzyme buffer, 1 M $MgSO_4$, and 2 units *Pfx* polymerase (Invitrogen) in a 50 µl reaction. Primer details and conditions are listed in specific methods sections.

2.2.8 Site directed mutagenesis

Site directed PCR-based mutagenesis was performed essentially as described (Hutchison et al., 1978). Primers were designed with at least 15 complementary nucleotides on either side of the centrally located desired substitution, with a G or C nucleotide at the 3' end. The mutants were produced in two PCR stages (figure 2.1).

Part 1: Two PCR amplifications were performed, each containing a single mutant primer and a primer complimentary to the vector sequence.

Part 2: To produce the full length mutated product, a final PCR amplification was performed using the two amplified fragments obtained from the initial two PCRs as template DNA together with vector specific primers.

2.2.9 Agarose gel electrophoresis of DNA

Electrophoresis of PCR products or restriction enzyme digested DNA was carried out in agarose gels (1-2% w/v) in 1x TAE buffer containing 0.6 µg/ml ethidium bromide. All gels were cast and run in horizontal gel tanks. PCR products or fragments were mixed with 10x gel loading buffer (to make 1x final concentration), loaded into the gel and separated by electrophoresis. DNA-ethidium bromide complexes were visualised under ultraviolet light at a wavelength of 302 nm.

2.2.10 Recovery of DNA from agarose gels

Following electrophoresis DNA bands were excised and purified using the Minelute kit (Qiagen) following manufacturer's instructions. Samples were resuspended in TE and stored at -20°C.

2.2.11 Alkaline Phosphatase treatment of vector DNA

Shrimp alkaline phosphatase (SAP) (Promega) was used to dephosphorylate 5' phosphates from vector DNA. Dephosphorylation was performed following manufacturer's guidelines: briefly one unit SAP per µg DNA was incubated at 37°C for 15 min in 1x SAP reaction buffer in a final volume of 30 µl. SAP was then inactivated by incubation at 65°C for 15 min. Dephosphorylated vector was used directly for ligations or frozen at -20°C.

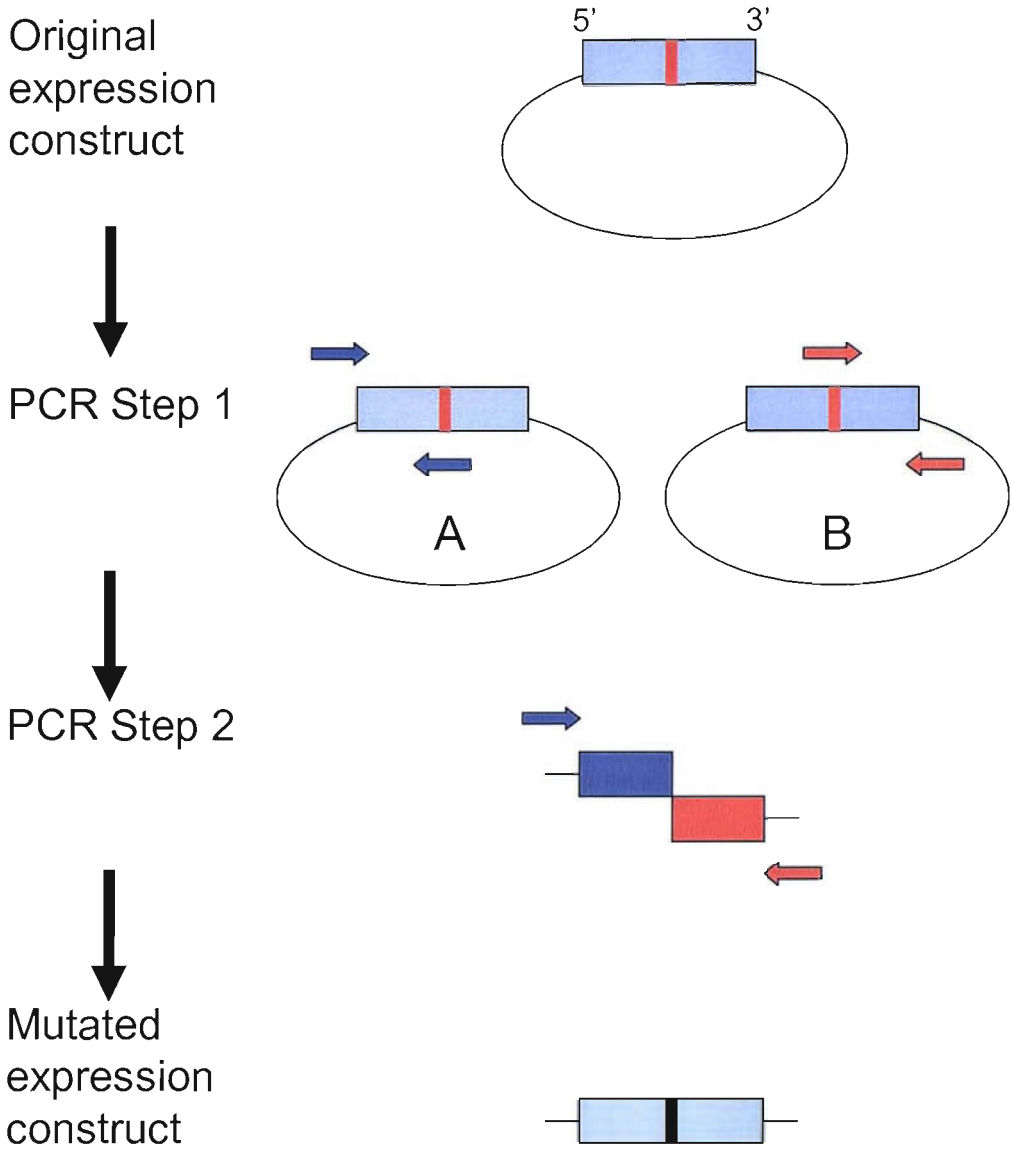


Figure 2.1 Schematic representation of the procedure used to produce mutated expression constructs.
The original expression construct is depicted as a light blue box, with the area to be mutated indicated by a red vertical line. The mutation was produced in two steps.
Step 1: Two PCRs were performed, with each PCR containing one vector specific primer and one of two primers that contained the desired mutation, but were otherwise complementary to the target sequence. The 5' portion of the gene was amplified using the blue primer pair and the 3' portion was amplified using the red primer pair. The amplified fragments were gel purified.
Step 2: A final PCR was performed using both products (5' fragment shown as blue box, 3' fragment shown as red box) and vector specific primers.
The mutated expression construct is depicted as a light blue box with the mutated area indicated by a black vertical line.

2.2.12 Ligation of DNA into vector

Ligation of PCR product, or restriction enzyme digested fragment, into restriction enzyme digested vector were carried out using 10 units of T4 DNA ligase (Promega), 1x ligation buffer, 25 ng vector, 10 ng PCR product or restriction enzyme digested fragment, in a final volume of 10 µl. Ligation reactions were incubated overnight at 14°C and used to transform competent bacterial cells or stored at -20°C.

2.2.13 Preparation of heat-shock competent *E.coli*

DH5α cells were streaked out onto LB agar plates and grown overnight at 37°C. A single colony was used to inoculate 2.5 ml LB medium and grown overnight at 37°C with shaking at 180 rpm. The overnight culture was used to inoculate 250 ml LB medium supplemented with 20 mM $MgSO_4$. The cells were grown until $O.D._{600}$ reached 0.4-0.6. The cells were centrifuged at 2,500 rpm for 10 minutes at 4°C. The cells were resuspended in 100 ml cold TFB1 and kept on ice for five minutes. The cells were centrifuged at 2,500 rpm for 10 minutes at 4°C and resuspended in 10 ml ice-cold TFB2. Following incubation on ice for one hour the cells were aliquotted into cold microfuge tubes and snap frozen in liquid nitrogen and stored at -80°C.

2.2.14 Transformation of *E. coli* by heat shock

Two µl of each ligation product was gently mixed with 50 µl of competent DH5α and incubated on ice for 30 minutes. Cells were transformed by heat shock at 42°C for 30 seconds and returned to ice for two minutes. 150 µl of LB was added and incubated at 37°C on a shaker at 180 rpm for one hour. Fifty µl of the culture was spread onto LB agar plates supplemented with appropriate selection antibiotic.

2.2.15 Purification of plasmid DNA

Individual colonies were picked into either four ml (miniprep) or 100 ml (maxiprep) LB supplemented with the appropriate selection antibiotic and grown on a shaker at 180 rpm overnight at 37°C. The cells were centrifuged at 2500 rpm and the supernatant aspirated. Plasmid DNA was purified using either Miniprep or Endo-free Maxiprep kits (Qiagen) following the manufacturer's instructions.

2.2.16 Automated DNA sequencing

All sequencing was performed using the Beckman CEQ capillary sequencer, following the manufacturer's guidelines. The sequencing reaction consisted of 1-2 µl DNA (~200 µg) and water being incubated at 85°C for five minutes, before five pmoles primer, three µl of Beckman QuickStart sequencing mix was added to make 10 µl final volume. Cycling conditions were as follows: 96°C for two minutes, followed by 30 cycles of 96°C for 30 s, 50°C for 30 s and 60°C for four minutes. Samples were precipitated, washed and resuspended in 40 µl sample loading solution following the manufacturer's instructions.

2.3 General protein and cellular methods

2.3.1 SDS PAGE

SDS PAGE gels were made using BioRad Mini-Protean II or III apparatus following the manufacturer's instructions. Laemmli polyacrylamide gels were prepared with 10-15% acrylamide/BIS (30% stock, Biorad) separating gel and a 4% acrylamide/BIS stacking gel. Prior to loading, the samples were mixed with LSB, heated at 95°C for three minutes and centrifuged at 16,000 g. Gels were run in 1x SDS PAGE running buffer.

2.3.2 Western blotting

Following SDS PAGE electophoresis, the stacking gel was removed, and the proteins transferred to a nitrocellulose filter (Hybond-C extra, Amersham) by western blotting using a BioRad Mini Trans-Blot Electrophoretic Transfer cell, following the manufacturer's instructions. The membranes were subsequently blocked with gentle agitation for at least one hour at room temperature or 4°C overnight, and screened with appropriate antibodies (tabulated below). Blots were incubated with antibody solutions for one hour at room temperature with gentle agitation. Following antibody incubations, blots were quickly rinsed two times and then washed three times with gentle agitation for at least five minutes with TBS-Tween. Following secondary antibody incubations, blots were washed as described in TBS-Tween, and then incubated with ECL western blotting detection reagents (Amersham) following manufacturer's instructions. The blots were exposed to film (either KAR-5 or Biomax Light-1, Kodak) and developed using an X-ograph Compact X2 processor.

Gene product / epitope tag	Blocking reagent	Primary antibody & dilution (in blocking reagent)	Secondary antibody: dilution & diluent
Myc	3% BSA	1:1,000 anti-myc (Autogen Bioclear)	goat α-mouse Ig* HRP: 1:4,000 in 10% milk
Flag	3% BSA	20 μg/ml anti-Flag M2	goat α-mouse Ig* HRP: 1:4,000 in 10% milk
BF2	10% milk	1 μg/ml or 1:50 Techno-Mouse supernatant F21-2 (Crone et al., 1985)	goat α-mouse Ig* HRP: 1:4,000 in 10% milk
Calnexin	3% BSA	1:2,000 anti-calnexin (Stressgen)	α-rabbit IgG HRP: 1:4000 in 3% BSA
Human tapasin	3% BSA	1:1,000 Giles 11/11‡	α-rabbit IgG HRP: 1:4000 in 3% BSA
V5	3% BSA	1:5,000 anti-V5 (Invitrogen)	goat α-mouse Ig* HRP: 1:4,000 in 10% milk
His_6	10% milk	1:100 anti-His_6 (Serotec)	goat α-mouse Ig* HRP: 1:4,000 in 10% milk
Chicken β_2m	5% milk	Rabbit anti-chicken β_2m antisera†	α-rabbit IgG HRP: 1:4000 in 3% BSA

* Sigma, pre-absorbed against human serum to remove cross reactivity.

‡ Giles 11/11 (a gift from Prof. T. Elliott), a rabbit anti-peptide antibody raised against the N-terminal 66 amino acids of human tapasin.

† Provided by Dr J. Kaufman, raised against chicken β_2m.

2.3.3 General protocol for cell lysis

A volume of cell suspension containing the appropriate number of cells was transferred to an eppendorf tube and centrifuged at 380 g for three minutes at 4°C. The supernatant was completely removed and the cell pellet lysed by the addition of ice cold RIPA (or N-P40) lysis buffer. The lysates were vortexed briefly and incubated on ice for 15 minutes. The nuclear debris was removed by centrifugation at 16,000 g for 10 minutes at 4°C, and the supernatant transferred to a new eppendorf tube.

2.3.4 Flow cytometry staining and acquisition

Cells ($5x10^5$ - $1x10^6$) were transferred to 96 well U-shaped plates, and washed twice in FACS wash with centrifugation at 320 g for 2.5 minutes. Cells were incubated on ice for 30-60 minutes in 200 µl of primary antibody diluted in FACS wash (tabulated below). Cells were then washed two or three times in FACS wash before incubation in 200 µl of five µg/ml of FITC conjugated polyclonal goat anti-mouse Ig $F(ab')_2$ (DAKO) on ice, in the dark, for 30-60 minutes. Cells were then washed two or three times in FACS wash before resuspending in FACS wash containing five µg/ml propidium iodide (PI). Cytometry was performed on Becton Dickson FACScalibur following the manufacturer's guidelines, with between 10,000 – 30,000 events acquired. Live cells were analysed, assessed by exclusion of PI and FSC vs. SSC profiles.

Primary Antibody, reference or source	Antigen & comments	Dilution or concentration
F21-2 (Crone et al., 1985)	Chicken MHC class I heavy chains*	15 µg/ml
F21-21 (Skjodt et al., 1986)	Chicken β_2m*	1:10 Techno-Mouse supernatant
Flag M2	DYKDDDDK epitope	15-20 µg/ml
Anti-P75 (American Type Culture Collection, hybridoma 8737)	Human nerve growth receptor	15 µg/ml
W6/32 (Parham et al., 1979)	HLA: β_2m heterodimers	1:100
BBM1 (Brodsky et al., 1979)	Human β_2m	15 µg/ml
Anti-HLA-B Ab-1 (clone Joan-1) (Neomarkers) †	A determinant present on HLA-B locus encoded gene products	1:100
TT4-A20 (Muller et al., 1989; Tahara et al., 1990) †	The HLA-Bw4 public determinant controlled by polymorphic residues 77 and 80-83. TT4-A20 is sensitive peptide to the cohort (Tan et al., 2002).	50 µl
HC10 (Stam et al., 1986) †	Preferentially recognises β_2m free HLA-B and HLA-C heavy chain molecules	2 µg/ml
TRT1	IgG1 isotype control. Specific for a turkey rhinotracheitis antigen	1/10

All dilutions were made in FACS wash

* detects all protein conformations

† Kindly provided by Dr A. Williams, University of Southampton.

2.3.5 Transfection of cell lines and selection of stable transfectant clones.

Transfections were performed using a Nucleofector (Amaxa), following the manufacturer's guidelines. Briefly, the cells were sub-cultured into fresh media the day prior to transfection. On the day of transfection, $5x10^6$ cells per transfection were centrifuged at 200 g for 10 minutes and the supernatant was completely aspirated. The cell pellet was resuspended in 100 µl per transfection of appropriate Amaxa solution, mixed with DNA (linearised for stable transfection attempts), and transfected using the indicated program (tabulated below). Following transfection, the cells were immediately diluted into pre-warmed medium to obtain the following cell concentrations and were plated into 96 well U-shaped plates at a volume of 100 µl/well for each cell dilution: at 10^5/ml (labelled as "A" dilution), at $5x10^4$/ml (labelled as "B" dilution) and at 10^4/ml (labelled as "C" dilution). One to two days later, 100 µl of media supplemented with appropriate concentration of selection antibiotic was applied. Plates were incubated in CO_2 supplemented incubators, with the media being replenished every week. Clones were picked only from plates on which less than one third of the wells were growing and transferred to 24 well plates with application of fresh selection media. Subsequently clones were screened for expression of transfected gene via SDS PAGE and western blotting or flow cytometry.

Cell	Amount of DNA	Amaxa solution	Amaxa program
.220 & derivatives	Seven µg	V	A24
TG15	Five µg	T	T01
CHO	Five µg	T	U27
293T	Five µg	V	A23

Gene	Vector	Enzyme used to linearise vector	Selection antibiotic and concentration
BF2	pcDNA 3.1+neo	*Pvu*I	2 mg/ml G418
BF2	pcDNA 6	*Pvu*I	10 µg/ml blasticidin
Human tapasin R48 mutants	pMCFR	Not linearised	5 µg/ml puromycin
Chicken tapasin	pcDNA 3.1+neo	*Sca*I	2 mg/ml G418
Chicken tapasin	pcDNA 6	*Sca*I	10 µg/ml blasticidin
Chicken β_2m	pCIpac	Not linearised	5 µg/ml puromycin

2.4 Specific methods for chapter three

2.4.1 PCR amplification of chicken tapasin from cDNA

cDNA encoding chicken tapasin was amplified using c1353 (overlapping the ATG codon) and c469 (in 3'utr, near stop codon) primers. PCR conditions: two µl cDNA (produced with c469 primer) was amplified with a 96°C two min initial step followed by 30 cycles of 96°C for 30 s, annealing at 57°C for 30 s and extending at 68°C for 90 s. Lastly one µl *Taq* polymerase (Invitrogen) was added and incubated at 68°C for 15 min to create A nucleotide overhangs on the PCR product. PCR products were separated on agarose gels and bands of the correct size were excised and purified, and T/A ligated into pIST vector. Following transformation into DH5α cells plasmid DNA was isolated.

2.4.2 PCR amplification of chicken tapasin from genomic DNA

Using stocks of genomic DNA available in the laboratory, genomic DNA was amplified using the following primer pairs:

Amplified Region	Primer	Location of primer binding site	Annealing Temp (°C)	Extension time (min)
Exons & introns of tapasin	C1353	Overlapping ATG codon	60	Four
	C469	after stop codon	60	Four
BLB2-Tapasin "upstream"	C1576	Tapasin intron 1/exon 1	63	One
	C1578	~850bp upstream of exon 1 of tapasin, near BLB2	63	One
BLB1-Tapasin "downstream"	C1571	~700bp downstream of exon 8 of tapasin, near BLB1	61	One
	C1573	Intron 7/Exon 8	61	One

PCR conditions: 200 ng genomic DNA was amplified using 20 pmol of appropriate primers with a 96°C five min initial step followed by 35 cycles of 96°C for 30 s, annealing at the indicated temperature for 30 s and extending at 68°C for the indicated time. Lastly one µl *Taq* polymerase was added and incubated at 68°C for 15 min in order to create A nucleotide overhangs on the PCR product. PCR products were separated on agarose gels and bands of the correct size were excised and purified, and T/A ligated into pIST vector. Following transformation into DH5α cells plasmid DNA was isolated.

2.4.3 Design of primers used for sequencing tapasin

Primers were designed using the AL023516 BF/BL region sequence.

2.4.4 Sequence alignment and phylogenetic analysis

Sequence alignments were produced using Pileup (GCG10), and optimised using GeneDoc. Phylogenetic trees were constructed using the Neighbour Joining method in the Phylip software package.

2.4.5 Protein modelling

The Tapasin*15 protein sequence was modelled using Swiss-Model (Peitsch, 1996; Guex and Peitsch, 1997) against the C (1QLRC, 1DN0C) and A (1QLRA, 1DN0A) chains of the Fab fragment of human IgM specific for cold agglutinin, and against an anti-HIV-1 epitope and antibody structure (1TJIL). H-2K^{b} (accession NM001001892) was modelled against H-2K^{d} (1VGKA) and H-2D^{d} (1S7XA, 1S7WA, 1S7VD and 1S7XD). BF2*14 was modelled against RT1-A^{a} (1ED3A, 1ED3D and 1KJMA) and HLA-B*2705 (1HSAA and 1HSAD).

2.4.6 Production of human tapasin position 48 mutant expression constructs.

Human tapasin position 48 mutants were produced by site directed mutagenesis, using human tapasin in pMCFR vector as template DNA, and the primers listed below.

Step 1: PCR conditions: One µl DNA (~100 ng) was amplified with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 60°C for 30 s and extending at 68°C for 30 s.

Step 2: PCR conditions: One µl of each purified PCR product (from step 1) was amplified with c1667 and c1668 primers with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 60°C for 30 s and extending at 68°C for 90 s.

Human tapasin mutant	PCR step	Forward Primer	Reverse Primer
R48E	1A	C1667	C1672
	1B	C1671	C1668
R48A	1A	C1667	C1674
	1B	C1673	C1668
R48L	1A	C1667	C1675
	1B	C1670	C1668

The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Kpn*I and *Cla*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into empty pMCFR vector (that had been digested with *Kpn*I and *Cla*I and SAP treated). Following transformation into DH5α cells plasmid DNA was isolated and sequenced. Following retransformation of DH5α cells with a correct sequence construct, Endofree Maxipreps were prepared.

2.4.7 Transient transfection of HLA-B*4402 with human tapasin constructs

.220B*4402 cells were transiently transfected with seven µg wild-type human tapasin or R48E, R48A or R48L human tapasin mutants (all in pMCFR expression constructs). The following day, the cells were resuspended and counted, with an aliquot of cells being taken for SDS PAGE and

western blot analysis, with the remainder being stained with W6/32 antibody before flow cytometry was performed.

2.4.8 Generation of .220B4402 human tapasin position 48 mutant stable transfectants.

.220B4402 cells were transfected with seven μg human tapasin R48E, R48A or R48L in pMCFR expression constructs. Antibiotic resistant clones were screened by western blot for expression of human tapasin, and designated as below:

.220B4402 h.Tpn R48E	. 220B4402 h.Tpn R48A	.220B4402 h.Tpn R48L
R48E1, R48E2	R48A1	none

2.4.9 Flow cytometry and western blot analysis of position 48 mutant tapasin clones (experiment 275)

.220B*4402, .220B*4402hTpn, .220B*4402hTpnR48E1, .220B*4402hTpnR48A1 were sub-cultured the day prior to the experiment. On the day of the experiment the cells were resuspended by gentle pipetting and centrifuged at 320 g for five min. The supernatant was aspirated, and the cells resuspended in PBSa and counted. For each clone $8.5x10^6$ cells were taken and were centrifuged at 320 g for five min, 4°C before the supernatant was aspirated. The cells were then resuspended at $5x10^6$ cells/ml in cold FACS wash.

Flow cytometry: For each clone, 100 μl aliquots ($5x10^5$ cells) were aliquotted to 12 wells of a 96 well U-shaped plate, and were stained in triplicate with either W6/32, anti HLA-B Ab-1 "Joan-1", TT4-A20 or HC10 antibodies.

Western blotting studies: Lysates were prepared and volumes equivalent to either $3.3x10^5$ cells or $3.3x10^4$ cells were separated by 12% SDS PAGE and western blotted.

2.5 Specific methods for Chapter four

2.5.1 Production of BF2 expression constructs.

The BF2 gene was amplified from cDNA (made using oligo dT) using c1622 (introduces 5' *Hind*III site and Kozak sequence prior to ATG codon) and c1623 (introduces *Xba*I site after stop codon) primers. PCR conditions: one µl cDNA was amplified with a 96°C two min initial step followed by 30 cycles of 96°C for 30 s, annealing at 57°C for 30 s and extending at 68°C for 90 s. The PCR products were separated on agarose gels, and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *Xba*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into pcDNA6 vector that had been digested with *Hind*III and *Xba*I and SAP treated. Following transformation into DH5α cells plasmid DNA was isolated and sequenced. Following retransformation of DH5α cells with a correct sequence construct, Endofree Maxipreps were prepared.

2.5.2 Description of chicken tapasin expression constructs.

Due to the absence of an anti-chicken tapasin antibody, epitope tags were incorporated into tapasin expression constructs at the C-terminus. To decrease the possibility of inhibiting the ER-retention function of the dilysine motif, the epitope tag was inserted immediately prior to the last four amino acids and stop codon (KKSQ*). Two types of expression constructs of tapasin were produced (Figure 2.2).

2.5.3 Production of Tapasin His_6 expression construct.

Tapasin His_6 expression constructs were produced in three steps (Figure 2.3):

Step 1: PCR was performed using Tapasin*12 cDNA as the template and c1466 (introduces 5' *Hind*III site and Kozak sequence) and c1597 (replaces stop codon with *Eco*RI site) primers. PCR conditions: one µl cDNA was amplified with a 96°C two min initial step followed by 30 cycles of 96°C for 30 s, annealing at 57°C for 30 s and extending at 68°C for 90 s. The PCR products were separated on agarose gels, and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *Eco*RI restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer.

Step 2: To create the His_6 epitope tag complete with 5' *Eco*RI site, 3' duplicated dilysine motif, stop codon (KKSQ*) and *Xba*I site, two oligos were annealed together:

```
C1598 5' AA TTC CAC CAT CAC CAC CAT CAT AAG AAA TCG CAG TGA T 3'
C1491 5'      G GTG GTA GTG GTG GTA GTA TTC TTT AGC GTC ACT AGA TC 3'
AA seq          H   H   H   H   H   H   K   K   S   Q   stop
```

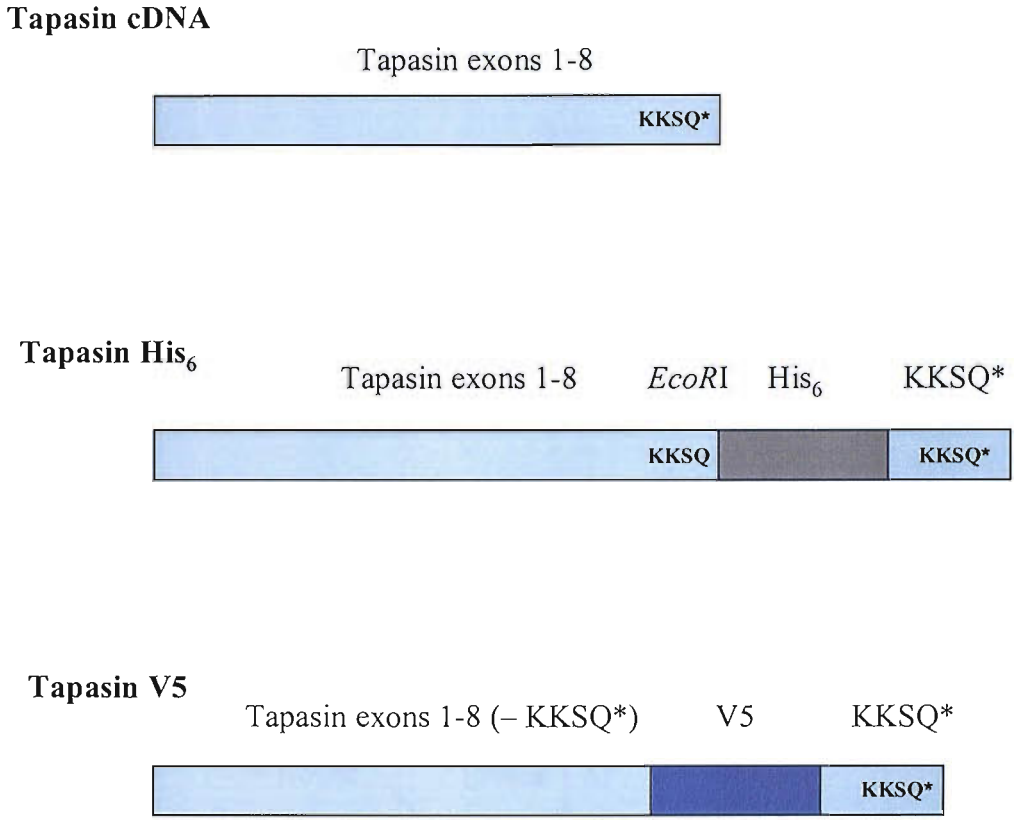


Figure 2.2 Chicken tapasin expression constructs.

Two expression constructs were produced, both having 5' *Hind*III restriction enzyme site and Kozak sequence and 3' *Xba*I restriction enzyme site:

Tapasin His_6: Tapasin exons 1-8 were amplified, replacing the stop codon with 3' *EcoR*I site.
Two oligos were annealed to create 5' *EcoR*I site: His_6 tag: duplicate of the last four amino acids: stop codon (KKSQ*): 3' *Xba*I site.
The PCR product was combined with annealed oligos and ligated into expression vector.

Tapasin V5: PCR was performed using Tapasin His_6 as template, replacing both copies of the duplicated KKSQ* and His_6 tag with a V5 epitope tag and a single copy of the last four amino acids and stop codon (KKSQ*).

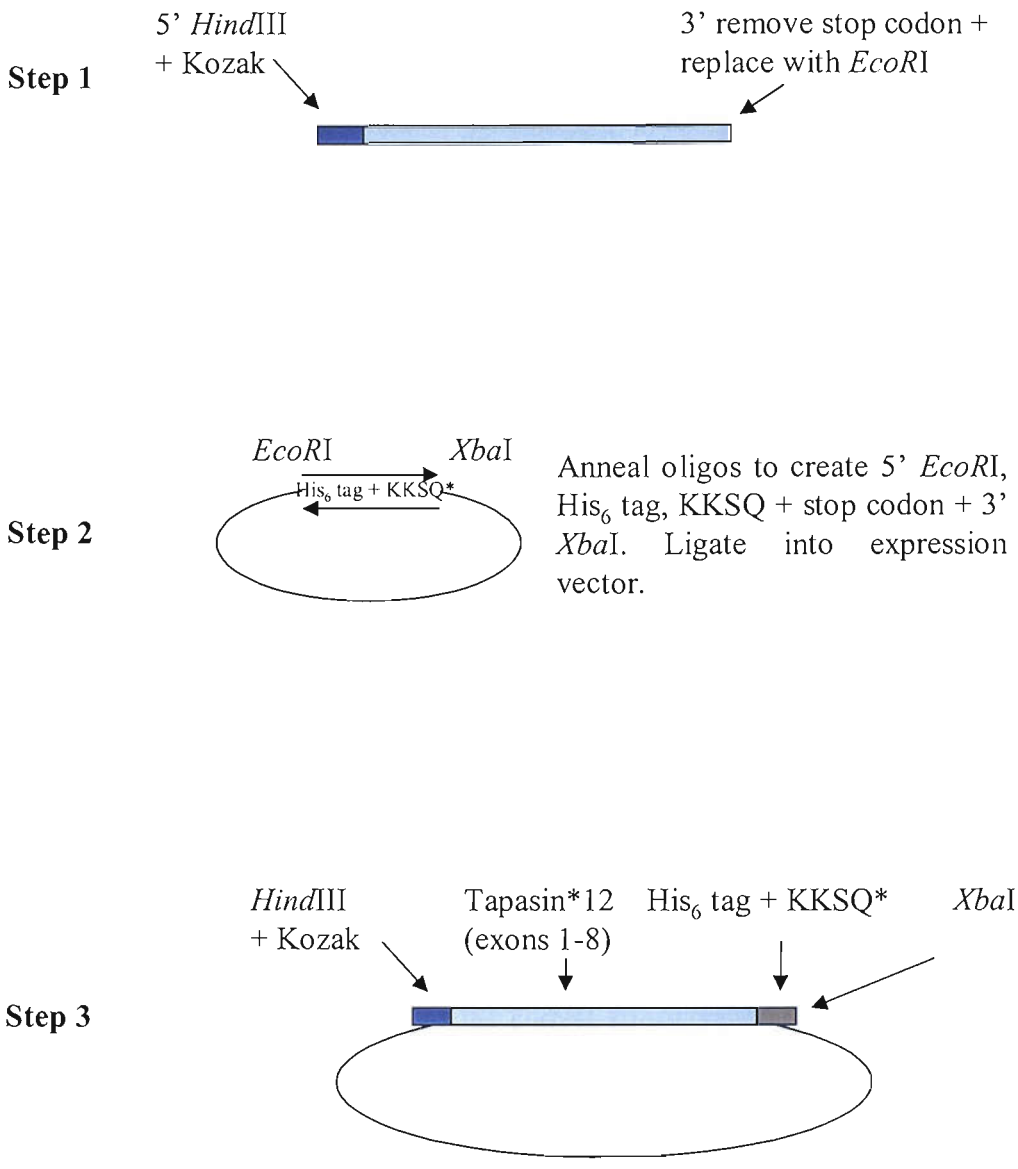


Figure 2.3 Production of Tapasin*12 His_6 expression construct.

Step 1: Tapasin*12 cDNA (exons 1-8) were amplified introducing 5' *Hind*III and Kozak sequence, and removing the stop codon and introducing 3' *Eco*RI restriction enzyme sites.

Step 2: Two oligos were annealed to create a His_6 tag complete with 5' *Eco*RI site and 3' KKSQ, stop codon and *Xba*I site. Annealed oligos were ligated into vector cut with *Eco*RI and *Xba*I enzymes.

Step 3: The His_6-containing-vector was digested with *Hind*III and *Eco*RI enzymes and ligated with the *Hind*III and *Eco*RI digested PCR product that was produced in step 1.

Fifty µl of 100 µM stocks of c1491 and c1598 primers were mixed and boiled for 10 min before being allowed to cool slowly. Dilutions of annealed primers were ligated into pcDNA3.1+neo that had been cut with *EcoR*I and *Xba*I restriction enzymes and SAP treated. Following transformation into DH5α cells plasmid DNA was purified and sequenced to verify the insertion of the annealed oligos into the expression vector.

Step 3: The digested PCR products (produced in step 1) were ligated with His_6 tag-containing-vector (produced in step 2) that had been digested with *Hind*III and *EcoR*I and SAP treated. Following transformation into DH5α cells plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

To create tapasin His_6 expression constructs for B2, B4, B14, B15 and B21 haplotypes, PCR amplifications were performed using tapasin cDNA clones for each haplotype with c1466 and c1597 primers as described in step 1 above. The PCR products were treated as detailed above.

2.5.4 Production of Tapasin V5 expression constructs.

To replace the His_6 tag and duplicated KKSQ motifs with a V5 epitope tag and single KKSQ motif (figure 2.2), PCR amplification was performed using tapasin His_6 constructs for each haplotype as templates and c1466 and c1757 primers. PCR conditions: one µl cDNA was amplified with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 63°C for 30 s and extending at 68°C for 90 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *Xba*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into pcDNA 3.1+neo or pcDNA6 vector that had been digested with *Hind*III and *Xba*I and SAP treated. Following transformation into DH5α cells plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

2.5.5 Production of chicken/human hybrid tapasin expression constructs.

Expression constructs consisting of the first five exons of chicken tapasin and the last three exons of human tapasin were produced in a similar fashion to that used to perform site directed mutagenesis. Exons 6-8 of human tapasin were amplified using c1791 and c1792 primers from the human tapasin in pMCFR expression construct. PCR conditions: one µg DNA was amplified with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 50°C for 30 s and extending at 68°C for 30 s. Exons 1-5 of chicken Tapasin*12 and Tapasin*15 were amplified using c1593 and c1790 primers from the Tapasin*12 or Tapasin*15 in pcDNA6 vector expression constructs. PCR conditions: one µg DNA was amplified with a 96°C two min initial step followed

by 20 cycles of 96°C for 30 s, annealing at 50°C for 30 s and extending at 68°C for 30 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. To fuse the human and chicken portions together, a third amplification was performed using c1593 and c1792 primers. PCR conditions: one µl of human tapasin and one µl of chicken Tapasin*12 or Tapasin*15 PCR products were amplified with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 50°C for 30 s and extending at 68°C for 30 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *Xba*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into pcDNA3.1+neo vector that had been digested with *Hind*III and *Xba*I and SAP treated. Following transformation into DH5α cells, plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

2.5.6 Production of chicken β_2m expression construct.

The chicken β_2m gene was amplified from B21 Line N cDNA (made using oligo dT) using c1648 (introduces 5' *Xho*I site and Kozak sequence prior to ATG codon) and c1650 (introduces *Kpn*I site after stop codon) primers. PCR conditions: one µl cDNA was amplified with a 96°C two min initial step followed by five cycles of 96°C for 30 s, annealing at 45°C for 30 s and extending at 68°C for 60 s and 20 cycles of 96°C for 30 s, annealing and extending at 68°C for 60 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Xho*I and *Kpn*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into pCIpac vector that had been digested with *Xho*I and *Kpn*I restriction enzymes and SAP treated. Following transformation into DH5α cells plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

2.5.7 Generation of .220BF2*12 transfectants

.220 cells were transfected with seven µg BF2*12 in pcDNA6. Antibiotic resistant clones were screened by western blot for expression of BF2.

.220BF2*12 clones
107B1
107C: 2,3,7,12

2.5.8 Generation and analysis of .220BF2*12Tapasin*12 or .220BF2*12Tapasin*15 transfectants

.220BF2*12 clone 107C12 was transfected with seven µg V5 epitope tagged Tapasin*12 or Tapasin*15 in pcDNA3.1+neo. Antibiotic resistant clones were screened by western blot for expression of V5 epitope tagged tapasin.

.220BF2*12Tpn*12 clones	.220BF2*12Tpn*15clones
224B: 2,3,5,9,11,13,24,28,32	225B: 11,15,16
224C: 1,2,3,6,9	225C: 7,8,11,12,17,18

Comparison of BF2*12 expressed in presence of Tapasin*12 or Tapasin*15 by SDS PAGE and western blotting studies: (experiment 227)

For each clone lysates were prepared, and volumes equivalent to 10^5 cells or 10^4 cells were loaded into two wells of a 12% SDS PAGE and western blotted.

Comparison of BF2*12 expressed in presence of Tapasin*12 or Tapasin*15 after overnight culture at 37°C or 27°C (experiment 230)

.220, .220BF2*12 clone 107C12, .220BF2*12Tpn*12 clones 224C1 and 224C2 and .220BF2*12Tpn*15 clones 225C7 and 225C8 were sub-cultured the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the medium aspirated. The cell pellets were resuspended in PBSa and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in RPMI supplemented with two mg/ml BSA (R-BSA) at $5x10^5$ cells/ml (except for clone 225C7, which was resuspended at $2.75x10^5$/ml). For each clone 400 µl ($2x10^5$ cells, apart from 225C7 where $1.1x10^5$ cells were used) were plated into individual wells on two 24 well plates. One plate was incubated overnight at 37°C in a CO_2 incubator, the other plate was incubated overnight at 27°C in a CO_2 incubator. The next day the cells were resuspended and transferred to a 96 well plate and flow cytometry was performed following staining with F21-2 as primary antibody.

2.5.9 Generation and analysis of .220hTpnBF2*12 transfectants

.220hTpn cells were transfected with seven µg BF2*12 in pcDNA6. Antibiotic resistant clones were screened by western blot for expression of BF2.

.220hTpnBF2*12 clones
108C: 12,34

Comparison of BF2*12 expressed in presence or absence of human tapasin by exogenous peptide/β2m stabilisation assays (experiment 129)

.220BF2*12 clone 107B1 and .220hTpnBF2*12 clone 108C12 were sub-cultured into fresh medium the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the medium aspirated. The cell pellets were resuspended in pre-warmed R-BSA and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in R-BSA at 10^6 cells/ml. For each clone 100 µl volumes (10^5 cells) were plated into the required numbers of wells of a 96 well plate.

For both clones a total of nine peptide/β_2m conditions were used, each in triplicate. To each well of the 96 well plate 100 µl of the appropriate peptide/β_2m solution (all diluted in R-BSA) was added to make the following final concentrations:

- Medium only
- One µM chicken β_2m*

- One μM human β_2m†
- One μM chicken β_2m* and 5,10 or 25 μM peptide B12-6 (LPACVLEV)
- One μM human β_2m† and 5,10 or 25 μM peptide B12-6 (LPACVLEV)

* Chicken β_2m was produced and purified previously: briefly an expression vector encoding the mature protein of chicken β_2m was expressed in BL21 bacteria. Following harvesting and sonication of the bacteria, inclusion bodies were isolated and solubilised. The resulting β_2m containing solution was diluted into an excess of cold PBSa and stirred for two days at 4°C. The solution was then concentrated and purified by ion exchange chromatography.

† Human β_2m was purchased from Europa Bioproducts and solubilised in PBSa.

The plate was then incubated at 37°C in a CO_2 incubator for 130 minutes. The cells were then stained with F21-2 antibody and analysed by flow cytometry.

Comparison of BF2*12 expressed in presence or absence of human tapasin after overnight culture at 37°C or 28°C

.220BF2*12 and .220hTpnBF2*12 clones were sub-cultured on the morning of the experiment. Seven hours later one ml aliquots of each clone was plated into two 24 well plates. One plate was incubated overnight at 37°C in a CO_2 incubator, the second plate was incubated overnight at 28°C in a non-CO_2 regulated incubator. The following day the cells were fixed in 1% PFA (diluted in FACS wash) before being stained with F21-2 antibody and analysed by flow cytometry.

2.5.10 Generation and analysis of .220BF2*12chβ_2m transfectants

.220BF2*12 clone 107C12 was transfected with seven μg chβ_2m in pCIpac. Antibiotic resistant clones were screened by western blot for expression of BF2.

.220BF2*12chβ_2m clones
.220BF2*12chβ_2m^{low} 235.1, .220BF2*12chβ_2m^{high} 235.2

Flow cytometry analysis of .220BF2*12chβ2m transfectants (experiment 235)

For the following clones: .220 cells, .220BF2*12 clone 107C12 and .220BF2*12chβ_2m clones 235.1 and 235.2, 5x10^5 cells were stained with TRT1, F21-21 or F21-2 antibodies before flow cytometry was performed.

2.5.11 Generation and analysis of .220BF2*12chβ_2mTapasin*12 transfectants

.220BF2*12chβ_2m^{low} clone 235.1 and .220BF2*12chβ_2m^{high} 235.2 were transfected with seven μg Tapasin*12 in pcDNA3.1+neo. Antibiotic resistant clones were screened by western blot for expression of V5 epitope tagged tapasin.

.220BF2*12chβ_2m^{low}Tapasin*12 clones	.220BF2*12chβ_2m^{high}Tapasin*12 clones
238.1, 238.2, 238.3, 238.4	239.1, 239.3, 239.5, 239.7

Analysis of .220BF2*12chβ2mTapasin*12 transfectants by flow cytometry and western blot analysis (experiment 243)

Flow cytometry studies: For the following clones: .220, .220BF2*12 clone 107C12, .220BF2*12Tapasin*12 clone 224C1 and .220BF2*12chβ_2m^{low}Tapasin*12 clones 238.1, 238.2, 238.4 and .220BF2*12chβ_2m^{high}Tapasin*12 clones 239.1, 239.3, 239.5, $5x10^5$ cells were stained with TRT1, F21-21 or F21-2 antibodies before flow cytometry was performed.

Western blotting studies: For each clone lysates were prepared from $5x10^5$ cells and were separated by 12% SDS PAGE and western blotted.

Analysis of .220BF2*12 transfectants by exogenous peptide MHC class I stabilisation assay (experiment 244)

.220, .220BF2*12 clone 107C12, .220BF2*12Tapasin*12 clone 224C1 and .220BF2*12chβ_2m^{high}Tapasin*12 clones 239.1 and 239.7 were sub-cultured into fresh medium the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the medium aspirated. The cell pellets were resuspended in pre-warmed R-BSA and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in R-BSA at $5x10^5$ cells/ml. For each clone 300 µl of cells ($1.5x10^5$ cells) were plated into the required numbers of wells of a 24 well plate and cultured overnight at 37°C in a CO_2 regulated incubator. The following day 100 µl of either R-BSA or a final concentration of 40 µM FANYIDKV peptide (diluted in R-BSA) was added to the appropriate wells and cultured for one hour at 37°C in a CO_2 regulated incubator. The cells were then resuspended, divided into two equal aliquots and transferred into a 96 well plate. The cells were stained with F21-21 or F21-2 antibodies before flow cytometry was performed.

Analysis of .220BF2*12 transfectants by culture with brefeldin A (experiment 247)

.220, .220BF2*12 clone 107C12, .220BF2*12Tpn*12 clone 224C1, .220BF2*12chβ_2m^{high} clone 235.2, .220BF2*12chβ_2m^{high}Tpn*12 clone 239.1 and .220BF2*15 cells were sub-cultured the day prior to the experiment. On the day of the experiment the cells were resuspended by gentle pipetting and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in R10 at $5x10^5$ cells/ml (except clones 235.2 and 239.1, where lower numbers of cells were available). For each clone one ml ($5x10^5$ cells, except for 235.2 where $1.6x10^5$, and for 239.1 where $3.6x10^4$ were used) was aliquotted into 12 wells of a 24 well plate. The cells were incubated at 37°C for 30 minutes in a CO_2 regulated incubator.

For each clone 10 µg of brefeldin A (10 µl of a one mg/ml stock, dissolved in methanol) was added to the first three wells of the 24 well plate and the plate returned to culture at 37°C in a CO_2 regulated incubator (these samples formed the t = 6 hour time point). This process was repeated for the next three wells two hours later (these samples formed the t = 4 hour time point), and repeated a further two hours later (these samples formed the t = 2 hour time point), and repeated for a final

time a further two hours later (these samples formed the t = 0 hour time point). The plate was then placed on ice. The cells were then resuspended and transferred to 15 ml tubes containing two ml cold FACS wash. The cells were centrifuged at 320 g and resuspended in 200 µl of FACS wash supplemented with 2% PFA. The cells were transferred to a 96 well plate and stored overnight on ice. The following day the cells were stained with F21-2 antibody before flow cytometry was performed.

2.5.12 Generation of .220BF2*15 transfectants

.220 cells were transfected with seven µg BF2*15 in pcDNA3.1+neo. Multiple antibiotic resistant wells grew only on the "A" dilution plates (plated at 10^5 cells/ml). These wells were screened initially by western blot for expression of BF2. Flow cytometry analysis of antibiotic resistant and BF2 expressing clones revealed that none of the clones were clonal cell populations, therefore these transfectants were cloned by limiting dilution analysis. Briefly .220 cells were used as feeder cells and were produced as follows: the .220 cells were washed in PBSa and then incubated on ice before being γ-irradiated by exposure to 3,000 rad. The irradiated cells were washed twice in R10 medium supplemented with 2 mg/ml G418, before 100 µl of irradiated .220 cells (at a concentration of $5x10^4$ cells/ml) were plated in 96 well plates. Selected antibiotic resistant and BF2 expressing cells were centrifuged at 320 g and the supernatant aspirated. The cell pellet was resuspended in cold PBSa and the cells were counted. The cells were diluted in R10 medium supplemented with G418 2mg/ml to obtain concentrations of three cells/ml and 100 µl volumes (equivalent to 0.3 cells/well) were added to the irradiated .220 feeder cells. The plates were cultured at 37°C in a CO_2 regulated incubator with the medium being replenished every week until clones had grown. Growing clones were transferred to 24 well plates and were screened by western blot for expression of BF2 and by flow cytometry. Five clonal BF2 expressing clones were generated as detailed below. Flow cytometry expression confirmed all expressed similar levels of BF2 at steady state after F21-2 staining (data not shown), therefore these transfectants will be labelled .220BF2*15 hereafter.

.220BF2*15 clones
.220BF2*15: 1,2,4,5,7

2.5.13 Generation and analysis of .220BF2*15Tapasin*15 clone

.220BF2*15 clone 7 was transfected with seven µg Tapasin*15 in pcDNA6. Antibiotic resistant clones were screened by western blot for expression of V5 epitope tagged tapasin. Only one clone was generated which was labelled .220BF2*15Tapasin*15.

Analysis of BF2*15 expressed in presence or absence of Tapasin*15 by flow cytometry and western blot analysis (experiment 237)

Flow cytometry studies: For each clone (.220, .220BF2*12 clone 107C12 and .220BF2*12chβ_2m clone 235.1, .220BF2*15 and .220BF2*15Tpn*15) $4x10^5$ cells were stained for flow cytometry with TRT1, F21-21 or F21-2 antibodies, before flow cytometry was performed.

Western blotting studies: For each clone lysates were prepared from $2x10^5$ cells and were separated by 12% SDS PAGE and western blotted.

2.5.14 Screening of chicken/human hybrid tapasin transfectants

.220BF2*12 clone 107C12 was transfected with seven μg of linearised chicken/human hybrid tapasin (Tapasin*12 or Tapasin*15) in pcDNA3.1+neo vector. Antibiotic resistant clones were screened by western blot for expression of hybrid tapasin using rabbit polyclonal antisera specific for the C-terminus of human tapasin. Two antibodies were used for screening, firstly: CSA-625 (Stressgen Bioreagents) which was raised against amino acids 430-448 of human tapasin; secondly, RbαSTC (Tan et al., 2002) (a kind gift from Dr Frank Momburg, German Cancer Research Centre, Heidelberg, Germany), which was raised against amino acid residues 418-428 of human tapasin. No hybrid tapasin expressing clones were identified.

2.6 Specific methods for Chapter five

2.6.1 Production of N- and C- terminally tagged BF2 expression constructs

To create BF2*4 and BF2*15 expression constructs with N-terminal Flag and C-terminal myc epitope tags (referred to as Flag-BF2*4 or Flag-BF2*15) a three step strategy was employed (figure 2.4).

Step 1

Full length BF2*4 and BF2*15 cDNA was amplified using c1622 and c1623 primers as described in specific methods for chapter four.

Step 2

To introduce the C-terminal myc epitope (EQKLISEENL) PCR was performed to replace the stop codon with an *EcoR*I site at the C-terminus. PCR conditions: one µl cDNA was amplified using c1622 (B15 and B4) and c1678 (B15) or c1700 (B4) primers with a 96°C two min initial step followed by 15 cycles of 96°C for 30 s, annealing at 55°C for 30 s and extending at 68°C for 90 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *EcoR*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. To create the myc epitope tag complete with 5' *EcoR*I site, and 3' stop codon and *Xba*I site, two oligos were annealed together (as below).

```
C1676 5' aattc gaa caa aaa ctc atc tca gaa gag gat ctg aat taa t     3'
C1677 3'     g ctt gtt ttt gag tag agt ctt ctc cta gac tta att agatc 5'
AA seq         E   Q   K   L   I   S   E   E   D   L     stop
```

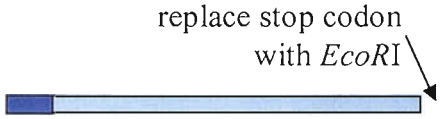
Fifty µl of 100 µM stocks of c1676 and c1677 primers were mixed, boiled for 10 min before being allowed to cool slowly. The annealed primers and digested PCR products were ligated into pcDNA3.1+neo that had been cut with *Hind*III and *Xba*I restriction enzymes and SAP treated. Following transformation into DH5α cells, plasmid DNA was purified and sequenced.

Step 3

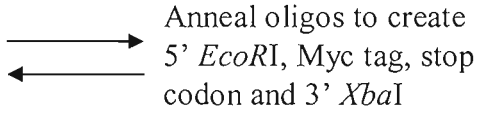
To introduce the Flag epitope tag (DYKDDDDK) between the exon 1 and 2 boundary PCR was performed essentially as described (Fulton et al., 1995). PCR conditions: one µl BF2*12 cDNA was amplified using c1622 and c1708 primers. PCR conditions: one µl (~100 ng) BF2*12 cDNA was amplified with a 96°C two min initial step followed by three cycles of 96°C for 30 s, annealing at 45°C for 30 s and extending at 68°C for 30 s followed by 20 cycles of 96°C for 30 s, annealing at 57°C for 30 s and extending at 68°C for 30 s. PCR products were separated on agarose gels and bands of correct sized excised and purified. Purified PCR products were digested with *Hind*III and *Sac*I (which cuts at the exon 1 and 2 boundary) restriction enzymes before DNA was precipitated and resuspended in TE buffer. Digested PCR products were ligated into pIST vector that had been digested with *Hind*III and *Sac*I and SAP treated. Following transformation into DH5α cells, plasmid DNA was purified and sequenced.

Step 1

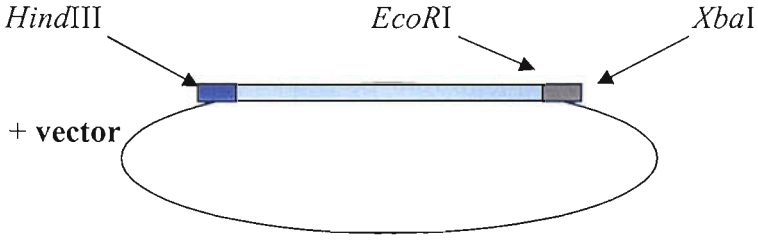
Step 2A



Step 2B



Combine 2A + 2B + vector



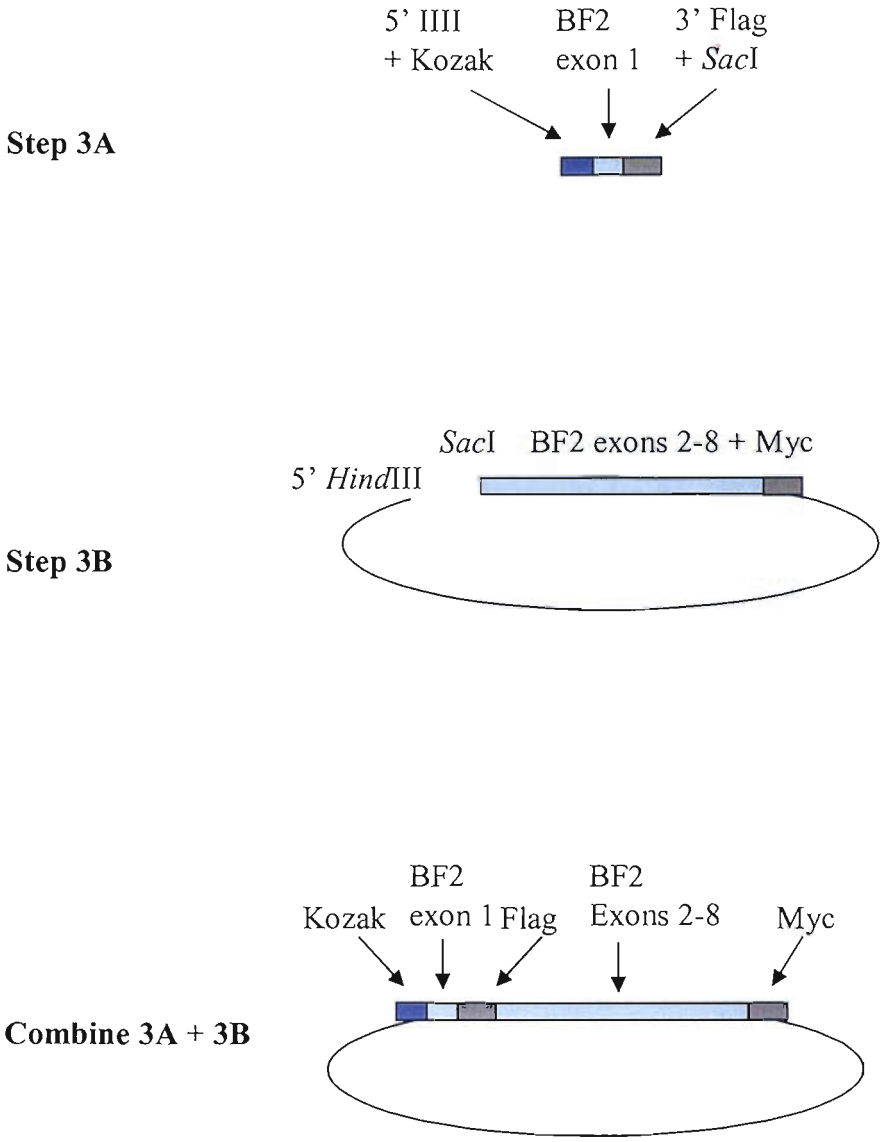


Figure 2.4 Production of Flag-Myc-BF2 constructs.

Step 1: BF2 cDNA (exons 1-8) was amplified introducing: 5' *Hind*III and Kozak sequence, and 3' *Xba*I restriction enzyme sites.
Step 2A: PCR was used to replace the stop codon with an *EcoR*I restriction enzyme site.
Step 2B: Two oligos were annealed to create a C-terminal myc epitope tag complete stop codon and 5' *EcoR*I and 3' *Xba*I sites. The annealed oligos and PCR product were ligated into vector cut with *Hind*III and *Xba*I enzymes.
Step 3A: Introduce Flag epitope tag between exons 1 and 2 by PCR. Digest exon 1 and Flag tag fragment with *Hind*III and *Sac*I enzymes.
Step 3B: Remove *Hind*III–*Sa* I fragment from construct made in step 2, replace with *Hind*III-*Sac*I fragment produced in Step 3A.

The exon 1 and Flag tag construct was excised from pIST vector by digestion with *Hind*III and *Sac*I and then separated by agarose gel electrophoresis. Bands of the correct size were excised and purified. The BF2 exon 1 and Flag tag construct was ligated along with *Sac*I and *Xba*I digested myc tagged BF2 constructs (produced in step 2) into pcDNA3.1+ vector that had been digested with *Hind*III and *Xba*I and SAP treated. Following transformation into DH5α cells plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

2.6.2 Flow cytometry studies and western blot comparison of Flag-BF2 clones
(experiment 180)

Cells were sub-cultured the day prior to the experiment. On the day of the experiment the cells were resuspended by gentle pipetting and counted. $3x10^5$ cells were taken for flow cytometry staining and $1x10^5$ cells taken for western blotting.

Flow cytometry: For every clone three samples each of 10^5 cells were transferred to wells of a 96 well U-shaped plate, and fixed in 200 µl 1% Paraformaldehyde (PFA) diluted in FACS wash at room temperature for 20 minutes before the cells were stained with either Flag or F21-21 antibodies, or left unstained before flow cytometry was performed.

Western blotting studies: For each clone lysates were prepared, and a volume equivalent to 10^4 cells was separated by 10% SDS PAGE and western blotted.

2.6.3 Surface versus intracellular flow cytometry (experiment 188)

Cells were sub-cultured on the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the medium aspirated. The cell pellets were resuspended in FACS wash and counted. For each clone 10^6 cells were taken and divided into two wells on separate 96 well plates. The cells were fixed in 3% paraformaldehyde (PFA) diluted in FACS wash for 20 minutes at room temperature, and then washed once with FACS wash.

Intracellular staining: cells were resuspended in FACS wash + 0.5% (w/v) saponin (FACS wash+Sap) for 30 minutes at room temperature. Cells were washed three times with FACS wash+Sap, and stained with either F21-21 or Flag M2 antibodies in FACS wash+Sap for 30 minutes on ice. Cells were washed three times in FACS wash+Sap before incubation with five µg/ml of FITC conjugated polyclonal goat anti-mouse Ig $F(ab')_2$ (DAKO) in FACS wash+Sap for 30 minutes on ice in the dark. Cells were washed twice with FACS wash+Sap before resuspending in FACS wash.

Surface staining: all steps were as above with the exception that saponin was not present in the buffers.

2.6.4 Exogenous peptide MHC class I stabilisation assay overnight at 28°C or 41°C (experiment 178)

Cells were sub-cultured the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the media aspirated. The cell pellets were resuspended in pre-warmed R-BSA and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in R-BSA at 10^6 cells/ml. For each clone 100 µl of cells (10^5 cells) were plated into six wells of two 96 well plates. The plates were then incubated at 41°C in a CO_2 incubator for four hours.

For both plates and for each clone, three of the six wells received 100 µl R-BSA and the remaining three wells received peptide solution (either KRLIGKRY for B15 clones or ADVEEYEE for B4 clones, peptides diluted in R-BSA, final peptide concentration of 20 µM). The cells were then incubated overnight either at 41°C in a CO_2 incubator or at 28°C in a non-CO_2 regulated incubator. The following day the cells were divided into two equal samples. Flow cytometry staining was performed using either F21-21 or Flag M2 antibodies as primary antibodies.

2.6.5 Comparison of three B4-specific peptides by exogenous peptide MHC class I stabilisation assay overnight at 41°C (experiments 216 and 219)

Cells were sub-cultured the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the medium aspirated. The cell pellets were resuspended in PBSa and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in pre-warmed R-BSA at 10^6 cells/ml. For each clone 200 µl of cells ($2x10^5$ cells) were plated into five wells of a 24 well plate. Into each well 600 µl of appropriate peptide solution was added (diluent, ADVEEYEE, LDVEDVKF, IDWFDGKE or KRLIGKRY, with each peptide diluted in R-BSA) to make a final peptide concentration of 30 µM. The cells were then incubated overnight at 41°C in a CO_2 incubator. The following day the cells were divided into three equal samples and flow cytometry staining was performed using either F21-21, Flag M2 or F21-2 antibodies as primary antibodies.

2.6.6 Generation of Flag-BF2*4Tapasin*4 transfectant

Flag-BF2*4 clone 2 was transfected with seven µg Tapasin*4 in pcDNA6. Antibiotic resistant clones were screened by western blot for expression of V5 epitope tagged tapasin. Only one clone was generated which was labelled Flag-BF2*4Tapasin*4.

2.6.7 Brefeldin A treatments (experiment 209)

Cells were sub-cultured on the day prior to the experiment. On the day of the experiment the cells were centrifuged at 320 g and the media aspirated. The cell pellets were resuspended in pre-warmed R-BSA and the cells were counted. The required number of cells were centrifuged at 320 g and resuspended in R-BSA at 10^6 cells/ml. For each clone one ml (10^6 cells) was pipetted into one

well of a 24 well plate. The plate was incubated at 41°C for 15 minutes in a CO_2 regulated incubator.

The cells were resuspended by gentle pipetting and two aliquots of 150 µl were withdrawn from each well and added to tubes containing cold PBSa and kept on ice until the last time point was collected. To the remaining cells brefeldin A was added to a final concentration of 10 µg/ml, and the plate returned to the incubator. One aliquot was taken from each well every two hours for eight hours as described above. After the last timepoints were collected, the cells were centrifuged at 320 g and the supernatant removed. Flow cytometry staining was performed with Flag M2 as primary antibody. The second aliquot taken at the first timepoint was stained with TRT1 antibody.

2.6.8 Production of peptide minigenes

Peptide minigenes comprising the BF2 signal sequence (the first 21 amino acids) preceding nucleotides encoding either the ADVEEYEE or KRLIGKRY peptide sequences were produced by PCR (figure 2.5). PCR conditions: one µl (~100 ng) BF2*12 cDNA in pcDNA3.1+neo was amplified using c1756 and c1753 (ADVEEYEE) or c1754 (KRLIGKRY) primers with a 96°C two min initial step followed by 30 cycles of 96°C for 30 s, annealing at 55°C for 30 s and extending at 68°C for 30 s. The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *BamH*I and *Xba*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into the CMVbipepΔNGFR (Duch et al., 1999; Tolstrup et al., 2001) or pcDNA6 vectors that had been digested with *BamH*I and *Xba*I restriction enzymes and SAP treated. Following transformation into DH5α cells, plasmid DNA isolated was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

2.6.9 Cell sorting of Flag-BF2*4 cells transfected with peptide minigenes

Three days after Flag-BF2*4 clone 4-3 was transfected with ADVEEYEE or KRLIGKRY: CMVbipepΔNGFR expression vectors, cell sorting was performed using anti-P75 antibody coated CELLection Pan Mouse IgG magnetic beads (Dynal) following the manufacturer's conditions. Briefly, prior to cell sorting the beads were washed and coated with 0.5 µg anti-P75 antibody per 10^7 beads before washing again. Cell sorting was performed using 10^7 beads/ml of cell suspension and a ratio of 5 beads/target cell. Following washing steps, the positively selected cells were released from the beads by incubation with DNase releasing buffer. The released cells were placed in 6 well plates and cultured at 41°C in a CO_2 regulated incubator. Flow cytometry analysis was performed using anti-P75 antibody on samples representing the positively, negatively and unsorted cell fractions.

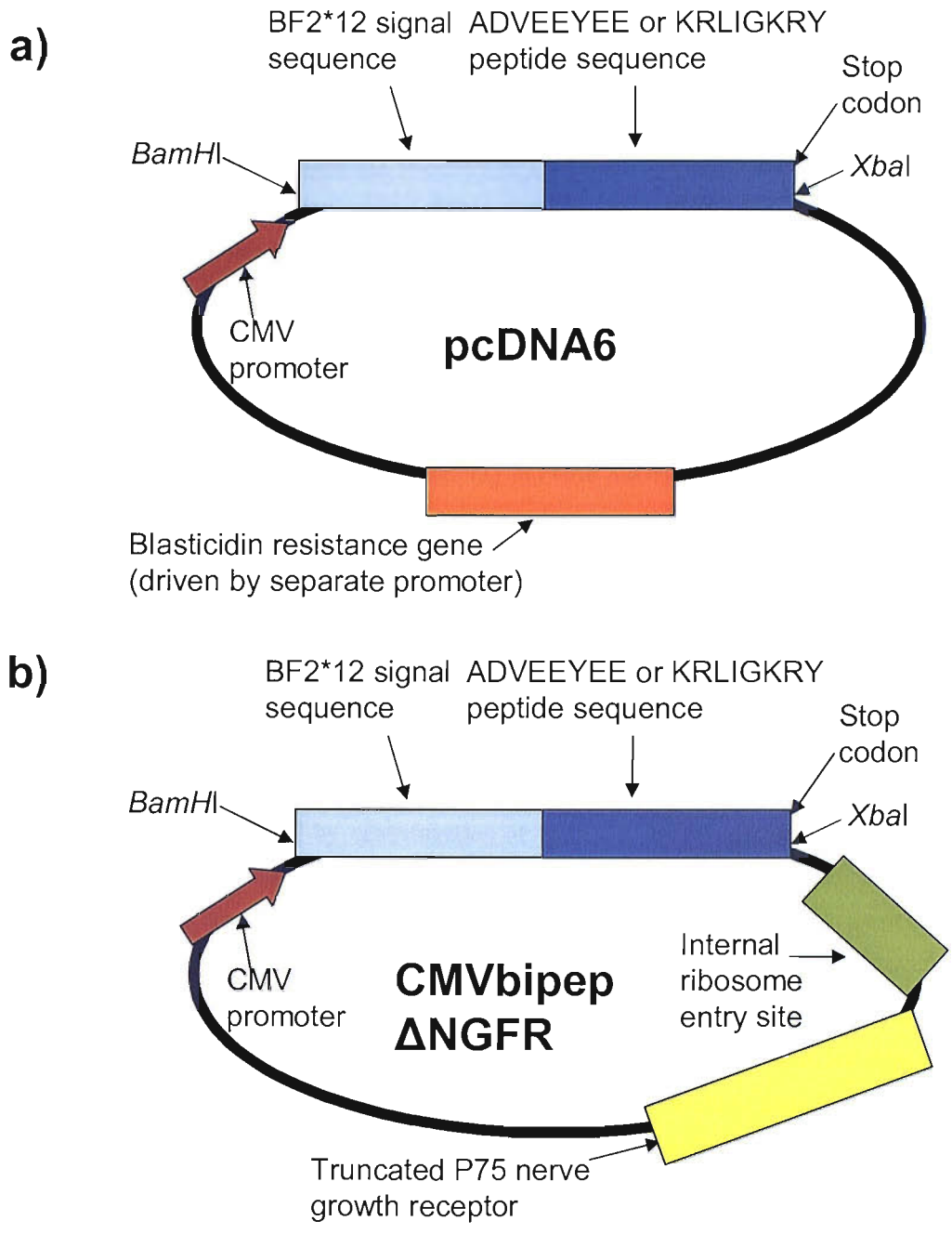


Figure 2.5 Cartoon schematic depicting peptide minigene expression constructs

Expression constructs comprised the signal sequence of BF2*12 (the N-terminal 21 amino acids) immediately preceding nucleotides encoding either ADVEEYEE or KRLIGKRY peptide sequences. A stop codon immediately followed the last amino acid of the peptide sequence. The peptide minigene were cloned into the expression vectors via *BamH*I and *Xba*I restriction enzyme sites.

a) pcDNA6 expression construct conferring blasticidin antibiotic resistance selection.

b) CMVbipepΔNGFR vector conferring the cell surface expressed truncated human P75 nerve growth receptor, whose expression is driven off the same promoter via an internal ribosome entry site.

2.6.10 ^{35}S pulse/chases, myc immunoprecipitations and endoglycosidase H digestions (experiment 197)

TG15 cells, Flag-BF2*15 clone 15-8 and Flag-BF2*4 clone 4-3 were sub-cultured the day prior to the experiment. On the day of experiment $5x10^6$ cells were taken and centrifuged at 320 g for five minutes. The supernatant was completely aspirated and the cell pellet resuspended in one ml pre-warmed cysteine-free and methionine-free RPMI supplemented with 10% dialysed FCS and antibiotics, and the cells were incubated at 41°C in a 5% CO_2 incubator for 30 minutes. Cells were pulse labelled by the addition of 100 µCi (3.7 MBq) of ^{35}S Promix (Amersham) for 15 minutes. The cells were then centrifuged at 380 g at room temperature for one minute and the supernatant completely removed. The chase was initiated by the addition of one ml pre-warmed and CO_2-equilibrated RPMI supplemented with 10% FCS and two mM cysteine and two mM methionine. Aliquots were taken immediately after resuspending in chase media and after a further two, four, eight or 22 hours. The aliquots were added to an eppendorf containing ice cold PBSa and kept on ice until the last time point was collected. The aliquots were then centrifuged at 380 g for three minutes at 4°C, and the supernatant completely removed, and the cell pellets stored at -20°C.

The cell pellets were thawed and lysed by the addition of 100 µl of ice cold RIPA buffer supplemented with 10 mM iodoacetamide and 4 mM 4-(2-Aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF) (Roche), and were vortexed briefly and incubated on ice for 15 minutes. The nuclear debris was removed by centrifugation at 16,000 g for 10 minutes at 4°C, and the supernatant transferred to a new eppendorf tube. Samples were precleared by the addition of 50 µl 50% sepharose 4B and rotated at 4°C for one hour. The precleared samples were centrifuged at 16,000 g for 10 minutes at 4°C and the precleared supernatant transferred to a new eppendorf tube. TG15 precleared supernatant was incubated with 15 µg F21-2 antibody and 35 µl 50% Sepharose 4B (suspended in PBSa). Precleared supernatants of Flag clones 15-8 and 4-3 were incubated with 50 µl anti-myc agarose (which had been washed five times in 100 µl aliquots each in one ml PBSa with centrifugation at 2,375 g for one minute between washes, before finally resuspending at 50% in RIPA buffer). Samples were rotated at 4°C for four hours. The tubes were then centrifuged at 2,375 g for one minute at 4°C and the supernatant discarded. The pellet was washed five times in 500 µl ice cold N-P40 lysis buffer, with centrifugation at 2,375 g for one minute at 4°C between washes. The immunoprecipitated proteins were eluted by the addition of 20 µl 100 mM sodium acetate pH5.4, 0.02% (w/v) SDS, 100 mM β-ME, 10 mM iodoacetamide and 4 mM AEBSF and heating at 85°C for five minutes. The tubes were centrifuged at 16,000 g and the supernatant removed and split equally to two eppendorf tubes. One µl (five mU) of endoglycosidase H (Roche) was added to one of these tubes, before all tubes were incubated at 37°C overnight. The samples were separated by electrophoresis on 10% SDS PAGE gels. Afterwards the gels were fixed in 40% methanol and 7.5% acetic acid for 30 minutes and then soaked in 1 M sodium salicylate + 20 ml/l glycerol for 30 minutes. The gels were then dried under vacuum and were subsequently exposed to film (MR1, Kodak) and stored at -80°C.

2.7 Specific methods for Chapter six

2.7.1 Production of C-terminal myc tagged BF2 expression constructs

BF2 expression constructs with C-terminal myc epitope tags were produced for B15, B19v1 and B4 haplotypes as described in specific methods for chapter five (steps 1 and 2), and are referred to as BF2myc.

2.7.2 Production of BF2myc*15 and BF2myc*19v1 mutant expression constructs

BF2myc point mutations were introduced by site directed mutagenesis, using C-terminal myc tagged BF2*15 or BF2myc*19v1 (in pcDNA3.1+neo vector) as template DNA, and the primers listed below.

Step 1 PCR conditions: ~100 ng DNA was amplified with the indicated primers with a 96°C two min initial step followed by 25 cycles of 96°C for 30 s, annealing at 52°C for 30 s and extending at 68°C for 60 s.

Step 2 PCR conditions: one µl of each purified PCR products was used as template DNA and was amplified with c1593 and c1594 primers with a 96°C two min initial step followed by 25 cycles of 96°C for 30 s, annealing at 52°C for 30 s and extending at 68°C for 60 s.

The PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Hind*III and *Xba*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into empty pcDNA3.1+neo vector that had been digested with *Hind*III and *Xba*I and SAP treated. Following transformation into DH5α cells, plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct sequence constructs, Endofree Maxipreps were prepared.

BF2myc*15 mutants	**PCR step**	**Forward primer**	**Reverse primer**
D126G	1A	C1593	C1826
	1B	C1594	C1825
D126G+Q220R (using G126D as template).	1A	C1593	C1828
	1B	C1594	C1827
S111R+D113Y and **S111R+D113Y+D126G+Q220R** (using D126G+Q220R DNA as template).	1A	C1593	C1830
	1B	C1594	C1829
BF2myc*19v1 mutants	**PCR step**	**Forward primer**	**Reverse primer**
G126D	1A	C1593	C1802
	1B	C1594	C1801
R220Q and **G126D+R220Q** (using G126D as template).	1A	C1593	C1804
	1B	C1594	C1803
R111S+Y113D	1A	C1593	C1806
	1B	C1594	C1805

2.7.3 ^{35}S pulse/chases, myc immunoprecipitations and endoglycosidase H digestions

Protocol for experiments 231, 233, 236, 240, 241 and 245

Experiments were performed as detailed for experiment 197, with the following exceptions:

- The cells were incubated in cysteine- and methionine-free medium for 45-60 minutes.
- Cells were pulse labelled with 100-200 μCi (7.4 MBq) of ^{35}S Promix for five minutes at 41°C.
- After the chase was initiated, aliquots were removed immediately and every 30 minutes thereafter for 210 minutes (120 minutes for experiment 231).
- After the last timepoint the cell pellets were immediately lysed by the addition of 100 μl ice cold RIPA buffer supplemented with 10 mM iodoacetamide and four mM AEBSF or Complete protease inhibitor cocktail (Roche).
- The clarified lysates were precleared in eppendorf tubes containing 50 μl of 50% protein G sepharose slurry (diluted in lysis buffer). The tubes were precleared by rotation at 4°C overnight.
- Anti-myc agarose was washed three times in PBSa and then twice in RIPA buffer supplemented with one mg/ml BSA, with centrifugation at 380 g between washes, and finally resuspended at 50% in RIPA buffer supplemented with one mg/ml BSA.
- The precleared supernatants were transferred to tubes containing anti-myc agarose (40-110 μl 50% slurry per immunoprecipitate), and rotated at 4°C for two hours.
- The immunoprecipitates were washed three-five times in 500 μl lysis buffer.
- The immunoprecipitated proteins were eluted by the addition of 15-25 μl 100 mM sodium acetate pH 5.4, 0.02% (w/v) SDS, 100 mM β-ME, 10 mM iodoacetamide and four mM AEBSF and heating at 85°C for five minutes.
- After overnight endo H digestion, the samples were separated by electrophoresis on 10 or 12% SDS PAGE gels. After fixation, the gels were soaked in Amplify (Amersham) for 30 minutes.

Protocol for experiments 277, 282 and 285

These experiments were performed as above, with the following exceptions:

Experiment 277:

- Cells were labelled with 100 μCi Promix for 30 minutes.
- The chase was initiated by diluting the labeling mix in ten-fold excess of pre-warmed non-radioactive media supplemented with two mM cysteine and two mM methionine. The chased lasted for 240 minutes.
- After washing, each immunoprecipitate was split: with one third being taken for mock endo H digestion, and two thirds being digested with endo H for three hours at 37°C.

Experiments 282 and 285:

- Cells were labeled with 140 µCi (experiment 285) or 200 µCi (experiment 282) Promix for a total of 20 minutes, which constituted 17 minutes at 41°C in a CO_2 regulated incubator, followed by three minutes centrifugation at 320 g at room temperature.
- The chase was initiated by removing labelling media and replacing with 10 ml pre-warmed non-radioactive media supplemented with two mM cysteine and two mM methionine.
- Two aliquots were removed at the onset of the chase (t = 0 minutes), with one sample subsequently being mock digested, whilst all other samples were endo H digested at 37°C for 90 minutes.
- Cells were lysed in 250 µl lysis buffer, and precleared at 4°C for 90 minutes and incubated with myc agarose overnight with rotation at 4°C.

2.7.4 Thermal-stability assays

Western blot thermal-stability assay (experiment 281)

TG15 BF2myc*15 clone C3 and TG15 BF2myc*19v1 clone C1 were sub-cultured one day prior to the experiment. The cells were centrifuged at 320 g for five minutes at 4°C before the supernatant was aspirated, and the cell pellets were resuspended in cold PBSa and counted. The cells were centrifuged again, and the cell pellets were lysed in RIPA lysis buffer at a concentration of 10^5 cells/µl for 15 minutes on ice. The lysate was clarified by centrifugation at 16,000 g for 10 minutes, 4°C. Fifty µl volumes ($5x10^6$ cells) were aliquotted into thin walled eppendorf tubes that were heated for one hour in a PCR block with the lid temperature being 20°C greater than the following temperatures (°C): 37, 48, 50, 52, 54, 56, 60, with one aliquot being kept on ice. The samples were then cooled and precleared by the addition of 30 µl 50% slurry of protein G sepharose (washed and diluted in RIPA buffer) with rotation for 30 minutes at 4°C. The samples were centrifuged at 16,000 g for 10 minutes at 4°C. Precleared supernatants were transferred to new eppendorf tubes containing 100 µl 50% slurry of washed myc agarose, and allowed to rotate at 4°C for 90 minutes. The immunoprecipitates were then washed five times in lysis buffer, including rotation at 4°C for 10 minutes. The immunoprecipitates were eluted by the addition of 30 µl non-reducing LSB, boiled for five minutes and then following centrifugation, separated by SDS PAGE and western blotted. The blots were blocked in 5% milk and then divided according to molecular weight markers. The top half of the blots, containing proteins of higher molecular weights, were incubated with an anti-myc-HRP antibody, whilst the bottom portion of the blots were incubated with rabbit anti-chicken β_2m antiserum, before washing and then incubated with anti rabbit Ig-HRP conjugate.

^{35}S pulse/chase thermal-stability assays (experiments 292)

TG15 BF2myc*15 clone C7 and TG15 BF2myc*19v1 clone C9 were radiolabelled with 100 µCi Promix for a total of 20 minutes, as described for experiment 282, and chased for two hours. Three ml aliquots were removed at the onset, after one hour, or after two hours of culture at 41°C, and lysed in one ml RIPA lysis buffer.

The clarified lysate was divided equally into five aliquots in thin walled eppendorf tubes that were then heated for one hour in a PCR block with the lid temperature being 20°C greater than the following temperatures (°C): 37, 41, 45, 50 with one aliquot being kept on ice. The samples were then cooled and precleared by the addition of 20 µl 50% slurry of protein G sepharose (washed and diluted in RIPA buffer) with rotation for one hour at 4°C. The samples were centrifuged at 16,000 g for 10 minutes at 4°C. Precleared supernatants were transferred to new eppendorf tubes containing 25 µl 50% slurry of myc agarose (previously washed in PBSa and then lysis buffer), and allowed to rotate at 4°C overnight. The immunoprecipitates were then washed three times in lysis buffer, including rotation at 4°C for 10 minutes. The immunoprecipitates were eluted by the addition of 15 µl non-reducing LSB, boiled for five minutes and then following centrifugation, separated by 15% SDS PAGE. The gels were fixed in 40% methanol and 7.5% acetic acid for 30 minutes and then soaked in Amplify for 30 minutes. The gels were then dried under vacuum and were subsequently exposed to phosphor screens.

Experiment 293 was conducted as for experiment 292, with three exceptions:

- The radiolabelling was conducted for 30 min
- The lysates were split into four aliquots that were heated at 41, 45, 50°C or kept on ice
- The immunoprecipitates were washed five times.

Analysis of phosphor images

Following scanning of the phosphor screens, the band volumes were measured for the immunoprecipitated BF2myc heavy chains, and for the co-precipitated β_2m molecules. The "β_2m: heavy chain ratio" was calculated by dividing the measured β_2m band volume by the corresponding heavy chain band volume (to account for differences in amounts of immunoprecipitated heavy chains).

The "Recovery %" was calculated firstly by setting the β_2m: heavy chain ratio determined for each clone following incubation on ice to 100% for each time point. The β_2m: heavy chain ratio that was recorded for the same clone after heat treatment at the same time point was then calculated as a percentage of this value.

2.7.5 Peptide: MHC class I time-dissociation assay (experiment 296)

TG15 BF2myc*15 clone C3 and TG15 BF2myc*19v1 clone C9 were radiolabelled for 30 minutes (otherwise conditions were as described for experiment 282). The radiolabelled cells were lysed in one ml RIPA lysis buffer, and immediately divided into five aliquots. To one aliquot a final concentration of 500 µM KRLIGKRY peptide was added.

The clarified lysates were then immediately incubated with 60 µl myc agarose (without a preclear step) and allowed to rotate at 4°C for a minimum of 20 minutes. Two aliquots, one of which was supplemented with peptide were then removed, quickly washed twice with lysis buffer and the pellets immediately frozen. Aliquots were removed and treated as above, after one or five hours, or after incubation overnight.

The samples then thawed and separated on 15% SDS PAGE gels. Afterwards the gels were fixed in 40% methanol and 7.5% acetic acid for 30 minutes and then soaked in Amplify for 30 minutes. The gels were then dried under vacuum and were subsequently exposed to phosphor screens. The β_2m: heavy chain ratio was calculated as described above.

2.7.6 Peptide stabilisation assay (experiment 295)

TG15 BF2myc*15 clone C7, TG15 BF2myc*19v1 clone C1 and TG15 BF2myc*19v1 G126D&R220Q clone A6 were radiolabelled (as described for experiment 282) with 100 µCi (7.4 MBq) of ^{35}S Promix for 30 minutes. The radiolabelled cells were then divided equally into five eppendorf tubes containing cold PBSa, and centrifuged at 380 g for three minutes at 4°C. The supernatant was discarded, and the cell pellets (equivalent to 10^6 cells) were resuspended in RIPA lysis buffer supplemented with either: 0, 100 nM, 500 nM, 1 µM or 10 µM KRLIGKRY peptide and incubated on ice for 15 min. Following clarification of the lysate, the lysates were precleared by the addition of 20 µl sepharose beads, and allowed to rotate at 4°C for 30 min. Following centrifugation at 16,000 g for five min, the precleared lysate was transferred to a new eppendorf containing 60 µl myc agarose, and allowed to rotate at 4°C for two hours. The immunoprecipitates were washed five times in lysis buffer, and were eluted in 10 µl of 100 mM sodium acetate pH 5.4, 0.02% (w/v) SDS, 100 mM β-ME, 10 mM iodoacetamide and Complete protease inhibitor cocktail and heated at 85°C for five minutes. The tubes were centrifuged at 16,000 g and the supernatant transferred to a new eppendorf tube containing one µl (five mU) of endoglycosidase H, before the tubes were incubated at 37°C for three hours. The samples were separated by electrophoresis on 15% SDS PAGE gels. Afterwards the gels were fixed in 40% methanol and 7.5% acetic acid for 30 minutes and then soaked in Amplify for 30 minutes. The gels were then dried under vacuum and were subsequently exposed to phosphor screens. The β_2m: heavy chain ratio was calculated as described above.

2.7.7 Production of human tapasin "5x" mutant expression construct

Human tapasin "5x" mutant expression constructs (in pMCFR expression vector) were produced by three rounds of site directed mutagenesis using the following conditions and primer pairs:

Round 1: G334T, P336S and Q341R "3x" mutant

Step 1 PCR conditions: ~100 ng of human tapasin in pMCFR vector was amplified with the indicated primers with a 96°C five min initial step followed by 30 cycles of 96°C for 30 s, annealing at 55°C for 30 s and extending at 68°C for 60 s.

Step 2 PCR conditions: one µl of each purified PCR products was used as template DNA and was amplified with c1667 and c1667 primers with the same PCR conditions as above.

Round 2: S189P, G334T, P336S and Q341R "4x" mutant

Step 1 PCR conditions: one µl of the final "3x" product produced in round one was amplified with the indicated primers with a 96°C five min initial step followed by 20 cycles of 96°C for 30 s, annealing at 60°C for 30 s and extending at 68°C for 60 s.

Step 2 PCR conditions: one µl of each purified PCR products was used as template DNA and was amplified with c1667 and c1667 primers with the same PCR conditions as above.

Round 3: S189P, G334T, P336S, Q341R and R379I "5x" mutant

Step 1 PCR conditions: one µl of the final "4x" product produced in round two was amplified with the indicated primers with a 96°C two min initial step followed by 20 cycles of 96°C for 30 s, annealing at 60°C for 30 s and extending at 68°C for 60 s.

Step 2 PCR conditions: one µl of each purified PCR products was used as template DNA and was amplified with c1667 and c1667 primers with the same PCR conditions as above.

Human tapasin mutants	PCR step	Forward primer	Reverse primer
3x: G334T, P336S Q341R	1A	C1667	C1832
	1B	C1831	C1668
4x: S189P, G334T, P336S Q341R	2A	C1667	C1834
	2B	C1833	C1668
5x:S189P, G334T, P336S, Q341R R379I	3A	C1667	C1836
	3B	C1835	C1668

The final PCR products were separated on agarose gels and bands of the correct size were excised and purified. The purified PCR products were digested with *Kpn*I and *Cla*I restriction enzymes before the DNA was ethanol precipitated and resuspended in TE buffer. The digested PCR products were ligated into empty pMCFR vector that had been digested with *Kpn*I and *Cla*I and SAP treated. Following transformation into DH5α cells, plasmid DNA was purified and sequenced. Following retransformation of DH5α with correct constructs, Endofree Maxipreps were prepared.

2.7.8 Flow cytometry and western blot analysis of .220B*4402 tapasin mutant clones

experiment 275

For the following clones (.220B*4402, .220B*4402hTpn, .220B*4402hTpnR48E1, .220B*4402hTpnR48A1, .220B*4402 Tpn 5x clones 1 and 2) $8.5x10^6$ cells were taken and were centrifuged at 380 g for five minutes at 4°C before the supernatant was aspirated. The cells were resuspended at $5x10^6$ cells/ml in cold FACS wash.

Flow cytometry: For each clone $5x10^5$ cells were aliquotted to 12 wells of a 96 well U-shaped plate. The cells were stained in triplicate with either W6/32, anti HLA-B Ab-1 "Joan-1", TT4-A20 or HC10 antibodies, before flow cytometry was performed.

Western blotting studies: For each clone lysates were prepared, and a volume equivalent to either $3.3x10^5$ cells or $3.3x10^4$ cells was separated by electrophoresis on a 12% SDS PAGE gel and western blotted.

List of primers

C469	CTCTTTATCTGGAGCTGAGGTGCT
C1353	CAGCGGAGAGCAATGGCTG
C1466	CAGAAGCTTGCAATGGCTGCGGGGCTG
C1571	TGACCTCCATGGTCATCCAGTGC
C1573	CCATGCCTTGCAGAAATCGCAGTG
C1576	GGTTCGGGTCTCACCCGCCAG
C1578	GGCCCTCGTCCCACCCGATGG
C1593	GGCTAACTAGAGAACCCAC
C1594	CAAACAACAGATGGCTGGC
C1622	GCGGAAGCTTGCGATGGGGCCGTGCGG
C1623	TCGTTCTAGATTAGATGGCGGGGTTGCTCC
C1648	AGCGCTCGAGGGAGCCATGGGGAAGGCGGCGG
C1650	GGGGGTACCTTAGAACTCGGGATCCCA
C1667	TTCTGCGCCGTTACAGATCTC
C1668	TTCACAAAGATCCGGGGATCTG
C1670	GGTGCACTGCTGTTGCTCCAGGGACCGGGGGAAC
C1671	GGTGCACTGCTGTTGGAGCAGGGACCGGGGGAAC
C1672	TTCCCCCGGTCCCTGCTCCAACAGCAGTGCACC
C1673	GGTGCACTGCTGTTGGCACAGGGACCGGGGGAAC
C1674	TTCCCCCGGTCCCTGTGCCAACAGCAGTGCACC
C1675	TTCCCCCGGTCCCTGGAGCAACAGCAGTGCACC
C1678	CTCTGAATTCGATGGCGGGGTTGCTCC
C1700	GTTCGAATTCGATGGTGGGGTTGCTCC
C1708	CAGGGTATGGAGCTCCTTGTCGTCGTCGTCCTTGTAGTCGGCCGCCGCCCCGCACAC
C1753	CATTCTAGATTACTCCTCATATTCTTCAACATCTGCGGCCGCCGCCCCGCACACGGC
C1754	CATTCTAGATTAATACCTCTTACCAATTAGTCTTTTGGCCGCCGCCCCGCACACGGC
C1756	TGGCTAGCGTTTAAACTTGGATCCGCGATGGGG

C1757	CCCTCTAGATTACTGCGATTTCTTCGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAG-GCTTACCGGTTTCCTCTTTGGG
C1790	GGAGGGCCCTGAAAGACCAGCCAGGAGCAGTCGGACGGAGAC
C1791	CGACTGCTCCTGGCTGGTCTTTCAGGGCCCTCCCTTGAG
C1792	TATTCTAGATTACTCTGCTTTCTTCTTTGAATC
C1801	ATTGCCTTCGACAAAGACACGATGACGTTCACT
C1802	AGTGAACGTCATCGTGTCTTTGTCGAAGGCAATG
C1803	AAGGACGGCGCGGTGCAAGGCCAGGACGCCCAG
C1804	CTGGGCGTCCTGGCCTTGCACCGCGCCGTCCTTC
C1805	ACCATCCGGGGGTATAGTCAGGACGCCTACGATGGGAGAG
C1806	TCTCCCATCGTAGGCGTCCTGGCTATACCCCCGGATGGTG
C1825	ATTGCCTTCGACAAAGGAACGATGACGTTCACTG
C1826	AGTGAACGTCATCGTTCCTTTGTCGAAGGCAATG
C1827	AAGGACGGCGCGGTGCGAGGCCAGGACGCCCAG
C1828	CTGGGCGTCCTGGCCTCGCACCGCGCCGTCCTTC
C1829	ACCATCCGGGGGTATCGACAGTATGCCTACGATGGGAGAG
C1830	TCTCCCATCGTAGGCATACTGTCGATACCCCCGGATGGTG
C1831	CTGCCCCTCGGCCTTCCGAGAGCGGCCCCCGCTGCCTGTCCGGAGTTCCCACTC
C1832	GAGTGGGAACTCCGGACAGGCAGCGGGGGCCGCTCTCGGAAGGCCGAGGGGCAG
C1833	AGATGAGGCGGCCTCGGGGGTGGGGGGCATGTAG
C1834	TACATGCCCCCCACCCCCGAGGCCGCCTCATCTC
C1835	AATTCGACAGGCATATATTGCCCCATGCTGCTC
C1836	GAGCAGCATGGGGCAATATATGCCTGTCGAATTC

3. Identification and analysis of chicken tapasin sequence polymorphisms

3.1 Aims

The compact nature and low rate of recombination within the chicken MHC means there is an extremely small chance of recombination separating the class I, TAP and tapasin genes. Therefore it is likely that these genes are kept together in the same haplotype for extended periods of time, which may have allowed these genes to co-evolve. Whilst polymorphisms in chicken class I and TAP genes have previously been documented, the extent of polymorphism in chicken tapasin gene was unknown at the onset of this project.

Specific aims of this chapter are:

- To determine the extent of allelic polymorphism and sequence diversity in the chicken tapasin gene in the seven MHC haplotypes that are available at IAH Compton. The potential implications of any polymorphisms will be discussed.
- To seek evidence whether tapasin is polymorphic in protein sequence, and to compare the tapasin protein sequence with the polymorphic class I and TAP protein sequences to determine whether there is any clear evidence of evolution occurring between these proteins. Efforts will be made to identify amino acids that may mediate such interactions and to determine the functional relevance of predicted sites of protein interaction.

3.1.1 Tapasin DNA sequencing strategy

The chicken tapasin gene was amplified, cloned and sequenced using four separate approaches from each of the nine different lines representing seven MHC haplotypes of chicken available at IAH Compton (figure 3.1b):

1. From cDNA, with a primer that overlaps the start (ATG) codon and a primer following the stop codon.
2. From genomic DNA, using the same primer pair, in order to amplify both the exons and introns.
3. From genomic DNA, 829bp "upstream" of tapasin exon one, a region likely to contain the promoter region of tapasin.
4. From genomic DNA, 747bp "downstream" of tapasin exon eight.

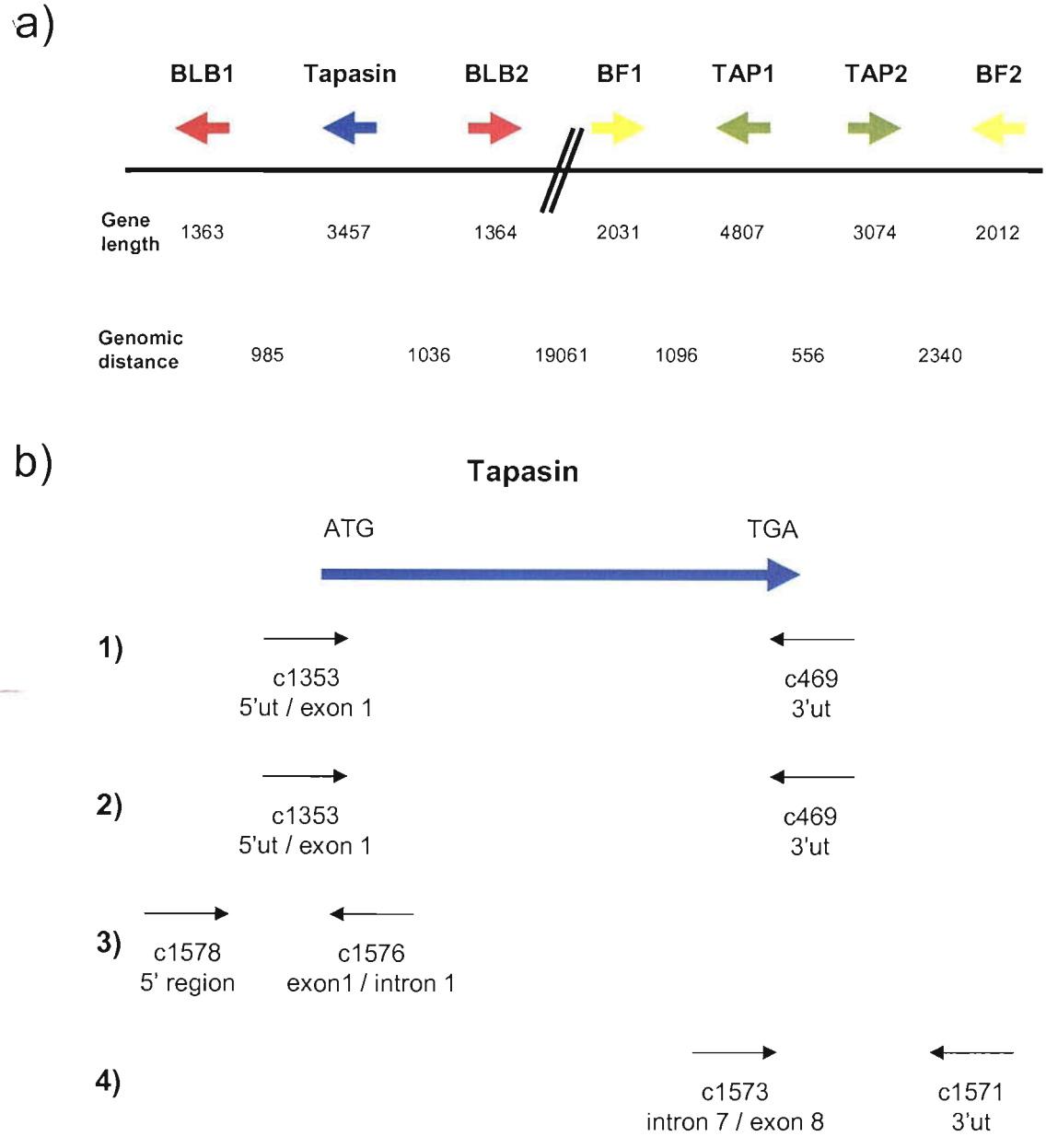


Figure 3.1 Schematic representing gene locations and PCR amplification strategies.

a) Schematic representing the locations of tapasin, class I major (BF2) and minor (BF1), TAP1 and TAP2, class IIB major (BLB2) and minor (BLB1). Not drawn to scale. Gene orientation, length and genomic distances between start and stop codons (in bp) are detailed (taken from AL023516).
b) Strategies used to amplify tapasin (not drawn to scale): 1) from cDNA using c1353 and c469 primers, 2) from genomic DNA using the same primer pair, 3) from genomic DNA using c1576 and c1578 primers, and 4) from genomic DNA using c1571 and c1573 primers.

- For each of the different lines and haplotypes that were sequenced, the genomic DNA derived from a different bird to that from which the cDNA was derived.
- All of the PCR fragments were cloned prior to sequencing.
- To confirm the validity of the results for each haplotype, at least two independent PCR amplifications were performed for each approach, with at least one clone completely sequenced per amplification.
- Any discrepancies that existed between clones were resolved by performing further PCR amplifications using the same template DNA, or sequencing alternate clones from existing amplifications.
- The B12 haplotype, which is the origin of the original MHC cosmid cluster sequence (accession number AL023516 (Kaufman et al., 1999b)), was included to confirm the fidelity of the sequence analysis.

3.2 Results

The tapasin gene was PCR amplified from cDNA and from genomic DNA for each of the nine lines representing seven MHC haplotypes available at IAH Compton (Figure 3.1b, and appendix 1). The nucleotide sequences of the protein-coding region of tapasin (from the initiating ATG to the stop codon, as determined from both cDNA and conceptual splicing of the genomic DNA sequence) from each of the seven haplotypes are shown in appendix 2. The deduced amino acid sequences of tapasin from the seven haplotypes are shown in appendix 3. The nucleotide sequences comprising the total genomic region sequenced are shown in appendix 4. For the B2 and B21 haplotypes two chicken lines were sequenced, B2: lines 6 and 7; B21: lines O and N. On each occasion that an entire haplotype was sequenced, the lines from the same haplotype gave identical sequences and therefore the sequences are shown only for the haplotype (except appendices 4 and 5).

3.2.1 Tapasin is polymorphic at the nucleotide level

The differences in the protein-coding region of tapasin in the seven haplotypes are summarised in tables 3.1 and 3.2, and the nucleotide and amino acid sequences are provided in appendices 2 and 3.

Analysis of the nucleotide sequences for the protein-coding region of tapasin from the seven haplotypes reveals all sequences contain the start ATG and stop TGA codons, and consist of 1293 nucleotides. There are six alleles of tapasin in the seven haplotypes, which share nucleotide identity of 98% or greater (table 3.3). The nucleotide sequences for the protein-coding region in the B12 and B19v1 haplotypes were found to be identical.

Analysis of the protein sequences reveals there are also six alleles of the tapasin protein (with the B12 and B19v1 alleles being identical, and referred to as Tapasin*12 hereafter), each 430 amino acids long and sharing an identity of 97% or greater (table 3.3). The polymorphic amino acid positions were dimorphic in the seven haplotypes sequenced.

Exon	Synonymous				Nonsynonymous				
	nt #	codon change	codon #	haplotypes	nt #	codon change	codon #	aa change	haplotypes
2	93	CCG→ CCA	31	B4	44	TCC→ TTC	15	S→ F	B14
					127	CGG→ CAG	43	G→ R	B14, B21
					173	CGC→ CTC	58	R→ L	B4
3	288	CCC→ CCT	96	B2, B15	463	GGA→ AGA	155	G→ R	B4
	417	CCG→ CCC	139	B4					
4	561	GCC→ GCT	187	B2, B12, B19v1	511	ACC→ GCC	171	T→ A	B21
	570	CCA→ CCG	190	B4, B14, B21	574	CCC→ TCC	192	P→ S	B2, B12, B19v1
	621	GTC→ GTG	207	B15	737	CGG→ CAG	246	R→ Q	B14, B21
	717	GCG→ GCA	239	B14, B21					
	735	GCG→ GCA	245	B14, B21					
5	870	GCC→ GCT	290	B21	878	ACG→ ATG	293	T→ M	B2, B4, B21
	873	CCG→ CCA	291	B15	949	GCC→ ACC	317	A→ T	B14, B15
	960	TCG→ TCA	320	B4	955	GGC→ AGC	319	G→ S	B14, B15, B21
	1050	GCG→ GCA	350	B4, B14	971	CGG→ CAG	324	R→ Q	B12, B19v1, B21
	1140	CGC→ CGT	380	B12, B14, B19v1	1073	CGC→ CAC	358	R→ H	B4
	1146	TCC→ TCT	382	B15	1093	GTC→ ATC	365	V→ I	B12, B19v1
	1155	CTG→ CTA	385	B21	1118	ACT→ ATT	373	T→ I	B14
7					1262	CGA→ CAA	421	R→ Q	B14
8	1287	TCG→ TCA	429	B4					
Total	16				15				

Table 3.1 List of nucleotide and amino acid polymorphisms in the tapasin protein-coding sequences (when compared to a consensus sequence). Nucleotide and amino acid differences in tapasin between seven haplotypes are shown. A consensus of the seven haplotypes was used as a reference sequence with changes from this sequence being shown; the haplotypes in which the changes occur are indicated in the final column. Numbering starts with the initiating ATG of the cDNA, or start methionine of the protein sequence. Changes are split into either synonymous (silent) or nonsynonymous (replacement) substitutions.

Exon	2			3	4			5							7
Residue	15	43	58	155	171	192	246	293	317	319	324	358	365	373	421
Tapasin*2	S	G	R	G	T	S	R	M	A	G	R	R	V	T	R
Tapasin*4	S	G	L	R	T	P	R	M	A	G	R	H	V	T	R
Tapasin*12	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*14	F	R	R	G	T	P	Q	T	T	S	R	R	V	I	Q
Tapasin*15	S	G	R	G	T	P	R	T	T	S	R	R	V	T	R
Tapasin*19	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*21	S	R	R	G	A	P	Q	M	A	S	Q	R	V	T	R

Table 3.2 Summary of the polymorphic amino acids in tapasin in the seven haplotypes.

The amino acid differences between the tapasin protein sequences found in the seven haplotypes are tabulated, with their locations indicated by exon and amino acid residue number (with the initiating methionine being +1). The amino acid present in most sequences is shown in blue text, with the alternate amino acid shown in red text.

Tapasin	B2	B4	B12	B14	B15	B19v1	B21
B2	100/100						
B4	99/99	100/100					
B12	99/99	98/98	100/100				
B14	98/97	98/97	98/97	100/100			
B15	99/99	98/98	99/98	98/98	100/100		
B19v1	99/99	98/98	100/100	98/97	99/98	100/100	
B21	98/98	98/98	98/98	99/98	98/98	98/98	100/100

Table 3.3 Comparison of nucleotide/amino acid percentage identities between haplotypes

The nucleotide and amino acid sequences of tapasin found in the seven haplotypes were compared. Percentage similarities are tabulated.

(nucleotide similarity % / amino acid similarity %).

3.2.2 Analysis of the ratio of synonymous and nonsynonymous nucleotide substitutions

The ratio of nonsynonymous to synonymous (N/S) nucleotide substitutions can be used to gain insight into whether a gene is polymorphic as a result of the pressure of selection, or is polymorphic but under no selective pressure. Under the pressure of selection a N/S ratio greater than 2:1 is expected, because generally only the first two bases of a codon need vary to alter the amino acid incorporated into the gene product. The third base is often called the "wobble base", as nucleotide substitutions here rarely change the amino acid incorporated into the gene product. If a gene is under no selective pressure then a N/S ratio of 2:1 or lower is expected, based upon a nucleotide substitution being equally likely at any nucleotide of a codon, with no selective pressure exerted on the resulting selection of amino acid.

In genes that are under enormous selective pressure, such as MHC class I loci, N/S ratios as high as 17:1 have been reported within exons comprising the peptide-binding groove of class I molecules, and as high as 8:1 within the whole gene (Parham et al., 1989; Hunt and Fulton, 1998) (table 3.4).

Exon	Haplotype													
	B2		B13(4)		B12		B14		B15		B19*		B21	
	N	S	N	S	N	S	N	S	N	S	N	S	N	S
1	0	0	0	0	1	1	n.d.	n.d.	1	0	0	0	1	1
2	1	0	9	3	4	2	12	4	21	3	19	3	13	2
3	6	0	19	2	17	1	7	0	6	2	2	2	8	3
4	0	0	0	0	0	0	n.d.	n.d.	3	2	2	1	1	2
5	1	1	0	0	1	1	n.d.	n.d.	0	1	0	1	0	1
6-8	0	0	1	1	0	0	n.d.	n.d.	0	0	0	0	1	0
Total	8	1	29	6	23	5	19	4	31	8	23	7	24	9
N/S	8.00		4.83		4.60		4.75		3.88		3.29		2.67	

Table 3.4 The distribution of nonsynonymous (N, replacement) and synonymous (S, silent) substitutions in exons of chicken class I major alleles. Data taken from (Hunt and Fulton, 1998). * B19 class I sequence is from the American B19 line (Briles et al., 1982), and not the Scandinavian H.B19 line, which is designated B19v1 (Kaufman et al., 1992).

In the seven haplotypes a total of 31 polymorphic nucleotide positions were found in the protein-coding region of tapasin. Sixteen nucleotide substitutions did not change the amino acid selection (synonymous), whilst 15 nucleotide substitutions did change the amino acid (nonsynonymous). However, the N/S ratios did not exceed 1.50 for the whole protein-coding region, with no individual exon displaying any more than three nonsynonymous substitutions when compared to a consensus tapasin sequence (table 3.5).

Exon number (length in nucleotides)	Haplotype													
	B2		B4		B12		B14		B15		B19v1		B21	
	N	S	N	S	N	S	N	S	N	S	N	S	N	S
1 (31)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2 (210)	0	0	1	1	0	0	2	0	0	0	0	0	1	0
3 (240)	0	1	1	1	0	0	0	0	0	1	0	0	0	0
4 (345)	1	1	0	1	1	1	1	3	0	1	1	1	2	3
5 (339)	1	0	2	2	2	1	3	2	2	2	2	1	3	2
6 (90)	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7 (26)	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8 (12)	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Total	2	2	4	6	3	2	7	5	2	4	3	2	6	5
N/S	1.00		0.67		1.50		1.40		0.50		1.50		1.20	

Table 3.5 The distribution of nonsynonymous (N, replacement) and synonymous (S, silent) substitutions in exons of tapasin alleles compared to the consensus sequence.

3.2.3 Analysis of polymorphism in the total genomic region sequenced

For each of the seven haplotypes a total genomic region exceeding 5,000 bp was sequenced (figure 3.1). This comprised: 3,457 bp of the tapasin genomic locus (from the start to stop codons); 829 bp upstream of tapasin exon 1 extending towards the BLB2 gene (a region likely to contain 5' untranslated region (UTR) of tapasin and proximal promoter elements); and 747 bp downstream of tapasin exon 8 extending towards the BLB1 gene (a region likely to contain the 3' UTR of tapasin) (appendix 4).

The sequencing of the tapasin locus from genomic DNA confirmed the results of the cDNA sequencing. Analysis of the genomic sequences revealed that the genomic regions flanking the tapasin gene and the introns were all polymorphic at the nucleotide level, but for each haplotype the intron donor/acceptor sites flanking the exons were all conserved. The functional relevance of nucleotide polymorphisms within introns and untranslated regions can be determined only by experimentation.

The genomic sequences for each of the haplotypes were aligned with the corresponding region of the AL023516 sequence, which was used as a reference sequence (appendix 4). It can be seen that the sequence from the B12 haplotype matched the AL023516 reference sequence exactly apart from one inserted nucleotide within intron 1 and one deleted nucleotide within intron 7, both of which occurred as part of long G/C stretches. This is perhaps a consequence of random slippage of the DNA polymerase enzyme. The sequence of the B19v1 haplotype matched the AL023516 reference sequence exactly apart from a single nucleotide polymorphism in the genomic region downstream of the gene.

For each haplotype an assessment was made of the extent and distribution of polymorphic nucleotides within the total genomic region sequenced. This was achieved by detailing, for each haplotype, the number and location of polymorphic nucleotides (as well as inserted or deleted nucleotides) that differed from the AL023516 reference sequence (table 3.6 and figure 3.2).

Analysis of the exons revealed that exons 1 and 6 were invariant, with exons 7 and 8 each having only a single nucleotide polymorphism in the seven haplotypes examined. The remaining polymorphisms were located in exons 2, 3, 4 and most notably 5. The nature and distribution of the polymorphic amino acids is discussed elsewhere in this chapter.

Region	Polymorphism	B2	B4	B12	B14	B15	B19v1	B21	Length
5' region	SNP	17	23	0	22	26	0	21	829
	INS	2	1	0	1	1	0	1	
	DEL	0	0	0	0	0	0	0	
Exon 1	SNP	0	0	0	0	0	0	0	31
Intron 1	SNP	0	1	0	1	2	0	3	97
	INS	0	0	1	0	0	0	0	
	DEL	4	1	0	3	1	0	0	
Exon 2	SNP	0	2	0	2	0	0	0	210
Intron 2	SNP	6	5	0	2	6	0	3	568
	INS	1	2	0	1	1	0	3	
	DEL	0	0	0	0	0	0	0	
Exon 3	SNP	1	2	0	0	1	0	0	240
Intron 3	SNP	11	11	0	10	14	0	5	682
	INS	0	0	0	0	0	0	0	
	DEL	3	3	0	11	3	0	0	
Exon 4	SNP	0	3	0	6	3	0	7	345
Intron 4	SNP	0	3	0	1	4	0	3	370
	INS	0	0	0	0	0	0	0	
	DEL	0	0	0	0	39	0	2	
Exon 5	SNP	4	7	0	6	7	0	6	339
Intron 5	SNP	1	1	0	0	1	0	1	179
	INS	0	0	0	0	0	0	0	
	DEL	0	0	0	0	0	0	0	
Exon 6	SNP	0	0	0	0	0	0	0	90
Intron 6	SNP	0	0	0	1	1	0	2	163
	INS	0	0	0	0	0	0	0	
	DEL	0	0	0	0	0	0	0	
Exon 7	SNP	0	0	0	1	0	0	0	26
Intron 7	SNP	2	1	0	2	7	0	3	106
	INS	0	0	0	0	0	0	0	
	DEL	5	9	1	1	1	0	8	
Exon 8	SNP	0	1	0	0	0	0	0	12
3' region	SNP	8	12	0	9	10	1	7	747
	INS	0	0	0	0	21	0	0	
	DEL	0	0	0	0	0	0	0	
	Total	65	88	2	80	149	1	75	5034

Total		B2	B4	B12	B14	B15	B19v1	B21
Exon	SNP	5	15	0	15	11	0	13
Intron	SNP	20	22	0	17	35	0	20
	INS	1	2	1	1	1	0	3
	DEL	12	13	1	15	44	0	10
5' region	SNP	17	23	0	22	26	0	21
	INS	2	1	0	1	1	0	1
	DEL	0	0	0	0	0	0	0
3' region	SNP	8	12	0	9	10	1	7
	INS	0	0	0	0	21	0	0
	DEL	0	0	0	0	0	0	0
	Total	65	88	2	80	149	1	75

Ratio		B2	B4	B12	B14	B15	B19v1	B21
Exon	SNP	0.004	0.012	0.000	0.012	0.009	0.000	0.010
Intron	SNP	0.009	0.010	0.000	0.008	0.016	0.000	0.009
	INS	0.000	0.001	0.000	0.000	0.000	0.000	0.001
	DEL	0.006	0.006	0.000	0.007	0.020	0.000	0.005
	INDEL	0.006	0.007	0.001	0.007	0.021	0.000	0.006
5' region	SNP	0.021	0.028	0.000	0.027	0.031	0.000	0.025
	INS	0.002	0.001	0.000	0.001	0.001	0.000	0.001
	DEL	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	INDEL	0.002	0.001	0.000	0.001	0.001	0.000	0.001
3' region	SNP	0.011	0.016	0.000	0.012	0.013	0.001	0.009
	INS	0.000	0.000	0.000	0.000	0.028	0.000	0.000
	DEL	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	INDEL	0.000	0.000	0.000	0.000	0.028	0.000	0.000
	Total	0.013	0.017	0.000	0.016	0.030	0.000	0.015

Table 3.6 Analysis of the distribution of nucleotide polymorphisms within the total genomic region sequenced in comparison to the AL023516 reference sequence. Using the AL023516 (B12 haplotype) accession sequence as a reference sequence, the number and location of polymorphisms (single nucleotide polymorphisms (SNP), inserted nucleotides (INS), deleted nucleotides (DEL) or a combination of both inserted and deleted nucleotides (INDEL)) was tabulated according to genomic location for each of the seven haplotypes. Genomic length is shown in nucleotides for the AL023516 sequence. Ratio refers to the number of nucleotide polymorphisms within a specific nucleotide length.

Similar analysis of the number and location of polymorphisms within the introns revealed that all introns contained polymorphisms; however introns 5 and 6 were the least polymorphic and did not have any inserted or deleted nucleotides. The coincidence of limited polymorphism in the introns surrounding the monomorphic 90 nucleotides of exon 6 may suggest this genomic region is under some pressure to remain invariant. This finding is all the more startling as this region follows the single most polymorphic exon (exon 5).

If the number of nucleotide substitutions within a genomic region is divided by the number of nucleotides constituting that region, then the resulting ratio (referred to here as the SNP ratio) may be used to directly compare the incidence of polymorphisms within and between exons, introns and the genomic regions surrounding the tapasin locus (table 3.6).

The SNP ratio of the exons revealed that the level of variation in the B2 sequence was the most similar to the B12 and B19v1 haplotypes, having a SNP ratio of 0.4%. The remaining haplotypes displayed a slightly greater level of polymorphism with a SNP ratio of between 0.9% and 1.2%.

The SNP ratio of the introns was broadly similar to that observed in the exons, being between 0.8% and 1.6%. Of note are the B2 and B15 haplotypes, where the SNP ratios of the introns were approximately double the SNP ratios of the exons.

The genomic region upstream of exon 1 displayed the highest SNP ratio, being between 2.1% and 3.1%. Further analysis of the polymorphisms within this genomic region is provided later in this chapter. The genomic region downstream of exon 8 displayed a similar SNP ratio to that observed in the introns, being between 0.9% and 1.6% in the haplotypes other than B12 and B19v1.

Similarly, if the numbers of inserted and deleted nucleotides within a genomic region is divided by the number of nucleotides constituting that region, then the resulting ratio (referred to here as the INDEL ratio) will allow comparison of the incidence of such polymorphisms. The introns displayed the highest level of INDELs, with a ratio of between 0.6% and 0.7% for the B2, B4, B14 and B21 haplotypes and a higher INDEL ratio of 2.1% for the B15 haplotype. In comparison the upstream genomic region had a slightly lower INDEL ratio of between 0.1% and 0.2%. The downstream genomic region was devoid of any INDELs apart from the B15 haplotype which displayed an INDEL ratio of 2.8% due to a 21 nucleotide insertion, which separates the stop codon and the site of polyadenylation. This 21 nucleotide insertion in the B15 haplotype was not specific for an individual bird, as it was confirmed from two independent sources of genomic DNA derived from different chickens of the B15 haplotype. The functional relevance of this 21 bp insertion remains to be determined, but it is possible that polymorphisms in the 3' untranslated region could influence polyadenylation, mRNA stability or translation efficiency (Grzybowska et al., 2001).

A graphical depiction of the distribution of the polymorphisms relative to the AL023516 reference sequence is made in figure 3.2. Six points can be made about this representation.

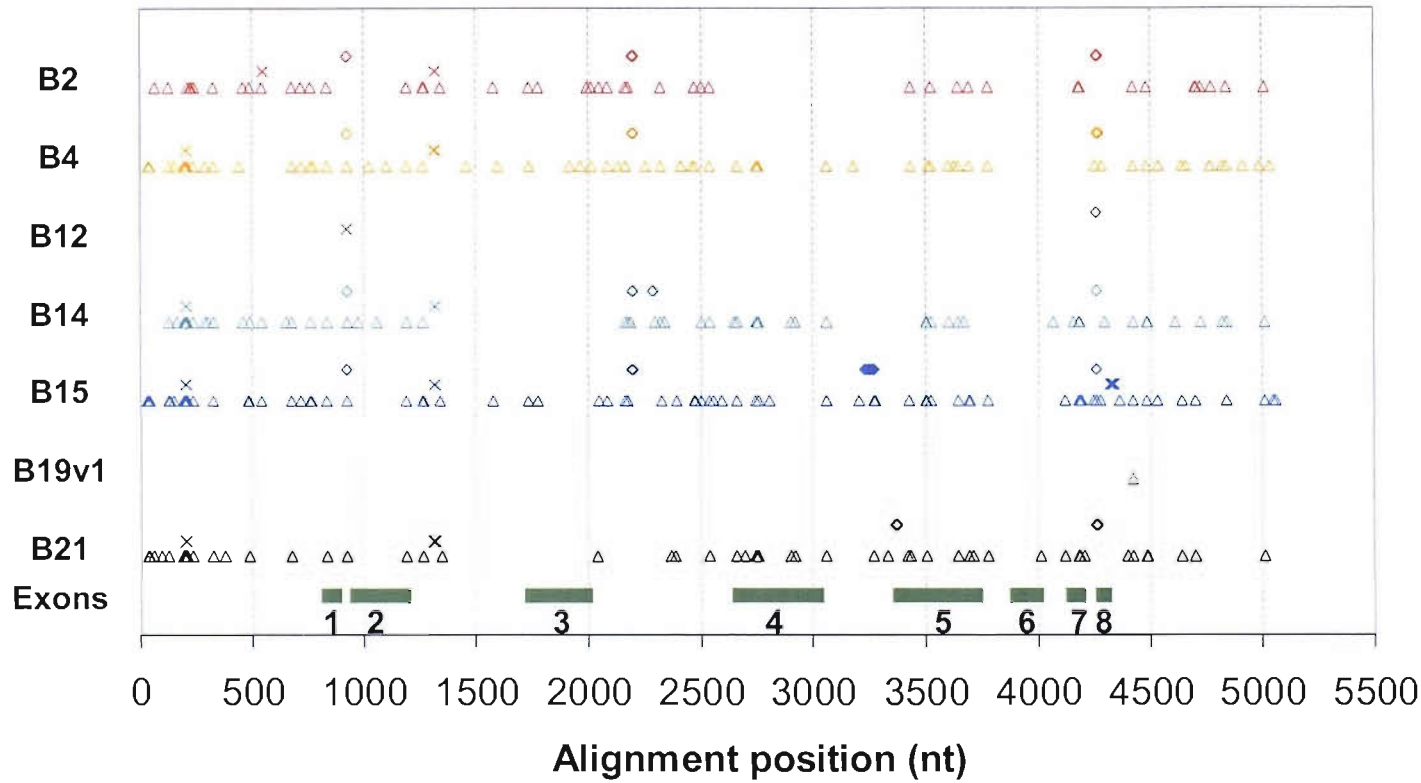


Figure 3.2 The distribution of nucleotide polymorphisms within the total genomic region sequenced.
Using the AL023516 accession sequence as a reference sequence, the position of nucleotide polymorphisms that differed between each of the seven haplotypes were plotted onto a graph as open triangles (Δ). The position of inserted nucleotides were shown as crosses (x). The position of deleted nucleotides were shown as open diamonds (◊). Each of the seven haplotypes is shaded a different colour for clarity. The positions of exons are indicated at the bottom by green rectangles, with exon 1 left-most. The x axis represents the position of the polymorphic nucleotide in the alignment (see Appendix 3).

1. This highlights how similar the B12 and B19v1 sequences are to each other, and to the reference AL023516 sequence.

2. The B14 and B21 sequences appear similar to each other especially at their 5' ends. Thereafter, intron 2 displays little polymorphism from the reference B12 sequence with exon 3 being identical between B12, B14, B19v1 and B21 haplotypes. From intron 3 onwards substitutions are largely scattered throughout the region sequenced in these haplotypes.

3. The B4 and B15 sequences appear reasonably similar, with single nucleotide polymorphisms, for the most part, being scattered throughout the region sequenced.

4. The B2 sequence is similar to the B4 and B15 sequences in the region approximately up to exon 4, where exon 4 and intron 4 are identical in the B4, B12 and B19v1 sequences. From exon 5 onwards the B2 sequence has polymorphisms scattered throughout most of the region sequenced.

5. Any additional or deleted nucleotides that were found in particular haplotypes were mostly concentrated in introns 1, 2, 3 and 7.

6. The last point to make is in agreement with the observation made earlier that intron 5, exon 6 and intron 6 are either invariant or display little polymorphism in the haplotypes studied.

3.2.4 The genomic region upstream of tapasin

In the B12 haplotype 1,036 nucleotides separate the start codons of tapasin and BLB2, with the genes lying in opposite transcriptional orientation (figure 3.1). In each of the seven haplotypes, 829 nucleotides upstream from the translational start site of tapasin were sequenced (appendix 4 and figure 3.2). The majority of the polymorphisms in this region are located within the first ~250 nucleotides at the 5' end of the region sequenced, residing roughly midway between the tapasin and BLB2 genes. There are fewer polymorphisms in the ~500 nucleotides immediately proximal to the translational start site of tapasin. It is possible that nucleotide polymorphisms within this region may influence the transcriptional control of tapasin, BLB2, or both genes.

In the initial characterisation of the chicken tapasin gene, a potential TATA box and SP1 binding site were predicted (appendix 4) (Frangoulis et al., 1999). In this study, the TATA box is well conserved between haplotypes; with only the B4 haplotype differing from the other haplotypes by one nucleotide (T→A). A large variety of A/T rich sequences are capable of functioning as TATA boxes, so it may be expected that the B4 sequence is capable of functioning as a TATA box in much the same fashion as all other haplotypes. The predicted SP1 binding site does not vary between the haplotypes examined.

Overall the genomic region that is upstream from tapasin exon 1 is not highly polymorphic in the haplotypes sequenced, with a relatively low incidence of polymorphic nucleotides immediately preceding the start of translation: a region most likely to contain proximal transcription factor binding sites and other promoter elements. Therefore it is likely (although unproven) that the level of transcription of tapasin will not vary between the haplotypes.

3.2.5 Comparison of B12 tapasin genomic sequences

An alignment was made of the genomic regions sequenced in this study with the AL023516 and AJ004999 sequences (appendix 5). The AJ004999 sequence differs from the AL023516 sequence and from the Tapasin*12 sequence obtained in this study in a number of places, with all differences (nucleotide polymorphisms, insertions or deletions) being unique to the AJ004999 sequence, and are not shared by the other haplotypes sequenced in this study (table 3.7).

The most important difference between the AJ004999 sequence and the other B12 sequences concerns exon 4, which was the only exon where polymorphisms between sequences were observed. Here two nucleotide polymorphisms were observed in the AJ004999 sequence (alignment positions 3068-9), which leads to a coding change in the protein sequence (appendix 6). In the AJ004999 sequence, histidine (codon CAC) is reported at residue 222, whilst in the AL023516 sequence and for all of the seven haplotypes that were sequenced in this study, threonine occupies residue 222 (codon ACC). The near complete identity between the AL023516 sequence and the Tapasin*12 sequence obtained in this study, which derive from independent sources of genomic DNA, confirms the authenticity of these sequences. It also suggests that the differences between the AJ004999 sequence and the other B12 sequences are most probably a consequence of errors made in the sequencing that constituted the AJ004999 sequence.

Genomic Location	Differences in the AJ004999 sequence in comparison to the AL023516 reference sequence	
	Alignment position(s)	Details
BLB2-Tpn	129, 140, 141, 147, 154, 156, 212	7 x SNP
Exon 4	3068-3069	coding change AJ004999 CAC = Histidine 222 AL023516 ACC = Threonine 222
Intron 3	2546-2606, 2656	61 x Del, 1 x Del
	2714, 2833, 2864	3 x SNP
Intron 4	3346	1 x Del
	3550	1 x SNP
Intron 7	4431	1 x Del
	4493-4	2 x Ins
Tpn-BLB1	4597, 4599, 4601, 4770, 4793-5, 4798-4802	12 x SNP
	4764	1 x Del
	4790	1 x Ins

Genomic Location	Differences in the Tapasin*12 sequence obtained in this study in comparison to the AL023516 reference sequence	
	Alignment position(s)	Details
Intron 1	1143	1 x Ins
Intron 7	4492	1 x Del

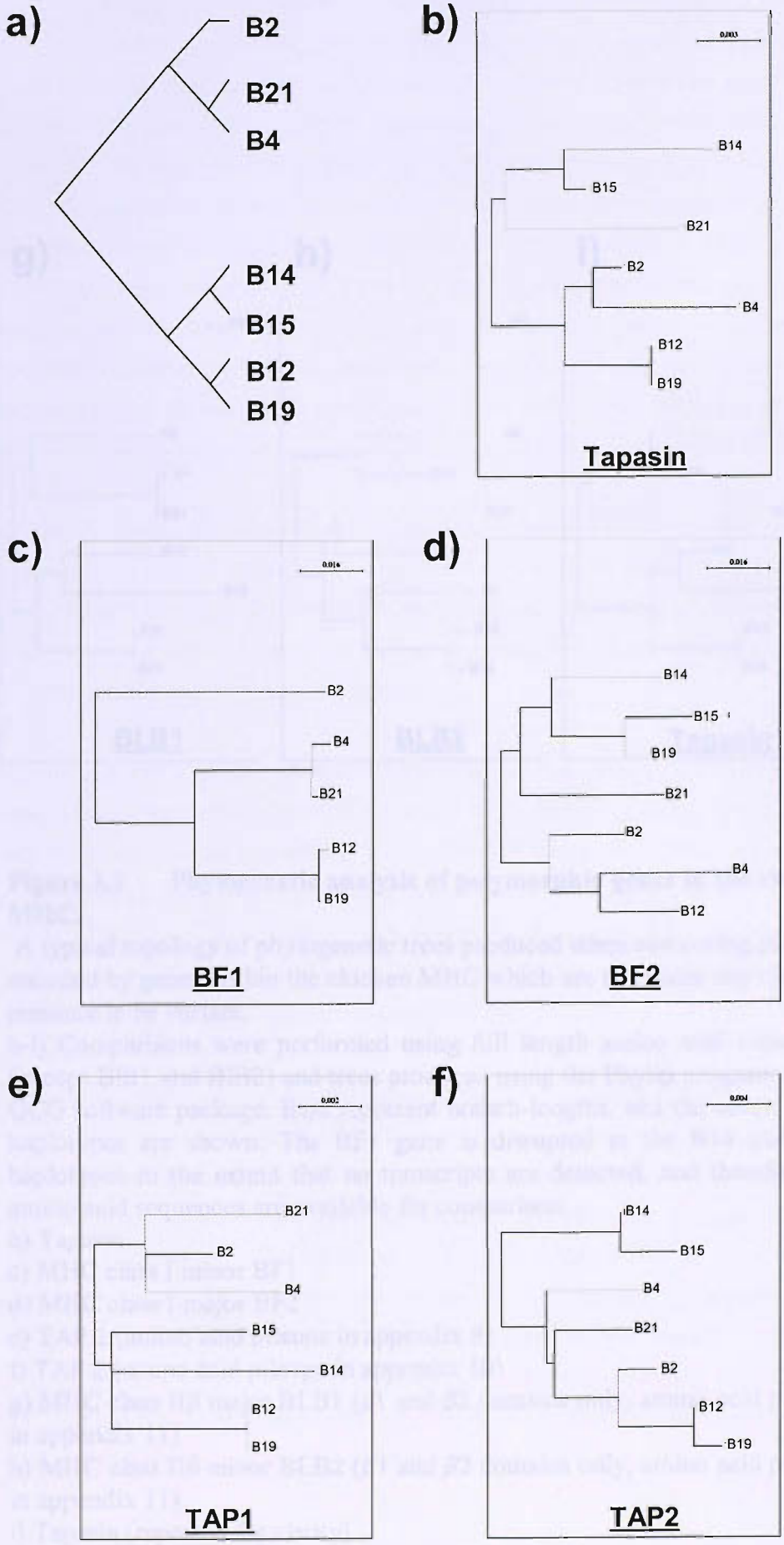
Table 3.7 Nucleotide differences between either the AJ004999 sequence, or the Tapasin*12 sequence obtained in this study, in comparison to the AL023516 reference sequence. For either sequence, the nature and location of nucleotide polymorphisms in comparison to the AL023516 reference sequence, was annotated according to genomic location.

3.2.6 Phylogenetic trees suggest that the polymorphic amino acids in tapasin are not the result of evolutionary drift, and may provide evidence that tapasin is evolving with the BF2 and TAP2 genes

In humans and mice, it is likely that only the class I and class II genes are under strong selective pressure to be divergent at the amino acid level. Although not fully investigated, most of the other genes found within the MHC would be expected to be under lower levels of selective pressure to be divergent at the amino acid level. Within the compact chicken MHC, the low incidence of recombination means that it is likely that the entire MHC (apart from where strong selective pressure is applied) evolves as a unit undergoing evolutionary drift. One way in which this can be visualised is by phylogenetic analysis of proteins that are not expected to be under strong pressures of selection.

The BF1 and BLB1 genes are referred to as the class I and class II minor genes respectively, because relative to the BF2 and BLB2 genes (the class I and class II major genes) transcripts from the minor genes generally represent ~10% of the total class I or II transcripts within the cell (Kaufman et al., 1999a; Jacob et al., 2000; Wallny et al., 2006). The low expression levels of the minor genes suggests that these genes are not under strong selective pressure to be divergent. The topology of the phylogenetic trees that are produced for these proteins is relatively constant and has been observed for other proteins that are encoded by genes in the chicken MHC (J. Kaufman, personal communication), and presumably represents evolutionary drift as the haplotypes have diverged from their last common ancestor. A representation of this relationship between genes that are not under strong selective pressure is shown in figure 3.3a. This shows the B12 and B19v1 sequences to be closely related, along with the closely related B14 and B15 sequences. In a separate branch of the tree, and presumably diverging from a common ancestor distinct from a B12/B19v1 and B14/B15 common ancestor, the B4 and B21 sequences are closely related, with B2 branching separately to the B4/B21 branch. This relationship is observed for the BLB1 and BF1 proteins (figure 3.3c and 3.6g, the B14 and B15 BF1 proteins are omitted as the genes have not been cloned; a likely consequence of a large DNA insertion, or other rearrangement, that leads to undetectable levels of mRNA transcription, J. Kaufman, personal communication).

This relationship is not observed when genes are under selective pressure to be divergent. This is evident when phylogenetic trees are constructed for the class I and class IIB major proteins, BF2 and BLB2, which are under strong selective pressure to be divergent (figure 3.3d and 3.6h). It is possible that molecules that are evolving with the BF2 or BLB2 molecules through co-evolution of polymorphic, interacting amino acid residues that mediate the specificity of the interaction, may share a common topology to the BF2 or BLB2 genes, which is distinct from the topology observed when genes follow haplotype descent.


a)
B2
B21
B4
B14
B15
B12
B19
b)
B14
B15
B21
B2
B4
B12
B19
Tapasin
c)
B2
B4
B21
B12
B19
BF1
d)
B14
B15
B19
B21
B2
B4
B12
BF2
e)
B21
B2
B4
B15
B14
B12
B19
TAP1
f)
B14
B15
B4
B21
B2
B12
B19
TAP2

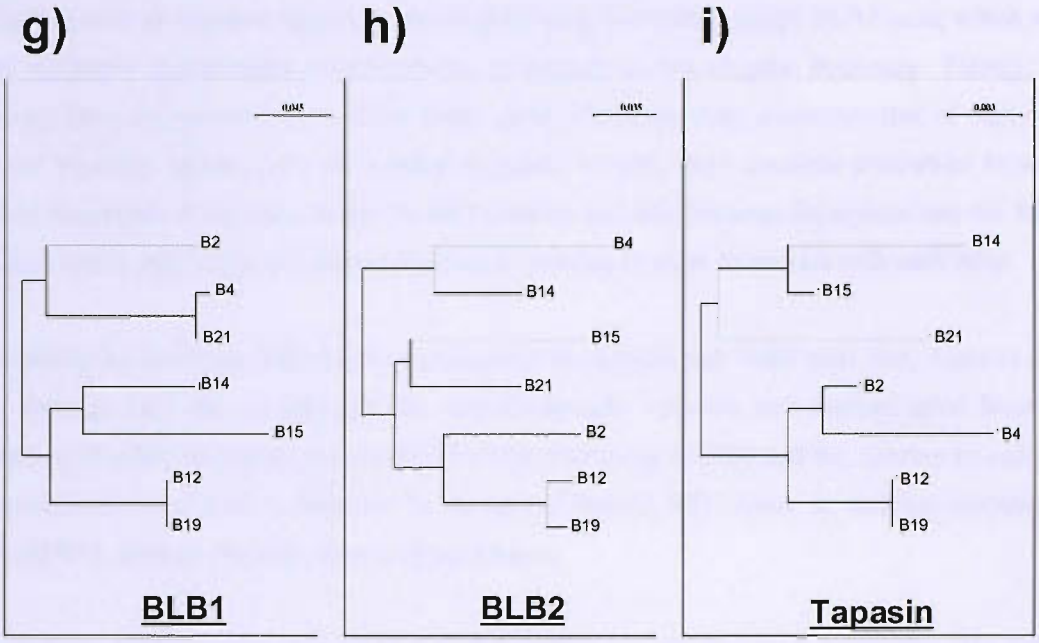


Figure 3.3 Phylogenetic analysis of polymorphic genes in the chicken MHC.

A typical topology of phylogenetic trees produced when comparing proteins encoded by genes within the chicken MHC which are not under any obvious pressure to be variant.

b-i) Comparisons were performed using full length amino acid sequences (except BlB1 and BlB2) and trees produced using the Phylip program of the GCG software package. Bars represent branch-lengths, and the seven MHC haplotypes are shown. The BF1 gene is disrupted in the B14 and B15 haplotypes to the extent that no transcripts are detected, and therefore no amino acid sequences are available for comparison.

b) Tapasin

c) MHC class I minor BF1

d) MHC class I major BF2

e) TAP 1 (amino acid pileups in appendix 9)

f) TAP 2 (amino acid pileups in appendix 10)

g) MHC class IIβ major BLB1 (β1 and β2 domains only, amino acid pileups in appendix 11)

h) MHC class IIβ minor BLB2 (β1 and β2 domains only, amino acid pileups in appendix 11)

i) Tapasin (repeated for clarity)

Phylogenetic tree analysis of the tapasin protein sequences (figure 3.3b) reveals four things: firstly, the topology suggests the tapasin sequences have not followed the pattern of haplotype descent that is evident in proteins (such as BLB1 and BF1) that are not under the pressures of selection. This suggests that there may be some selective pressure for tapasin to be divergent at the amino acid level. Secondly, the topology does not resemble that of the BLB2 gene (which is distinct from all others). This suggests the polymorphism in tapasin is not simply a consequence of genetic "hitch-hiking", caused by selection applied to the neighbouring and tightly linked BLB2 gene, which may cause randomly accumulated polymorphisms in tapasin to rise in gene frequency. Thirdly, the topology does not resemble that of the TAP1 gene, whose topology resembles that of haplotype descent. Fourthly, despite some differences in branch lengths, there are clear similarities between the tree topologies of the tapasin and the BF2 proteins and also between the tapasin and the TAP2 proteins, which may imply that these proteins are evolving in order to interact with each other.

Considering the similarity between the topologies of the tapasin and TAP2 trees first, it can be seen that, although they are not identical, the trees do broadly resemble one another, apart from the position of the B21 haplotype. For TAP2, B21 branches along with B2 and B4, sharing an ancestor common to B12 and B19v1. However in the case of tapasin, B21 shares an ancestor common to B14 and B15, distinct from the other four haplotypes.

The tapasin and BF2 trees also share broadly similar, but not identical, topologies. The major difference between trees concerns the placing of the B19v1 haplotype. In the case of tapasin, the B12 and B19v1 proteins are identical and accordingly branch together, whilst in the case of BF2, B19v1 is located on a totally separate branch, being most similar to B15. The similarity between BF2*15 and BF2*19v1 is a consequence of a recombination event (which is discussed in a later chapter), and therefore the placement of the BF2*19v1 allele is likely to be a reflection of a recombination event, rather than potential co-evolution.

In an effort to determine whether specific regions of the BF2 molecules exhibit more similarities in the topologies of their phylogenetic trees to that observed for the tapasin alleles, phylogenetic trees were constructed using only specific regions of the BF2 molecules which may interact directly or otherwise influence the association with tapasin (figure 3.4). However, trees constructed using polymorphic amino acids found only in the potential tapasin-binding α2 and α3 loops, or within the F pocket only, or a combination of these regions, did not generate phylogenetic trees with topologies more similar to that of the tapasin alleles; for example, in all cases B12 and B19v1 did not branch together. This suggests that residues that were considered separately in this analysis, and other residues that were not included in this analysis, combine to generate the similar phylogenetic tree topologies that were constructed for the full-length tapasin and BF2 proteins.

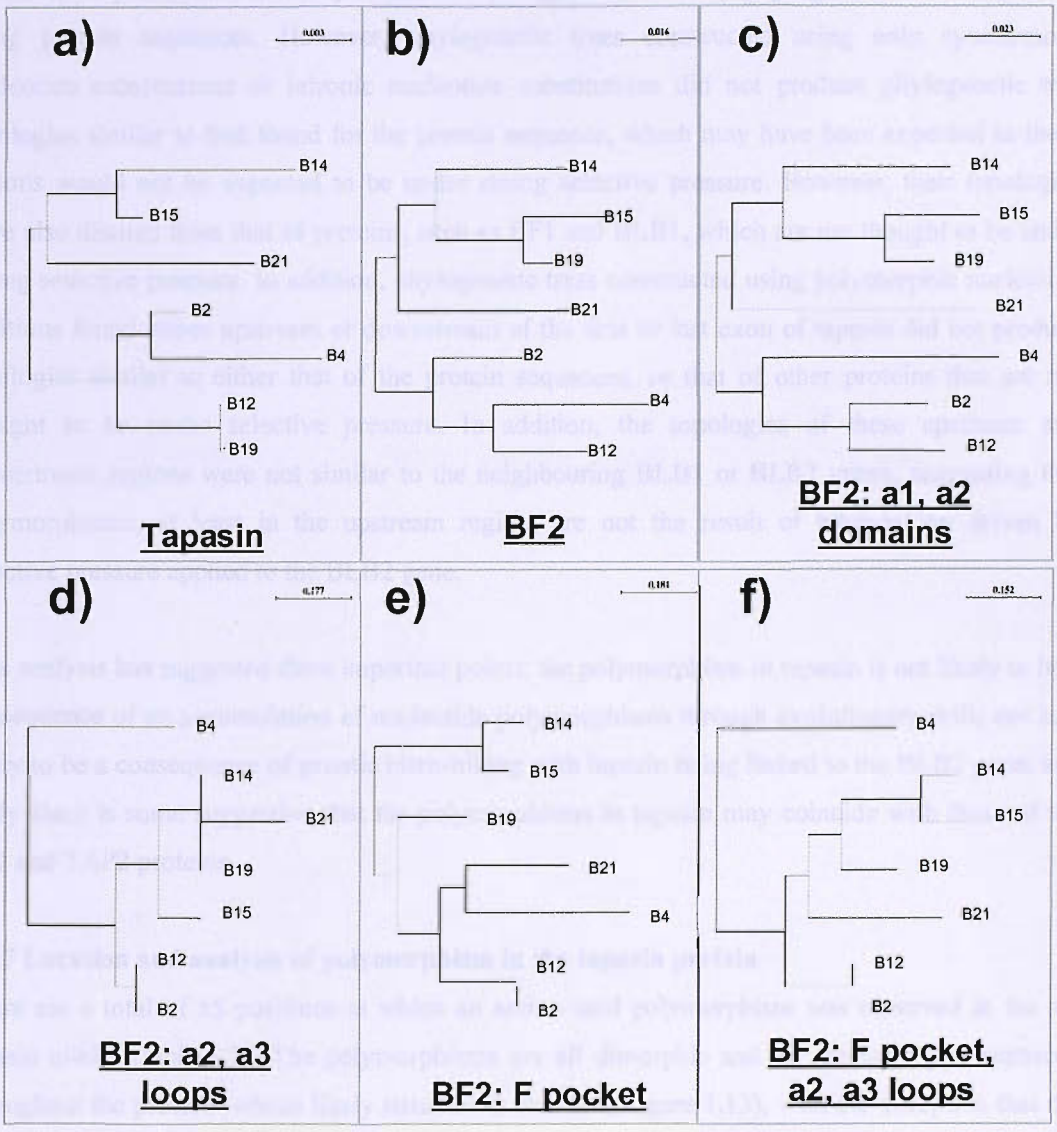


Figure 3.4 Phylogenetic analysis of portions of the BF2 proteins.
a-f) Comparisons were performed using amino acid sequences of tapasin and BF2 proteins. Trees were produced using the Phylip program of the GCG software package. Bars represent branch-lengths, and the seven MHC haplotypes are shown.
a) Tapasin: full length protein
b) BF2: full length protein
c) BF2: just alpha 1 (a1) and alpha 2 (a2) domains
d) BF2: just polymorphic residues in the potential tapasin binding loops in the a2 and a3 domains
e) BF2: polymorphic residues around the F pocket (positions 76, 80, 95, 97, 111 and 113 – equivalent to HLA-A2 77, 81, 97, 99, 114 and 116).
f) BF2: polymorphic positions around the F pocket and in potential tapasin binding loops in the a2 and a3 domains

Phylogenetic trees were also constructed using polymorphic nucleotide positions within specific regions of the complete tapasin sequences obtained in this study (figure 3.5). As may have been expected, the topology of the phylogenetic tree constructed using nonsynonymous nucleotide substitutions produced a virtually identical topology to that of the phylogenetic tree constructed using protein sequences. However, phylogenetic trees constructed using only synonymous nucleotide substitutions or intronic nucleotide substitutions did not produce phylogenetic tree topologies similar to that found for the protein sequence, which may have been expected as these regions would not be expected to be under strong selective pressure. However, their topologies were also distinct from that of proteins, such as BF1 and BLB1, which are not thought to be under strong selective pressure. In addition, phylogenetic trees constructed using polymorphic nucleotide positions found either upstream or downstream of the first or last exon of tapasin did not produce topologies similar to either that of the protein sequences, or that of other proteins that are not thought to be under selective pressure. In addition, the topologies of these upstream and downstream regions were not similar to the neighbouring BLB1 or BLB2 genes, suggesting that polymorphisms, at least in the upstream region, are not the result of hitch-hiking driven by selective pressure applied to the BLB2 gene.

This analysis has suggested three important points: the polymorphism in tapasin is not likely to be a consequence of an accumulation of nucleotide polymorphisms through evolutionary drift; nor is it likely to be a consequence of genetic hitch-hiking with tapasin being linked to the BLB2 gene; and lastly there is some suggestion that the polymorphisms in tapasin may coincide with those of the BF2 and TAP2 proteins.

3.2.7 Location and analysis of polymorphism in the tapasin protein

There are a total of 15 positions at which an amino acid polymorphism was observed in the six tapasin alleles (table 3.2). The polymorphisms are all dimorphic and for the most part scattered throughout the protein (whose likely structure is shown in figure 1.13), with the exception that the Ig-C domain encoded by exon 5 appeared to have a noticeable clustering of polymorphisms. Between the haplotypes there were three polymorphic amino acid residues in exon 2, one in exon 3, three in exon 4, seven in exon 5, and one in exon 8. The polymorphic amino acids in chicken tapasin do not occupy equivalent locations to the amino acid polymorphisms in the mouse tapasin sequences (figure 3.6); however, the single amino acid polymorphism in human tapasin (arginine/threonine at position 260) does occupy an equivalent position to a polymorphic amino acid in chicken tapasin in this alignment (arginine/glutamine at position 246), within the likely Ig-V domain encoded by exon 4. The coincidence of the location and nature of this polymorphism in both chicken and human tapasin proteins may suggest a functional significance for this polymorphism; however, it must be noted that the chicken and human proteins may differ in structure around this polymorphism, especially with regard to the close proximity of the N-linked glycan that is found in mammalian, but not chicken tapasin sequences.

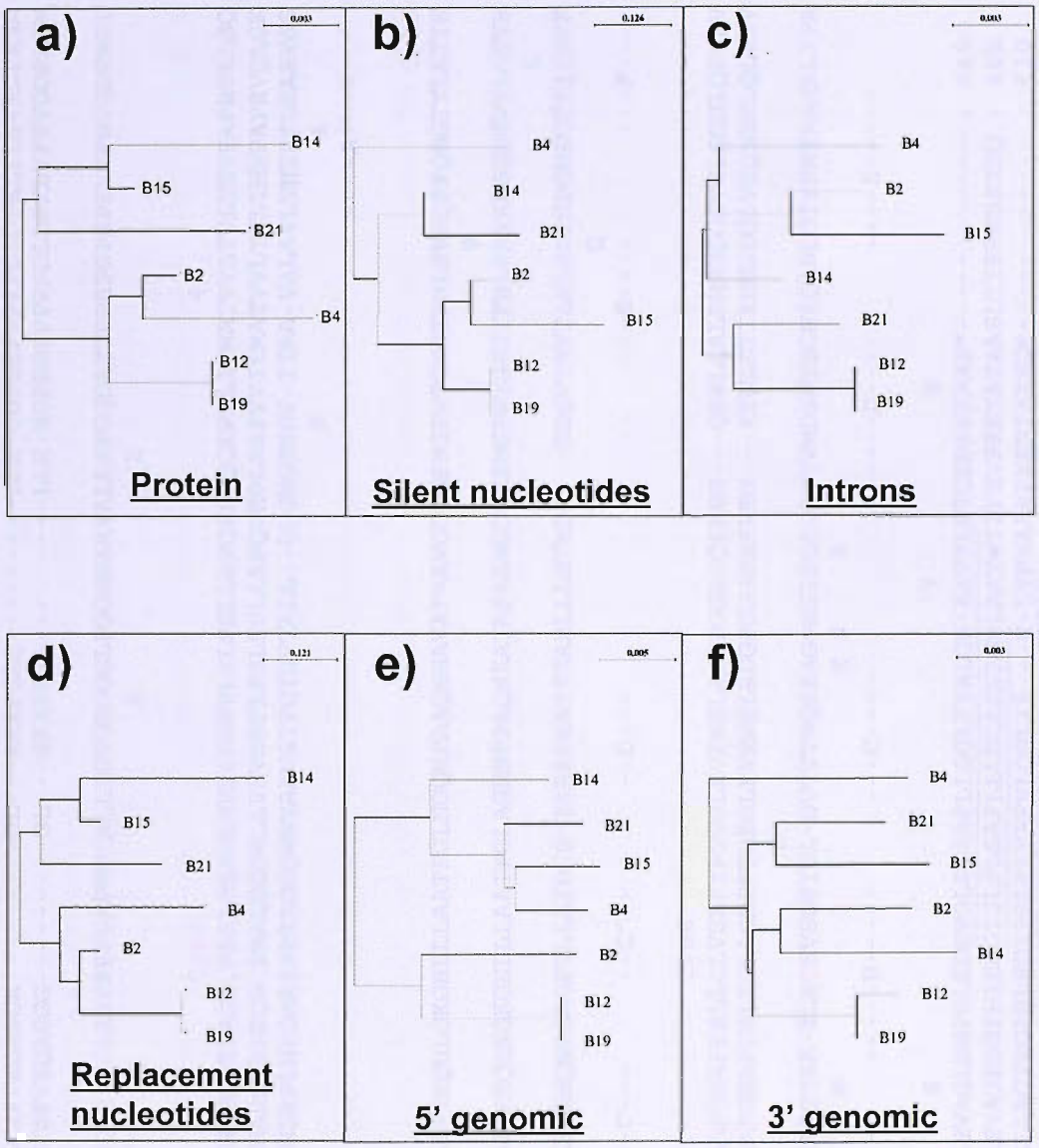


Figure 3.5 Phylogenetic analysis of tapasin.

a-f) Comparisons of tapasin were performed using full length amino acid sequences. Trees were produced using the Phylip program of the GCG software package. Bars represent branch-lengths, and the seven MHC haplotypes are shown.

a) Full length protein

b) Synonymous nucleotide (silent) substitutions only

c) Intronic genomic sequences

d) Non-synonymous nucleotide (replacement) substitutions only

e) The genomic region sequenced upstream of exon 1

f) The genomic region sequenced downstream of exon 8

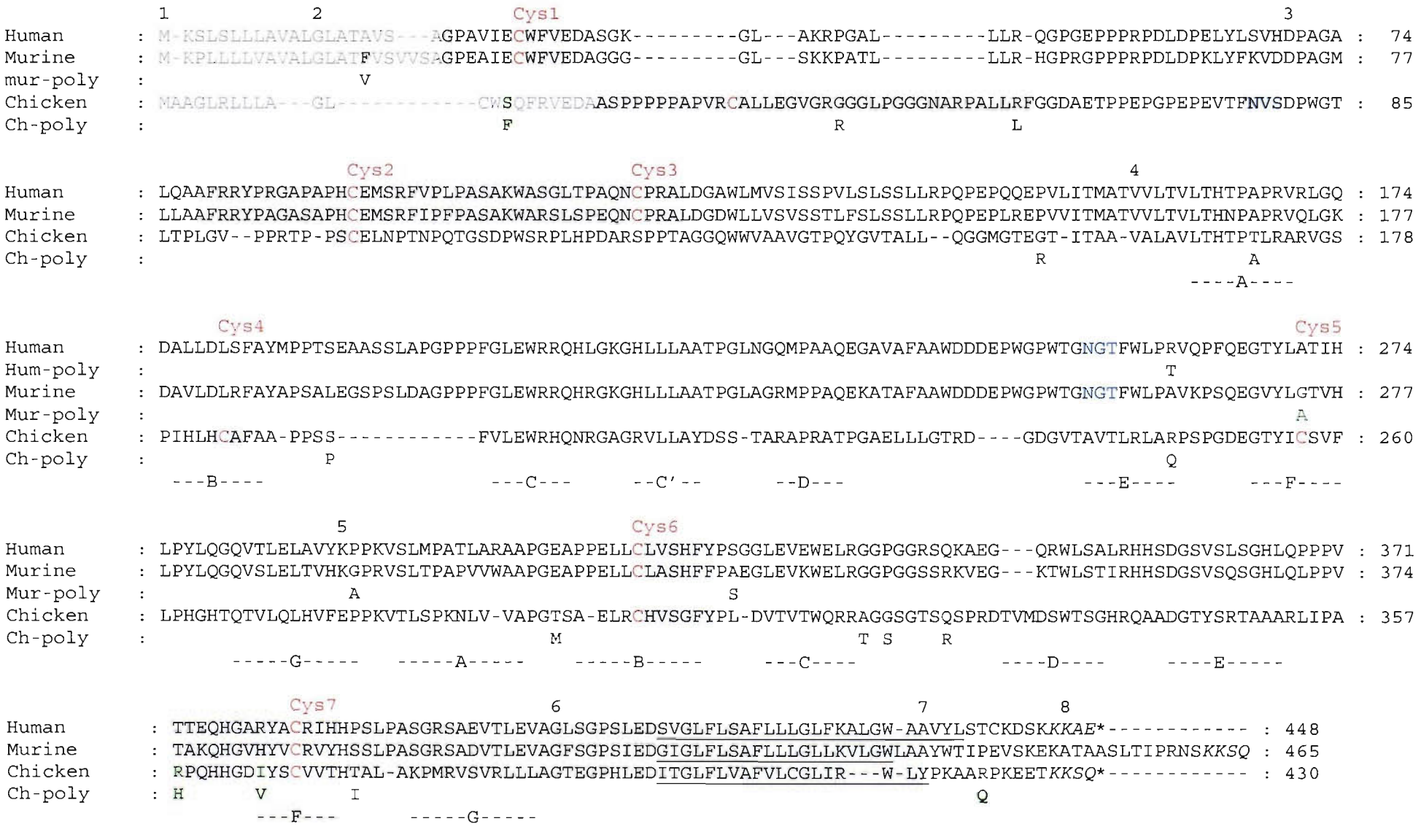

1
2
Cys1
3
Human : M-KSLSLLLAVALGLATAVS---AGPAVIECWFVEDASGK---------GL---AKRPGAL---------LLR-QGPGEPPPRPDLDPELYLSVHDPAGA : 74
Murine : M-KPLLLLVAVALGLATFVSVVSAGPEAIECWFVEDAGGG---------GL---SKKPATL---------LLR-HGPRGPPPRPDLDPKLYFKVDDPAGM : 77
mur-poly : V
Chicken : MAAGLRLLLA---GL-----------CWSQFRVEDAASPPPPPAPVRCALLEGVGRGGGLPGGGNARPALLRFGGDAETPPEPGPEPEVTFNVSDPWGT : 85
Ch-poly : F R L
Cys2
Cys3
4
Human : LQAAFRRYPRGAPAPHCEMSRFVPLPASAKWASGLTPAQNCPRALDGAWLMVSISSPVLSLSSLLRPQPEPQQEPVLITMATVVLTVLTHTPAPRVRLGQ : 174
Murine : LLAAFRRYPAGASAPHCEMSRFIPFPASAKWARSLSPEQNCPRALDGDWLLVSVSSTLFSLSSLLRPQPEPLREPVVITMATVVLTVLTHNPAPRVQLGK : 177
Chicken : LTPLGV--PPRTP-PSCELNPTNPQTGSDPWSRPLHPDARSPPTAGGQWWVAAVGTPQYGVTALL--QGGMGTEGT-ITAA-VALAVLTHTPTLRARVGS : 178
Ch-poly : R A
----A----
Cys4
Cys5
Human : DALLDLSFAYMPPTSEAASSLAPGPPPFGLEWRRQHLGKGHLLLAATPGLNGQMPAAQEGAVAFAAWDDDEPWGPWTGNGTFWLPRVQPFQEGTYLATIH : 274
Hum-poly : T
Murine : DAVLDLRFAYAPSALEGSPSLDAGPPPFGLEWRRQHRGKGHLLLAATPGLAGRMPPAQEKATAFAAWDDDEPWGPWTGNGTFWLPAVKPSQEGVYLGTVH : 277
Mur-poly : A
Chicken : PIHLHCAFAA-PPSS-----------FVLEWRHQNRGAGRVLLAYDSS-TARAPRATPGAELLLGTRD----GDGVTAVTLRLARPSPGDEGTYICSVF : 260
Ch-poly : P Q
---B---- ---C--- --C'-- --D--- ---E---- ---F----
5
Cys6
Human : LPYLQGQVTLELAVYKPPKVSLMPATLARAAPGEAPPELLCLVSHFYPSGGLEVEWELRGGPGGRSQKAEG---QRWLSALRHHSDGSVSLSGHLQPPPV : 371
Murine : LPYLQGQVSLELTVHKGPRVSLTPAPVVWAAPGEAPPELLCLASHFFPAEGLEVKWELRGGPGGSSRKVEG---KTWLSTIRHHSDGSVSQSGHLQLPPV : 374
Mur-poly : A S
Chicken : LPHGHTQTVLQLHVFEPPKVTLSPKNLV-VAPGTSA-ELRCHVSGFYPL-DVTVTWQRRAGGSGTSQSPRDTVMDSWTSGHRQAADGTYSRTAAARLIPA : 357
Ch-poly : M T S R
-----G----- -----A----- -----B----- ---C---- ----D---- ----E-----
Cys7
6
7
8
Human : TTEQHGARYACRIHHPSLPASGRSAEVTLEVAGLSGPSLEDSVGLFLSAFLLLGLFKALGW-AAVYLSTCKDSKKKAE*---------- : 448
Murine : TAKQHGVHYVCRVYHSSLPASGRSADVTLEVAGFSGPSIEDGIGLFLSAFLLLGLLKVLGWLAAYWTIPEVSKEKATAASLTIPRNSKKSQ : 465
Chicken : RPQHHGDIYSCVVTHTAL-AKPMRVSVRLLLAGTEGPHLEDITGLFLVAFVLCGLIR---W-LYPKAARPKEETKKSQ*---------- : 430
Ch-poly : H V I Q
---F--- -----G-----

Figure 3.6 Alignments of human, murine and chicken tapasin protein sequences, with polymorphic amino acids indicated.

The alignment was produced in two steps: firstly, human and murine sequences were aligned according to (Grandea et al., 1998); secondly, chicken was aligned according to (Frangoulis et al., 1999). Gaps are introduced to facilitate alignment, however note that other alignments are possible. The start of a new exon is indicated by numbers above the text. Dilysine motifs are shown in italics; sites of N-linked glycosylation are shown in blue. Predicted signal sequences are shown in grey, human (Ortmann et al., 1997), murine (Grandea et al., 1998), chicken (this study, although a 15 amino acid signal sequence was predicted by (Frangoulis et al., 1999)). Predicted transmembrane TM domains are underlined, human (Ortmann et al., 1997), murine (Grandea et al., 1998), chicken (Frangoulis et al., 1999). Cysteine residues within the predicted ER luminal domains are shown in red, and are sequentially numbered. The position of immunoglobulin β strands is shown beneath the alignment, and was aligned using the chicken tapasin sequence (Teng et al., 2002).

Polymorphisms that are reported between tapasin sequences or were determined in this study are shown in green, bold text, with the alternate amino acid shown in the line beneath (labeled either hum/mur/ch – poly).

A polymorphism in human tapasin is described at position 260 (R/T) (Copeman et al., 1998), and polymorphisms exist between two published murine tapasin sequences at positions 17 (F/V), 274 (G/A), 294 (G/A), and 326 (A/S) (Grandea et al., 1998; Li et al., 1999). Polymorphic amino acids in chicken tapasin that were found in the seven MHC haplotypes that were sequenced in this study are indicated, with the alternate choice of amino acid shown beneath.

Human: AF009510, murine: AF043943 (Grandea et al., 1998) is shown, which differs to AF106278 (Li et al., 1999) at four positions, chicken: AL023516 (Kaufman et al., 1999).

The first polymorphic amino acid in the haplotypes sequenced is within the second exon at residue 15. In all haplotypes serine is found here (uncharged, polar side chain), apart from in the B14 haplotype where phenylalanine is present (non-polar, hydrophobic side chain). Frangoulis *et al* used their Tapasin*12 protein sequence to predict a signal peptidase cleavage site occurring between residues 15 and 16 (Frangoulis et al., 1999). This may mean a polymorphism at this location may affect signal sequence cleavage. An alteration in signal peptide cleavage may have profound consequences; the defect in the tapasin gene in the human tapasin-deficient 721.220 cell line not only causes the splicing out of exon 2 (a domain likely to be critically important for the interaction with class I molecules), but an alteration in signal peptide sequence that results in the truncated protein being poorly translocated into the ER (Copeman et al., 1998; Bangia et al., 1999).

Alignments of chicken and mammalian tapasin protein sequences (figure 3.6) reveals their N termini to be quite divergent, preventing accurate comparisons of the sites of signal peptide cleavage between species from being made. Therefore in an effort to determine the potential effect that a polymorphism at this position may have on signal peptide cleavage, the Tapasin*12 protein sequence was submitted for signal peptide analysis using the SignalP 3.0 program (University of Denmark) (Nielsen et al., 1997). The most likely signal peptidase cleavage site was predicted to occur between residues 22 and 23, however the second most likely cleavage site was predicted to occur between residues 15 and 16. When similar analysis was conducted using the Tapasin*14 peptide sequence, the most likely cleavage site was again predicted to occur between residues 22 and 23. However, the likelihood of a signal peptidase cleavage site occurring between residues 15 and 16 was significantly reduced. Signal sequences primarily contain a core sequence of five to ten hydrophobic amino acids. The first 22 amino acids are predominantly composed of hydrophobic amino acids; therefore it is unlikely that the substitution of a polar residue (serine) for an additional hydrophobic amino acid (phenylalanine) would affect the translocation of the protein into the ER.

The other polymorphic amino acids located in the mature tapasin protein could affect the ability of tapasin to interact productively with the class I and TAP molecules such that interactions between molecules may only occur optimally when the molecules are encoded by the same haplotype. If this is true, the distribution and nature of amino acid polymorphisms in tapasin may offer insight into the molecular interactions.

It is likely that the transmembrane and/or cytoplasmic domains of tapasin mediate the interaction with the TAP heterodimer, whilst the ER luminal domains of tapasin are likely to interact with the class I molecule (Lehner et al., 1998; Bangia et al., 1999). Therefore the presence of only a single polymorphic amino acid in the transmembrane and cytoplasmic domains of tapasin (exons 6-8) suggests that the interaction between tapasin and TAP is likely to be mediated by invariant amino acids. The possible exception is the B14 haplotype, where glutamine replaces the arginine found in the other haplotypes at residue 421. However some of the amino acid polymorphisms in the ER

luminal domains of tapasin may contribute to mediating a specific interaction with the TAP heterodimer that is encoded from the same haplotype, a possibility which is supported by the phylogenetic analysis of tapasin and TAP2 protein sequences. The prominent clustering of polymorphic residues at the base of the membrane proximal Ig-C domain encoded by exon 5 appear the most likely candidates to mediate such a potential interaction (figure 3.7).

If tapasin is evolving to interact with the class I and TAP molecules in a haplotype-specific fashion, then the specificity of the interaction may be mediated by a small number of polymorphic interacting residues. Therefore the nature and location of polymorphic amino acids in tapasin, class I major (BF2) and minor (BF1) and both TAP proteins was tabulated and analysed for regions of haplotype-specific complementarities (figure 3.8). In particular, patterns of charge complementarities were sought as these are the most easily identified, and the nature of such interactions might be highly perturbed by amino acid polymorphism. This analysis revealed many patterns of complementarities between tapasin and other proteins (summarised in figure 3.8). However, only those potential interactions which involve at least one charged residue, and are located within domains of the proteins that are likely to interact are discussed.

Two patterns of complementarities were evident involving tapasin and the class I major proteins. Firstly (potential interaction 1a), there is a symmetry in the polymorphisms concerning basic and hydrophobic amino acids which occurs at residue 58 of tapasin (residues numbered from initiating methionine in tapasin) where arginine (basic side chain) is found in all haplotypes apart from B4, where leucine (non-polar, hydrophobic side chain) is found. This coincides with a unique polymorphism in the BF2 protein at residue 128 (residues numbered from first residue of mature protein for BF proteins), where all haplotypes have methionine (non-polar, hydrophobic side chain) apart from B4, where lysine (basic side chain) is present.

Potential interaction 1a

Haplotype	Tapasin	BF2
B2, B12, B14, B15, B19v1, B21	R58	M128
B4	L58	K128

When chicken BF2 proteins are aligned with mammalian class I proteins (appendix 7) it can be seen that BF2 residue 128 occupies an equivalent position to mammalian residue 131 within a loop in the α2 domain that is likely to mediate an interaction with the mammalian tapasin protein (figure 3.9a). Indeed a mutation of this residue in the H-2L^{d} protein (K131D) lead to increased association with the PLC and enhanced loading of endogenous peptides (Yu et al., 1999).

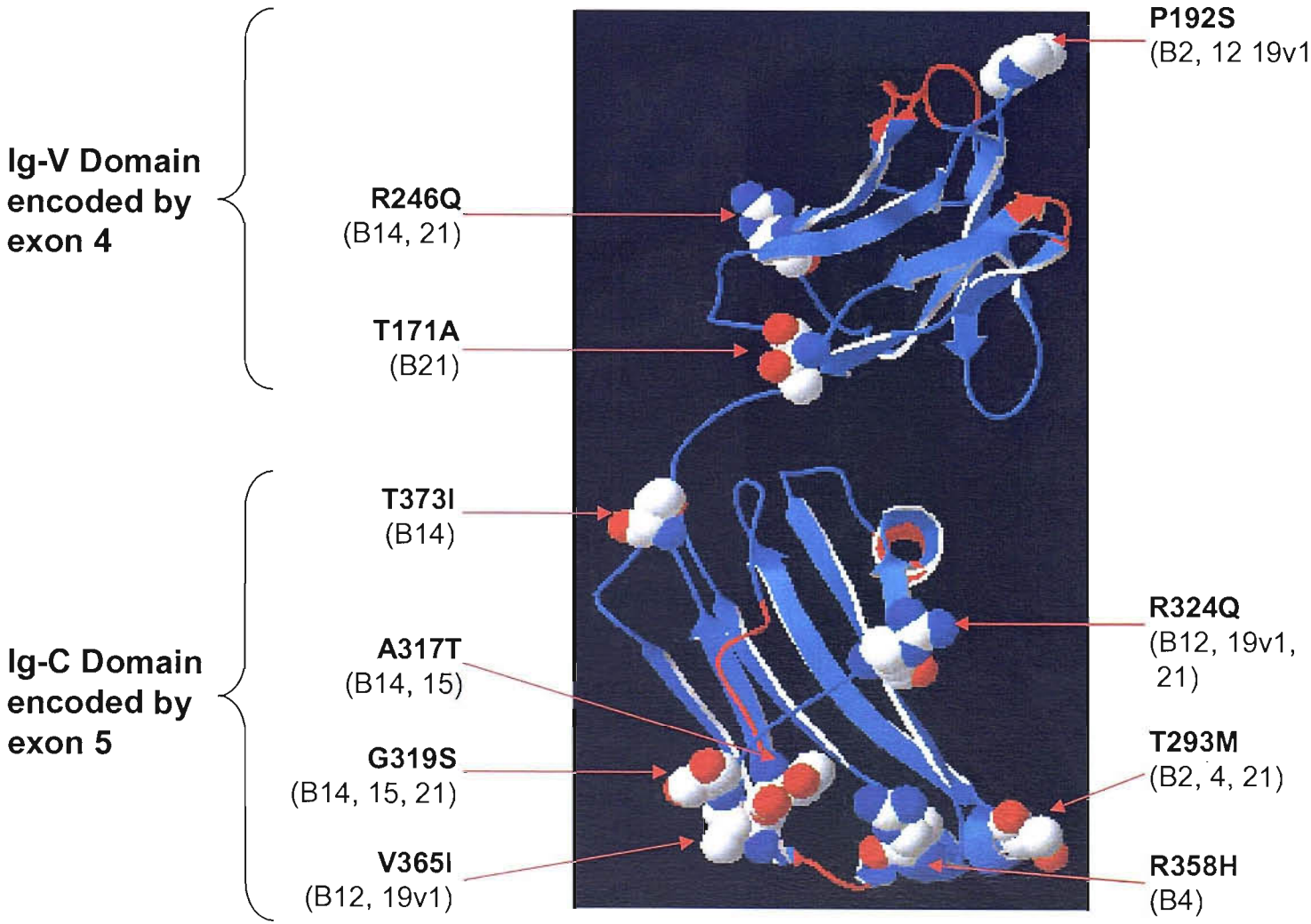


Figure 3.7 Location of polymorphic amino acid residues in the domains encoded by exons 4 and 5 of chicken tapasin.

Tapasin*15 protein sequence was modelled, with the position of amino acid polymorphisms found in any of the seven MHC haplotypes sequenced indicated according to their location in the primary sequence. The haplotypes in which the variant amino acids are present is indicated in parentheses. Polymorphic amino acids are shown in space filling form, with β strands and α helices shown as ribbons or coils. Red loops indicate poor alignment with template structures.

Figure 3.8 Comparison of the variant amino acid changes between chicken tapasin, MHC class I and TAP molecules.

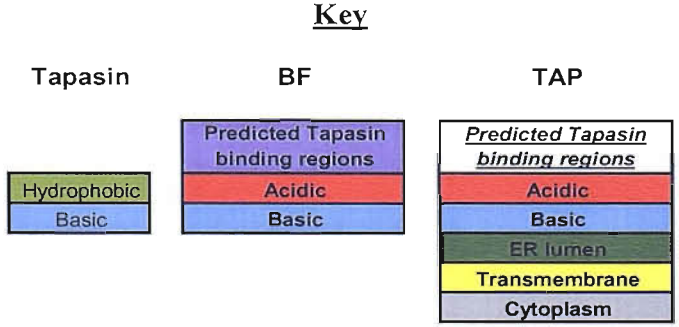
Each table shows only the amino acids which vary between haplotypes in tapasin, class I major (BF2), class I minor (BF1), TAP1 and TAP2. Amino acids positions are numbered in the top row of each table, with residue 1 being the initiating methionine for tapasin, TAP 1 and TAP2, and with residue 1 being the first residue of the mature protein for BF1 and BF2. MHC haplotypes are indicated in the left-hand column.

Positively charged (basic) amino acids are shown on a blue background, and negatively charged (acidic) amino acids are shown on a red background for all proteins. Hydrophobic amino acids are shown on a green background for tapasin. Amino acid positions in TAP1 and TAP2 predicted to reside within the ER lumen, in transmembrane domains, in the cytoplasm, or predicted to interact with tapasin are detailed as below. TAP sequence information taken from (Walker et al., 2005). MHC class I residues that align with residues 128-136 and 222-227 of mammalian MHC class I molecules were designated as regions predicted to bind tapasin (see appendix for alignment).

The last table provides a summary of possible interactions between tapasin and partner proteins.

Amino acid sequences can be found in the appendix as detailed below:

Tapasin:	Appendix 3,
BF2:	Appendix 7,
BF1:	Appendix 8,
TAP1:	Appendix 9,
TAP2:	Appendix 10



Exon	2			3	4			5							7
Allele	15	43	58	155	171	192	246	293	317	319	324	358	365	373	421
Tapasin*2	S	G	R	G	T	S	R	M	A	G	R	R	V	T	R
Tapasin*4	S	G	L	R	T	P	R	M	A	G	R	H	V	T	R
Tapasin*12	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*14	F	R	R	G	T	P	Q	T	T	S	R	R	V	I	Q
Tapasin*15	S	G	R	G	T	P	R	T	T	S	R	R	V	T	R
Tapasin*19	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*21	S	R	R	G	A	P	Q	M	A	S	Q	R	V	T	R

Allele	9	19	22	24	34	43	50	53	54	60	61	62	65	66	68	69	73	75	76	78	79	80	81
BF2*2	R	L	Y	D	V	Y	I	K	A	D	G	Q	I	G	G	N	D	E	N	G	I	L	Q
BF2*4	R	Q	F	T	V	Y	I	N	T	D	G	Q	I	G	L	N	N	E	N	G	I	R	Q
BF2*12	Q	Q	F	T	V	Y	I	K	A	D	G	Q	I	G	G	N	D	E	N	G	I	L	Q
BF2*14	Q	Q	F	T	V	V	M	N	T	N	G	Q	I	V	G	N	D	D	D	G	T	L	Q
BF2*15	S	Q	Y	D	T	A	I	N	T	D	S	E	T	S	R	T	D	D	G	G	T	L	Q
BF2*19	S	Q	F	D	T	A	I	N	T	D	S	E	T	S	R	S	D	D	G	G	I	L	Q
BF2*21	R	L	F	D	M	A	I	N	T	D	R	E	I	V	G	S	N	E	N	D	I	L	R

Allele	95	97	104	106	111	113	121	123	126	127	128	130	147	148	149	150	152	153	155
BF2*2	W	Y	G	P	Y	M	T	F	G	T	M	F	G	D	Y	A	G	L	Q
BF2*4	W	F	D	T	R	S	I	L	D	M	K	F	E	S	E	P	R	W	N
BF2*12	W	Y	G	P	Y	M	T	F	G	T	M	F	E	S	E	P	R	W	N
BF2*14	L	Y	D	T	S	D	I	F	G	T	M	F	G	D	Y	A	G	L	Q
BF2*15	L	Y	D	T	S	D	I	F	D	T	M	F	G	D	Y	A	G	L	Q
BF2*19	W	Y	D	T	R	Y	I	F	G	T	M	F	G	D	Y	A	G	L	Q
BF2*21	W	S	D	T	H	A	V	F	G	T	M	L	G	G	Y	A	G	L	Q

Allele	210	220	225	294	300	333
BF2*2	V	R	H	M	I	A
BF2*4	V	R	H	V	I	T
BF2*12	V	R	H	M	I	A
BF2*14	V	R	Q	V	T	A
BF2*15	A	Q	Q	V	I	A
BF2*19	A	R	Q	V	I	A
BF2*21	V	R	Q	V	I	S

Tapasin

Hydrophobic

Basic

BF

Predicted Tapasin binding regions

Acidic

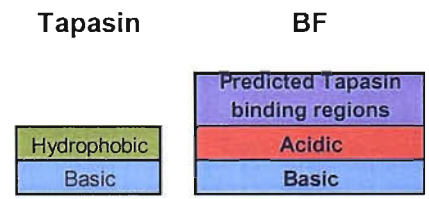
Basic

Exon	2			3	4			5							7
Allele	15	43	58	155	171	192	246	293	317	319	324	358	365	373	421
Tpn*2	S	G	R	G	T	S	R	M	A	G	R	R	V	T	R
Tpn*4	S	G	L	R	T	P	R	M	A	G	R	H	V	T	R
Tpn*12	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tpn*14	F	R	R	G	T	P	Q	T	T	S	R	R	V	I	Q
Tpn*15	S	G	R	G	T	P	R	T	T	S	R	R	V	T	R
Tpn*19	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tpn*21	S	R	R	G	A	P	Q	M	A	S	Q	R	V	T	R

Allele	4	8	9	19	22	34	43	50	53	54	61	62	66	68	71	72	73	74	75	76	78	79	80	82	84
BF1*2	T	I	S	Q	Y	T	A	I	N	T	R	E	A	G	Q	I	D	R	E	N	D	I	R	Q	H
BF2*4	T	I	H	Q	Y	V	Y	M	K	A	G	Q	G	R	R	S	V	K	V	S	D	T	L	E	Y
BF1*12	S	V	H	L	F	V	Y	M	N	T	G	Q	G	G	R	S	V	E	V	S	N	T	L	E	Y
BF1*19	S	V	H	L	F	V	Y	M	N	T	G	Q	G	G	R	S	V	K	V	S	N	T	L	E	Y
BF1*21	T	I	H	Q	Y	V	Y	M	K	A	G	Q	G	R	R	S	V	K	V	S	D	T	L	E	Y

Allele	93	95	97	111	113	115	118	123	124	125	126	127	128	149	152	153	155	164
BF1*2	A	W	Y	H	M	C	R	L	A	E	D	M	K	Y	R	K	Q	G
BF2*4	V	W	F	R	V	Y	K	F	D	K	D	M	K	V	G	W	S	W
BF1*12	V	L	Y	H	T	Y	R	F	D	K	G	T	M	V	R	W	S	G
BF1*19	V	L	Y	H	T	Y	R	F	D	K	G	T	M	V	R	W	S	G
BF1*21	V	W	F	R	V	Y	K	F	D	K	D	M	K	V	G	W	S	W

Allele	210	230	245	301	304	321	333
BF1*2	V	V	A	L	C	D	S
BF2*4	A	M	V	I	R	G	S
BF1*12	A	V	A	I	R	G	A
BF1*19	A	V	A	I	R	G	A
BF1*21	V	M	A	I	R	G	S



Exon	2			3	4			5							7
Allele	15	43	58	155	171	192	246	293	317	319	324	358	365	373	421
Tapasin*2	S	G	R	G	T	S	R	M	A	G	R	R	V	T	R
Tapasin*4	S	G	L	R	T	P	R	M	A	G	R	H	V	T	R
Tapasin*12	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*14	F	R	R	G	T	P	Q	T	T	S	R	R	V	I	Q
Tapasin*15	S	G	R	G	T	P	R	T	T	S	R	R	V	T	R
Tapasin*19	S	G	R	G	T	S	R	T	A	G	Q	R	I	T	R
Tapasin*21	S	R	R	G	A	P	Q	M	A	S	Q	R	V	T	R

Allele	3	54	92	116	131	162	179	289	327	437	446	455	487	537	579	580	582-584	578-580
TAP1*2	T	E	R	M	E	V	R	R	A	R	R	R	R	Q	del	S	del	
TAP1*4	T	K	R	M	K	A	R	R	A	R	R	R	R	Q	del	S	del	
TAP1*12	K	E	R	M	E	A	R	E	A	R	Q	R	R	E	I	A	VMD	RGW
TAP1*14	T	E	H	M	E	A	R	E	A	R	R	L	Q	Q	del	S	del	
TAP1*15	T	E	R	T	E	A	Q	E	A	R	R	R	R	Q	del	S	del	
TAP1*19	K	E	R	M	E	A	R	E	A	R	Q	R	R	E	I	A	VMD	RGW
TAP1*21	T	E	R	M	E	A	R	R	T	W	R	R	R	Q	del	S	del	

Allele	*9*	*37*	*64*	216	220	248	263	325	351	378	398	406
TAP2*2	*R*	*W*	*H*	R	I	A	A	D	H	R	R	S
TAP2*4	*R*	*W*	*R*	R	V	A	A	D	H	Q	R	S
TAP2*12	*R*	*W*	*R*	R	I	A	A	D	H	Q	H	G
TAP2*14	*H*	*W*	*H*	G	V	A	T	D	H	Q	R	S
TAP2*15	*H*	*S*	*H*	G	V	S	T	D	H	Q	R	S
TAP2*19	*R*	*W*	*R*	R	I	A	A	D	H	Q	H	G
TAP2*21	*R*	*W*	*H*	R	V	A	A	N	Y	Q	R	S

Allele	416	436	446	450	458	466	478	494	570	579	622	671
TAP2*2	K	A	N	A	V	V	R	G	E	R	H	I
TAP2*4	K	V	D	P	V	I	R	S	E	R	R	V
TAP2*12	K	A	N	A	V	V	R	G	E	R	H	I
TAP2*14	N	A	D	A	M	I	R	G	K	R	H	I
TAP2*15	N	A	D	A	M	I	R	G	E	R	R	I
TAP2*19	K	A	N	A	V	V	C	G	K	R	R	V
TAP2*21	K	V	D	A	V	V	R	G	E	K	R	V

Tapasin

Hydrophobic

Basic

TAP

Predicted Tapasin binding regions

Acidic

Basic

ER lumen

Transmembrane

Cytoplasm

Tapasin residues	Interacting protein	Residue	Comments
43, 246	BF2	B14: 76 B21: 78	Possible, see text
43, 246	BF2	B14: 60 B21: 61, 81	Possible, but alpha 1, and imperfect charge correlations
58, 155, 358	BF2	80	Possible, but alpha 1
58, 155, 358	BF2	128	Possible, see text
58, 155, 358	TAP1	54	Possible, but TAP1 Tapasin binding domain absent
58, 155, 358	TAP1	131	Dismissed, Tapasin residues ER lumenal, whilst TAP1 131 cytoplasmic,
58, 155, 358	TAP2	436	Dismissed, Tapasin residues ER lumenal, whilst TAP2 436 cytoplasmic
171	BF2	61, 78, 81	Possible, but alpha 1, and Tapasin 171 not charged
171	BF2	130	Possible, but Tapasin 171 not charged
192	BF1	53, 54, 68, 97	Possible, but Tapasin 192 not charged, and alpha 1, and B14 and B15 sequences absent
192	BF1	111, 118, 164, 230	Possible, but Tapasin 192 not charged, and BF1 residues outside of Tapasin binding domain and B14 and B15 sequences absent
293	BF1	4, 8, 19, 22, 78, 95	Possible, but alpha 1, and Tapasin 293 not charged, and B14 and B15 sequences absent
293	BF1	126-128	Possible, see text
293	BF1	333	Possible, but outside of Tapasin binding domain, and Tapasin 293 not charged, and B14 and B15 sequences absent
317	BF2	113	Possible, but Tapasin 317 not charged, and BF2 113 within peptide binding groove
317	TAP2	9, 416	Dismissed, Tapasin 317 not charged and ER lumenal whilst TAP2 residues transmembrane
317	TAP2	216	Dismissed, Tapasin 317 not charged and ER lumenal whilst TAP2 216 cytoplasmic
317	TAP2	458	Dismissed, Tapasin 317 not charged and ER lumenal whilst TAP2 458 cytoplasmic
365	TAP1	3, 446, 537, 579-end	Dismissed, TAP1 Tapasin binding domain absent, and Tapasin 365 not charged and ER lumenal, whilst all TAP1 residues are cytoplasmic
365	TAP2	398, 406	Possible, but outside of N-terminal domain, and not charged interaction
365	BF1	4, 8, 19, 22, 78, 95	Possible, but alpha 1, and Tapasin 365 not charged and B14 and B15 sequences absent
365	BF1	126-128	Possible, see text
365	BF1	333	Possible, but outside of Tapasin binding domain, and Tapasin 365 not charged, and B14 and B15 sequences absent
373	BF2	76	Possible, but alpha 1, and Tapasin 373 not charged
373	TAP1	92, 455, 487	Dismissed, TAP1 tapasin binding domain absent, Tapasin 373 ER lumena
421	TAP1	92, 455, 487	Possible, but TAP1 tapasin binding domain absent
421	TAP1	92, 455, 487	Possible, but TAP1 tapasin binding domain absent
421	BF2	76	Dismissed, Tapasin 421 cytoplasmic, whilst BF2 76 ER lumenal

Whilst a potential haplotype-specific interaction involving these residues may occur between the tapasin and BF2 proteins, the details of such an interaction are not known. It is possible that an interaction between these proteins is mediated through a combination of a complementary charge interaction occurring between amino acids on opposing proteins, and a second interaction involving a hydrophobic residue that is dependent on the charged interaction taking place. For example, for all haplotypes apart from B4, tapasin Arg58 may interact with an acidic residue on the BF2 molecule, with BF2 hydrophobic Met128 interacting with a specific residue on tapasin that is only accessible when the charge interaction occurs. It is possible that in the B4 haplotype a different interaction occurs, which may involve BF2 Lys128 interacting with an acidic residue on tapasin, with tapasin Leu58 interacting with a specific residue on BF2 that is only accessible when the charge interaction occurs.

Alternatively, it is possible that the basic charge of tapasin Arg58 could normally be neutralised by another amino acid within tapasin, perhaps with an acidic side chain, which may be in close proximity to Arg58 in the folded protein conformation. Met128 in the BF2 protein may interact with this neutralised region of the tapasin protein. In the Tapasin*4 protein the absence of the basic side chain at residue 58 may create an acidic surface which mediates a specific interaction with the complementary charge of Lys128 in the BF2*4 protein.

The polymorphism in the BF2 protein at residue 128 also coincides with unique polymorphisms in tapasin at residue 155, where the B4 haplotype has arginine instead of glycine (non-polar side chain), and at residue 358, where the B4 haplotype has histidine (weakly basic side chain) instead of strongly basic arginine.

Potential interactions 1b

Haplotype	BF2	Tapasin	Tapasin
B2, B12, B14, B15, B19v1, B21	M128	G155	R358
B4	K128	R155	H358

In contrast to the potential complementarities between residue 58 in tapasin and residue 128 in BF2 discussed above (potential interaction 1a), an interaction between residue 128 in BF2 and residue 155 in tapasin appears less likely; the occurrence of basic residues in the B4 haplotype at these positions in both proteins does not suggest a specific interaction occurs between these residues. The replacement of a strongly basic side chain with a weakly basic side chain at position 358 also appears unlikely to mediate an interaction with residue 128 in the BF2 protein.

A second pattern of complementarities exists (potential interaction 2a) involving tapasin residue 43, which is glycine in all haplotypes apart from B14 and B21, where arginine is found. This coincides with polymorphisms in the BF2*14 protein at residue 76, where either asparagine (uncharged, polar side chain) or glycine is replaced by aspartic acid (acidic side chain), or residue 78 in the BF2*21 protein where glycine is replaced by aspartic acid.

Potential interaction 2a

Haplotype	Tapasin	BF2	BF2
B2, B4, B12, B15, B19v1	G43	N76 / G76	G78
B14	R43	D76	G78
B21	R43	N76	D78

It is possible that in the B14 and B21 haplotypes, the basic side chain of tapasin Arg43 mediates an interaction with the complementary charge of Asp76 or Asp78 in the BF2 protein. Whilst in the other haplotypes, an interaction between tapasin and BF2 proteins may be mediated through the hydrophobic side chain of Gly43 of tapasin interacting with the hydrophobic side chains of Gly76 and Gly78, or perhaps with the uncharged, but polar side chain of Asn76 that is found in the B2, B4 and B12 haplotypes.

Alternatively (potential interaction 2b), the residues present in the BF2 protein at positions 76 and 78 may interact with residue 246 of tapasin (which is the amino acid that is also polymorphic in human tapasin). However, the nature of this potential molecular interaction is not as clear as with the previous potential interaction (2a). In the B14 and B21 haplotypes, glutamine (uncharged, polar side chain) is present at position 246 of tapasin, which may interact with the acidic side chain of Asp76 or Asp78 that are present in the BF2*14 or BF2*21 proteins. However, in the other haplotypes the interaction is less obvious, but may be mediated through the basic side chain of Arg246 in tapasin interacting with the BF2 protein either through the uncharged, polar side chain of Asn76 in the B2, B4 and B12 haplotypes, or the non-polar side chain of Gly76 that is present in the B15 and B19v1 haplotypes, or with Gly78 that is present in all haplotypes other than the B21 haplotype.

Potential interaction 2b

Haplotype	Tapasin	BF2	BF2
B2, B4, B12, B15, B19v1	R246	N76 / G76	G78
B14	Q246	D76	G78
B21	Q246	N76	D78

However, residues 76 and 78 lie in the α1 domain of the BF2 protein, and occupy positions that have so far not been implicated in the molecular interaction between mammalian tapasin and class I proteins (figure 3.9a). When the BF2*14 protein sequence was modelled against an assortment of mammalian class I molecules (figure 3.9b), it is evident that residue 76 points upwards from the α1 helix, whereas residue 78 points into the peptide-binding groove from the α1 helix. It is thought that the predominant site of interaction between class I and calreticulin proteins occurs via the N-linked glycan attached to the nearby asparagine 85 residue, and so it is unlikely (but not impossible) that the tapasin protein can associate with these residues in the class I molecule.

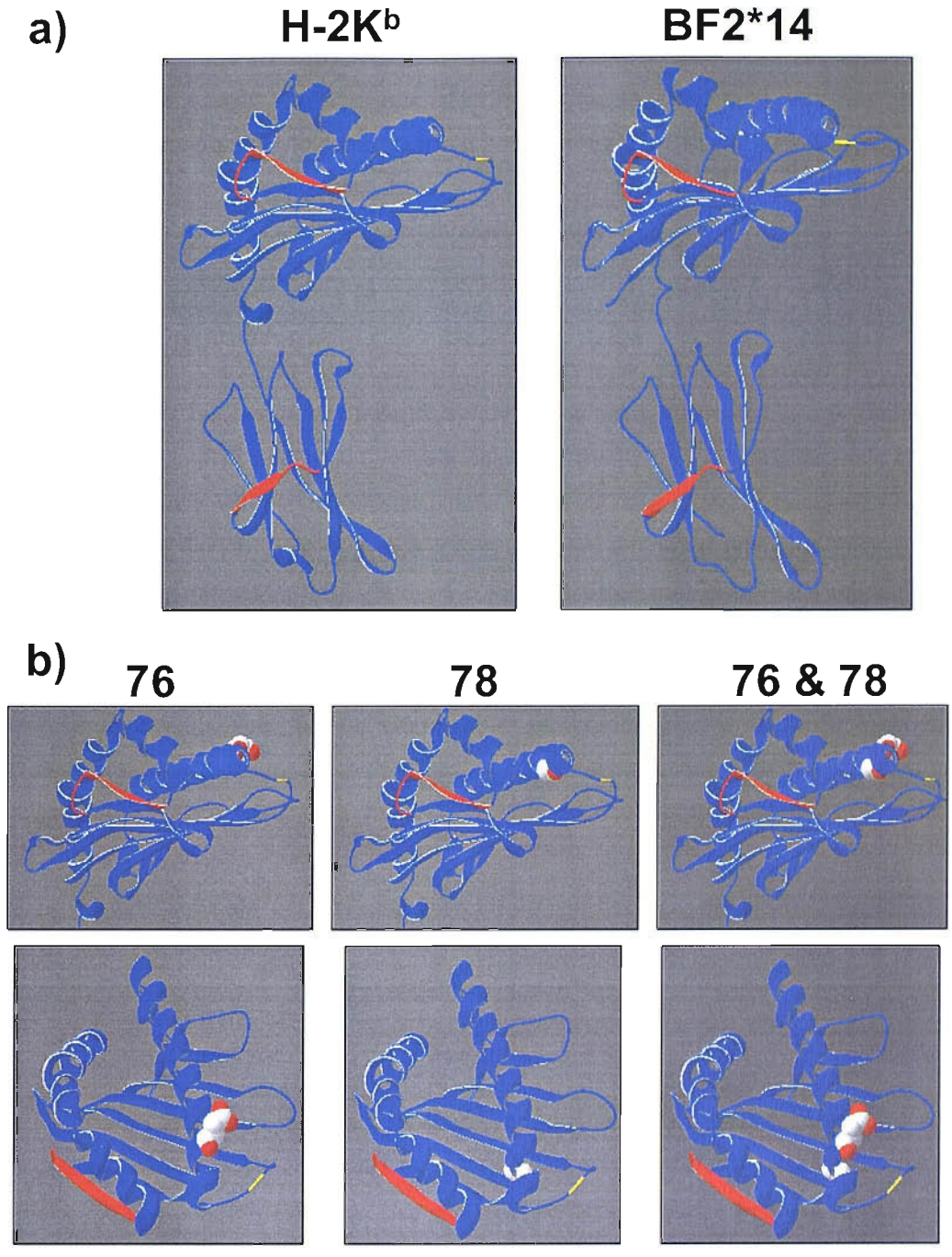


Figure 3. 9 3D structures of mouse and chicken MHC class I proteins

a) Models of two class I molecules in ribbon format: H-2K^b (left) and BF2*14 (right).

H-2K^b: Asn86 shown in yellow, $\alpha 2$ residues 128-136 and $\alpha 3$ residues 222-227 shown in red.

BF2*14: Asn85 is shown in yellow, $\alpha 2$ residues 125-133 and $\alpha 3$ residues 218-223 are shown in red.

b) BF2*14 structures. The location of residues 76 and 78 are shown individually or together (in space filling form). The top panels show the peptide binding groove viewed from the side, and the bottom panels show the view from above.

No obvious charge correlations exist between tapasin and the class I minor (BF1) gene. The BF1 genes are disrupted in the B14 and B15 haplotypes to the extent that no mRNA transcripts are produced, so these haplotypes were excluded from this analysis. However, two potential correlations exist between the two proteins, which both concern the B12 and B19v1 haplotypes.

Potential interaction 3a

Haplotype	Tapasin	BF1	BF1	BF1
B2, B4, B21	V365	D126	M127	K128
B12, B19v1	I365	G126	T127	M128

In the first example, residue 365 of tapasin is valine (non-polar, hydrophobic side chain) in all haplotypes other than B12 and B19v1, where isoleucine (non-polar, hydrophobic side chain) is present. This coincides with polymorphisms unique to the B12 and B19v1 haplotypes at positions 126, 127 and 128 of the BF1 protein, which occupy positions within the predicted tapasin-binding domain. The polymorphisms at these residues in the BF1 protein all involve the replacement of charged residues such as aspartic acid or lysine with either non-polar glycine or non-polar, hydrophobic methionine, or the replacement of methionine with uncharged, polar threonine. However, it appears unlikely that the conservative substitution of two non-polar hydrophobic amino acids in tapasin at residue 365 would mediate a specific interaction with these residues in the BF1 protein.

Perhaps a more likely interaction involving these residues of the BF1 protein concerns residue 293 of tapasin (potential interaction 3b).

Potential interaction 3b

Haplotype	Tapasin	BF1	BF1	BF1
B2, B4, B21	M293	D126	M127	K128
B12, B19v1	T293	G126	T127	M128

In the tapasin proteins that are encoded by the B12 and B19v1 (as well as the B14 and B15) haplotypes, the non-polar, hydrophobic side chain of Met293 is replaced by the uncharged, polar side chain of Thr293. This residue may mediate an interaction with some or all of the side chains of the residues that are present at positions 126, 127 and 128 of the BF1 protein. Of these potential interactions, an interaction between residue 293 of tapasin and residue 127 of BF1 is the most promising possibility, as there is a perfect correlation of amino acid replacements; such that in the B2, B4 and B21 haplotypes a hydrophobic interaction occurs, whilst in the B12 and B19v1 haplotypes a polar interaction occurs with residue 127.

The portions of the mammalian TAP heterodimer that mediate the interaction with tapasin have been mapped to the first four or three transmembrane loops of the TAP1 and TAP2 proteins respectively (the so called N-terminal domains) (Koch et al., 2004). The chicken TAP1 gene has a deletion which comprises this predicted tapasin-binding domain (Walker et al., 2005), and

therefore chicken tapasin may interact directly with only the TAP2 protein, although it is possible that chicken tapasin may interact with chicken TAP heterodimers in a different manner. However, no specific patterns of complementarities were observed between the tapasin and TAP proteins, which occurred within regions of the proteins that are likely to interact.

This approach has therefore predicted points of interaction between tapasin and the BF2 and BF1 proteins. However, tapasin might mediate specific interactions with the class I and TAP molecules that are not easily predicted using this approach. For example, tapasin may use the variable numbers of basic residues within the Ig domain encoded by exon 5 to mediate an interaction with either the class I or luminal portions of the TAP proteins.

3.2.8 Mutation of human tapasin at residue 48 modestly affects HLA-B*4402 surface expression characteristics

Of all of the correlations of polymorphic amino acids between tapasin and other MHC-encoded proteins, the most promising potential interaction involved residue 58 of tapasin and residue 128 of BF2. Residue 128 of the BF2 protein is likely to associate with tapasin directly, and the N-terminus of tapasin may contact this region, as deletion of the N-terminal 50 amino acids of human tapasin resulted in class I expression characteristics that were comparable to tapasin-deficient cells (Bangia et al., 1999). To determine whether this residue of tapasin mediates a specific interaction with the class I molecule, the residue in tapasin was mutated, and the expression of a class I molecule peptide-loaded in the presence of the mutated, wild-type or no tapasin was compared.

Tapasin-deficient chicken cell lines are not available, which prohibits the use of chicken cell lines for such transfection experiments; as all cells will express endogenous wild-type tapasin proteins. So as an alternative, the human tapasin-deficient .220 cell line stably transfected with HLA-B*4402 (.220B*4402) was chosen to determine whether the mutation of the equivalent residue of human tapasin affected the expression of B*4402, a class I allele that is highly dependent upon tapasin for its efficient surface expression.

From the alignment of chicken and human tapasin proteins (figure 3.6) it can be seen that arginine 48 is the human tapasin residue equivalent to arginine 58 in chicken tapasin. Therefore this residue was mutated to leucine (the variant amino acid identified in the chicken haplotypes), glutamic acid (the opposing charge), or alanine (a neutral amino acid) and expression constructs comprising the mutated human tapasin constructs (labelled hTpn R48E, R48A or R48L) were transfected into the .220B*4402 cell line.

Initially the .220B*4402 cell line was transiently transfected with either wild-type human tapasin or the position 48 mutants, and the increase in HLA-B*4402 surface expression level was assessed the next day by W6/32 staining (that detects folded HLA: β_2m heterodimers) and flow cytometry.

Table 3.8 reveals the W6/32 staining levels that were found in four duplicate experiments, and figure 3.10 shows the relative expression levels of tapasin protein in the transiently transfected cells. Between the experiments there was some considerable variation in the extent that the position 48 tapasin mutants increased W6/32 staining in comparison to wild-type human tapasin transfection. This may be potentially explained by the variable tapasin protein expression levels (assuming that mutation to tapasin residue 48 does not affect recognition by the anti-human tapasin antibody that was used in this experiment; which was raised against the N-terminal 66 amino acids). In all of the experiments, it appeared that the R48E mutant was over-expressed in comparison to the wild-type protein. Despite this, in three of the four experiments, the R48E mutant increased W6/32 staining levels to a slightly lesser extent than the wild-type protein, and acted comparably to wild-type tapasin in the other experiment. The R48A mutant appeared to be expressed at more similar levels to the wild-type protein, and increased W6/32 staining levels to a slightly lesser extent than wild-type tapasin in two experiments, but acted comparably to wild-type tapasin in the other two experiments. The R48L mutant was expressed in smaller amounts than wild-type tapasin in three experiments, which led to decreased W6/32 staining levels in two experiments, but comparable staining levels to wild-type tapasin in the third, which was also found when R48L was expressed in similar or greater amounts to wild-type tapasin in the remaining experiment. Clearly accurate analysis of these results is hampered by the variable expression of tapasin in the transiently transfected cells, but it can be concluded that all of the position 48 mutants did increase the expression of class I heterodimers to levels that were not grossly different from the wild-type protein.

To characterise the effect of these mutations in more detail, efforts were made to generate stable position 48 mutant human tapasin transfectants. However, few stable transfectants were generated: only two clones expressing hTpn R48E, and one clone expressing hTpn R48A survived antibiotic selection. No clones were generated expressing hTpn R48L.

Flow cytometry and western blot analysis was conducted using these stable transfectants. All tapasin transfectants, including the R48E and R48A stable transfectants, displayed massively increased levels of HLA: β_2m heterodimers (as detected by the W6/32 antibody) in comparison to the tapasin-deficient .220B*4402 cell line (table 3.9). The W6/32 expression level of the position 48 mutants appeared greater than that of the .220B*4402hTpn clone expressing wild-type tapasin, which may be explained by the greater expression level of the R48E and R48A proteins in these transfectants (figure 3.11) than of wild-type tapasin protein in the .220B*4402hTpn cell line (both R48E clones gave comparable results in terms of W6/32 staining and protein expression level, data not shown).

Expt 145					
Clone	**DNA transfected**	**Number of events**	**W6/32 MFI**	**Fold increase relative to hTpn transfection**	
.220B*4402hTpn untransfected		n.d.	n.d.	n.d.	**Estimated tapasin protein expression level relative to hTpn t/f**
.220B*4402	mock	47	663.47	0.60	
	7ug hTpn	2026	1110.37	1.00	
	7ug hTpn R48E	1875	982.17	0.88	Overexpressed
	7ug hTpn R48A	1450	939.11	0.85	Similar?
	7ug hTpn R48L	1587	824.54	0.74	Underexpressed
Expt 147					
.220B*4402hTpn untransfected		7804	1695.68	2.34	**Estimated tapasin protein expression level relative to hTpn t/f**
.220B*4402	mock	72	426.7	0.59	
	7ug hTpn	1338	723.48	1.00	
	7ug hTpn R48E	2792	684.93	0.95	Overexpressed
	7ug hTpn R48A	1924	549.16	0.76	Similar?
	7ug hTpn R48L	994	513.66	0.71	Underexpressed
Expt 148					
.220B*4402hTpn untransfected		7284	786.81	1.76	**Estimated tapasin protein expression level relative to hTpn t/f**
.220B*4402	mock	121	301.81	0.67	
	7ug hTpn	3673	448.07	1.00	
	7ug hTpn R48E	5767	392.08	0.88	Overexpressed
	7ug hTpn R48A	4782	444.42	0.99	Similar / overexpressed?
	7ug hTpn R48L	4603	409.48	0.91	Similar / overexpressed?
Expt 152					
.220B*4402hTpn untransfected		10392	443.9	2.20	**Estimated tapasin protein expression level relative to hTpn t/f**
.220B*4402	mock	89	91.92	0.46	
	7ug hTpn	2092	201.5	1.00	
	7ug hTpn R48E	1804	222.73	1.11	Overexpressed
	7ug hTpn R48A	2210	200.38	0.99	Similar?
	7ug hTpn R48L	1175	194.09	0.96	Underexpressed
	Untransfected	107	98.24	0.49	

Combined experiments		Average fold increase relative to hTpn t/f	Std Dev
.220B*4402hTpn untransfected		2.10	0.31
.220B*4402	mock	0.58	0.09
	hTpn	1.00	0.00
	hTpn R48E	0.95	0.11
	hTpn R48A	0.90	0.12
	hTpn R48L	0.83	0.12

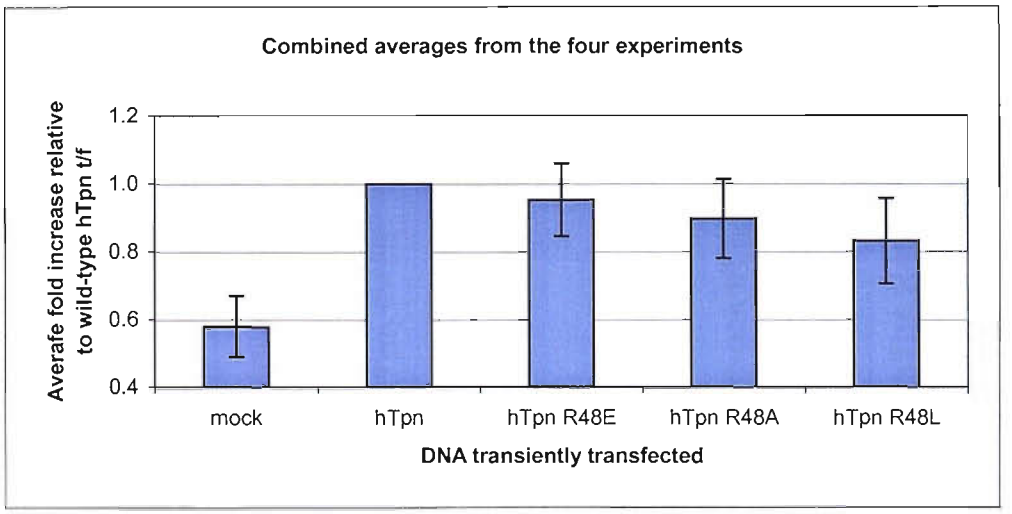


Table 3.8 Surface expression levels of HLA: β_2m heterodimers following transient transfection with either wild-type or position 48 point mutants of human tapasin.

Four experiments were performed in which the .220B*4402 cell line was transiently transfected with either wild-type or position 48 point mutants of human tapasin, or were mock transfected. In each experiment, only those cells that gave greater than background fluorescence following W6/32 staining and flow cytometry were included in the analysis; with the number of cells (events), and their mean fluorescence being tabulated. In each experiment, the expression level of the transfected tapasin protein was estimated following SDS PAGE and western blotting (figure 3.12). A table and graph summarising the results of the four experiments is provided.

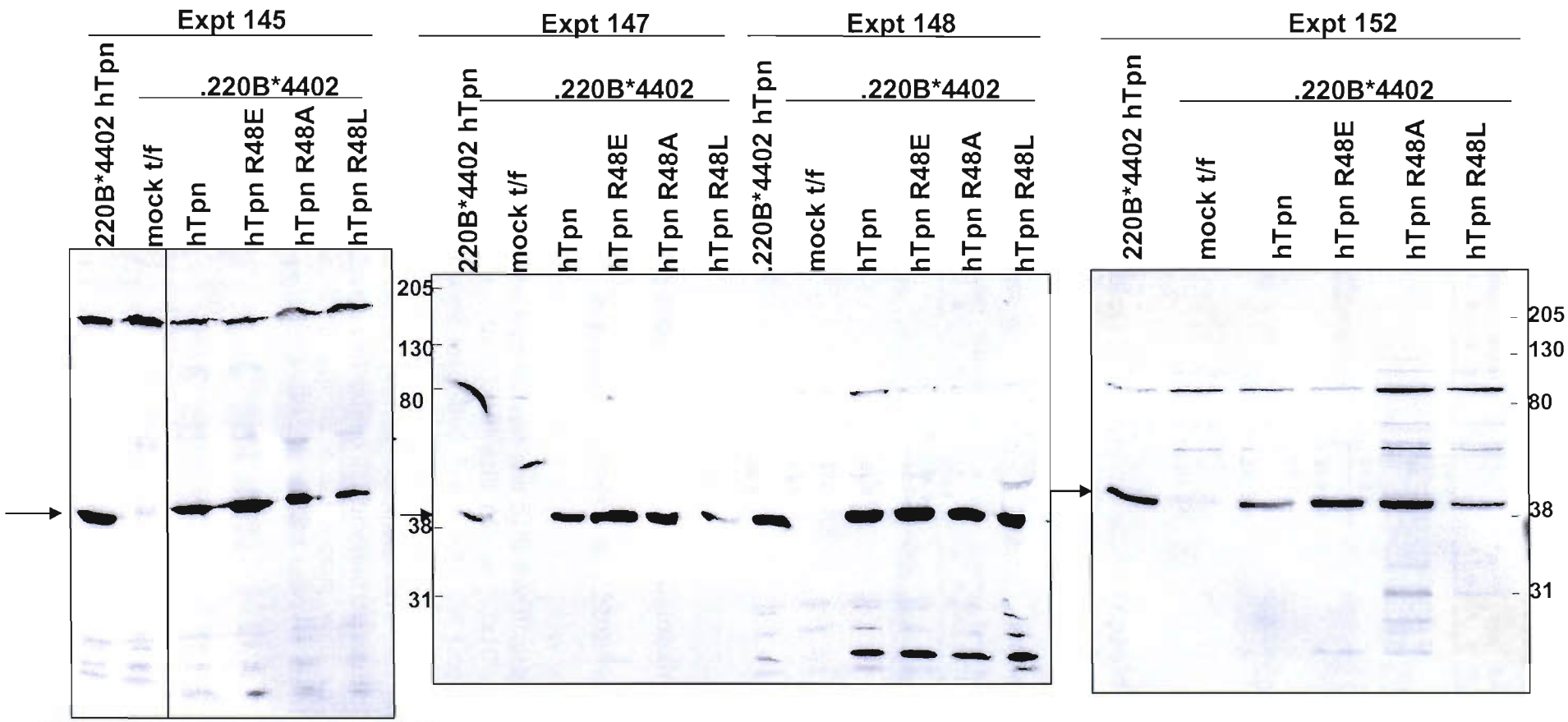


Figure 3.10 Western blots depicting expression level of human tapasin protein following transient transfection

In each experiment lysates were made, and a volume equivalent to 10^4 cells was loaded into each lane and separated by SDS PAGE (except for experiment 145, where a volume equivalent to 10^5 cells was used). Following western blotting, the blot was incubated with anti human tapasin antibody and developed. Arrowheads indicate human tapasin protein. Molecular weight markers are indicated on one side of the blots.

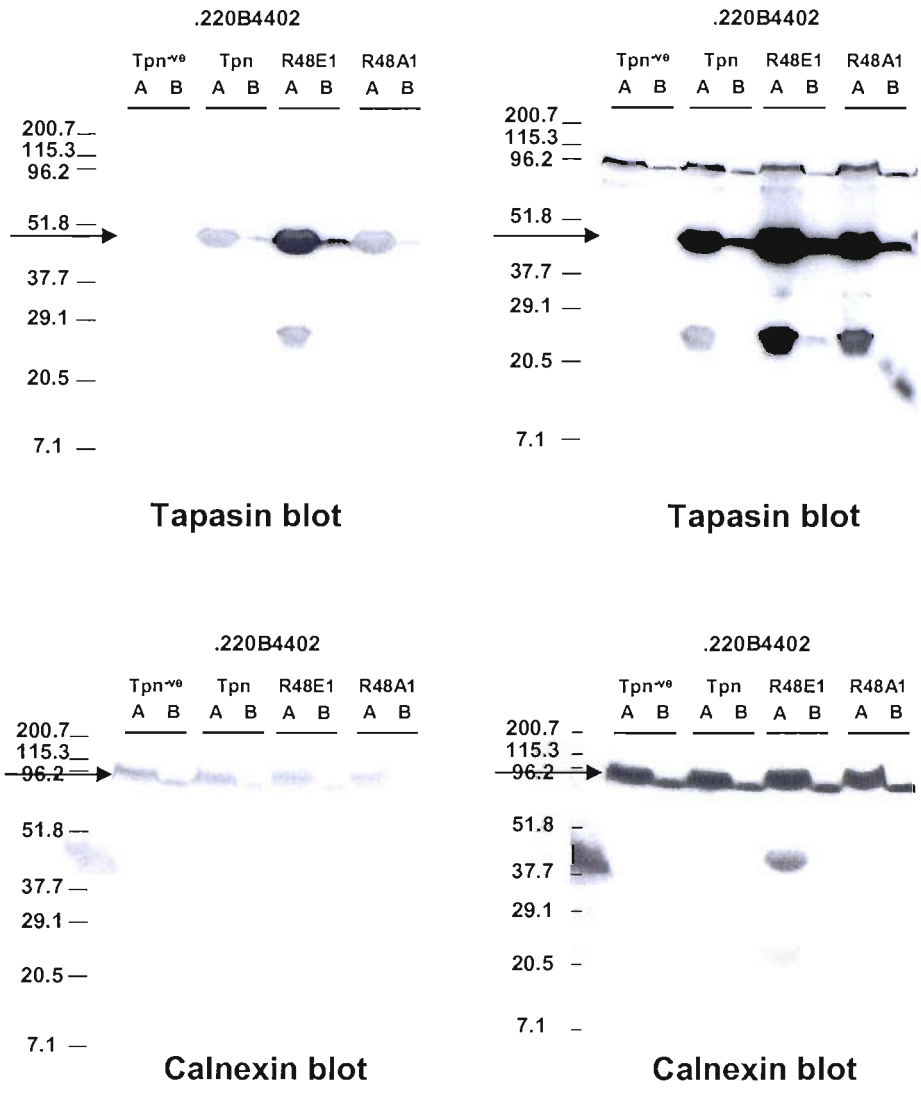


Figure 3.11 Tapasin expression levels in .220B*4402 tapasin transfectants.

Lysates were made of the transfectants and SDS PAGE gels were loaded with cell lysate equivalent to $3.3x10^5$ cells (A) or $3.3x10^4$ cells (B) and western blotted. Blots were incubated firstly with the Giles 11/11 anti-human tapasin antibody and developed. Two exposures are shown, with an arrowhead marking the tapasin protein. The blots were then subsequently incubated with anti-calnexin antibody to demonstrate the amount of protein loaded into each lane. Two exposures are shown, with an arrowhead marking the calnexin protein. Molecular weight markers are shown with masses in kDa.

Tpn^{-ve} = .220B*4402

Tpn = .220B*4402hTpn

R48E1 = .220B*4402hTpn R48E clone 1

R48A1 = .220B*4402hTpn R48A clone 1

Clone	W6/32 Average MFI	W6/32 St Dev	Joan-1 Average MFI	Joan-1 St Dev	TT4-A20 Average MFI	TT4-A20 St Dev	HC10 Average MFI	HC10 St Dev
.220B4402	13.59	0.11	3.09	0.02	2.41	0.02	3.33	0.04
.220B4402 hTpn	134.14	1.44	35.70	0.79	93.96	1.46	5.58	0.11
.220B4402 hTpn_R48E1	160.45	3.89	53.96	1.32	143.73	3.11	8.48	0.23
.220B4402 hTpn_R48A1	149.63	1.21	50.11	0.10	148.94	1.38	8.39	0.04

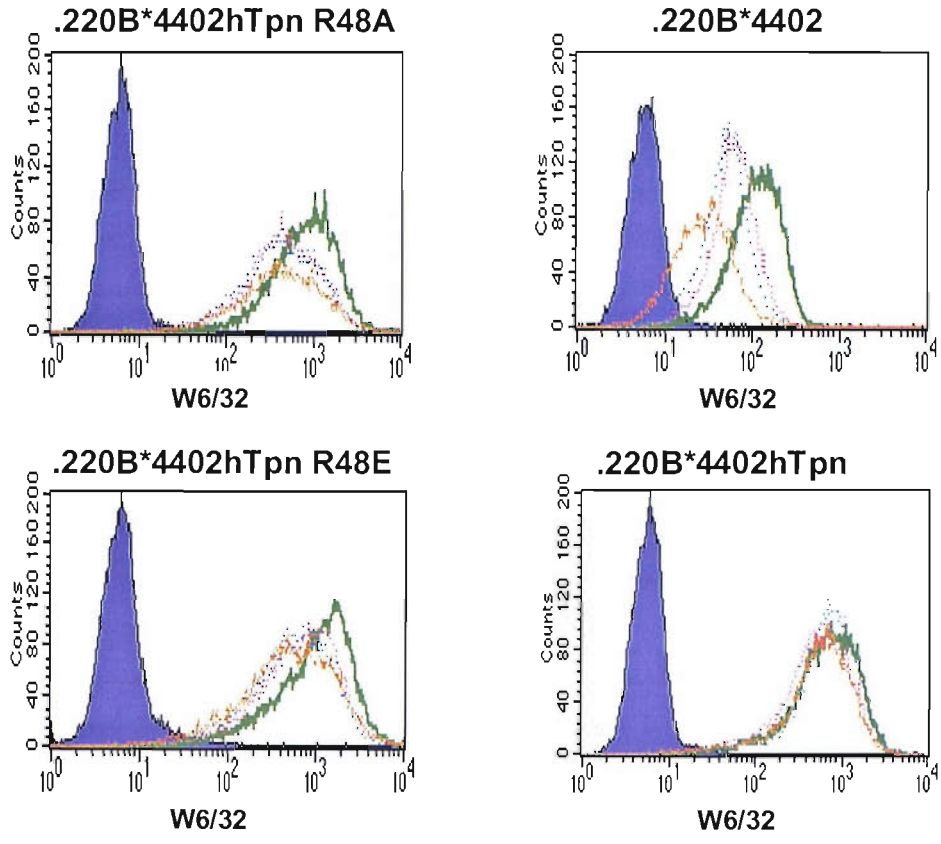
Table 3.9 Analysis of the surface class I expression levels of the position 48 mutant human tapasin stable transfectants.

As the W6/32 antibody detects all HLA: β_2m heterodimers, it will detect the HLA-C allele that is endogenous to the .220 cell line, and whose expression increases approximately ten-fold following tapasin expression (Tony Williams, personal communication). Therefore in order to study the effects of the R48E and R48A tapasin mutations on the expression level of HLA-B*4402, two antibodies specific to HLA-B: Joan-1 and TT4-A20 were used to stain these transfectants. TT4-A20 staining has previously been observed to be sensitive to the peptide repertoire loaded into these molecules (Tan et al., 2002). Table 3.9 reveals that for both antibodies the steady state expression level of B*4402 on either R48E or R48A expressing transfectants was greater than that seen on wild-type tapasin expressing clone.

The cells were also stained with HC10, an antibody that reacts with β_2m-free heavy chains. The expression level of free heavy chains appeared to be modestly greater in the R48E and R48A expressing transfectants than in the wild-type tapasin expressing clone, which was increased marginally above that of the tapasin-deficient clone. It would therefore appear that the expression level of free heavy chain correlates with the total expression level of HLA molecules.

These findings were confirmed and extended by Dr. Tony Williams (whose results have been kindly provided, figure 3.12). Using these stable transfectants, he confirmed the slightly increased surface expression levels of class I heterodimers on the R48E and R48A expressing clones (as detected by W6/32) relative to the clone expressing wild-type tapasin. However, with both of the position 48 mutant tapasin expressing clones, he found that during the first four hours of culture with brefeldin A, the rate of decay of surface HLA-B*4402 molecules was much greater than was evident in the wild-type tapasin expressing clone, and was comparable to tapasin-deficient cells. However, during the next 12 hours of culture with BFA, the position 48 mutant expressing clones decayed to a significantly smaller extent than the tapasin-deficient cells, which was nevertheless much greater than was apparent for the wild-type tapasin expressing clone. This suggests that the repertoires of peptides that were loaded in the presence of the position 48 mutant tapasin proteins was a substantial improvement over the peptide repertoire loaded in the absence of tapasin, but induced a lower level of stability on the class I molecules than the repertoire loaded by the wild-type tapasin protein. This finding is consistent with the possibility that the human tapasin position 48 mutant proteins bridged the HLA-B*4402 molecules to TAP, but optimising their peptide repertoire to a lesser extent than wild-type tapasin.

a)



b)

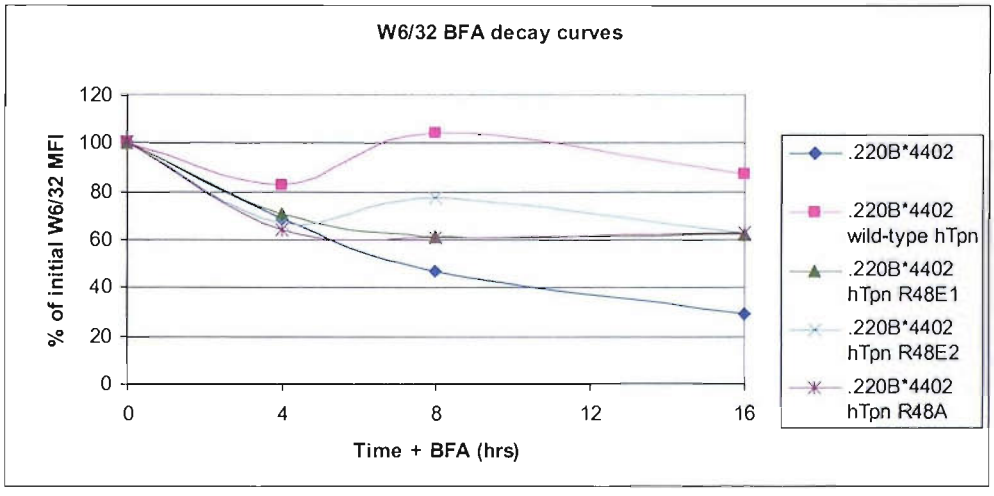


Figure 3.12 .220B*4402 brefeldin A surface decay rates

The indicated clones were cultured in the presence of brefeldin A for various lengths of time.

a) Histograms depict W6/32 expression levels when cultured for 0 (green, thick), 4 (pink dotted), 8 (blue dotted) or 16 (orange) hours.

b) Graph showing W6/32 decay rates as a percentage of the initial MFI when the individual clones were cultured for 0 hrs in BFA.

Data kindly provided by Dr. Tony Williams (University of Southampton)

3.3 Discussion

At the onset of this investigation it was not known whether chicken tapasin was polymorphic. Therefore, this chapter concerned the determination of the extent of allelic polymorphism and sequence diversity in the chicken tapasin gene and protein. Nucleotide sequences were obtained for the tapasin gene from the seven haplotypes available at IAH Compton; analysis of these sequences revealed that in contrast to the limited polymorphism in tapasin evident in typical mammals such as humans, the chicken tapasin gene in these seven haplotypes differs at the nucleotide level, and also at the amino acid level. In total, there are 15 polymorphic amino acid positions, with the tapasin proteins encoded by these haplotypes being distinct from each other (apart from the B12 and B19v1 tapasin proteins which are identical), and differ from a consensus sequence by up to seven amino acids,.

3.3.1 Mechanisms for the generation and maintenance of polymorphisms in tapasin and other MHC encoded genes

It is likely that the polymorphisms in chicken tapasin were generated through random mutational errors that occurred during meiosis (for example through errors in DNA replication or repair, or following DNA recombination). However it is not enough for a polymorphism to simply arise in one animal; somehow the new allele must rise in gene frequency to avoid becoming lost from the population by chance. There are at least three principle routes through which this could have occurred.

The first mechanism is that of natural selection, where the mutation allows the product of the new allele to function in a more optimal fashion. For example, new class I alleles may encode proteins that are able to bind and present pathogen-derived peptides that the products of other class I alleles were unable to. This may mean animals carrying the new allele would be able to clear a pathogen infection more efficiently, and are therefore more likely to survive, to be able to breed in the future. An example of how natural selection could lead to the positive selection of amino acid polymorphisms in tapasin could be that selective pressure is applied upon either the class I molecule to bind different peptides, or the TAP heterodimer to translocate different peptides, with the new class I or TAP molecule that is generated additionally having the drawback of being unable (or less able) to interact with tapasin productively. Therefore new tapasin alleles that subsequently appear that could restore the molecular interactions would be positively selected. Alternatively it is possible that new tapasin alleles may have been positively selected in order to escape pathogen subversion; there are several examples of pathogens using the mammalian PLC as a means of decreasing MHC class I restricted antigen presentation (Bennett et al., 1999; Park et al., 2004).

A second mechanism for the maintenance of new alleles in MHC-encoded genes may be through genetic "hitch-hiking". For example, a new tapasin allele may increase in gene frequency as a

consequence of being genetically linked to another polymorphic gene that is under strong selective pressure (taking a class I gene as an example). In such a scenario it could be envisaged that any beneficial polymorphisms in the linked class I gene would be positively selected. If polymorphisms were also to occur, or had previously done so in the tapasin gene, then due to the low incidence of recombination between the linked genes, the new tapasin allele will also rise in gene frequency. This would make no assumption about the effect of the polymorphisms in the linked gene that is under less selective pressure than the "driving" polymorphic gene: the polymorphisms may be neutral, beneficial or even adversely affect the ability of the protein to function. In the chicken MHC, genes such as the class I major BF2, or class IIB major BLB2, may be expected to be under high levels of selection, and may drive increases in gene frequency of other linked genes through such hitch-hiking.

A third mechanism by which new alleles can rise in gene frequency is through genetic drift, or processes like the "founder effect". For example, a particular chicken carrying a new polymorphism may have, by chance, been heavily used to breed a new stock of birds, forcing the new allele to rise in frequency. In this situation, protein polymorphisms may be neutral, adverse or beneficial.

However, it is possible that a combination of the above mechanisms may have collectively sustained a new allele in the population; whereby a portion of polymorphisms may have been maintained through processes such as genetic drift or hitch-hiking, that may have the net effect of not affecting how the protein functions, whilst other polymorphisms may have been positively selected in order for the gene product to function more optimally.

3.3.2 Potential implications concerning amino acid polymorphisms in chicken tapasin

Exactly how tapasin enhances class I peptide-loading is yet to be determined. The novel polymorphic amino acid residues in chicken tapasin may identify functionally important regions of the molecule that may reveal how tapasin functions.

However, it is important to note that the polymorphic amino acid residues in chicken tapasin may be located in regions of the molecule that are permitted to vary and which do not affect the ability of tapasin to function. Such functionally neutral polymorphisms are likely to be maintained in the population through processes other than natural selection. If the amino acid polymorphisms do not alter the ability of tapasin to enhance peptide-loading, then the products of different tapasin alleles would be expected to function interchangeably, and would at the least, reveal regions of tapasin that are not functionally relevant.

However, some (or all) of the polymorphic amino acid residues in chicken tapasin may have a profound influence on the ability of tapasin to facilitate peptide-loading. Arguments concerning how polymorphic amino acids in tapasin might influence class I peptide-loading are made below.

Given the compact size and low incidence of recombination within the chicken MHC, and the polymorphism in tapasin, TAP and class I molecules, it is possible that alleles of these genes have co-evolved. In such a scenario, polymorphic amino acids in tapasin may mediate a specific interaction with the class I and TAP molecules that are encoded by alleles from the same haplotype. The consequence of tapasin participating in a "haplotype-specific" interaction may be an advantage in optimal peptide-loading (i.e. preferential loading of those peptides that have an appropriate length and specificity to bind to the class I molecule with high affinity and perhaps more crucially a slow dissociation rate). For example, this may involve an enhanced ability of tapasin to chaperone the class I molecule, and edit the peptide repertoire, producing class I molecules with long half lives. Alternatively, this may involve an enhanced ability of tapasin to interact with TAP, leading to increased import of peptides into the ER. If the combination of alleles of tapasin, TAP and class I that are usually found in a haplotype were to become mixed, such as would occur following a recombination event, then the interaction between molecules may either not occur or be impaired to such an extent that peptide-loading of class I molecules is adversely affected.

Alternatively, the polymorphic amino acid residues in chicken tapasin may influence the ability of tapasin to associate with another protein (other than the class I or TAP molecules) which has a role in class I peptide-loading. This intermediary protein may be the chicken orthologues of calreticulin or ERp57 which both perform critical functions in mammalian class I peptide-loading. However, it would seem reasonable to assume that in order to specifically interact with the products of tapasin alleles, the intermediary protein would also have to be encoded by a gene within, or linked to, the chicken MHC. The genes encoding chicken calreticulin and ERp57 have not been located precisely in the chicken genome to date, but would be presumed not to be within or closely linked to the MHC, as is the case in mammals. However, polymorphic amino acids in tapasin may influence an association with another protein whose contribution to class I peptide-loading has not been described yet, which may be unique to chickens.

However, the polymorphic amino acid residues in chicken tapasin may have a profound influence on the association between tapasin and an intermediary molecule which is not a protein, but may instead be a peptide. For example, tapasin may fulfil its role in class I peptide-loading by acting as a peptide-binding chaperone, which physically transfers peptides directly into the peptide-binding grooves of class I molecules. In support of this is the evidence that tapasin can be immunoprecipitated bound to a cross-linkable, radioactive peptide (Li et al., 1997), however this finding has not been reproduced by other laboratories. If tapasin is responsible for physically

transferring peptides into the peptide-binding grooves of class I molecules, polymorphisms in tapasin may influence the repertoires of peptides that are loaded into class I proteins, and may be specifically tailored to suit the peptide-binding requirements of the dominantly expressed class I molecule. It would seem reasonable to assume that the N-terminal domain of tapasin would be most likely to be involved in such a process, as deletion of the N-terminal 50 amino acids of human tapasin resulted in a stable protein that was unable to facilitate peptide-loading (Bangia et al., 1999). However, there is relatively little polymorphism evident in the exons encoding the N-terminal domain of chicken tapasin, which may suggest that if such a mechanism of action occurs, then amino acid polymorphisms in tapasin may not have a significant influence on class I peptide-loading.

An alternate idea concerning how tapasin facilitates peptide-loading is that tapasin may act in a role similar to that of the invariant chain in class II peptide-loading. It is possible that part of tapasin physically occupies the peptide-binding groove of empty class I molecules, and is only displaced by high affinity peptides that can form stable interactions with the class I molecule. Potentially, polymorphic amino acids in chicken tapasin may have a direct influence on the ability of tapasin to enter or remain within the peptide-binding groove of class I molecules, and may have evolved to complement the dominantly expressed class I molecule. However, if this mechanism of action occurs, then it might have been expected that there would be a greater incidence of tapasin-derived peptides being eluted from class I molecules, as is the case concerning invariant chain-derived peptides being eluted from class II molecules. In addition, if such a mechanism were to operate it is again reasonable to assume that the N-terminal domain of tapasin would be the most likely region of tapasin to be involved, and it is likely that a greater extent of polymorphism than is evident would be needed in order for tapasin to enter the peptide-binding grooves of class I alleles with different binding specificities. Therefore this possible mechanism is unlikely to operate.

Therefore, of the possibilities discussed above, the strongest two appear to be: either that the polymorphic amino acids do not affect the ability of tapasin to enhance peptide-loading; or that they mediate haplotype-specific interactions with the class I and TAP molecules.

3.3.3 Can we determine if polymorphic amino acids in tapasin influence class I peptide-loading?

Given the low rate of recombination within the chicken MHC, any new tapasin allele that rises in gene frequency and is maintained in the population through any mechanism (including natural selection, genetic drift or hitch-hiking) is unlikely to be exchanged by recombination and potentially lost from the population. Therefore the presence of polymorphic amino acid residues does not automatically imply that they have been positively selected because they conferred an improved ability to function. Therefore three approaches were taken in an attempt to gain insight into the potential consequence of amino acid polymorphisms in tapasin.

The first approach was to assess the ratio of nonsynonymous to synonymous (N/S) nucleotide substitutions. Regardless of the mechanism by which tapasin enhances peptide-loading, if the polymorphisms in different tapasin molecules allow a peptide-loading function that is beneficial for the class I molecules expressed by the same haplotype, then these polymorphisms are likely to have been positively selected, and therefore a N/S ratio in excess of 2:1 would be evident (as nucleotide polymorphisms in the first two positions of a codon would be expected to be more highly selected that at the third "wobble" position). However, the N/S ratios were found to be less than 2:1 in all the tapasin alleles. This suggests that the majority of the tapasin gene is under selective pressure to remain invariant (so called purifying selection).

The low N/S ratio may be a consequence of the necessity for the majority of the tapasin protein to remain invariant. This may be in order to form the framework required for the correct conformation, or may constitute an invariant site of interaction with another protein. Alternatively, the low N/S ratio may be due to nucleotide substitutions that do not alter the protein sequence that accumulate through processes other than natural selection, giving a N/S ratio that is lower than that expected of genes under strong selective pressure. However, a small number of polymorphic amino acid residues may still dramatically change the ability of tapasin to function (in any of the possible ways discussed above). In such a situation, the total N/S ratio may not exceed 2:1, but some of the nucleotide substitutions that alter the protein sequence may still have a dramatic influence on class I peptide-loading.

The second approach involved a phylogenetic analysis of the tapasin alleles. This revealed that the products of the tapasin alleles produced a different topology to that observed for other MHC-encoded proteins which are not thought to be under the pressures of selection (such as the BF1 class I and BLB1 class II proteins). This suggests the tapasin alleles have not arisen through an accumulation of random polymorphisms from the last common ancestor of the haplotypes through evolutionary drift, as is thought to have occurred in most of the chicken MHC-encoded genes. In addition, the topologies of the phylogenetic trees produced for tapasin and BLB2 proteins were dissimilar, suggesting the polymorphisms in tapasin are not the product of genetic hitch-hiking driven by the neighbouring BLB2 gene.

Some similarities were evident in the topologies of the tapasin protein tree and those of the BF2 class I or TAP2 proteins. This may indicate that the majority of the polymorphic residues in these proteins have co-evolved, and that these proteins are involved in a haplotype-specific interaction. In each case however, the topologies were not identical. It is likely that in order for phylogenetic trees constructed for co-evolving proteins to share absolutely identical topologies, all of the polymorphisms within the proteins must be involved in mediating the specificity of the molecular co-evolution. However, it is perhaps more likely that this might not be the case; whilst some

polymorphic residues may mediate specific protein interactions (where perhaps as few as one polymorphic residue is critically important for the interaction), other polymorphic residues (perhaps many more) may exist that do not affect molecular function or protein interactions (that could be maintained in the population through other mechanisms, such as evolutionary drift). Therefore, this might explain why the topologies of phylogenetic trees do not match exactly, even if the proteins specifically interact through co-evolution of polymorphic amino acid residues.

However, the possibility that tapasin is involved in a haplotype-specific interaction with the TAP2 protein is not directly supported by the distribution of the polymorphic amino acid residues in these proteins. Assuming chicken tapasin and TAP interact through the same sites as has been determined in mammals, there appears to be an insufficient amount of polymorphic residues in these regions to mediate such an interaction. However, it remains possible that residues besides those previously identified can contribute to a haplotype-specific interaction occurring between the chicken tapasin and TAP2 proteins. If so, the clustered polymorphic amino acid residues at the base of the membrane proximal Ig domain encoded by exon 5 may be likely to mediate such an interaction.

The phylogenetic trees that were constructed for the TAP1, BF1 and BLB proteins did not share similar topologies to that of tapasin, suggesting that these proteins are not involved in a haplotype-specific interaction with tapasin. As the predicted tapasin-binding N-domain of the chicken TAP1 protein is deleted, the dissimilarity in the topologies of the phylogenetic trees that were constructed for TAP1 and tapasin strengthens the argument that specifically interacting proteins share similar phylogenetic tree topologies.

These two approaches cumulatively provide clues that are consistent with a model where the majority of tapasin does not vary and may be needed for the protein to function correctly (accounting for the low N/S ratio). In addition, there are other polymorphic residues of which most, but possibly not all, are involved in mediating specific molecular interactions with the class I (and possibly the TAP2) molecules that are encoded from alleles present on the same haplotype (explaining the less than perfect correlation of phylogenetic tree topologies).

3.3.4 Prediction of amino acid residues that may mediate a specific interaction with other MHC-encoded molecules

As a third approach to gain insight into the potential consequence of amino acid polymorphisms in tapasin, an analysis of the polymorphic amino acids in tapasin, BF1, BF2, TAP1 and TAP2 proteins was conducted. This approach could potentially reveal individual residues that may mediate a specific interaction between polymorphic proteins encoded by the same haplotype. Although many patterns of complementarities were found, two patterns of charge

complementarities were observed between tapasin and the BF2 class I proteins that were particularly interesting.

The first of these involved a correlation of complementary charges between tapasin and residues 76 and 78 of the BF2 protein, which are located in the α helix of the α1 domain in close proximity to the site of N-linked glycosylation at residue asparagine 85. However, three points can be made why it could be considered unlikely that these residues mediate a potential interaction with tapasin. Firstly, although not proven, it seems unlikely that tapasin interacts with this portion of the class I molecule; as tapasin appears to bind predominantly to the α2 and α3 surfaces and not with the α1 domain. Secondly, if calreticulin does bind predominantly to the mono-glucosylated N-linked glycan attached to Asn85, it may be difficult for tapasin to physically reach these residues due to steric hinderance. Thirdly, there has only been one report of a peptide sequence derived from tapasin being eluted from a class I molecule suggesting that tapasin does not enter the peptide-binding groove (Turnquist et al., 2001). Therefore it appears unlikely that tapasin does interact with residues 76 and 78 that are within, or are in intimate proximity with the peptide-binding groove.

The second potential point of interaction between tapasin and BF2 molecules involved tapasin residue 58 and BF2 residue 128. This possibility is particularly attractive for two reasons. Firstly, the N-terminal domain of tapasin may associate with the class I molecule directly, as deletion of the N-terminal 50 amino acids of human tapasin resulted in failure to facilitate peptide-loading (Bangia et al., 1999). Secondly, BF2 residue 128 occupies a position within a loop in the α2 domain that is predicted to mediate a direct interaction with the mammalian tapasin protein. Indeed when the equivalent residue was mutated in H-2L^{d} (K131D), an increased association was observed between the K131D mutant and the endogenous peptide-loading complex (Yu et al., 1999). This also led to the K131D mutant being less receptive to loading with exogenous peptides, which suggests the increased association with the PLC led to an enhanced loading of endogenous peptides.

As it was not directly possible to determine whether chicken tapasin residue 58 and BF2 residue 128 are involved in a specific interaction, the equivalent residue in human tapasin (arginine 48) was mutated to leucine, glutamic acid, or alanine. Whilst the position 48 mutant tapasin proteins were capable of increasing the surface expression level of class I: β_2m heterodimers to an extent similar to the wild-type tapasin protein, the repertoires of peptides that were loaded in the presence of the position 48 mutant tapasin proteins appeared to stabilise the class I heterodimers to a slightly lesser extent than the wild-type protein. However it is important to note that the position 48 mutant tapasin proteins substantially improved peptide-loading in comparison to tapasin-deficient cells, suggesting that human tapasin position 48 mutants were capable of optimising the peptide repertoire, albeit less efficiently than wild-type tapasin.

A more prominent phenotype may have been observed if a different class I allele were used: B*4402 enjoys a very strong association with the PLC, and therefore a more noticeable effect may have been observed with a class I allele that associates to a lesser extent with the human PLC. In addition, the residue that occupies the equivalent position in B*4402 to BF2 residue 128 is neither hydrophobic (such as methionine) nor basic (such as lysine), and is instead polar (serine 131). In human class I molecules, amino acid residue 131 is dimorphic (serine or arginine). It is possible that mutations to residue 48 in human tapasin affect class I molecules with basic arginine at this position to a greater extent than class I molecules with polar serine. Therefore it would be of interest to determine the effect that position 48 mutant human tapasin proteins have on tapasin-dependent R131 alleles, such as HLA-B*0801, or on a HLA-B*4402 S131M mutant. Additionionally, in the chicken, tapasin alleles might co-evolve with the class I alleles, whereby a combination of polymorphic residues may be needed to shape the correct conformation of the tapasin protein. Therefore residue 58 in chicken tapasin may mediate a specific interaction with residue 128 in the BF2 molecule only in this setting, which is not conserved in the tapasin and class I proteins of other species, such as the human.

3.3.5 Conclusion

This chapter has shown chicken tapasin to be much more polymorphic than its mammalian counterparts, with six protein alleles being found in seven MHC haplotypes. There are a total of 15 polymorphic amino acids which may influence the ability of tapasin to function. In these haplotypes the protein sequences differ from a consensus protein by up to seven amino acids, with the majority of the polymorphic residues being located in the ER luminal domains, with a notable clustering of polymorphic residues in the Ig domain encoded by exon five.

The ratio of nonsynonymous: synonymous nucleotide substitutions suggests there may be pressure for the majority of tapasin to remain invariant, but despite this it remains possible that a small number of polymorphic amino acids result in the products of tapasin alleles functioning distinctly. The topology of the phylogenetic tree constructed using the tapasin protein sequences suggests the polymorphisms are not the result of genetic drift occurring over the entire MHC region, and are not likely to be a consequence of hitch-hiking driven by the neighbouring BLB2 gene. Comparison of the topologies of phylogenetic trees constructed for the tapasin, BF2 and TAP2 proteins suggests that tapasin may be evolving to interact with these molecules in a haplotype-specific interaction. Additionally an analysis of complementary patterns of polymorphic amino acids between MHC-encoded proteins revealed potential points of interaction between the tapasin and BF2 protein; however, an attempt to prove one such point of interaction was not conclusive.

3.3.6 In later chapters

Having found sequence polymorphisms in tapasin, attention was then focused on determining whether these polymorphic amino acids influence the ability of tapasin to facilitate optimal

peptide-loading of class I molecules. The interaction between tapasin and class I was chosen because the majority of amino acid polymorphisms in tapasin are located within the ER luminal domains, and are therefore more likely to influence the interaction with the class I molecule as opposed to the TAP heterodimer, which is thought to extend minimally into the ER lumen.

The molecular consequences of comparing haplotype-matched or haplotype-mismatched tapasin and class I alleles was initially studied using the human tapasin-deficient .220 cell line (chapter four). Whilst this approach was ultimately unsuccessful in determining the functional relevance of the polymorphisms in tapasin, it is included in this thesis as several findings regarding molecular incompatibilities were made that offer insight into the differences between chicken and human class I antigen presentation.

In chapter five attempts were made to reconstitute the appropriate alleles of tapasin, TAP, or both tapasin and TAP into a chicken MHC homozygous cell line transfected with a haplotype-mismatched epitope-tagged class I allele; which may reveal the influence that each protein has on peptide-loading the haplotype-mismatched class I allele. This approach proved unsuccessful; due to difficulties transfecting the appropriate TAP or tapasin molecules into the cells. However, it is included as several observations were made that influenced the design of the next research chapter.

Chapter six also concerns a MHC homozygous chicken cell line transfected with a haplotype-mismatched epitope-tagged class I allele. However in this case, the choice of haplotypes was guided by the presumed sequence of events that led to the generation of a rare recombinant haplotype. In these transfectants, the haplotype-mismatched class I allele was expected to be compatible with the endogenous TAP transport specificity, but may have been incompatible with the endogenous tapasin allele. This chapter ultimately provides evidence that specific residues of the class I molecules, which are naturally polymorphic in chickens, influence the ability of the class I molecule to associate with tapasin, and that peptide-loading occurs sub-optimally if the haplotypes are mismatched.

4. Transfection of the human tapasin-deficient .220 cell line with MHC-matched or MHC-mismatched chicken class I and tapasin alleles

4.1 Introduction

In the last chapter, the chicken tapasin gene from seven MHC haplotypes was sequenced and shown to encode six moderately diverse alleles of the tapasin protein, with only the B12 and B19v1 haplotypes encoding identical proteins. This chapter concerns an attempt to characterise the functional relevance of these polymorphisms. Polymorphic amino acid residues may affect the function of tapasin in a number of ways, some of which are discussed below.

Firstly, the polymorphic residues may be located in regions of tapasin that can vary without influencing the ability of the protein to function. In this case the polymorphic residues would be functionally neutral, and tapasin alleles would function interchangeably.

Secondly, some or all of the polymorphic residues may be located where they dramatically alter the ability of the protein to function irrespective of MHC haplotype. For example a polymorphic residue may constitute part of a site of interaction with a non-MHC-encoded protein. Therefore the products of certain tapasin alleles may be capable of facilitating peptide-loading to a greater extent than other alleles, leading to tapasin alleles that are superior in function to other alleles.

Thirdly, the compact size and low rate of recombination within the chicken MHC makes it possible that each tapasin allele has evolved to facilitate the optimal peptide-loading of the class I allele encoded by the same haplotype, via the proteins interacting in a "haplotype-specific" interaction. If the tapasin, TAP and class I proteins are involved in such a specific interaction, the interaction between products of haplotype-mismatched alleles may either not occur, or lead to less optimal peptide-loading than occurs between the products of haplotype-matched alleles. This possibility is supported by the phylogenetic analysis of the tapasin, TAP2 and BF2 proteins that was conducted in the last chapter, which suggested that a majority of the polymorphic residues in these alleles may reflect co-evolution occurring between tapasin, class I and TAP molecules.

Therefore in this chapter the expression characteristics of class I molecules that were expressed with either haplotype-matched or haplotype-mismatched tapasin alleles were compared. This approach should allow the determination of whether tapasin alleles function interchangeably (i.e. the polymorphic residues are functionally neutral), or whether certain tapasin alleles facilitate optimal peptide-loading to a greater extent than other alleles (i.e. one or more "superior" tapasin allele), or whether the products of tapasin alleles participate in a haplotype-specific interaction leading to optimal peptide-loading of haplotype-matched combinations of alleles.

The ideal method to determine this would be to use a tapasin and/or class I deficient chicken cell line that could be reconstituted by transfection with different combinations of tapasin and class I alleles. Unfortunately such cell lines have not yet been generated, and therefore all chicken cell

lines will express endogenous alleles of tapasin, TAP and class I molecules which may complicate the analysis of any proteins transfected into such cells.

Therefore this chapter concerns the use of the 721.220 human tapasin-deficient cell line (referred to as .220 hereafter) to investigate the functional consequences of molecular interactions occurring between haplotype-matched or haplotype-mismatched combinations of chicken tapasin and class I molecules. The interaction between chicken tapasin and class I proteins was chosen because the majority of polymorphic amino acids in tapasin are located within the ER luminal domains, which are more likely to influence the interaction with the class I molecule as opposed to the TAP heterodimer.

4.1.1 The 721.220 cell line

The human tapasin-deficient .220 cell line provided significant insight into the function of tapasin, which was later extended with tapasin knock-out mice. The .220 cell line was generated by γ-irradiation and selection against surface expression of HLA-A and –B alleles (Greenwood et al., 1994). The cell line is hemizygous for chromosome 6 (which encodes the human MHC); the remaining chromosome 6 has HLA–A and –B loci deleted, and crucially has a single nucleotide polymorphism in tapasin located in the 5' splice site of the second intron resulting in exon 2 being spliced out of the protein (Copeman et al., 1998) (figure 4.1).

The .220 cell line does express a tapasin molecule with an altered signal sequence that is consequently poorly translocated into the ER lumen (Copeman et al., 1998). The absence of exon 2 results in a truncated tapasin protein missing the first N-terminal 49 amino acids of the mature protein, which has a negligible ability to enhance peptide-loading (Bangia et al., 1999).

Without tapasin, human class I molecules that have been transfected into .220 cells do not associate with the TAP heterodimer, and do not gain a cargo of optimal peptides. The majority of human class I molecules are retained in the ER, destined for degradation; the few class I molecules that reach the surface are unstable, and are more receptive to stabilisation by exogenous peptides than class I molecules loaded with the assistance of tapasin (Greenwood et al., 1994; Grandea et al., 1995; Copeman et al., 1998). Transfection of human tapasin cDNA into .220 cells increases class I surface expression levels and stability, and restores antigen presentation to normal (Ortmann et al., 1997). Some human class I alleles expressed in .220 cells are less dependent upon tapasin to achieve high expression levels and can present endogenous peptides to CTL (Peh et al., 1998); however, transfection of human tapasin qualitatively improves the stability of the class I molecules, and their ability to present peptides to CTL.

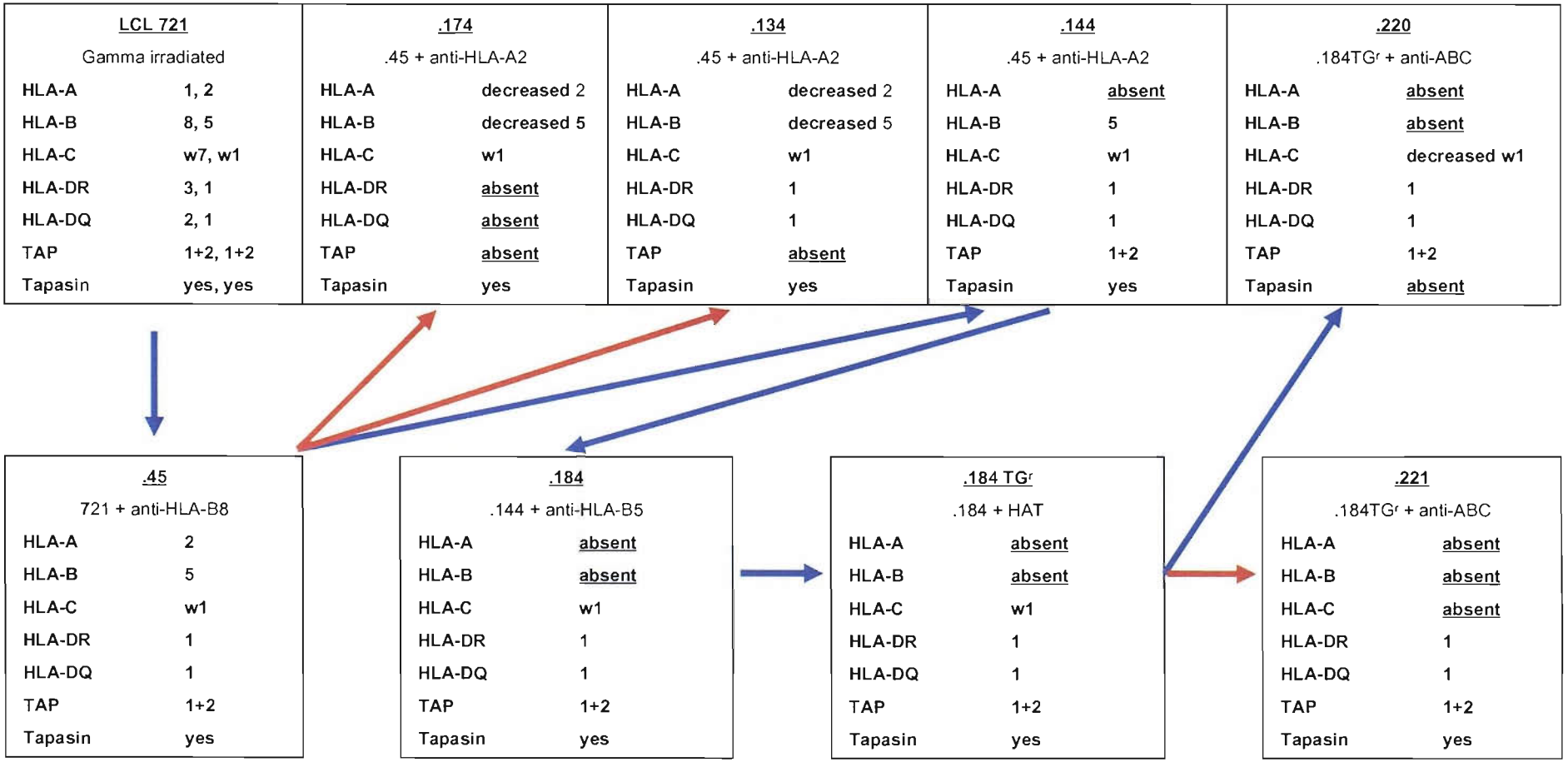


Figure 4.1 Schematic depicting the derivation of the 721.220 cell line

There are examples of murine class I molecules being expressed successfully in the .220 cell line. For example in one study, H-2L^d and epitope-tagged H-2K^d, H-2K^b molecules were successfully expressed in .220 cells and the effects of tapasin determined (Myers et al., 2000). In a separate study the effect that tapasin had on the expression characteristics of H-2K^b was determined (Barnden et al., 2000). In this study murine H-2K^b molecules were not retained within the ER (as human class I molecules generally are in .220 cells), but rapidly progressed to the cell surface, to be expressed at high levels. However, the surface expressed H-2K^b molecules exhibited defects consistent with sub-optimal peptide-loading, as described previously. The subsequent transfection of human tapasin slowed the rate of maturation of H-2K^b, and rectified the defects evident in its absence. These studies provide evidence that class I molecules from other species (at least from the mouse) can be expressed in .220 cells, and that the effects of tapasin can still be determined. However, it may be important to note that the experimental situation does not produce identical results to what may be the case if the proteins were expressed in cell lines of the correct species: for example murine class I molecules expressed in murine tapasin-deficient cells egress to the cell surface at the same rate as in tapasin-expressing murine cells.

4.1.2 Strategy

The experimental strategy employed in this chapter was to generate transfectants of .220 cells expressing a particular BF2 allele, such as BF2*12. The resulting transfectant cell line was then reconstituted with either the appropriate haplotype-matched allele of tapasin, such as Tapasin*12, or an inappropriate haplotype-mismatched tapasin allele, such as Tapasin*15. A summary of the transfectant cell lines that were produced is provided in figure 4.2, and is discussed in detail in the next part of this chapter.

Various approaches were taken that may reveal quantitative or qualitative differences in the repertoires of peptides loaded into the class I molecules in the presence of haplotype-matched or haplotype-mismatched tapasin alleles. The experimental approaches that were taken included:

- **Comparison of surface expression levels**: expression of the appropriate allele of tapasin may have resulted in an increase in the surface expression level of BF2 relative to the tapasin-negative cell line. Comparison of the surface expression levels evident in the absence of tapasin, or presence of the appropriate allele of tapasin, to that obtained with cells expressing the haplotype-mismatched allele of tapasin was conducted to reveal whether the haplotype-mismatched tapasin allele could function comparably to the haplotype-matched allele.

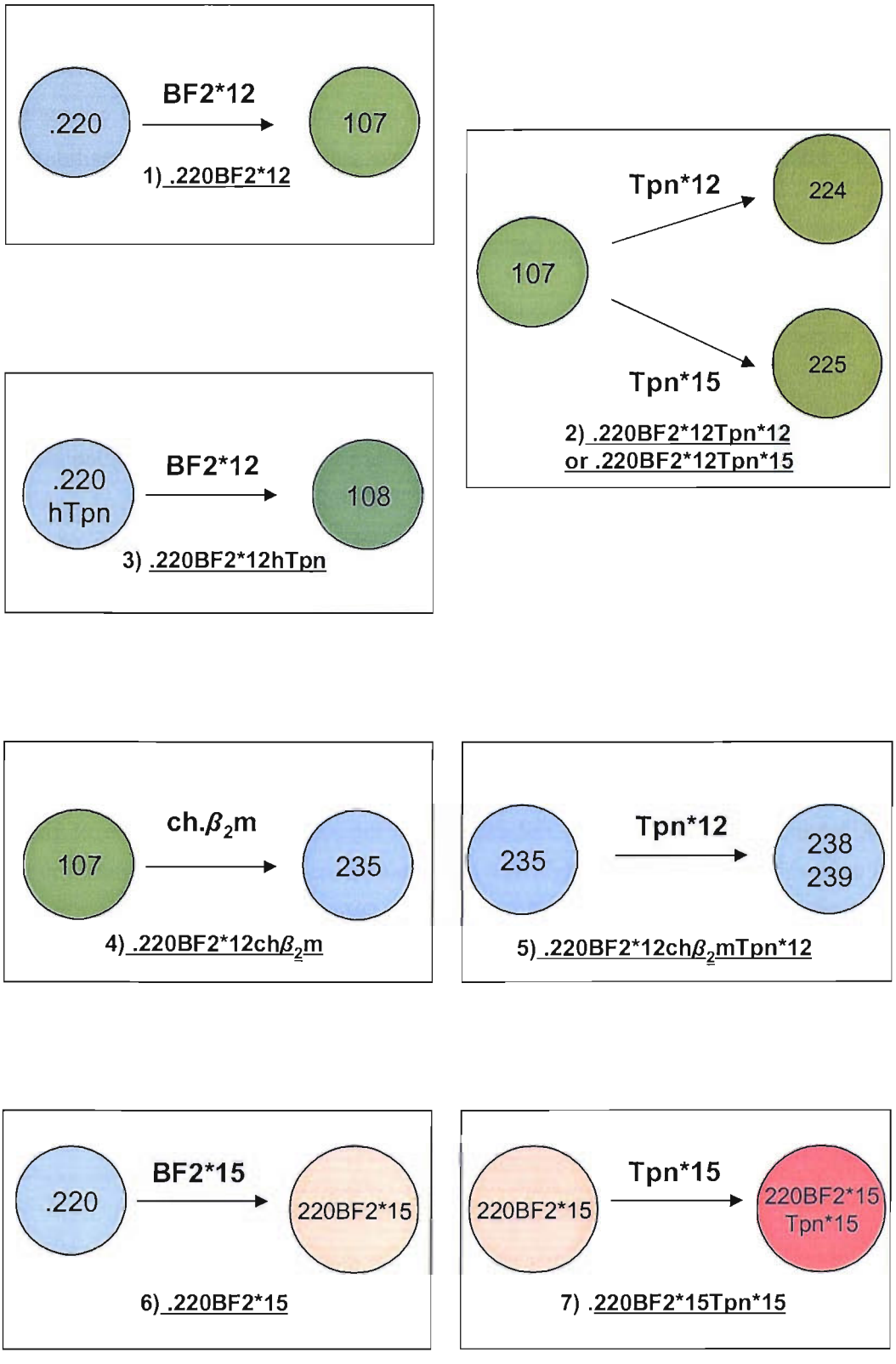


Figure 4.2 Summary of the transfectants that were produced.
Cells are indicated by circles, with cell line suffix shown inside the circle (e.g. 107 = .220 transfected with BF2*12). The transfectants are presented in the order that they are discussed in this chapter.

- **Exogenous peptide stabilisation assays**: whereby exogenous high affinity peptides are added to cells in culture. Surface expressed class I molecules that have exited the ER with a low affinity peptide which exhibits a fast dissociation rate are likely to exchange the existing low affinity peptide for the exogenous high affinity peptide. This may stabilise class I molecules that would otherwise dissociate, increasing the surface expression level, and could reveal qualitative differences in peptide repertoires loaded by either the haplotype-matched or haplotype-mismatched tapasin protein.

- **Comparison of surface expression levels following culture at 37°C versus room temperature**: when cells are cultured at room temperature this can permit the export from the ER to the cell surface of class I molecules that are loaded with low affinity peptides that would otherwise not populate the cell surface at 37°C (Ljunggren et al., 1990; De Silva et al., 1999). Therefore a comparison of the proportion of molecules that populate the cell surface at 37°C as opposed to room temperatures was conducted that could reveal qualitative differences in the peptide-loading conducted by haplotype-matched or haplotype-mismatched tapasin proteins.

- **Rates of surface class I molecule decay following incubation with brefeldin A** (BFA): BFA blocks the transport of newly synthesised proteins to the cell surface. Therefore when cells are cultured with BFA their class I molecules disappear from the surface as the existing molecules dissociate but are not replaced by new molecules. Class I molecules that dissociate at a rapid rate are likely to represent unstable molecules loaded with low affinity peptides, as opposed to the slower degradation of stable class I molecules loaded with higher affinity peptides. Therefore BFA treatments were conducted that may reveal a qualitative difference in peptide-loading conducted by haplotype-matched or haplotype-mismatched tapasin proteins.

4.2 Results

This chapter concerns the transfection of the human tapasin-deficient .220 cell line with chicken proteins, in an attempt to determine whether the polymorphic residues identified in chicken tapasin in the previous chapter are functionally relevant.

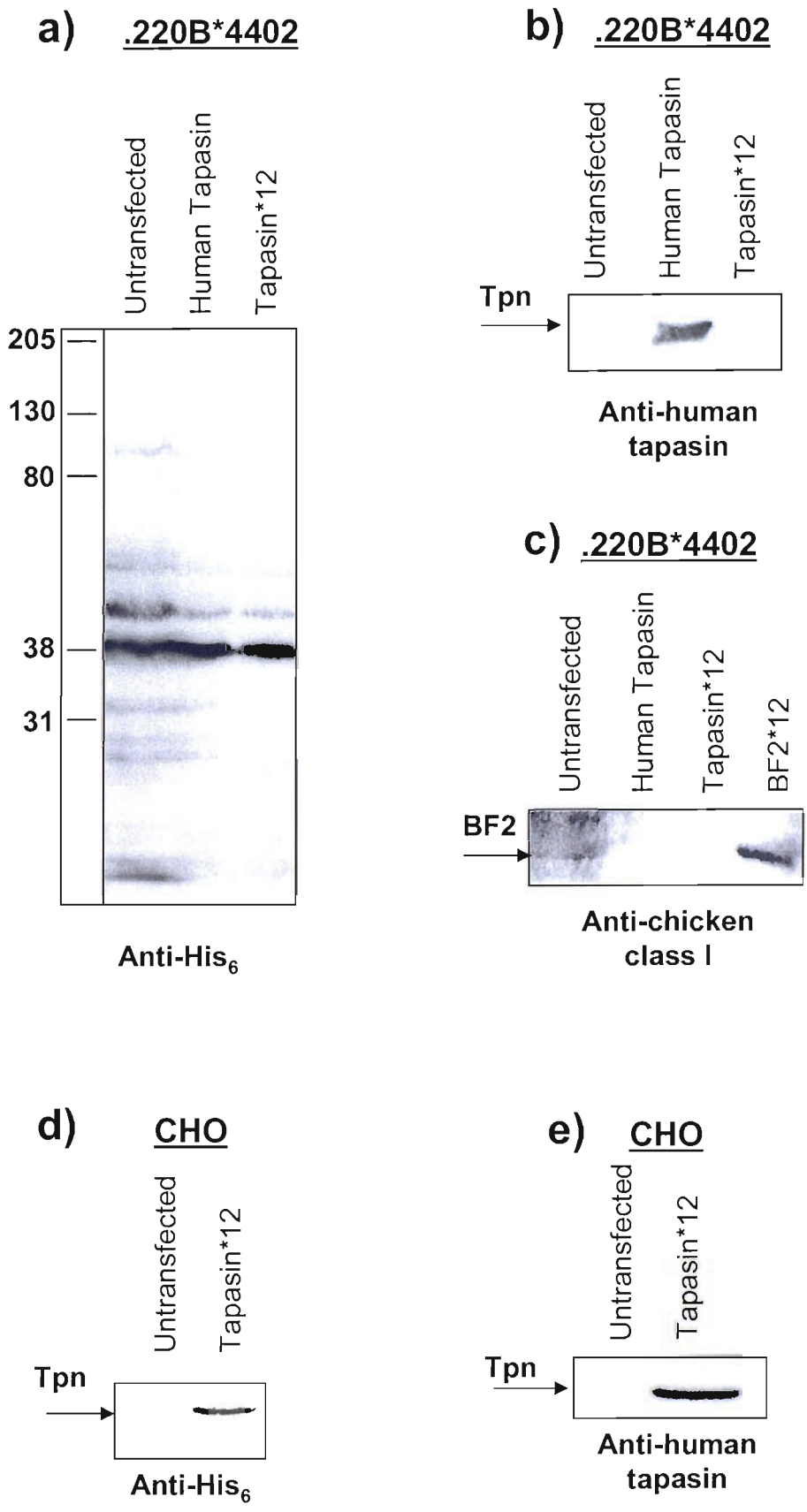
Various expression constructs were made comprising the full length cDNA of chicken BF2 alleles, chicken tapasin alleles, or chicken β_2m. Due to the lack of an antibody specific for chicken tapasin, an epitope tag was inserted into the chicken tapasin gene. Initially a 6x histidine tag (His_6) was used. As the N-terminal domains of tapasin are likely to interact with class I molecules, the epitope tag was inserted within the C-terminal domain. In order to maintain the ER retention function of the dilysine motif (KKxx) at the C-terminus of tapasin, the His_6 epitope tag replaced the stop codon, with the dilysine motif and a new stop codon (KKSQ*) duplicated after the epitope tag (see methods for detail). The research presented in this chapter was conducted over a prolonged period of time (approximately 18 months), and in the later experiments different chicken tapasin expression constructs were made that replaced the His_6 tag and duplicated dilysine motifs with a V5 epitope tag which was inserted immediately prior to a single copy of the dilysine motif.

4.2.1 Transient transfections of gene constructs into .220 cell lines

Transient transfection of .220 cells was initially chosen as it would enable rapid expression of combinations of chicken class I and chicken tapasin alleles, and would sample all transfectants, rather than concentrating on a small number of cloned stable transfectant cell lines.

At the onset of the project, preliminary experiments were conducted to determine whether chicken tapasin molecules could functionally replace their mammalian orthologues, and lead to the upregulation of surface expression of the tapasin-dependent HLA-B*4402 molecule. During the course of these experiments, multiple attempts (in excess of 20 independent experiments) were made to transiently transfect chicken Tapasin*12 His_6 expression constructs into the .220B*4402 cell line (.220 cells stably transfected with HLA-B*4402). However in all cases western blot analysis showed a lack of detectable chicken tapasin His_6 protein expression as detected by an anti-His_6 antibody (a representative example is shown in figure 4.3a).

To demonstrate that the transfection protocol was not the cause of the failure to transiently express chicken tapasin proteins in .220 cells, expression constructs encoding either human tapasin or chicken BF2*12 were separately transfected into the .220 cell line. In either case, expression of the transfected chicken class I or human tapasin proteins was readily detectable by western blot analysis (figure 4.3b and 4.3c).


a)
.220B*4402
Untransfected
Human Tapasin
Tapasin*12
205
130
80
38
31
Anti-His6
b)
.220B*4402
Untransfected
Human Tapasin
Tapasin*12
Tpn
Anti-human tapasin
c)
.220B*4402
Untransfected
Human Tapasin
Tapasin*12
BF2*12
BF2
Anti-chicken class I
d)
CHO
Untransfected
Tapasin*12
Tpn
Anti-His6
e)
CHO
Untransfected
Tapasin*12
Tpn
Anti-human tapasin

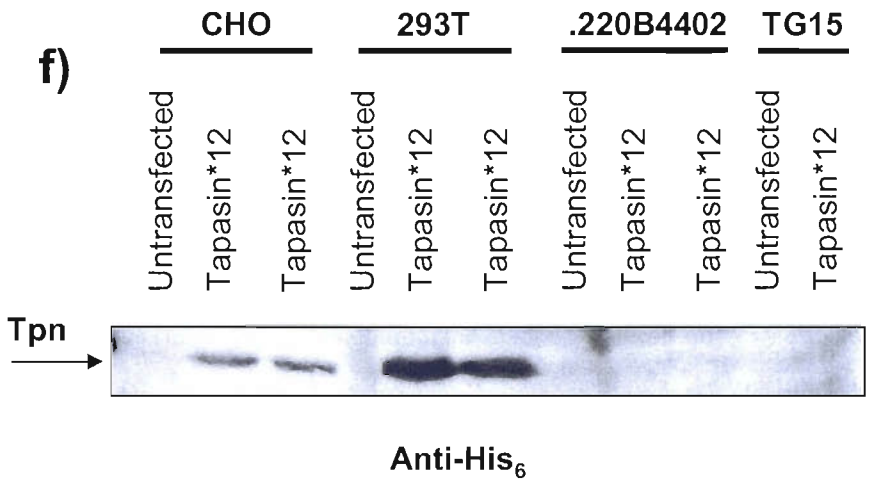


Figure 4.3 Western blot analysis following transient transfections reveals chicken Tapasin*12 His_6 protein expression can not be detected the day following transfection in .220B*4402 cells.

Bands representing tapasin (Tpn) or chicken class I (BF2) are indicated by a labelled arrowhead. The other bands represent non-specific staining. Molecular weight markers are shown on the left hand side of the blots. The antibodies used for each blot are indicated beneath.

a-c) **.220B*4402 cells**

.220B*4402 cells were transiently transfected with either human tapasin, chicken Tapasin*12 His_6, chicken BF2 (c only), or left untransfected

d-e) **CHO**

CHO cells were transiently transfected chicken Tapasin*12 His_6, or left untransfected.

f) **Various cells**

Hamster CHO, human 293T, human .220B*4402 or chicken TG15 cell lines were transiently transfected with either of two chicken Tapasin*12 His_6 expression constructs (in different vectors), or left untransfected.

To ensure that the choice of vector that comprised the backbone of the chicken tapasin His_6 expression construct was not the cause of this failure, various chicken Tapasin*12 His_6 expression constructs were made using different vectors including pcDNA 3.1+ neo, pcDNA6, pMCFR, pCIneo and pCIpac. However in all cases there remained a lack of detectable chicken tapasin His_6 protein expression following transient transfection into either .220 cells or their derivatives (figure 4.3f: .220B*4402 lanes, and data not shown).

In order to demonstrate the potential to express chicken tapasin His_6 molecules, the Chinese hamster CHO cell line was transiently transfected with these constructs, which resulted in detectable chicken tapasin His_6 protein expression by anti-His_6 antibody and western blot analysis (figure 4.3d). In the course of these experiments it was noted that the Giles 11/11 anti-human tapasin antibody cross-reacted with chicken tapasin His_6 proteins that were transiently expressed in the hamster CHO cell line (figure 4.3e). In addition, the human 293T cell line was also successfully transfected with chicken Tapasin*12 His_6 constructs, demonstrating the potential to express chicken tapasin His_6 molecules in a human cell line (figure 4.3f).

Therefore it appears that neither the expression constructs nor the transfection protocol caused the failure to detect chicken tapasin His_6 proteins following transient transfection of .220 cells. In addition, attempts to generate stable cell lines expressing chicken Tapasin*12 His_6 protein were unsuccessful.

Regardless of the cause of the failure to detect the expression of chicken tapasin His_6 proteins following transient transfection, attempts to generate clonal stable transfectant cell lines expressing V5 epitope-tagged chicken tapasin or other molecules did prove successful. Therefore in the rest of this chapter stable transfectant cell lines were used.

4.2.2 BF2*12 is expressed at a low level on the cell surface of .220 cells

The B12 class I peptide motif (table 4.1) is dominated by valine or isoleucine at position five (P5) and valine at P8 (Kaufman et al., 1995), and is likely to suit the human proteasome cleavage specificity and to be transported by the human TAP transporters. Therefore an expression construct encoding the BF2*12 gene was transfected into the .220 cell line. Five stable transfectant clones were generated and analysed by flow cytometry after staining with the F21-2 monoclonal antibody (figure 4.4). The F21-2 antibody reacts with the products of all chicken class I alleles with equal affinity, and is likely to be specific for a linear epitope as the antibody works in all assays including flow cytometry, immunoprecipitation, and in western blots.

A.

	B4								B12									B15								
position	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
Anchor		**D**			**D**			**E**					**V** **I**			**V**			**R**						**Y**	**y**
strong		E	I	P	E	V	D			A	N	K		G	R				K		F	P				
			F	Q		K	N			D	F			A	K						L	I				
			T	E			E			P	Q			N	Q						H					
			Y	K			Q				E				E						N					
			A	L			F														Y					
							G																			
							L																			
							M																			
							T																			
Weak		A		N	M		I	A	L		V		I	L		W		L			R		A	E	G	I	E	F	f
		V		L			A	K	I		K		M	P				I					Q	V	R	Q		W	w
		I						P			L			W									V	A					
								Y																L					

B.

1	A	D	V	E	E	Y	E	E	4	G	T	V	P	V	G	R	V		9	R	R	H	A	G	K	K	G	K	G	Y
2	L	D	V	E	D	V	K	F	5	K	A	A	A	V	V	L	V	T	10	R	R	A	L	R	E	G	Y			
3	I	D	W	F	D	G	K	E	6	F	A	N	Y	I	D	K	V		11	K	R	L	I	G	K	R	Y			
									7	W	Q	K	L	V	R	V	V		12	K	R	N	P	R	Q	I	N	Y		
									8	M	Q	I	F	V	K	T	L		13	R	R	F	F	P	Y	Y	V	Y		

Table 4.1 Class I peptide motifs for three MHC haplotypes

A. Sequences of peptides bound to class I molecules isolated from three chicken haplotypes determined from peptide pools showing anchor, strong and weak signals. Basic residues in blue, acidic residues in red, all others in black.

B. Sequences of individual peptides, with potential source proteins: 1, chicken tubulin β6 (P09207; peptide found in draft chicken genome sequence in contig 296.1 on chromosome 1); 2, chicken DEAD-box helicase (NP990158, contig 7.29 on 1); 3, FUS proto-onc gene (P35637 representative of peptide sequence in many species); 4, elongation factor α chain (Q6DYT7 representative; 52.85 on 21); 5, chicken β_2m (P21611; 9.294 on 6); 6, vimentin (Q9MZA9 representative; 2.352 on 2); 7, chicken haemoglobin β chain (P02112; 1785.2 on 1); 8, chicken ubiquitin (Q9PST8; 37.143 on 15); 9, chicken MHC class I α chain (Q31412; contig 368.19 on 16); 10, chicken calcium binding protein (Q90940; 71.39 on 10); 11, chicken hsp70 (P08106, similar to 7.547 on 1); 12, chicken 60S ribosomal protein L24 with terminal W (P83731, 34.64 on 1); 13, proteasome subunit beta type (Q59GN1 representative, 64.178 on 3)

Table taken with permission from (Wallny et al., 2006)

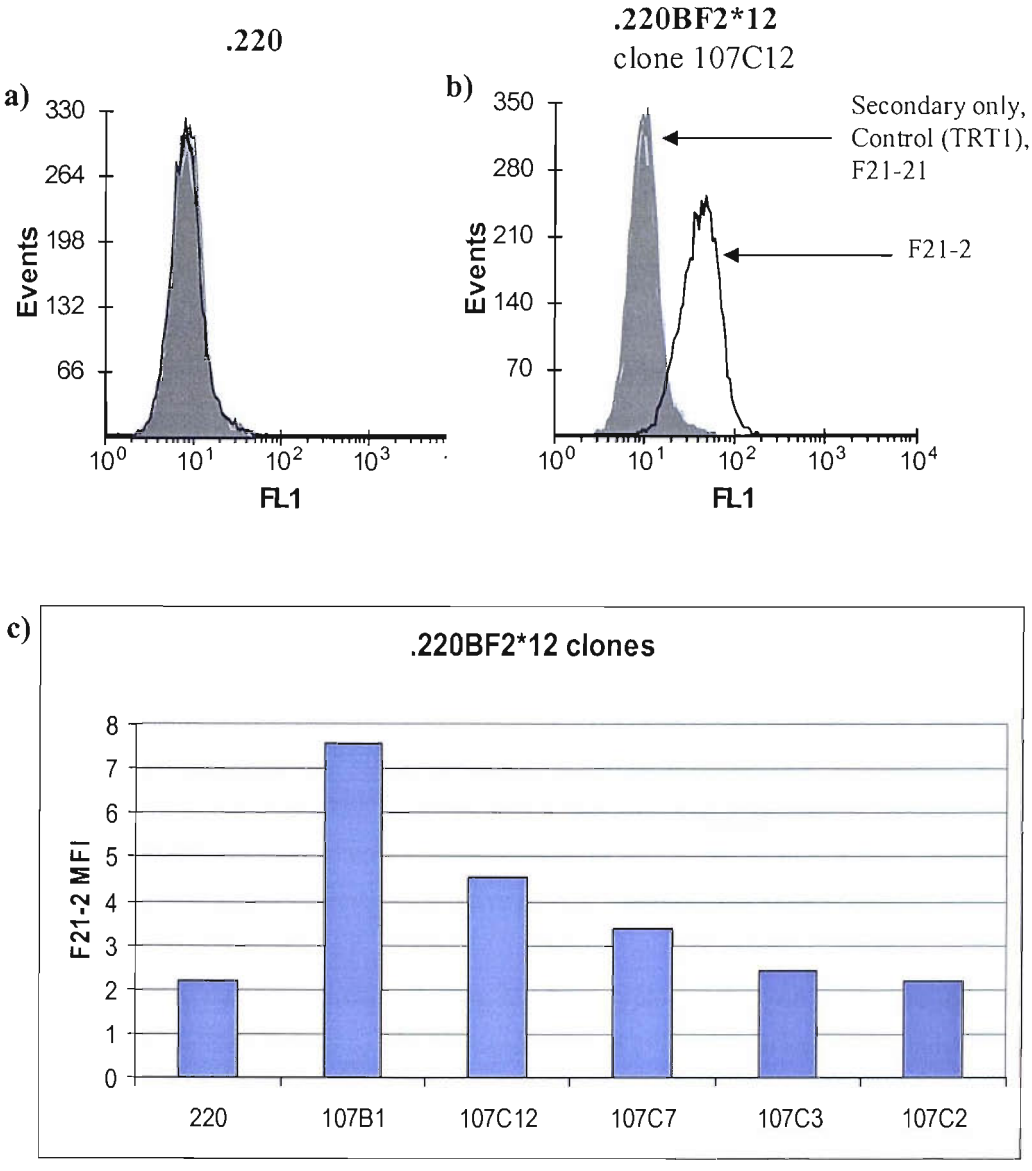


Figure 4.4 Flow cytometry analysis of .220BF2*12 clones

a) Histogram depicting surface expression levels of .220 cells stained with either secondary antibody only (light grey, filled), TRT1 (IgG1 isotype matched control, grey, filled), F21-21 (anti-chicken b2m, grey, open) or F21-2 (anti-chicken class I heavy chain, black, open) antibodies.

b) Histogram depicting surface expression levels of .220BF2*12 clone 107C12 stained as above.

c) Graph depicting the F21-2 MFI for .220 cells and all .220BF2*12 clones.

Figure 4.4b shows the amount of BF2*12 expressed at the cell surface of .220BF2*12 clone 107C12 to be less than one log higher than control antibody staining, in comparison to being two to three logs higher than control antibody staining of B12 erythrocytes (Kaufman et al., 1995). Figure 4.4c reveals that in comparison to clone 107C12, the amount of BF2*12 expressed at the cell surface was slightly lower in three of the other clones, and approximately two-fold higher in clone 107B1.

4.2.3 Expression of chicken tapasin does not increase BF2*12 surface expression levels

Certain mammalian class I alleles (for example HLA-B*4402) are expressed at low levels at the surface of tapasin-deficient cells. The subsequent expression of mammalian tapasin can produce large increases in their surface expression levels. In order to determine whether there is an increase in BF2*12 surface expression when the chicken Tapasin*12 protein is expressed, and whether such an increase occurs when a haplotype-mismatched chicken tapasin protein is expressed, the .220BF2*12 clone 107C12 was stably transfected with expression constructs encoding V5 epitope-tagged chicken tapasin molecules derived from either the B12 or B15 haplotypes (the transfected proteins are referred to as Tapasin*12 or Tapasin*15 hereafter).

Numerous stable transfectants of .220BF2*12Tpn*12 (clones were designated with the prefix 224) and .220BF2*12Tpn*15 (clones were designated with the prefix 225) were generated. Figure 4.5 reveals all clones express comparable amounts of the transfected chicken tapasin protein relative to the amount of protein detected when the blots were subsequently incubated with an anti-calnexin antibody which served to demonstrate the relative protein levels loaded in each lane.

For all transfectants the expression of either the haplotype-matched or haplotype-mismatched chicken tapasin allele did not lead to a large increase in BF2*12 surface expression relative to the tapasin-deficient parental .220BF2*12 clone 107C12 (table 4.2 and figure 4.6). Whilst some chicken tapasin-expressing clones (such as 224C1, C2, C3 and 225C7, C12 and C17) did exhibit a marginal increase in BF2*12 expression levels, which occurred for either chicken tapasin allele, other chicken tapasin-expressing clones (such as 224C6 and C9 and 225C8 and C11) exhibited negligible increases in BF2*12 surface expression levels relative to the tapasin-deficient clone, which suggests there was no major alterations in the amount of BF2*12 molecules expressed at the cell surface following chicken tapasin expression.

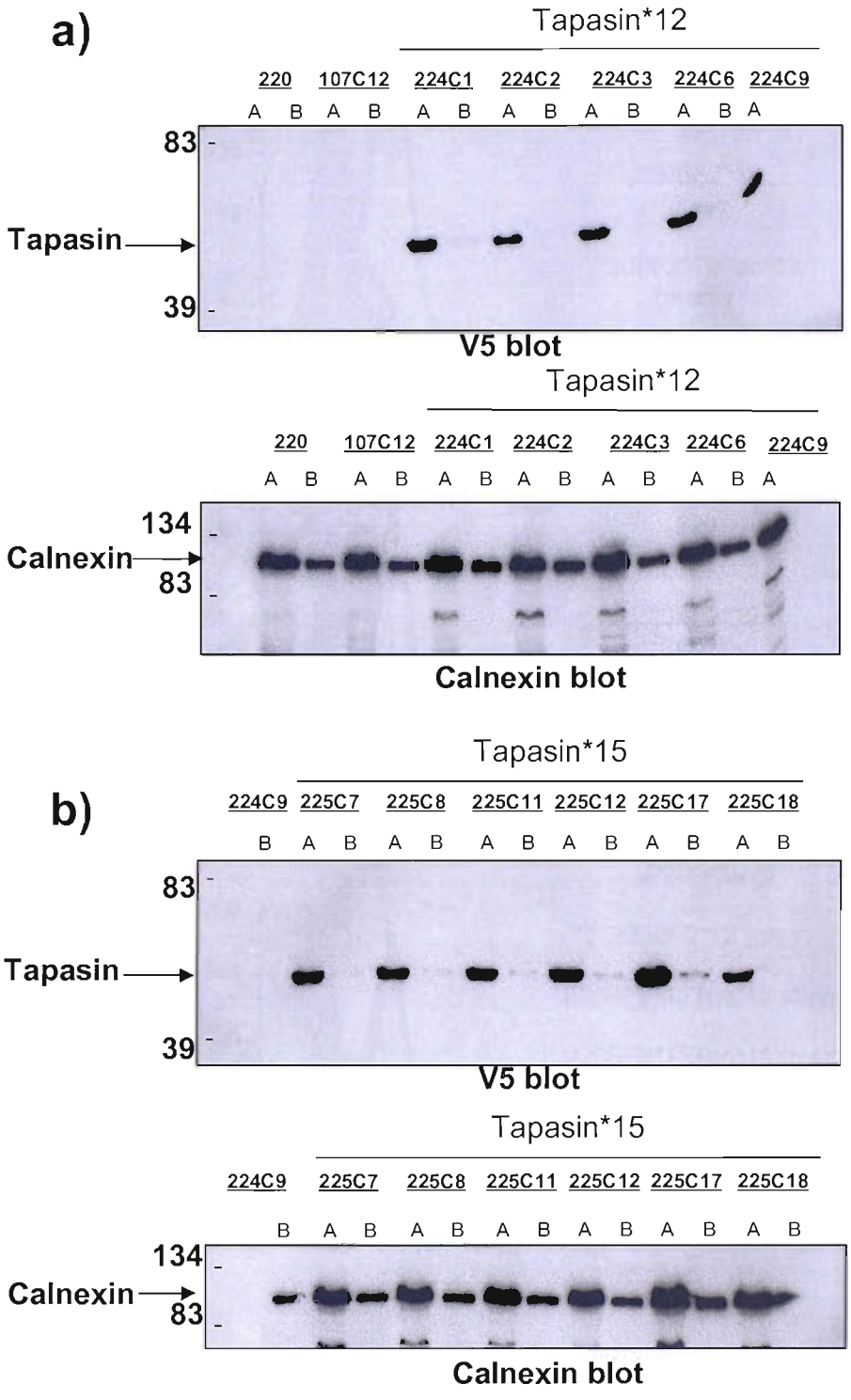


Figure 4.5 Western blots confirming Tapasin expression in .220BF2*12 Tapasin transfectants.

Lysates were prepared and SDS PAGE gels were loaded with cell lysate equivalent to 10^5 cells (A) or 10^4 cells (B) and western blotted. Blots were incubated with anti-V5 antibody and developed, then incubated with anti-calnexin antibody and developed. Arrowheads mark proteins of interest.

a) .220BF2*12Tpn*12 (prefix 224) transfectants.

b) .220BF2*12Tpn*15 (prefix 225) transfectants.

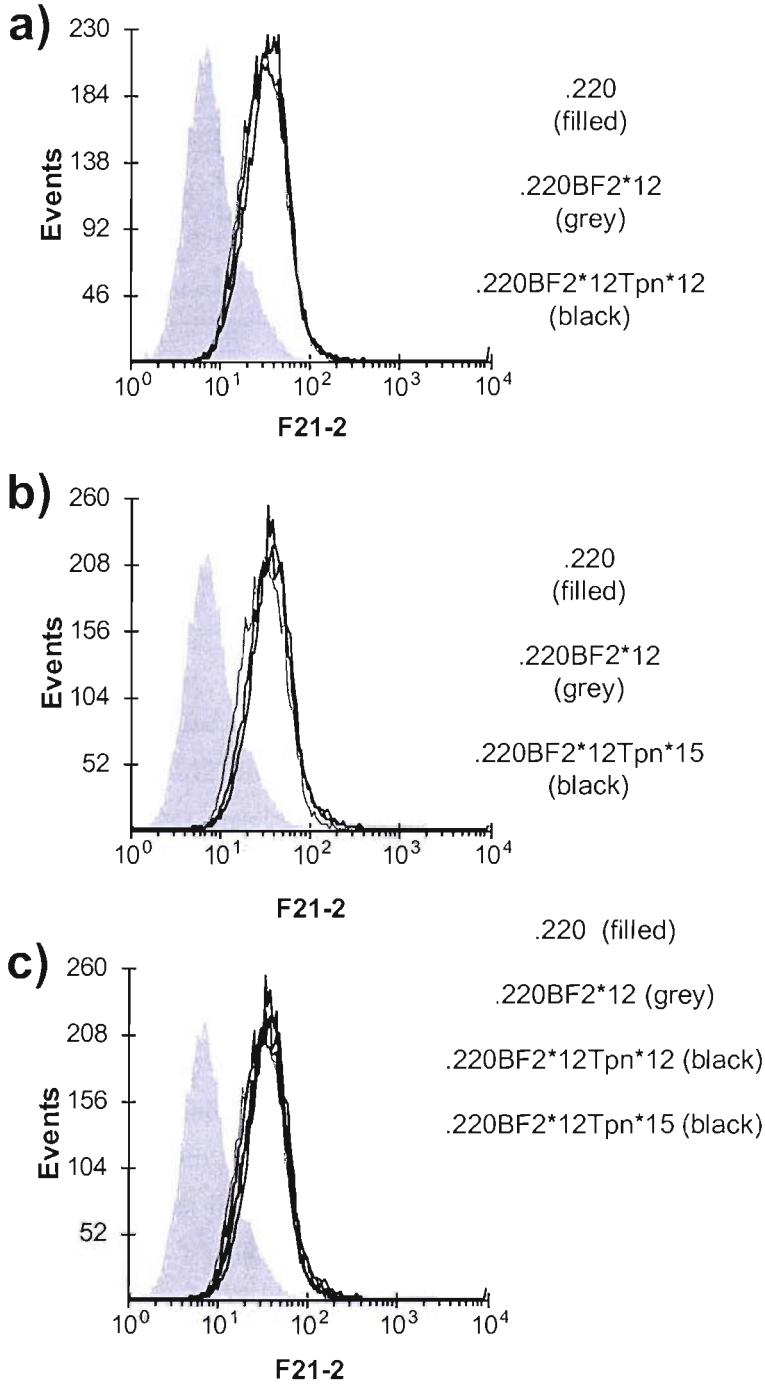


Figure 4.6 The effect of chicken Tapasin*12 or Tapasin*15 expression upon BF2*12 surface expression

Histograms depicting surface chicken class I (F21-2) expression levels on the clones listed below:

a) .220 cells (grey, filled), .220BF2*12 clone 107C12 (grey, open), .220BF2*12Tpn*12 clones 224C1 and 224C2 (black, open).

b) .220 cells (grey, filled), .220BF2*12 clone 107C12 (grey, open), .220BF2*12Tpn*15 clones 225C7 and 225C8 (black, open)

c) .220 cells (grey, filled), .220BF2*12 clone 107C12 (grey, open), .220BF2*12Tpn*12 clones 224C1 and 224C2 (black, open) .220BF2*12Tpn*15 clones 225C7 and 225C8 (black, open).

BF2*12	Tapasin	Clone	F21-2 MFI	BBM1 MFI
-	-	.220	7.10	24.11
+	-	107C12	25.58	19.57
+	12	224C1	30.36	19.78
+	12	224C2	30.57	15.76
+	12	224C3	29.59	16.68
+	12	224C6	26.84	17.88
+	12	224C9	26.76	14.95
		Mean MFI	28.82	17.01
+	15	225C7	30.02	16.19
+	15	225C8	28.59	17.61
+	15	225C11	26.78	15.5
+	15	225C12	32.13	21.54
+	15	225C17	30.44	16.93
+	15	225C18	29.87	20.73
		Mean MFI	29.64	18.08

Table 4.2 BF2*12 (F21-2) and human β_2m (BBM1) surface expression levels on .220BF2*12 transfectants

In agreement (table 4.2) there were no major differences in the amount of human β_2m molecules expressed at the cell surface of .220BF2*12 clone 107C12 and its chicken tapasin-expressing transfectants, assessed by BBM1 staining and flow cytometry. There was also no correlation between those clones which had the highest BF2*12 expression level, and those clones with the highest human β_2m expression levels.

4.2.4 Chicken tapasin expression does not affect the proportion of BF2*12 molecules expressed at the cell surface after incubation overnight at 37°C or 27°C

Whilst chicken tapasin protein expression did not significantly increase the amount of surface expressed BF2*12 molecules relative to tapasin-deficient cells, it was possible that chicken tapasin protein expression did lead to an improvement in the peptide-loading of the BF2*12 molecule which was not detected by F21-2 staining and flow cytometry. For example, the F21-2 antibody reacts with all conformations of chicken class I molecules, and therefore reacts with both denatured, free heavy chain and optimally peptide-loaded class I heterodimers. Therefore it was possible that in the .220BF2*12 transfectants, the surface expression level of BF2*12 may represent predominantly free heavy chains (that could have dissociated either en route to, or at the cell surface, due to being loaded with a repertoire of sub-optimal peptides), whilst the surface expression level of BF2*12 in the chicken tapasin-expressing transfectants may represent predominantly optimally peptide-loaded class I heterodimers.

Such an alteration in the quality of BF2*12 molecules may be revealed by differences in the proportions of BF2*12 molecules that can be expressed at the cell surface when cultured at 27°C or 37°C in chicken tapasin-expressing and tapasin-deficient cells. Culturing cells at reduced temperatures has been demonstrated to increase the surface expression levels of some class I

molecules (Ljunggren et al., 1990), which is likely to reflect an increased proportion of sub-optimally loaded molecules progressing to the cell surface, which are unstable at 37°C (De Silva et al., 1999).

Table 4.3 reveals that for the tapasin-deficient .220BF2*12 clone 107C12 the amount of BF2*12 molecules expressed at the cell surface following culture at 27°C was increased to 174% of the amount expressed following culture at 37°C overnight. The clones expressing either of the chicken tapasin alleles exhibited virtually identical ratios in BF2 surface expression to that observed for .220BF2*12 clone 107C12. This suggests that the presence of either chicken tapasin protein did not quantitatively or qualitatively improve BF2*12 peptide-loading in .220 cells. This finding was later confirmed by other assays which were performed after further transfectants were produced, and are discussed later in this chapter (tables 4.8 and 4.9).

BF2*12	Tapasin	Clone	Temp	Average F21-2 MFI	st dev	Ratio 27:37
-	-	.220	37	8.34	0.18	1.08
			27	8.98	0.08	
+	-	107C12	37	29.77	0.63	1.74
			27	51.80	0.76	
+	12	224C1	37	31.75	0.45	1.69
			27	53.52	1.00	
+	12	224C2	37	33.81	1.08	1.71
			27	57.82	1.13	
+	15	225C7	37	37.66	0.55	1.68
			27	63.23	2.06	
+	15	225C8	37	36.90	1.21	1.67
			27	61.56	0.21	

Table 4.3 BF2*12 (F21-2) expression levels following overnight culture at 27°C or 37°C.

4.2.5 Human tapasin does not increase BF2*12 surface expression levels

Having found that the expression of chicken tapasin did not quantitatively or qualitatively improve BF2*12 surface expression, it was of interest to determine the effect that human tapasin has on BF2*12 expression. Therefore, to allow direct comparison with .220BF2*12 cells, BF2*12 was transfected into the human tapasin-expressing .220hTpn cell line.

Only two stable transfectant clones were generated which were analysed by F21-2 staining and flow cytometry (figure 4.7a). The amount of BF2*12 expressed at the cell surface of these human tapasin-expressing clones was less than one log higher than .220 cells, which express no chicken class I molecules. The absence of a large increase in surface BF2*12 expression levels in these cells relative to tapasin-deficient cells suggests that the low surface BF2*12 expression level in these cells (in comparison to chicken PBLs) is not simply a consequence of a limited supply of peptides: as the expression of human tapasin is likely to have increased both the amount of TAP proteins present in the cells, and the import of peptides into the ER relative to the tapasin-deficient cells (Bangia et al., 1999; Garbi et al., 2003).

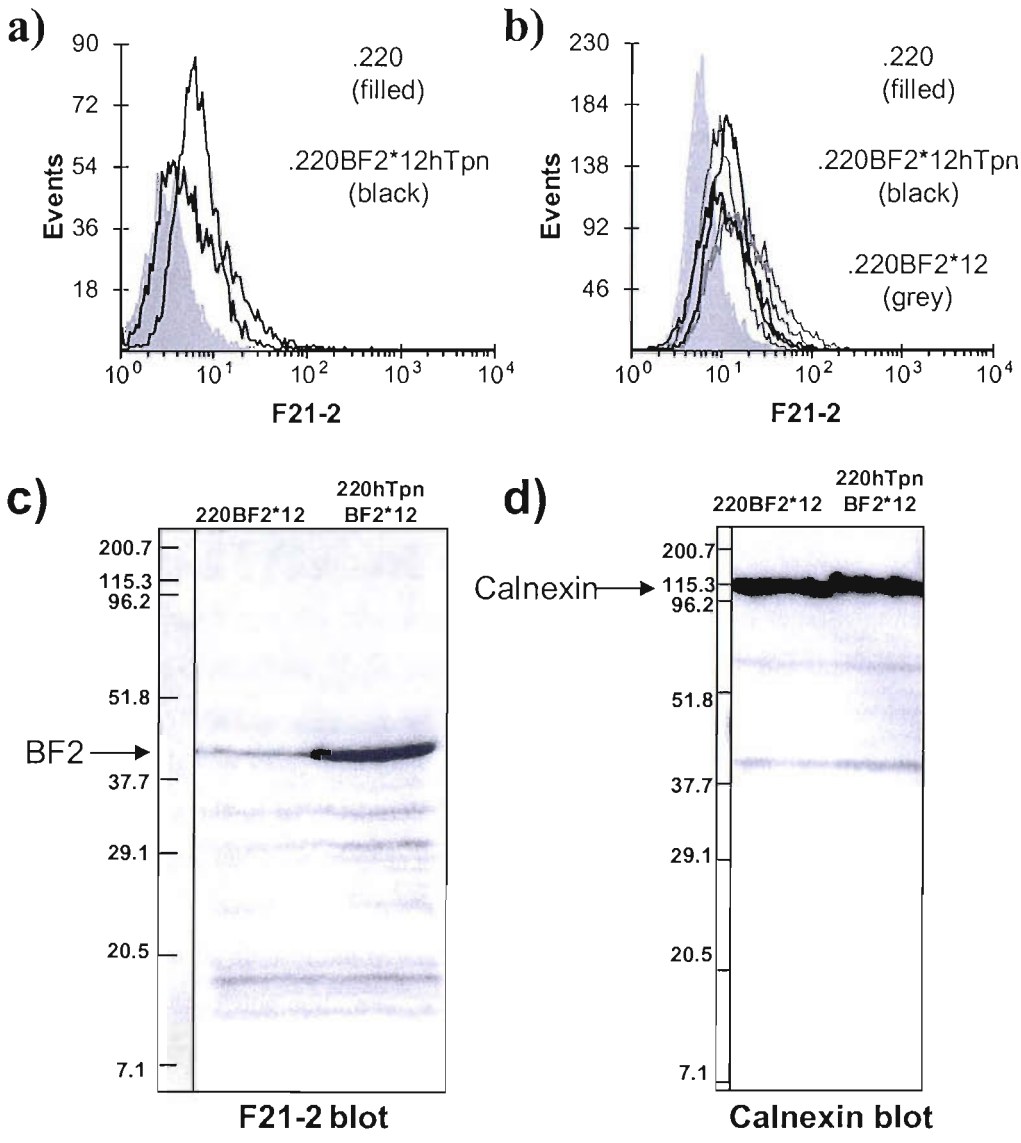


Figure 4.7 Comparison of BF2*12 expression levels in the presence or absence of human tapasin

a) Histogram depicting surface chicken class I (F21-2) expression level upon .220 (grey, filled) and .220hTpnBF2*12 clones 108C12 and 108C34 (black, open).

b) Histogram depicting F21-2 surface expression upon .220 (grey, filled), .220BF2*12 clones 107C7, 107C12 and 107B1 (grey, open), and .220hTpnBF2*12 clones 108C12 and 108C34 (black, open).

c-d) Lysates were made of .220BF2*12 clone 107C12 and .220hTpnBF2*12 clone 108C12, and SDS PAGE gels were loaded with lysate equivalent to 10^6 cells (c) or $2x10^5$ cells (d) and western blotted. Blots were incubated with the indicated antibodies and developed. Arrowheads mark the bands of interest. Molecular weight markers (kDa) are shown on the left hand side of each gel.

However in this experiment the amounts of BF2*12 molecules expressed at the cell surface of the human tapasin-expressing clones were slightly lower than in the tapasin-deficient clones (figure 4.7b). This modest decrease in surface expression of BF2*12 molecules in the human tapasin-expressing clones is particularly noteworthy in light of the overall greater total BF2*12 protein expression level (figure 4.7c-d) revealed by western blot comparison of lysates made from a human tapasin-expressing clone and a tapasin-deficient clone. This suggests that the human tapasin protein may have retained a proportion of BF2*12 molecules within the cell, which would otherwise have reached the cell surface.

4.2.6 Human tapasin does not affect the proportion of BF2*12 molecules expressed at the cell surface after incubation overnight at 37°C or 28°C

Having found a greater proportion of the total BF2*12 molecules are expressed at the cell surface in tapasin-deficient cells than in human tapasin-expressing cells; two approaches were taken to determine whether there is a qualitative difference in BF2*12 molecules that occurs as a result of peptide-loading occurring in the presence or absence of the human tapasin protein. The first attempt concerns a comparison of the proportions of BF2*12 molecules that egress to the cell surface following culture at 37°C or 28°C in the tapasin-deficient clones or human tapasin-expressing clones. It was likely that there would have been a significantly larger increase in the surface expression levels of BF2*12 molecules in the human tapasin-expressing clones following culture at 28°C, than would be evident in the tapasin-deficient clones.

Table 4.4 shows that for four of the five .220BF2*12 clones, the proportion of BF2*12 expressed on the cell surface following culture at 28°C was between 138-143% of that following culture at 37°C. For the human tapasin-expressing clones there was a similar increase in cell surface expression (140-153%) following culture at 28°C. Therefore despite there appearing to be a smaller proportion of BF2*12 molecules that reached the cell surface at 37°C in the human tapasin-expressing clones, there was no significant alteration to the proportion of molecules that could be bought to the cell surface by culturing the cells at 28°C relative to tapasin-deficient cells. This suggests that the majority of the BF2*12 molecules which do not reach the cell surface in human tapasin-expressing cells may not be competent for peptide-loading; it is possible there is a larger pool of free BF2*12 heavy chain in these clones than is evident in the tapasin-deficient clones.

BF2*12	human Tapasin	Clone	28°C F21-2 MFI	37°C F21-2 MFI	Ratio 28:37
-	-	220	2.20	2.19	1.00
+	-	107B1	10.55	7.55	1.40
+	-	107C12	6.45	4.52	1.43
+	-	107C7	4.76	3.38	1.41
+	-	107C3	3.34	2.42	1.38
+	-	107C2	2.29	2.17	1.06
-	-	220	3.39	3.28	1.03
+	+	108C12	8.48	6.06	1.40
+	+	108C34	9.96	6.51	1.53

Table 4.4 BF2*12 (F21-2) expression levels following overnight culture at 28°C or 37°C.

For .220BF2*12 clone 107C12 the increase in the proportion of BF2*12 molecules expressed at the cell surface recorded in this experiment (table 4.4) was lower than was observed previously (table 4.3). At least two possible reasons can account for this difference: firstly, in the first experiment (table 4.3) the cells were cultured in a CO_2 regulated incubator, whereas in the second experiment (table 4.4) the cells were cultured in a non-CO_2 regulated incubator: the change in pH which occurs when cells are cultured outside of a CO_2 regulated environment may have stressed the cells to some extent; secondly, the temperature of the culture varied: 27°C for the first experiment, 28°C for the second; however other variables in the environment may have also have accounted for the differences that were observed.

4.2.7 Exogenous peptide and β_2m stabilisation assays reveal no difference in BF2*12 molecules expressed in .220BF2*12 and .220hTpnBF2*12 cells

The second attempt that was made to determine whether there is a qualitative difference in BF2*12 molecules that occurs as a result of peptide-loading occurring in the presence or absence of the human tapasin protein involved a class I stabilisation assay. In this assay combinations of exogenous peptide, chicken β_2m or human β_2m were added to the clones, before F21-2 staining and flow cytometry was performed to determine whether the surface expression levels were increased differentially in the tapasin-deficient and human tapasin-expressing clones.

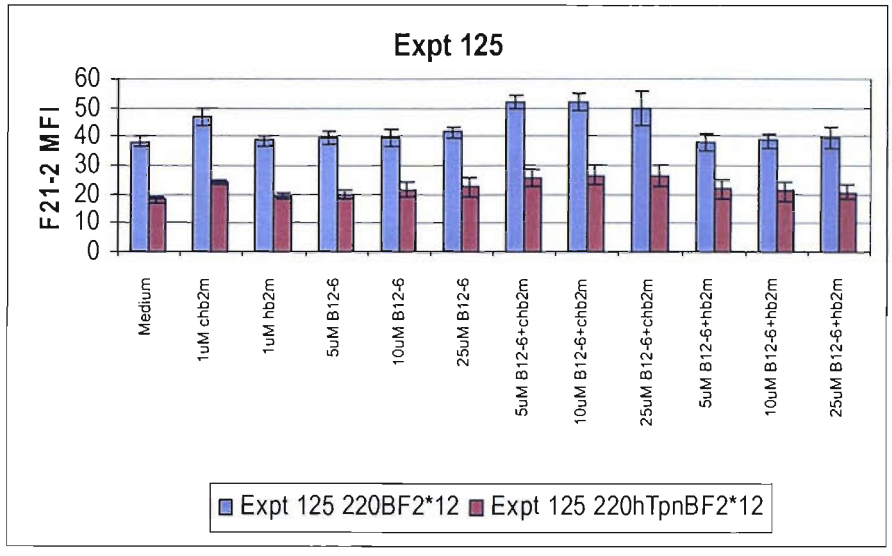
Table 4.5 and figure 4.8 show the result of two experiments (experiment 125 and experiment 129). Despite some minor variation between the experiments, the BF2*12 expression levels of either the human tapasin-expressing clone or the tapasin-deficient clone were comparably affected by the combinations of exogenous peptide and/or β_2m that were added to the cultured cells.

When chicken β_2m was added to either clone, the BF2*12 expression level was increased by 22-35% compared to the expression levels when cultured in media alone. However when exogenous human β_2m was added, there was no such increase in the amount of BF2*12 expressed for either clone.

		Expt 125					Expt 129			
Human tapasin	BF2*12	Clone	Treatment	mean F21-2 MFI	std dev	relative to medium MFI	Treatment	mean F21-2 MFI	std dev	relative to medium MFI
-	+	107B1	secondary ab only				secondary ab only	13.72	0.40	0.33
			Medium	38.40	1.84	1.00	Medium	41.87	0.49	1.00
			1uM ch.b2m	46.66	2.97	1.22	1uM ch.b2m	52.83	1.63	1.26
			1uM h.b2m	38.53	2.06	1.00	1uM h.b2m	43.27	1.87	1.03
			5uM B12-6	39.39	2.07	1.03	5uM B12-6	44.78	1.58	1.07
			10uM B12-6	39.73	2.78	1.03	10uM B12-6	45.92	2.18	1.10
			25uM B12-6	41.47	1.80	1.08	25uM B12-6	48.20	2.30	1.15
			5uM B12-6+1uM ch.B2m	52.02	2.33	1.35	5uM B12-6+1uM ch.B2m	58.46	2.29	1.40
			10uM B12-6+1uM ch.B2m	52.02	3.11	1.35	10uM B12-6+1uM ch.B2m	58.92	1.12	1.41
			25uM B12-6+1uM ch.B2m	50.04	5.78	1.30	25uM B12-6+1uM ch.B2m	57.40	0.94	1.37
			5uM B12-6+1uM h.B2m	38.09	3.05	0.99	5uM B12-6+1uM h.B2m	44.17	0.97	1.05
			10uM B12-6+1uM h.B2m	38.65	2.66	1.01	10uM B12-6+1uM h.B2m	47.09	0.60	1.12
			25uM B12-6+1uM h.B2m	39.44	3.58	1.03	25uM B12-6+1uM h.B2m	48.67	0.78	1.16
+	+	108C12	secondary ab only				secondary ab only	10.58	0.21	0.44
			Medium	18.07	0.98	1.00	Medium	23.92	1.05	1.00
			1uM ch.b2m	24.39	0.73	1.35	1uM ch.b2m	29.51	1.28	1.23
			1uM h.b2m	19.15	1.21	1.06	1uM h.b2m	24.74	0.78	1.03
			5uM B12-6	20.04	1.42	1.11	5uM B12-6	25.36	1.40	1.06
			10uM B12-6	21.45	2.79	1.19	10uM B12-6	25.93	0.77	1.08
			25uM B12-6	22.36	3.34	1.24	25uM B12-6	26.51	0.66	1.11
			5uM B12-6+1uM ch.B2m	25.71	2.80	1.42	5uM B12-6+1uM ch.B2m	31.33	1.24	1.31
			10uM B12-6+1uM ch.B2m	26.64	3.16	1.47	10uM B12-6+1uM ch.B2m	31.16	1.35	1.30
			25uM B12-6+1uM ch.B2m	26.43	3.92	1.46	25uM B12-6+1uM ch.B2m	31.73	1.18	1.33
			5uM B12-6+1uM h.B2m	21.61	3.04	1.20	5uM B12-6+1uM h.B2m	25.70	0.44	1.07
			10uM B12-6+1uM h.B2m	20.95	3.12	1.16	10uM B12-6+1uM h.B2m	25.94	0.52	1.08
			25uM B12-6+1uM h.B2m	20.59	2.61	1.14	25uM B12-6+1uM h.B2m	26.33	0.82	1.10

Table 4.5 BF2*12 (F21-2) expression levels following culture of the clones in the indicated conditions

a)



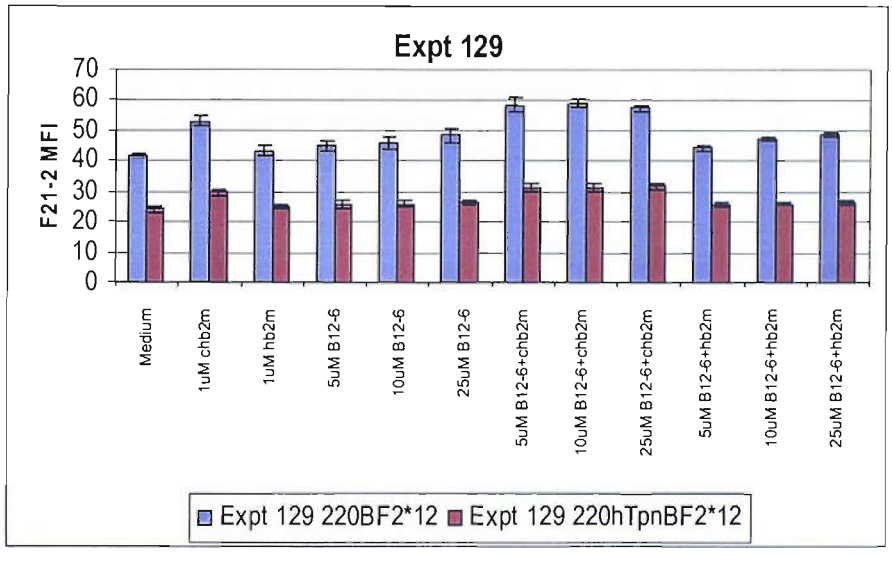
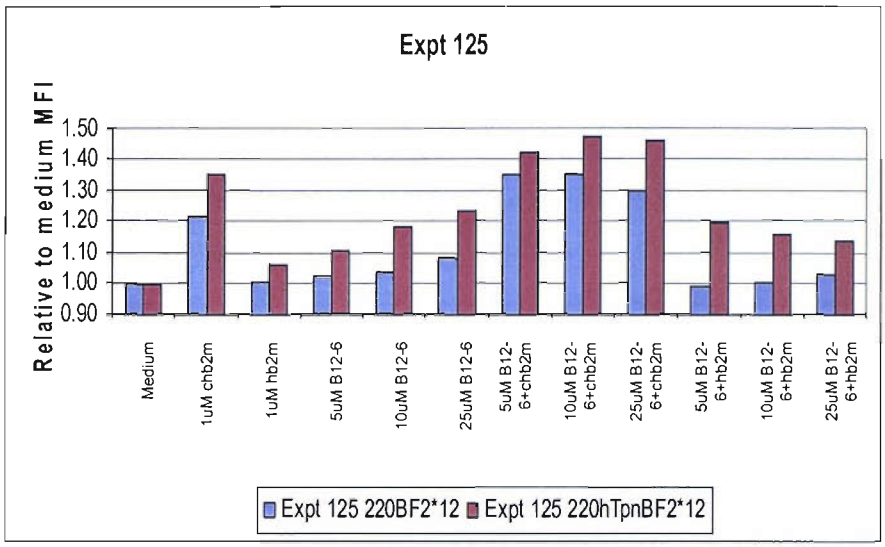


Figure 4.8 Peptide and/or β2m stabilisation assays

a) Graph depicting the mean F21-2 MFI values for .220BF2*12 clone 107B1 and .220hTpnBF2*12 clone 108C12 after culture in combinations of peptide and/or β2m in experiments 125 and 129

Continued on next page

b)



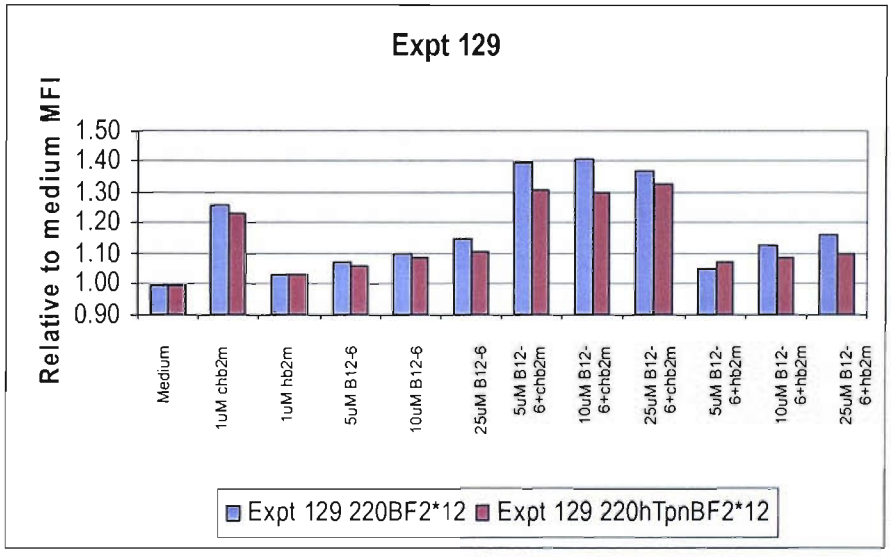


Figure 4.8 Peptide and/or β2m stabilisation assays

b) Graph depicting the mean F21-2 MFI relative to mean F21-2 MFI recorded after culture in media only in experiments 125 and 129

When the peptide LPACVLEV (labelled B12-6, and known to bind with high affinity to B12 class I molecules (Wallny et al., 2006)) was added to either clone, there was a small dose dependent increase in BF2*12 surface expression levels, with a maximal increase of 24% when 25 μM peptide was added. When the same peptide concentration range was added in addition to exogenous chicken β_2m, then for either clone there was a larger increase in BF2*12 expression of between 30-47%, which was not seen when the cells were cultured with either peptide or chicken β_2m alone. However there was no peptide dose dependent increase in BF2*12 expression over the peptide concentrations tested when chicken β_2m was also added.

When exogenous human β_2m was added in addition to the peptide concentration range, the results in experiment 129 were directly comparable to those obtained when the peptide concentration range was added independently of exogenous β_2m. This agrees with the lack of stabilisation of BF2*12 molecules when exogenous human β_2m was added independently to either clone. In experiment 125 the results appeared more variable, but at least for .220BF2*12 clone 107B1 were largely comparable to those reached in experiment 129.

Therefore it appears that for either .220BF2*12 or .220hTpnBF2*12 clone, there is a proportion of BF2*12 molecules which can be stabilised by exogenous chicken β_2m, exogenous peptide, or both, which leads to an increased cell surface expression level. The proportion of BF2*12 molecules that could be stabilised appeared to be roughly equal for either clone. This finding combined with those discussed previously suggest that the BF2*12 molecules were loaded with repertoires of peptides that were of equal quality in either the tapasin-deficient or human tapasin-expressing clones, and it appears that the presence of human tapasin did not qualitatively affect the peptide-loading of the BF2*12 molecules. The apparent differences in the proportions of the total BF2*12 molecules that reach the cell surface at 37°C in either tapasin-deficient clones or human tapasin-expressing clones (figure 4.7) may be a consequence of an increased proportion of free BF2*12 heavy chain.

4.2.8 Chicken β_2m protein expression does not increase BF2*12 surface expression

The observation that the surface expression level of BF2*12 molecules was increased following culture of the BF2*12 transfectants with exogenous chicken β_2m suggests that BF2*12 molecules may associate poorly with human β_2m; as the presence of exogenous chicken β_2m is likely to have stabilised BF2*12 molecules associated with human β_2m that would otherwise have dissociated. Molecular incompatibilities between mammalian proteins have been described before and therefore set a precedent (Peh et al., 1998). Also in support is the observation that the expression of BF2*12 molecules (albeit at a relatively low level at the cell surface) did not increase the amount of human β_2m expressed at the cell surface above that of the parental .220 cells (table 4.2). The possibility of an incompatibility occurring between chicken BF2 and human β_2m molecules may also explain why the expression of chicken Tapasin*12 (or human tapasin) did not improve the expression of BF2*12 molecules relative to tapasin-deficient cells; as chicken tapasin may only be able to

interact productively with BF2 molecules that are in an appropriate conformation, which may not be possible when the BF2 molecules are coupled to human β_2m. Therefore to determine how the presence of chicken β_2m affects BF2*12 surface expression levels more precisely, the .220BF2*12 clone 107C12 was transfected with an expression construct encoding chicken β_2m.

The two clones that were generated express different amounts of chicken β_2m on their cell surfaces (table 4.6 and figure 4.9); clone 235.1 expresses a relatively low amount of chicken β_2m in comparison to clone 235.2. Surprisingly however, the presence of chicken β_2m did not increase the amount of BF2*12 molecules expressed upon the cell surface in comparison to the parental .220BF2*12 clone 107C12. This suggests BF2*12 expression is not limited solely by the absence of chicken β_2m.

BF2*12	ch b2m	Clone	secondary only MFI	TRT1 MFI	F21-21 (-TRT1) MFI	F21-2 (-TRT1) MFI
-	-	.220	8.22	8.27	0.63	0.1
+	-	107C12	10.21	9.83	0.39	31.25
+	low	235.1	9.77	9.71	6.11	32.17
+	high	235.2	ND	10.84	26.31	27.67

Table 4.6 Cell surface expression levels of chicken β_2m (F21-21) and BF2*12 (F21-2) on the indicated clones. TRT1 represents isotype control staining.

4.2.9 BF2*12 surface expression levels can be increased by the presence of high levels of chicken β_2m and Tapasin*12

Having found that expression of chicken Tapasin*12 or chicken β_2m proteins did not affect the cell surface expression of BF2*12 molecules, it was of interest to determine what effect the expression of both chicken tapasin and chicken β_2m had on BF2*12 molecules. Therefore the two chicken β_2m-expressing .220BF2*12 clones (235.1 and 235.2) were separately transfected with an expression construct encoding V5 epitope-tagged chicken Tapasin*12, and stable transfectant clones were generated.

Clones that derived from .220BF2*12chβ_2m^{low} clone 235.1 were designated with the prefix 238, clones that derived from .220BF2*12chβ_2m^{high} clone 235.2 were given the prefix 239. The clones that were generated expressed variable amounts of chicken tapasin protein (figure 4.10f) relative to the protein levels revealed by incubating the blots with anti-calnexin antibody.

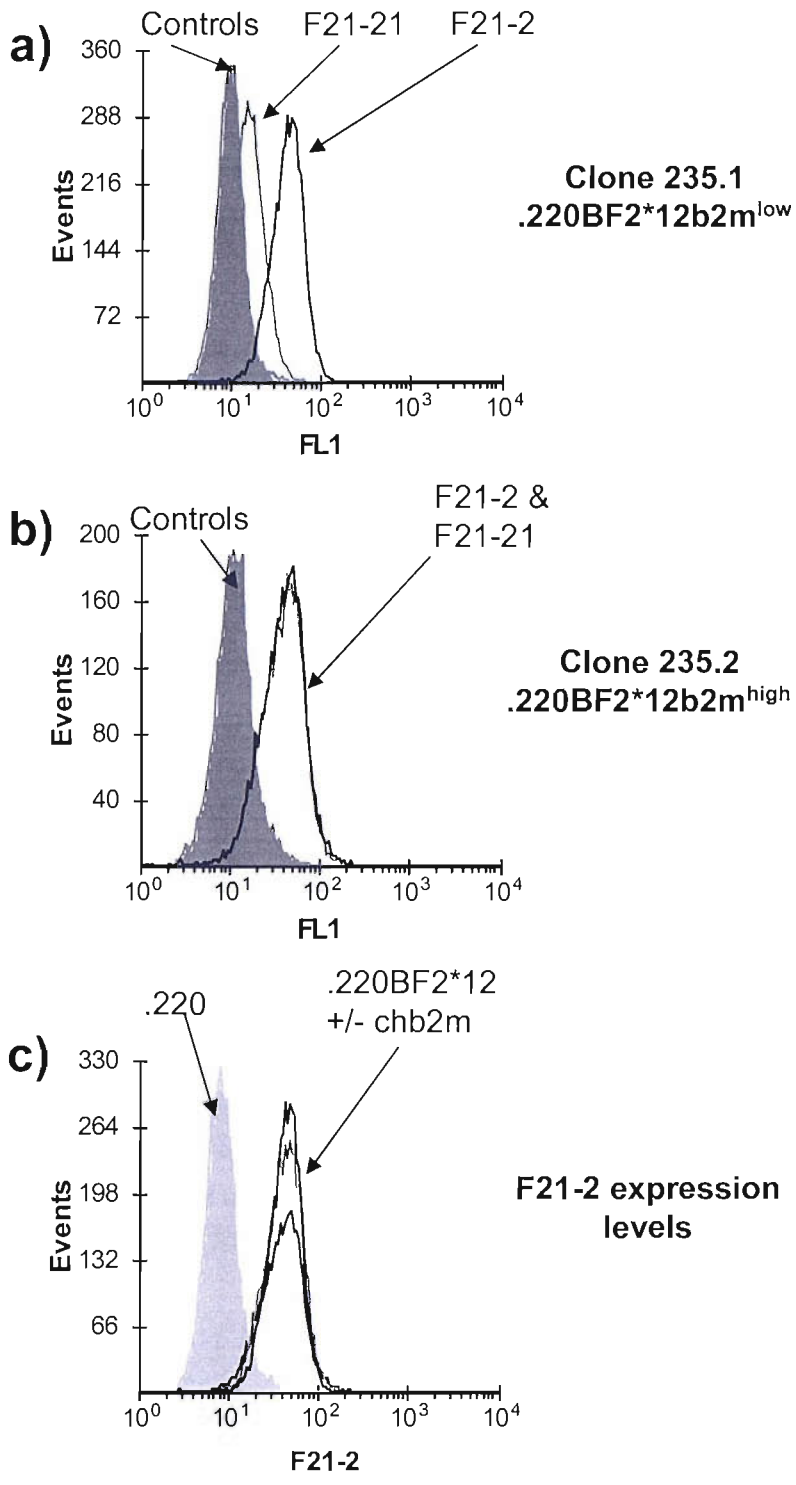


Figure 4.9 The effect of chicken β2m expression upon BF2*12 surface expression

a-b) Histograms depicting surface expression levels of .220BF2*12chb2m clones stained with either secondary antibody only (light grey, filled), TRT1 (dark grey, filled), F21-21 (grey, open) or F21-2 (black, open) antibodies.

c) Histogram depicting F21-2 surface expression levels upon .220 cells (light grey, filled), .220BF2*12 clone 107C12 (grey, open), and .220BF2*12ch.2m clones 235.1 and 235.2 (black, open).

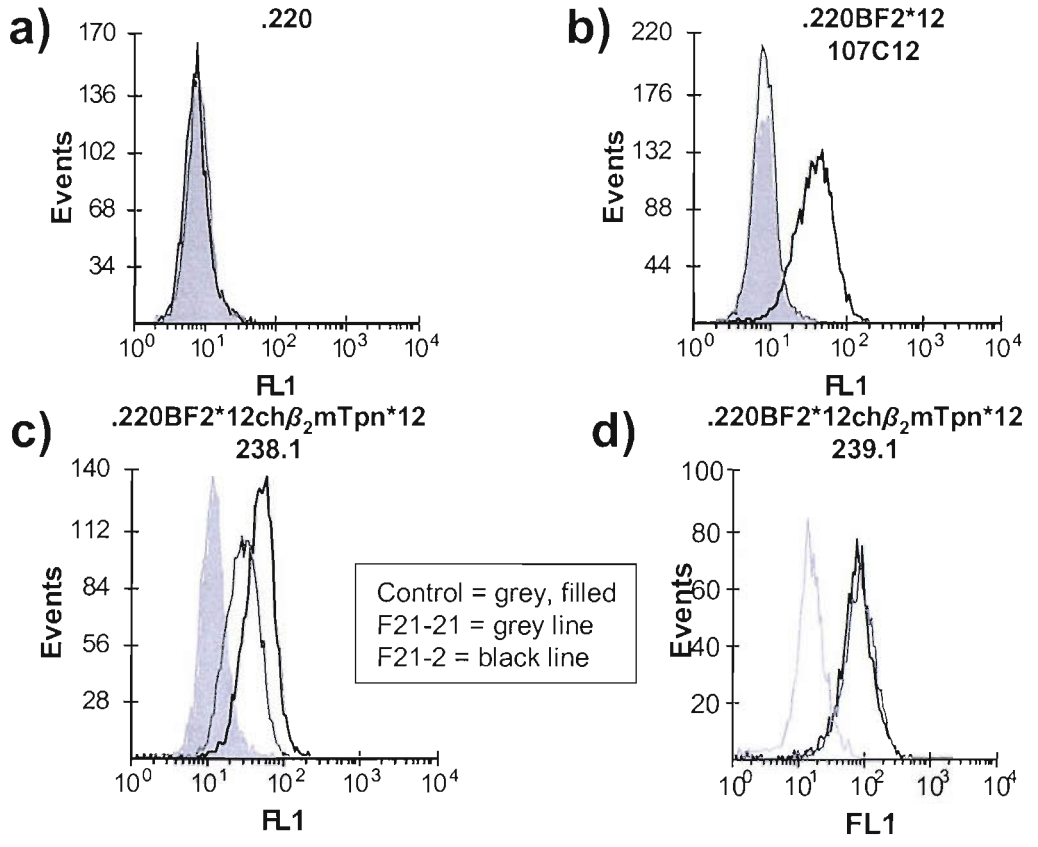
Table 4.7 and figure 4.10 shows that the amount of BF2*12 molecules expressed at the cell surface can be increased approximately two-fold by expression of both chicken β_2m and chicken Tapasin*12 proteins. For the clones that express a low amount of chicken β_2m (prefix 238) the surface BF2*12 expression level can be increased by almost double when a large amount of chicken Tapasin*12 is expressed, for example clone 238.4. However if less chicken Tapasin*12 is present, for example clone 238.1, the surface BF2*12 expression level is only moderately increased above that of tapasin-deficient or chicken β_2m-deficient cells.

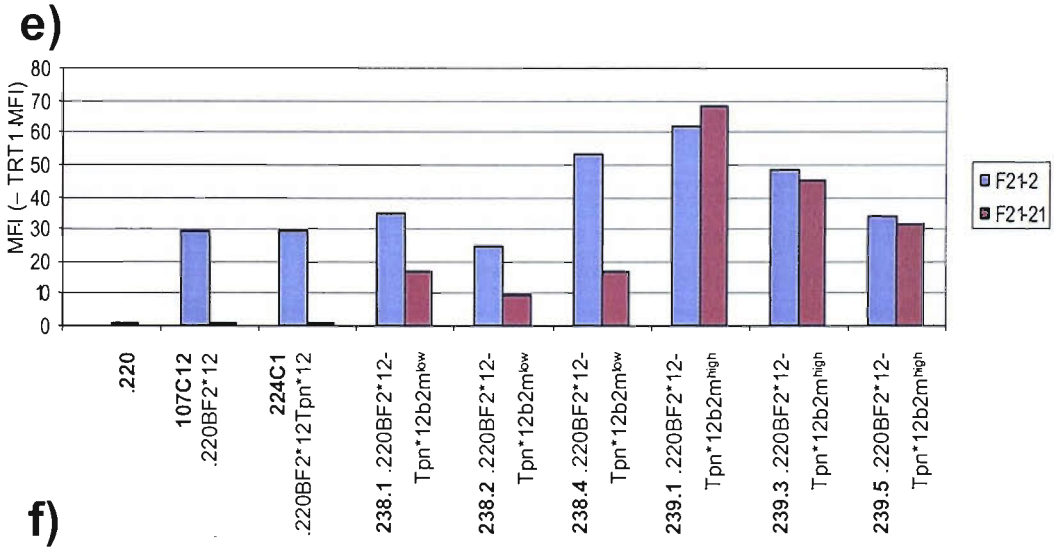
Clone	Transfectant proteins			TRT1	F21-2	F21-21
	BF2*12	Tapasin*12	ch b2m	MFI	MFI-TRT1	MFI-TRT1
.220	-	-	-	7.41	0.05	0.72
107C12	+	-	-	8.15	29.43	0.4
224C1	+	+	-	22.1	29.04	0.58
238.1	+	+	low	12.11	35.2	16.42
238.2/3	+	+	low	11.13	24.49	9.23
238.4	+	+	low	16.59	53.09	16.26
239.1	+	+	high	14.21	61.87	68.29
239.3	+	+	high	14.08	48.46	45.21
239.5	+	+	high	14.76	33.72	31.39

Table 4.7 Cell surface expression levels of BF2*12 (F21-2) and chicken β_2m (F21-21) on the indicated clones. TRT1 represents isotype control staining.

For the clones that express large amounts of chicken β_2m (prefix 239) the surface BF2*12 expression level doubled for clone 239.1 which had the highest chicken Tapasin*12 expression level of these cells. For clones 239.3 and 239.5 which express decreasing amounts of chicken Tapasin*12 protein, the surface BF2*12 expression level decreased in a chicken Tapasin*12 protein concentration dependent fashion to just above that observed for .220BF2*12 clone 107C12 or .220BF2*12Tpn*12 clone 224C1.

This suggests that the amount of BF2*12 molecules expressed at the cell surface can be moderately increased by combinations of chicken β_2m and chicken Tapasin*12 proteins. Interestingly it would appear high amounts of chicken tapasin can not fully compensate for lower amounts of chicken β_2m: for example .220BF2*12chβ_2m^{low} clone 238.4 expresses the highest amount of chicken tapasin protein of all the clones, but has a lower surface BF2*12 expression level than .220BF2*12chβ_2m^{high} clone 239.1 which expresses much lower amounts of chicken tapasin protein. This suggests that chicken β_2m molecules can limit the surface expression of BF2*12 molecules.





f)

Clone	.220	107C12	224C1	238.1	238.2	238.4	239.1	239.3	239.5
BF2*12	-	+	+	+	+	+	+	+	+
Tapasin*12	-	-	+	+	+	+	+	+	+
ch b2m	-	-	-	low	low	low	high	high	high

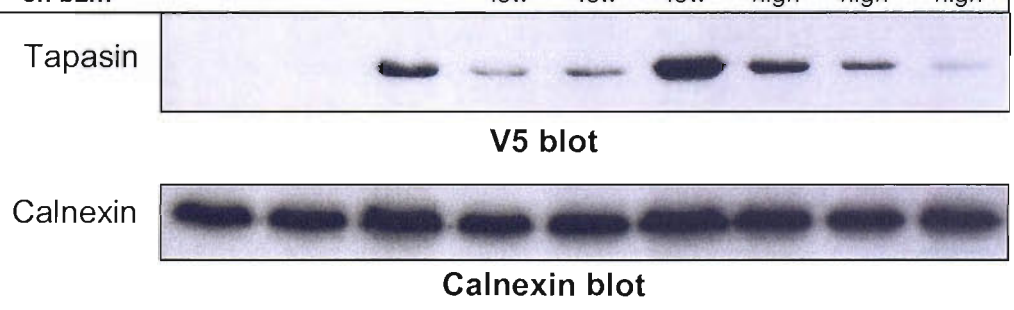


Figure 4.10 Flow cytometry and western blot analysis of BF2*12 transfectants .220BF2*12chβ_2mTapasin*12 transfectants

a-d) Histograms depicting surface expression levels for the indicated clones following staining with either a control antibody (TRT1, shown by grey, filled histogram), anti-chicken b2m F21-21 (grey line), or anti-chicken class I F21-2 (black line).
e) Graph depicting F21-2 and F21-21 MFI values (corrected by subtraction of TRT1 MFI) for the transfectants.
f) Lysates were made of the transfectants and SDS PAGE gels were loaded with cell lysate equivalent to $5x10^4$ cells and western blotted. Blots were incubated with the following antibodies and developed. Arrowheads mark the band of interest.

The amount of chicken β_2m detected on the cell surface by F21-21 staining and flow cytometry for the .220BF2*12chβ_2m^{high} clones was in agreement with the F21-2 staining results: there was a positive correlation between chicken Tapasin*12 protein concentration and the amount of surface expressed BF2 and chicken β_2m molecules. This suggests that as the amount of chicken Tapasin*12 protein decreases, fewer BF2: chicken β_2m heterodimers are loaded with repertoires of peptides that permit subsequent surface expression. For the .220BF2*12chβ_2m^{low} clones there was no such agreement between BF2 and chicken β_2m surface expression levels. This may be a consequence of the lower amount of chicken β_2m protein expressed by these cells: in effect all the chicken β_2m that these cells produce may still be associated with the greater amounts of BF2 molecules.

4.2.10 Peptide stabilisation assays and brefeldin A surface decay rates reveal no qualitative differences in BF2*12 molecules

Having found expression of both chicken Tapasin*12 and chicken β_2m proteins can increase the surface expression levels of BF2*12 molecules, two separate attempts were made to determine if the surface expressed molecules differed qualitatively. Firstly, the various BF2*12-expressing transfectants were cultured with B12 binding peptide FANYIDKV or in media alone in a class I stabilisation assay. Table 4.8 reveals that all clones were comparably affected by peptide addition, in that regardless of the presence of either chicken Tapasin*12, chicken β_2m or both proteins, there was a negligible stabilisation of BF2*12 molecules (4% or less) following peptide addition. This suggests the surface expressed BF2*12 molecules on the various transfectants did not differ in receptivity towards exogenous peptide stabilisation.

				F21-2		
Clone	**BF2*12**	**Tapasin*12**	**ch b2m**	**- peptide**	**+ FANYIDKV**	**Fold Increase**
.220	-	-	-	7.31	7.60	1.04
107C12	+	-	-	40.02	41.60	1.04
224C1	+	+	-	38.71	38.57	1.00
239.1	+	+	high	70.58	72.76	1.03
239.7	+	+	high	67.91	69.13	1.02

				F21-21		
Clone	**BF2*12**	**Tapasin*12**	**ch b2m**	**- peptide**	**+ FANYIDKV**	**Fold Increase**
.220	-	-	-	7.07	7.28	1.03
107C12	+	-	-	8.74	9.03	1.03
224C1	+	+	-	13.28	13.74	1.03
239.1	+	+	high	70.17	76.60	1.09
239.7	+	+	high	67.45	66.71	0.99

Table 4.8 BF2*12 (F21-2, top) and chicken β_2m (F21-21, bottom) expression levels following culture of the indicated clones in the presence or absence of specific exogenous peptide. Fold increase = MFI+peptide/MFI-peptide.

Independently of the above experiment, the various .220BF2*12 transfectants were cultured with brefeldin A, a drug which blocks ER/Golgi transport. Under these conditions, the time-courses in survival of surface BF2*12 molecules (table 4.9 and figure 4.11) were very similar among the transfectants, with BF2*12 expression levels reaching 72-81% of their initial levels after six hours of culture with BFA.

	Clone	BF2*12	Tapasin*12	ch b2m	Hours + BFA	Mean F21-2 MFI	St dev	% of 0hrs
.220	.220	-	-	-	0	11.66	0.21	100.0
					2	13.04	0.09	111.9
					4	12.31	0.19	105.6
					6	11.57	0.68	99.3
.220BF2*12	107C12	+	-	-	0	17.43	0.26	100.0
					2	14.25	0.07	81.7
					4	13.35	0.29	76.6
					6	12.93	0.16	74.1
.220BF2*12Tpn*12	224C1	+	+	-	0	21.97	0.36	100.0
					2	17.62	0.38	80.2
					4	15.58	0.39	70.9
					6	15.86	0.40	72.2
.220BF2*12b2m-high	235.2	+	-	high	0	18.83	0.48	100.0
					2	15.88	0.17	84.3
					4	15.75	0.23	83.7
					6	15.17	0.62	80.6
.220BF2*12Tpn*12b2m-high	239.1	+	+	high	0	22.49	0.27	100.0
					2	20.80	1.09	92.5
					4	19.43	0.27	86.4
					6	18.04	0.89	80.2

Table 4.9 BF2*12 (F21-2) expression levels following culture of the indicated clones with brefeldin A for the indicated periods of time.

The broadly equivalent disappearance of surface BF2*12 molecules in the various transfectants suggests that irrespective of the presence of chicken β_2m and/or chicken Tapasin*12 proteins, the BF2*12 molecules were loaded within the ER with repertoires of peptides that conferred equal levels of stability to the molecules.

The combination of these two experiments suggests that whilst the number of BF2*12 molecules expressed at the cell surface may have varied: a consequence of the presence of both chicken β_2m and chicken Tapasin*12 proteins; the peptides that were loaded into the molecules led to comparable stabilisation of the surface expressed BF2*12 molecules.

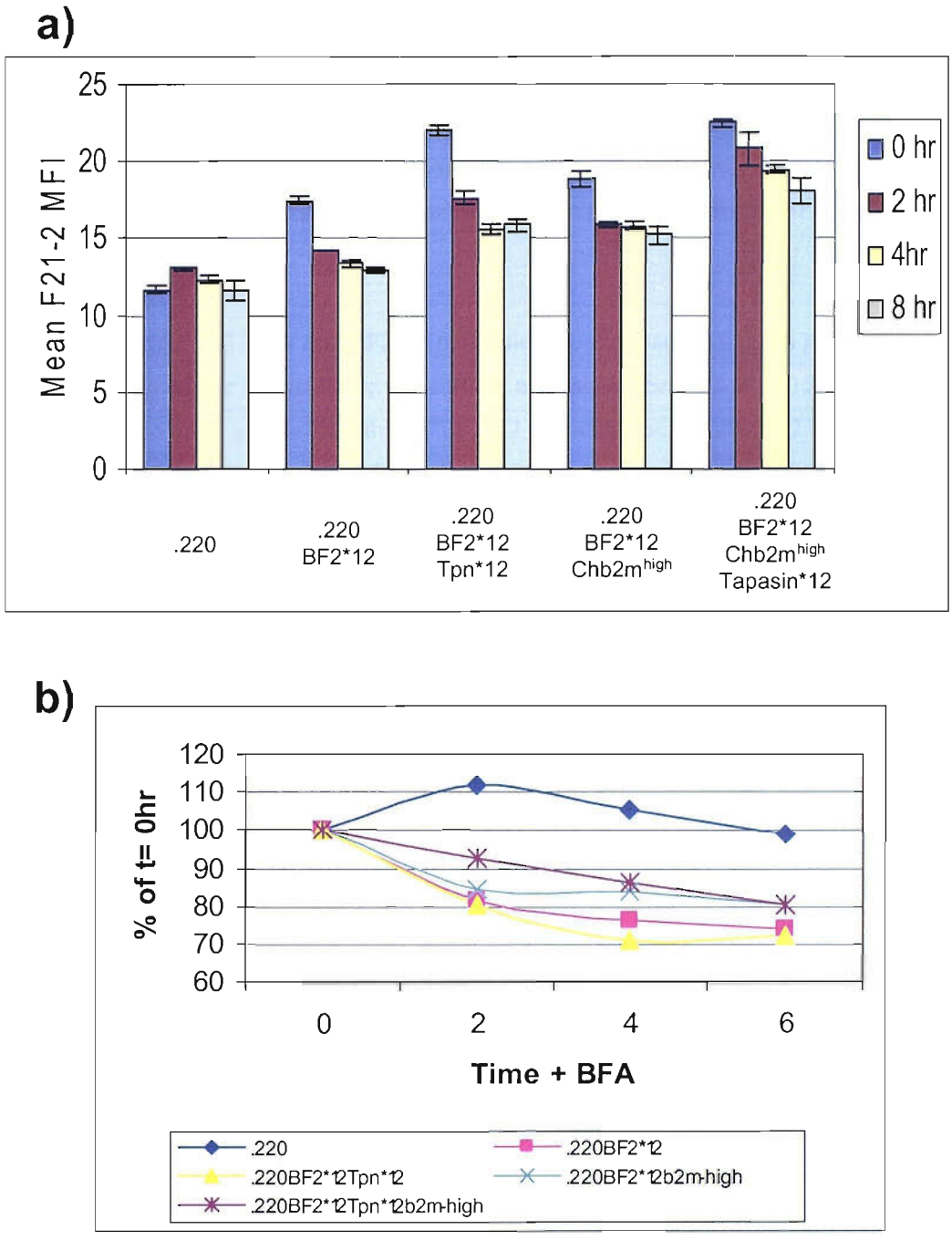


Figure 4.11 Surface BF2*12 expression levels upon .220BF2*12 transfectants following culture with Brefeldin A (BFA)

a) Graph depicting mean F21-2 MFI expression levels for .220BF2*12 transfectants following culture with BFA (mean taken of triplicate samples). Error bars depict +/- 1 standard deviation.
b) Graph depicting mean F21-2 MFI for .220BF2*12 transfectants relative to mean F21-2 MFI recorded at t = 0hr.

4.2.11 BF2*15 molecules are expressed in greater amounts than BF2*12 at the cell surface of .220 cells, but the expression level is not increased by chicken Tapasin*15 protein expression

Whilst the transfections and the experiments that are discussed above were being conducted, a different class I allele was separately transfected into the .220 cell line. This was conducted to determine the effect of chicken tapasin expression on a different chicken class I molecule, and was conducted prior to the characterisation of the apparent incompatibility between BF2*12 and human β_2m molecules.

Table 4.10 and figure 4.12 show that BF2*15 molecules are expressed at a much greater level than BF2*12 molecules at the cell surface; the surface expression level of BF2*15 molecules was approximately nine times greater than that of BF2*12 molecules on .220BF2*12 clone 107C12 in this experiment. Western blot analysis (figure 4.12) of lysates made from the different transfectants shows that more BF2 molecules are present in .220BF2*15 cells than in .220BF2*12 clone 107C12, relative to the amount of protein detected by incubating the blots with an anti-calnexin antibody, which demonstrates roughly equivalent amounts of protein were loaded into each lane.

Cell	mAb	MFI	MFI - TRT1
.220	TRT1	10.23	
	F21-2	10.16	-0.07
	F21-21	10.48	0.25
.220BF2*15	TRT1	19.81	
	F21-2	273.8	253.99
	F21-21	19.68	-0.13
.220BF2*15Tpn*15	TRT1	22.81	
	F21-2	202.71	179.90
	F21-21	22.61	-0.20
.220BF2*12 clone 107C12	TRT1	12.23	
	F21-2	41.09	28.86
	F21-21	12.4	0.17

Table 4.10 BF2 (F21-2), chicken β_2m (F21-21) and control isotype (TRT1) cell surface expression levels for the indicated clones.

To determine how surface BF2*15 expression levels were affected by expression of the chicken Tapasin*15 protein, .220BF2*15 cells were transfected with an expression construct encoding V5 epitope-tagged chicken Tapasin*15. Only one clone which expressed chicken Tapasin*15 was generated (figure 4.12), and in this clone the amount of BF2*15 molecules expressed at the cell surface (table 4.10 and figure 4.12) was not increased above that observed for .220BF2*15 cells; in fact the amount of BF2*15 expressed was marginally decreased following tapasin expression, which was evident both at the cell surface and in the total BF2*15 protein level. The significance of this modest decrease in BF2*15 expression can only be assessed accurately if more clones are generated, which would determine whether this observation is a clone-specific artefact.

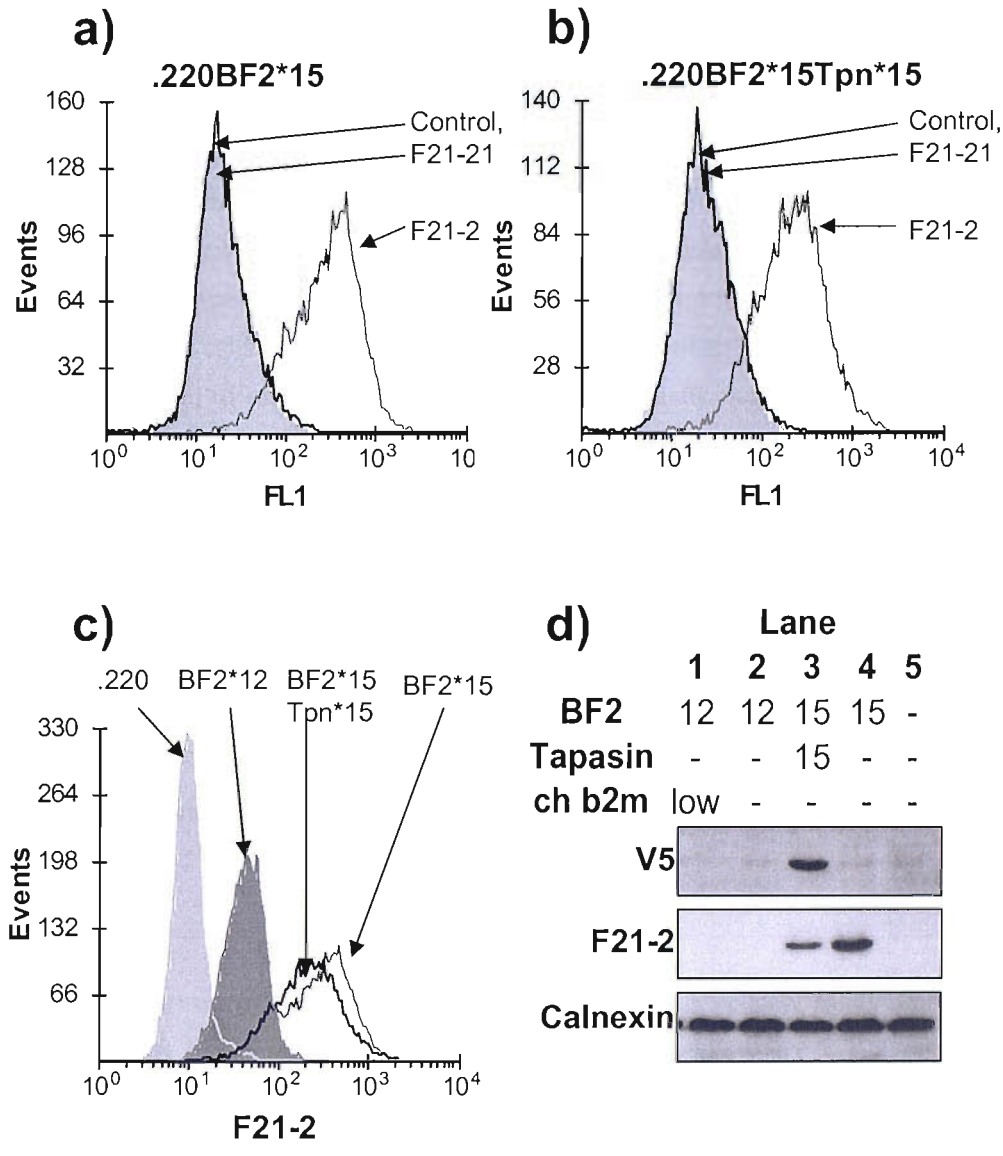


Figure 4.12 Flow cytometry analysis of .220BF2*15 and .220BF2*15Tpn*15 clones

a-b) Histograms depicting surface expression levels when the indicated transfectants were stained with either control antibody (TRT1 shown by grey, filled histogram), anti-chicken b2m F21-21 (black line) or anti-chicken class I F21-2 (grey line) antibodies.

c) Histogram depicting F21-2 surface expression upon .220 (light grey, filled), .220BF2*12 clone 107C12 (grey, filled), .220BF2*15 cells (grey, open) and .220BF2*15Tpn*15 (black, open).

d) Western blot of the various transfectants:

Lane 1: .220BF2*12ch.β_2m clone 235.1;

Lane 2: .220BF2*12 clone 107C12;

Lane 3: .220BF2*15Tpn*15;

Lane 4: .220BF2*15;

Lane 5: .220 cells.

Each lane was loaded with lysate equivalent to $2x10^5$ cells. The western blot was sequentially incubated with the listed antibodies and developed.

Nevertheless it appears that similar to BF2*12, expression of the haplotype-matched chicken tapasin allele did not lead to large alterations in the amount of surface expressed BF2*15 molecules.

In support of the apparent incompatibility between BF2 and human β_2m molecules is the finding that the surface expression of either BF2*12 or BF2*15 molecules did not substantially increase the amount of human β_2m molecules expressed at the cell surface (table 4.11), assessed by staining with the BBM1 antibody and flow cytometry, relative to either .220 cells (that express low levels of HLA-C) or .220B*4402 cells (that additionally express low levels of HLA-B*4402). In contrast the surface expression level of human β_2m molecules was substantially increased on human tapasin and HLA-B*4402-expressing .220B*4402hTpn cells.

	secondary only		BBM1		F21-2		W6/32	
Clone	**Average**	**St dev**	**Average**	**St dev**	**Average**	**St dev**	**Average**	**St dev**
.220	3.72	0.09	73.43	2.02	5.14	0.04	79.43	2.71
.220BF2*12	3.85	0.08	53.03	0.20	6.62	0.06	59.10	0.20
.220BF2*15	4.91	0.08	89.80	1.33	25.06	0.84	77.99	1.51
.220B4402	4.20	0.08	77.37	0.39	4.55	0.23	89.62	1.02
.220B4402hTpn	3.60	0.10	1198.69	13.13	4.20	0.06	1523.23	23.79

Table 4.11 Cell surface expression levels of human β_2m (BBM1), chicken class I molecules (F21-2), and HLA: human β_2m heterodimers (W6/32).

No further experiments were conducted with these BF2*15-expressing clones, as during the course of these experiments the potential incompatibility between chicken class I and human β_2m molecules became apparent, and other areas of research were being conducted that are discussed in the following two chapters.

4.3 Discussion

The aim of the research presented in this chapter was to determine whether the sequence polymorphisms identified in chicken tapasin alleles lead to functional differences in the ability of tapasin to function. To achieve this, the human tapasin-deficient .220 cell line was reconstituted with combinations of chicken genes. Unfortunately this aim was ultimately not answered, a consequence of the problems encountered resulting from apparent molecular incompatibilities between chicken and human proteins. However several intriguing observations were made during the course of these experiments.

A major handicap in this approach proved to be the failure to detect the expression of chicken tapasin His_6 protein following transient transfection in .220 cells. This appeared to be chicken tapasin His_6-specific as other proteins were successfully expressed following transient transfection, and also host cell-specific as chicken tapasin His_6 protein expression was detected following transient transfection in other cell lines (including a human cell line), which may be a consequence of higher transfection efficiency in these cells. Attempts to produce stable transfectant .220 cell lines expressing chicken tapasin did subsequently prove successful, however in these cases a V5 epitope tag was inserted into the C-terminus of chicken tapasin instead of the His_6 tag initially used. It is likely that the detection of V5-tagged chicken tapasin is more efficient than the detection of His_6 tagged chicken tapasin; alternatively it is possible that the His_6 tag, or duplicated dilysine motifs, may have destabilised chicken tapasin molecules in .220 cells, leading to their rapid degradation. V5 epitope-tagged chicken Tapasin*12 protein expression can be detected following transient transfection of .220 cell lines (data not shown), however, no further transient transfection experiments were performed for two reasons: firstly, the transient transfection efficiency was very low; and secondly that stable V5-tagged chicken tapasin-expressing transfectants were quickly generated, which offered the advantage of being clonal. The transfectants that were produced and a summary of the key findings are provided in figure 4.13.

The majority of the research that was conducted concentrated on the transfection of .220BF2*12 clone 107C12 with additional chicken gene expression constructs. In retrospect, this clone may have been a poor choice; as the clone exhibited low BF2*12 expression levels, which was evident at the cell surface and in total cellular lysates. The low BF2*12 molecule expression level may be a reflection of: a low gene copy number; the site of integration into the genome; sub-optimal peptide-loading (perhaps due to the absence of chicken tapasin, or other chicken molecules), or perhaps limited by inadequate peptide supply, which may have led to the rapid degradation of the BF2*12 protein.

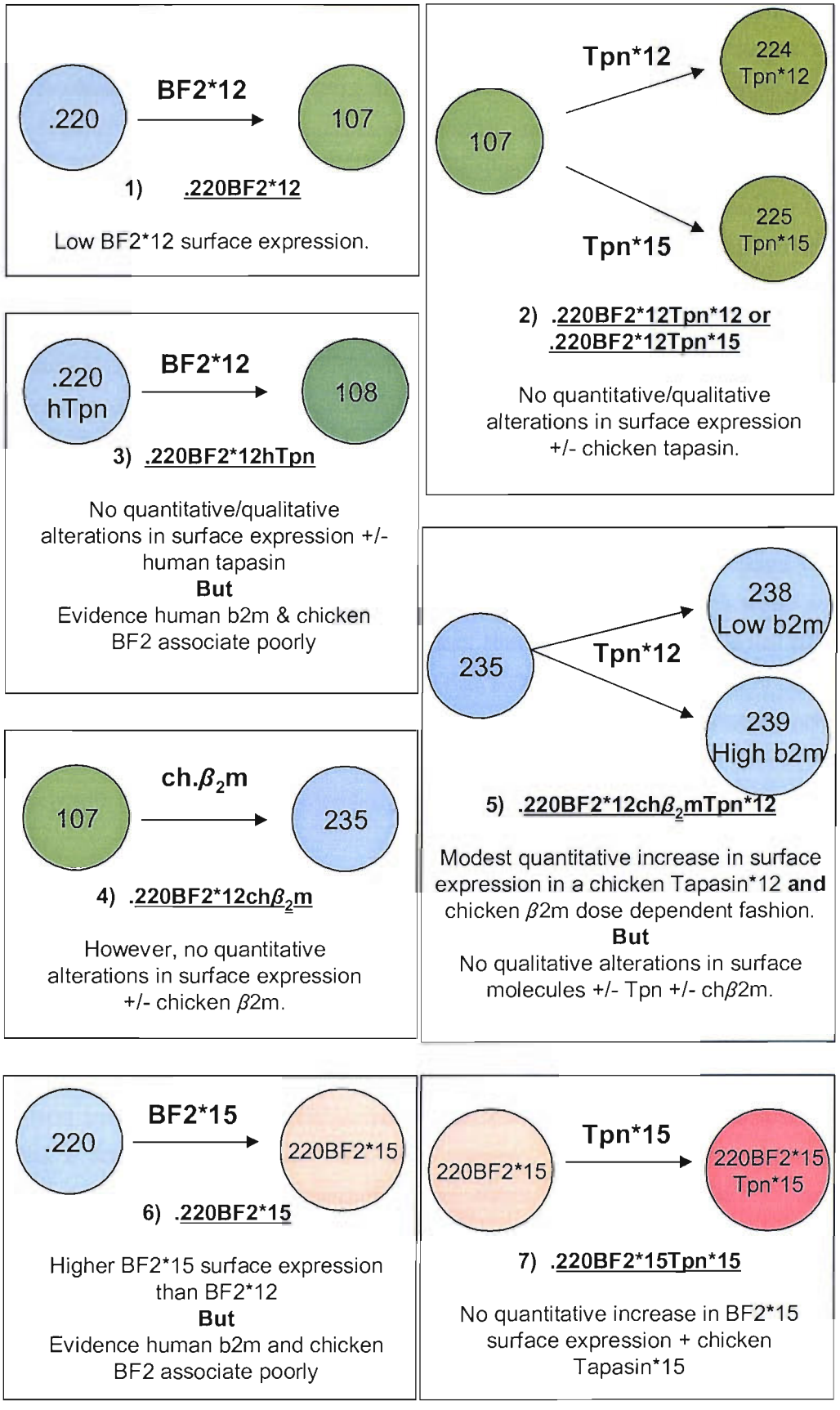


Figure 4.13 Summary of transfectants and key findings.
The transfectants are presented in the order that they are discussed in this chapter.

Having obtained BF2*12 expressing clones, attempts were made to determine whether a haplotype-mismatched combination of tapasin and class I alleles functioned comparably to the haplotype-matched combination, whereby .220BF2*12 clone 107C12 was transfected with V5 epitope-tagged chicken Tapasin*12 or a haplotype-mismatched chicken tapasin allele, Tapasin*15. However steady state flow cytometry, exogenous peptide stabilisation assays, reduced temperature culture assays, and brefeldin A culture assays all suggest the presence of either the haplotype-matched or haplotype-mismatched chicken tapasin allele did not quantitatively or qualitatively affect the surface expression of BF2*12 molecules in comparison to the tapasin-negative parental clone.

Transfectants of .220 cells expressing a different chicken class I allele, BF2*15, were generated for two reasons: to compare the expression of BF2*15 to BF2*12 and to determine how chicken Tapasin*15 affected BF2*15 expression. BF2*15 was expressed in much greater amounts than BF2*12, both at the cell surface and in cellular lysates. It is possible that the copy number or the location that the BF2*15 gene integrated into the genome may explain the greater BF2*15 protein and surface expression levels, which could be determined by performing southern blots (which would reveal the gene copy number) and by ^{35}S pulse/chase experiments (which would reveal differences in the rates of synthesis of BF2 molecules). However it is also possible that BF2*15 molecules are somehow stabilised to a greater extent than BF2*12 molecules are in .220 cells. For example the human TAP molecules may supply BF2*15 molecules with a more suitable repertoire of peptides than they supply BF2*12 molecules, alternatively BF2*15 may interact more productively with human chaperone molecules, perhaps calnexin, calreticulin, ERp57 or even β_2m.

However, in agreement with the previous findings concerning BF2*12 and chicken Tapasin*12, expression of the haplotype-matched chicken tapasin allele, Tapasin*15, failed to quantitatively alter BF2*15 surface expression levels. Therefore for both class I alleles examined, there was no chicken tapasin-dependent alterations in the surface expression of BF2 molecules.

To determine whether human tapasin affects the expression of BF2*12 molecules, the .220hTpn cell line was transfected with BF2*12. Despite the presence of a larger total pool of BF2*12 molecules, a combination of steady state flow cytometry, exogenous peptide and/or β_2m stabilisation assays and reduced temperature culture assays all suggest that BF2*12 surface expression was comparable to that of tapasin-deficient cells, a finding which allows two conclusions to be drawn. Firstly, it suggests the peptide supply was not solely limiting BF2*12 surface expression (as human tapasin expression is likely to improve the supply of peptides into the ER); and secondly, that human tapasin does not mediate a chaperone or peptide-loading/editing effect beneficial to the surface expression of BF2*12 molecules, which may indicate that the chicken class I molecule can not interact with the human orthologue of tapasin.

The unaltered surface expression characteristics of BF2 molecules in tapasin-expressing cells in comparison to tapasin-negative parental cells has many possible explanations. One explanation may be that, at least in .220 cells, BF2*12 and BF2*15 molecules can efficiently load with optimal repertoires of peptides without the assistance of chicken (or human) tapasin. However other explanations are also possible, and in the course of the analysis of the various transfectants, two observations were made that suggest that chicken class I molecules may not associate productively with human β_2m, which may prevent a productive association with tapasin. Firstly, it was observed that the addition of exogenous chicken β_2m led to an increase in the surface expression level of BF2*12 molecules, and secondly, that the surface expression level of human β_2m was not increased (relative to untransfected .220 cells) by the transfection and surface expression of BF2*12 or BF2*15 molecules. Such molecular incompatibilities are not without precedent: for example when human class I molecules are transfected into murine cell lines, human β_2m often must to be transfected in addition in order to achieve surface expression levels comparable to those observed in human cells (Alexander et al., 1989; Shields et al., 1998; Shields et al., 1999). There are other examples of molecular incompatibilities which occur in such species-switching experiments (Peh et al., 2000).

Therefore, to analyse this potential molecular incompatibility in more detail, chicken β_2m was transfected into the .220BF2*12 clone 107C12. Whilst the expression of chicken β_2m alone did not quantitatively increase BF2*12 surface expression, the expression of both chicken Tapasin*12 and chicken β_2m proteins in the same cell did increase surface BF2*12 expression levels, in a chicken β_2m and tapasin protein concentration dependent fashion. However the increase in surface expression was modest, with expression levels being doubled at best. Exogenous peptide stabilisation assays and brefeldin A culture assays failed to reveal qualitative differences in the surface expressed BF2*12 molecules irrespective of the presence of chicken β_2m and/or chicken tapasin. This suggests that the modest increase in the amount of surface expressed BF2*12 molecules is a consequence of an increase in the number of BF2*12 molecules that were loaded with peptides within the ER, which however conferred comparably stability to that achieved by the fewer number of surface expressed molecules that were loaded in the absence of chicken β_2m or tapasin.

The finding that large amounts of chicken β_2m are needed for BF2*12 surface expression levels to increase in a chicken tapasin protein dependent fashion appears to confirm that human β_2m can not completely replace the function of its chicken orthologue. In conditions when chicken β_2m is limiting (such as in the .220BF2*12chβ_2m^{low}Tpn*12 clones), large amounts of chicken tapasin could not increase BF2*12 surface expression levels to the same extent as .220BF2*12chβ_2m^{high}Tpn*12 clones that express lower amounts of chicken tapasin. In mammalian cells few mammalian class I molecules associate with the PLC in the absence of β_2m, and so it appears likely that in the absence of large amounts of chicken β_2m, few BF2 molecules are able to

interact with the chicken tapasin molecule. It is likely that BF2*12 heterodimers adopt slightly different conformations when coupled with either human or chicken β_2m molecules, and that chicken tapasin molecules can only associate with the conformation conferred by the chicken β_2m molecule.

It is possible that other molecular incompatibilities exist which limit a more comprehensive chicken tapasin-dependent and chicken β_2m-dependent increase in BF2*12 surface expression levels. For example, it has been documented that murine tapasin and human tapasin mutants can not form a fully functional PLC in human cells (Tan et al., 2002), so it is likely that chicken tapasin molecules also fail in recruiting chicken class I: chicken β_2m heterodimers to form a fully functional PLC in human cells. Attempts were therefore made to determine whether chicken tapasin could associate with the human TAP proteins, and whether such an association could increase BF2*12 surface expression levels. Firstly, attempts were made to immunoprecipitate the human TAP molecules in order to determine whether chicken tapasin could be co-precipitated, however these experiments have so far proven unsuccessful (data not shown). Secondly, hybrid tapasin molecules were produced, comprising the ER luminal portion of the chicken tapasin molecule, and the transmembrane and cytoplasmic domains of the human tapasin molecule. Unfortunately no hybrid tapasin expressing transfectants were generated in two independent transfections of the .220BF2*12 clone 107C12.

The requirement not only for generating triple transfectants of .220 cells (chicken β_2m, BF2 and chicken tapasin alleles), combined with the low efficiency of generating stable transfectants, and also for accurately determining the expression level of each protein (especially chicken β_2m and chicken tapasin) ultimately undermined attempts to determine whether haplotype-mismatched chicken tapasin alleles could fully replace the function of the haplotype-matched tapasin allele.

The findings learnt through these experiments suggest that it may be advantageous to start afresh, producing the triple transfectants in a different order and fashion. It would be wise to produce a chicken β_2m expressing clone of .220 that could be the basis of the subsequent experiments, into which different BF2 alleles could next be transfected. Lastly, into these double transfectant clones, different chicken tapasin alleles could be introduced. One additional improvement may be to use expression vectors that allow accurate assessment of intracellular protein levels via a surface expressed receptor whose expression is driven off the same promoter as the gene of interest, making use of an internal ribosome entry site, similar to those used by Howarth et al., 2004. However, these experiments were not pursued further, in most part due to pursuing more profitable lines of research detailed in the following chapters. However, the apparent molecular incompatibilities unearthed in these experiments may not be the full extent of the potential problems that could complicate the analysis of chicken class I peptide-loading in a human cell line.

5. Reconstitution of a chicken MHC-homozygous cell line transfected with MHC-mismatched proteins

5.1 Introduction

The previous chapter described attempts to determine whether polymorphic amino acid residues in chicken tapasin affect the surface expression of class I molecules when transfected into a human tapasin-deficient cell line. Unfortunately, a combination of problems undermined any detailed analysis from being made, but provided some interesting findings. For example, it became apparent that chicken BF2 molecules do not associate productively with human β_2m molecules. Therefore as described in this chapter, a different approach was undertaken to achieve the same goal, without any species-specific incompatibilities occurring. The approach is outlined in figure 5.1.

5.1.1 Strategy

The approach involved the transfection of a chicken B15 MHC-homozygous cell line with combinations of chicken proteins. As a starting point, the expression of two epitope-tagged class I (BF2) alleles was compared, that were either encoded from the same haplotype as the endogenous cell line (BF2*15) and so should be optimally loaded with peptides, or were encoded from a separate haplotype that may consequently be impaired in peptide-loading (step 1 in figure 5.1).

In order to distinguish the transfected BF2 molecules from the endogenous class I molecules, a Flag epitope tag was placed at the N-terminus of the class I heavy chain. This approach has been used successfully for chicken class I molecules without preventing antigen presentation to CTLs or alloantibody recognition (Fulton et al., 1995). In addition, mutant H-2K^d molecules with 20 additional amino acids preceding the α1 domain were expressed at the cell surface comparably to wild-type molecules, as assessed by flow cytometry following staining with either of two specific antibodies, and were recognised by allogeneic CTL (Hedley et al., 1989). In addition to the N-terminal Flag epitope tag, a myc epitope tag was placed at the C-terminus of the class I molecules, as C-terminally tagged class I molecules have been made without adversely affecting their association with the PLC or cell surface expression (Marguet et al., 1999; Spiliotis et al., 2000).

BF2*4 was chosen as the mismatched class I allele as this molecule has a totally disparate peptide-binding motif from the BF2*15 molecule (table 4.1). The expression of the mismatched BF2*4 molecule is likely to be severely compromised by a scarcity of suitable peptides; as the endogenous TAP*15 heterodimers have a restrictive transport specificity and are unlikely to transport peptides suitable for BF2*4 molecules (figure 5.2, B. Walker, and J. Kaufman personal communication). Therefore, in effect the transfection of BF2*4 molecules into the B15 homozygous cell line is likely to create a situation akin to the "cim" phenomenon. However, in addition to the scarcity of suitable peptides, the mismatched BF2*4 molecule may be compromised by a molecular incompatibility with the endogenous Tapasin*15 molecule. In order to determine the elements required for efficient surface expression of BF2*4 molecules, the approach relied upon generating three further transfectant cell lines, as described below.

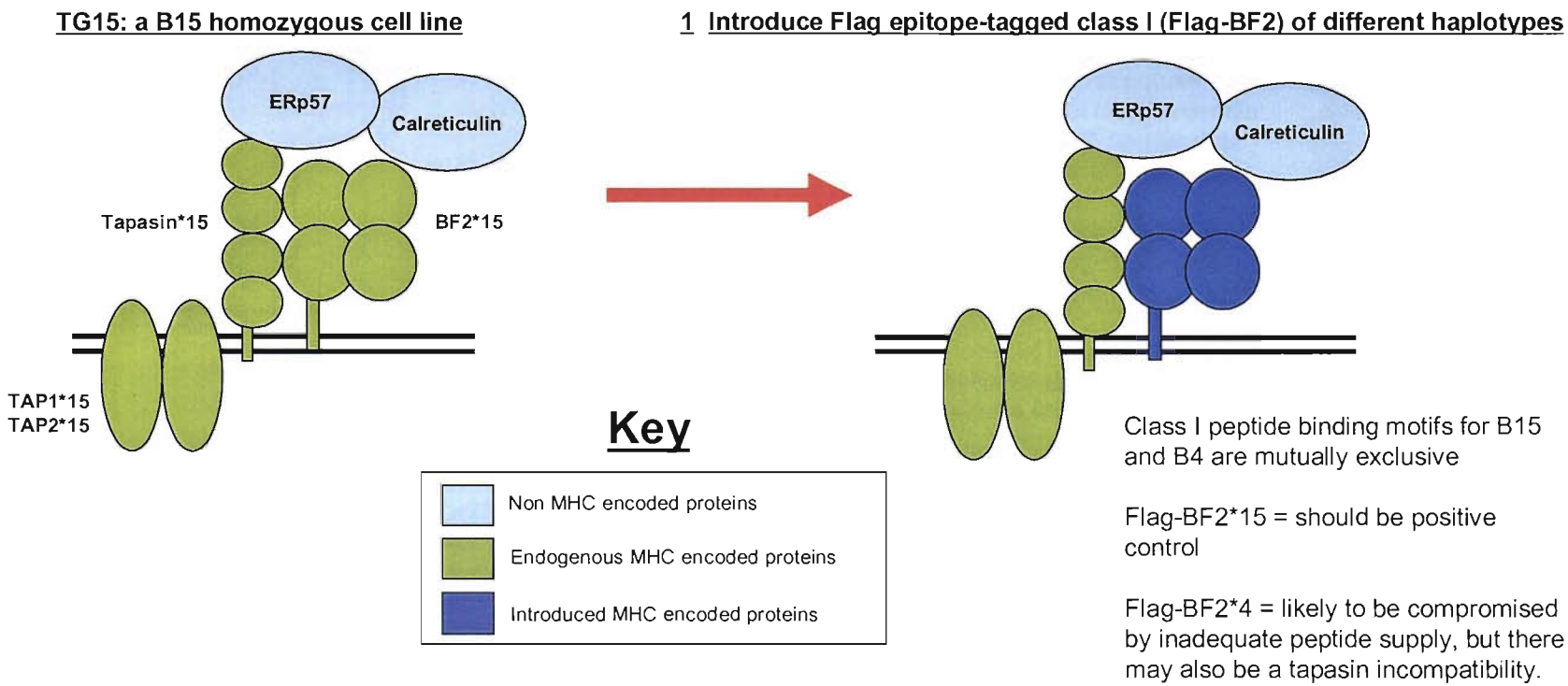


Figure 5.1 Part one of the experimental strategy.

Cartoons depict the likely peptide loading complexes present in the transfectants, with potential incompatibilities detailed. At all stages endogenous molecules may compete with transfected molecules (not shown). In all cases, steady state surface expression levels was used to compare the expression of Flag-BF2*4 to Flag-BF2*15, which was used as a reference of an optimally loaded molecule.

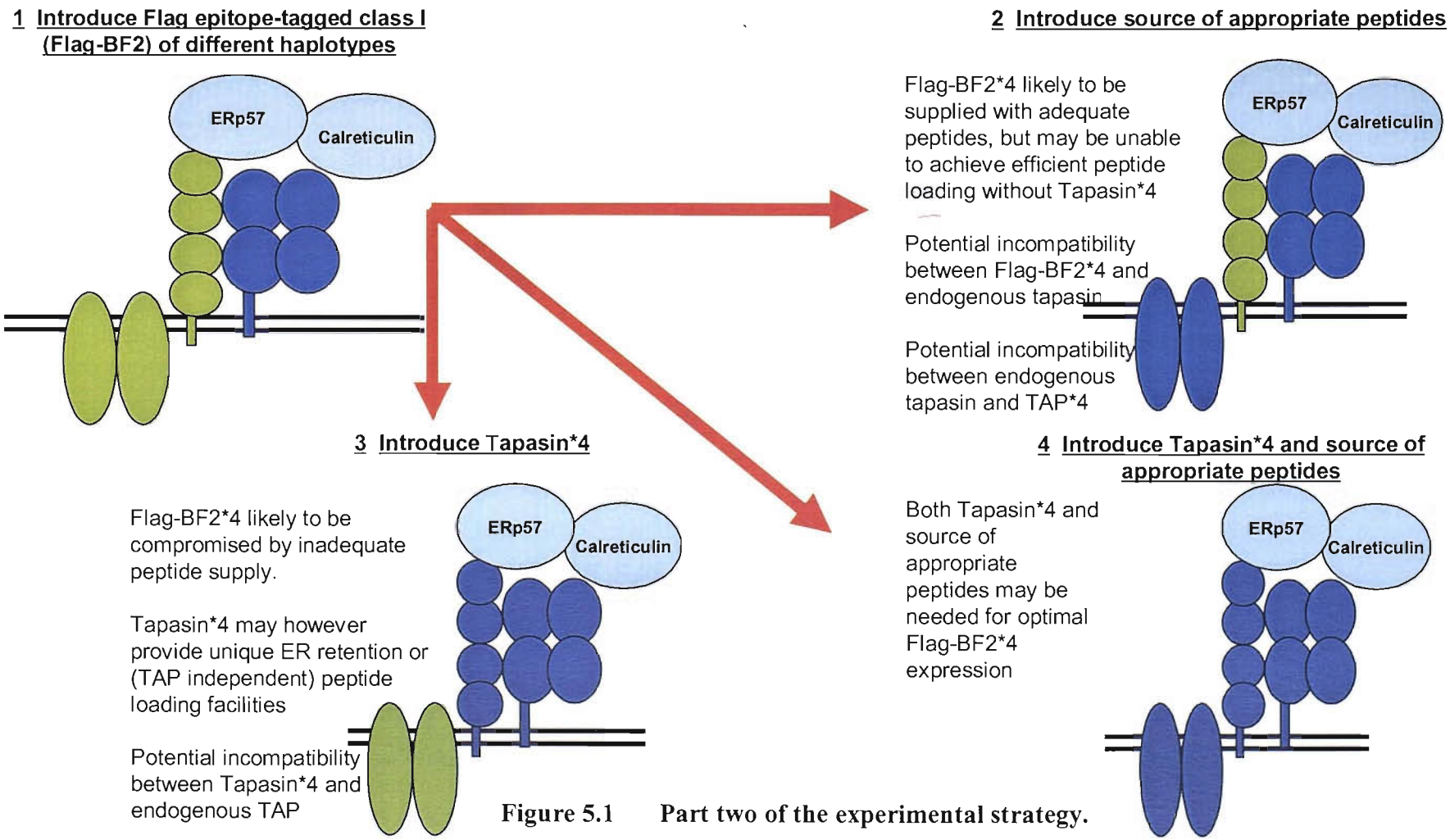


Figure 5.1 Part two of the experimental strategy.

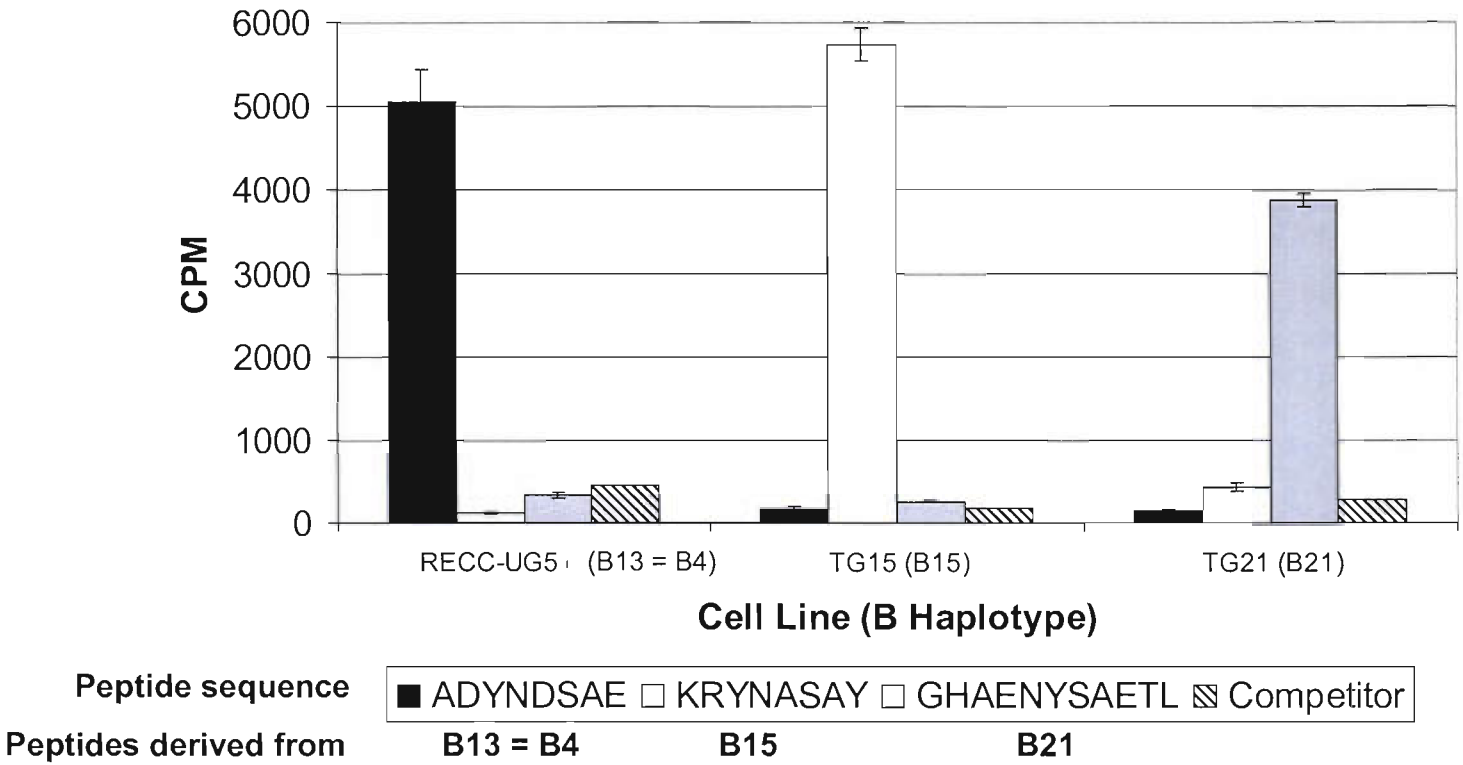


Figure 5.2 The restrictive TAP transport specificity evident in different MHC haplotypes

The indicated peptides were iodinated and individually cultured with ATP and microsomes prepared from each of the cell lines (note the B13 haplotype is identical to the B4 haplotype), before being cooled and lysed by the addition of N-P40 lysis buffer. Glycosylated peptides were then recovered by the addition of Con-A sepharose and the amount of radioactivity recovered determined by γ-counting. **Data kindly provided by Drs. B. Walker and J. Kaufman (unpublished).**

Firstly, attempts were made to reconstitute the BF2*4-expressing transfectants with a source of peptides suitable for the BF2*4 molecule (either by transfection of both TAP1*4 and TAP2*4, or by the transfection of peptide "minigenes", step 2 in figure 5.1). In such transfectants, there should be an appropriate source of peptides, but a molecular incompatibility with the endogenous Tapasin*15 allele may still compromise peptide-loading.

Secondly, attempts were made to reconstitute the BF2*4-expressing transfectants with Tapasin*4 (step 3 in figure 5.1). In such transfectants, the expression of BF2*4 molecules is still likely to be limited primarily by the absence of suitable peptides, but Tapasin*4 proteins may provide ER retention or (TAP-independent) peptide-loading facilities that are not provided by endogenous Tapasin*15 proteins.

Thirdly, it was aimed to reconstitute the BF2*4-expressing transfectants with both Tapasin*4 and an appropriate source of peptides (step 5 in figure 5.1). This would determine whether Tapasin*4 molecules facilitate peptide-loading to an extent beyond that when there is an abundant supply of suitable peptides, but only the endogenous Tapasin*15 molecules.

Ultimately this approach failed as it was not possible to generate all of the necessary transfectants; it proved impossible to transfect an appropriate source of peptides suitable for BF2*4 molecules, without which the BF2*4 molecules were severely compromised in their ability to be expressed at the cell surface. Without suitable peptides, only a small proportion of BF2*4 molecules reached the cell surface, but a combination of assays suggested these molecules primarily represented free heavy chains that had dissociated from β_2m. However, during the analysis of these transfectants, it became apparent that the N-terminal Flag epitope-tag compromised the ability of the BF2*15 molecules to load efficiently with peptides. This revealed that N-terminal Flag epitope-tagged class I molecules are not fully interchangeable with wild-type molecules, a finding which further compromised this strategy. Despite these failings, this line of research is presented in this thesis as key findings are of significance for the following chapter.

5.2 Results

In order to investigate the potential implications of mismatching class I, tapasin and TAP alleles, the TG15 cell line, a B15 MHC haplotype homozygous chicken cell line, was transfected firstly with epitope-tagged BF2 alleles from the haplotype-mismatched B4 or the haplotype-matched B15 haplotypes (step 1, figure 5.1). The transfected class I molecules were Flag tagged at their N-terminus and myc tagged at their C-terminus, and are referred to as Flag-BF2*4 or Flag-BF2*15 hereafter. Whilst attempts were made to generate the necessary transfectants for the ultimate success of this strategy (steps 2, 3 and 4, figure 5.1), the expression characteristics of the two Flag-BF2 alleles were compared in detail to determine the extent to which the haplotype-mismatched Flag-BF2*4 molecules were compromised, relative to the haplotype-matched Flag-BF2*15 molecules, when expressed in a B15 cell line.

5.2.1 Flag-BF2*4 is expressed at a lower level than Flag-BF2*15 in TG15 cells.

Four stable Flag-BF2*4-expressing transfectants were generated (clones Flag 4-2, 4-3, 4-5 and 4-6), whilst five stable Flag-BF2*15-expressing transfectants were generated (clones Flag 15-4, 15-7, 15-8, 15-9 and 15-11). Figure 5.3a-b shows a representative example of the steady state surface expression levels of the Flag-BF2 molecules as assessed by Flag epitope staining and flow cytometry. In other experiments minor fluctuations in the expression levels was observed between transfectants (as the experiments were conducted over a period of time during which the cells were frozen and thawed on numerous occasions), however in all cases the general trend remained that Flag-BF2*15 molecules were expressed at a much higher level than Flag-BF2*4 molecules at the cell surface. The Flag-BF2*4 expression level was nevertheless greater than background level, a finding which was unexpected based upon the restrictive nature of the TAP*15 peptide transporters (B. Walker, J. Kaufman, personal communication). The difference in surface expression levels of Flag-BF2 molecules was also evident in western blot analysis of the transfectants (figure 5.3c).

5.2.2 Flag-BF2*4 molecules are present in larger amounts than Flag-BF2*15 molecules within TG15 cells

It was possible that the difference in surface Flag-BF2 expression levels observed between Flag-BF2*15 and Flag-BF2*4 transfectants was merely a reflection of the lower total Flag-BF2*4 protein expression level. Therefore, the cellular distribution of Flag-BF2 molecules and β_2m-associated molecules was compared by performing flow cytometry on the transfectants that had been permeabilised by saponin treatment or left intact; reasoning that if a large proportion of Flag-BF2*4 molecules are starved of suitable peptides, there would be a larger intracellular pool of Flag-BF2*4 molecules than in the Flag-BF2*15 transfectants.

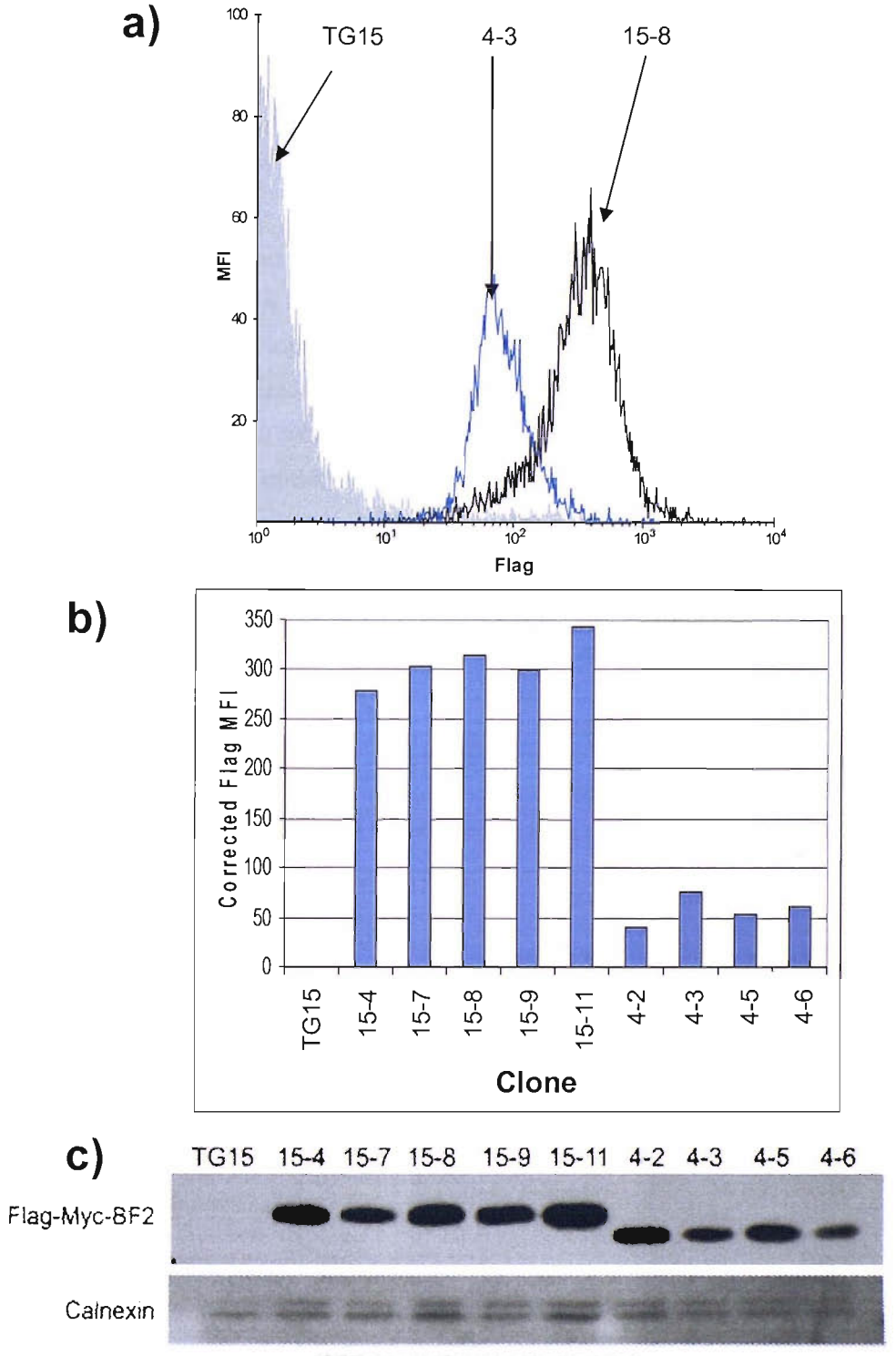


Figure 5.3 Comparison of Flag surface expression levels.
a) Representative histogram depicting Flag surface expression levels. TG15 (grey, filled), Flag clone 15-8 (black line) and Flag clone 4-3 (blue line).
b) Graph of Flag surface expression levels for every clone (corrected by subtraction of MFI following secondary antibody staining only).
c) Western blot of TG15 and Flag-BF2 clones. Each lane was loaded with a volume of lysate equivalent to 10^4 cells, separated by SDS PAGE and western blotted. The blot was subsequently incubated with 10 μg/ml Flag M2 or 1:2,000 anti-calnexin antibodies.

The five Flag-BF2*15 transfectants recorded 2.35-4.14 greater Flag staining following saponin permeabilisation (table 5.1 and figure 5.4), whilst the four Flag-BF2*4 transfectants recorded a ratio of 6.30-7.09 greater Flag staining following permeabilisation. The increased proportion of intracellular Flag-BF2*4 is likely to be a consequence of the scarcity of appropriate peptides (or other molecular incompatibility) limiting the surface expression of Flag-BF2*4 molecules.

There was an increase of between 1.08-1.44 in β_2m staining following saponin permeabilisation for seven of the transfectants; whilst for three transfectants the β_2m staining was moderately decreased (between 0.75-0.88) following saponin permeabilisation. This suggests that most β_2m molecules, which presumably represents predominantly class I heterodimers, are present at the cell surface, with only a small pool of β_2m molecules present within the cell, which may constitute class I heterodimers in the process of peptide-loading or otherwise en-route to the cell surface.

Expt 188					
Clone	**Treatment**	**F21-21 MFI**	**Flag MFI**	**Fold Increase F21-21**	**Fold Increase Flag**
TG15	Surface	530.37	ND	1.26	ND
	Permeabilised	670.74	174.66		
15-4	Surface	341.27	155.23	1.09	2.58
	Permeabilised	371.38	399.77		
15-7	Surface	1066.16	269.53	1.08	2.35
	Permeabilised	1152.13	632.64		
15-8	Surface	453.69	178.73	1.09	3.09
	Permeabilised	492.88	553.07		
15-9	Surface	425.74	145.00	1.12	4.14
	Permeabilised	476.49	600.23		
15-11	Surface	464.58	197.02	0.88	3.06
	Permeabilised	409.29	602.77		
4-2	Surface	796.89	66.91	0.84	6.69
	Permeabilised	669.61	447.87		
4-3	Surface	802.42	91.22	0.75	6.30
	Permeabilised	597.82	575.11		
4-5	Surface	535.87	83.92	1.44	6.63
	Permeabilised	772.39	556.73		
4-6	Surface	479.77	67.02	1.28	7.09
	Permeabilised	615.06	475.22		
Expt 186					
TG15	Surface	211.20	90.94	1.53	3.63
	Permeabilised	322.39	330.17		

Table 5.1 Expression levels of β_2m (F21-21) or Flag-BF2 (Flag) molecules following staining of intact or saponin permeabilised cells. Fold increase = MFI+saponin / MFI-saponin

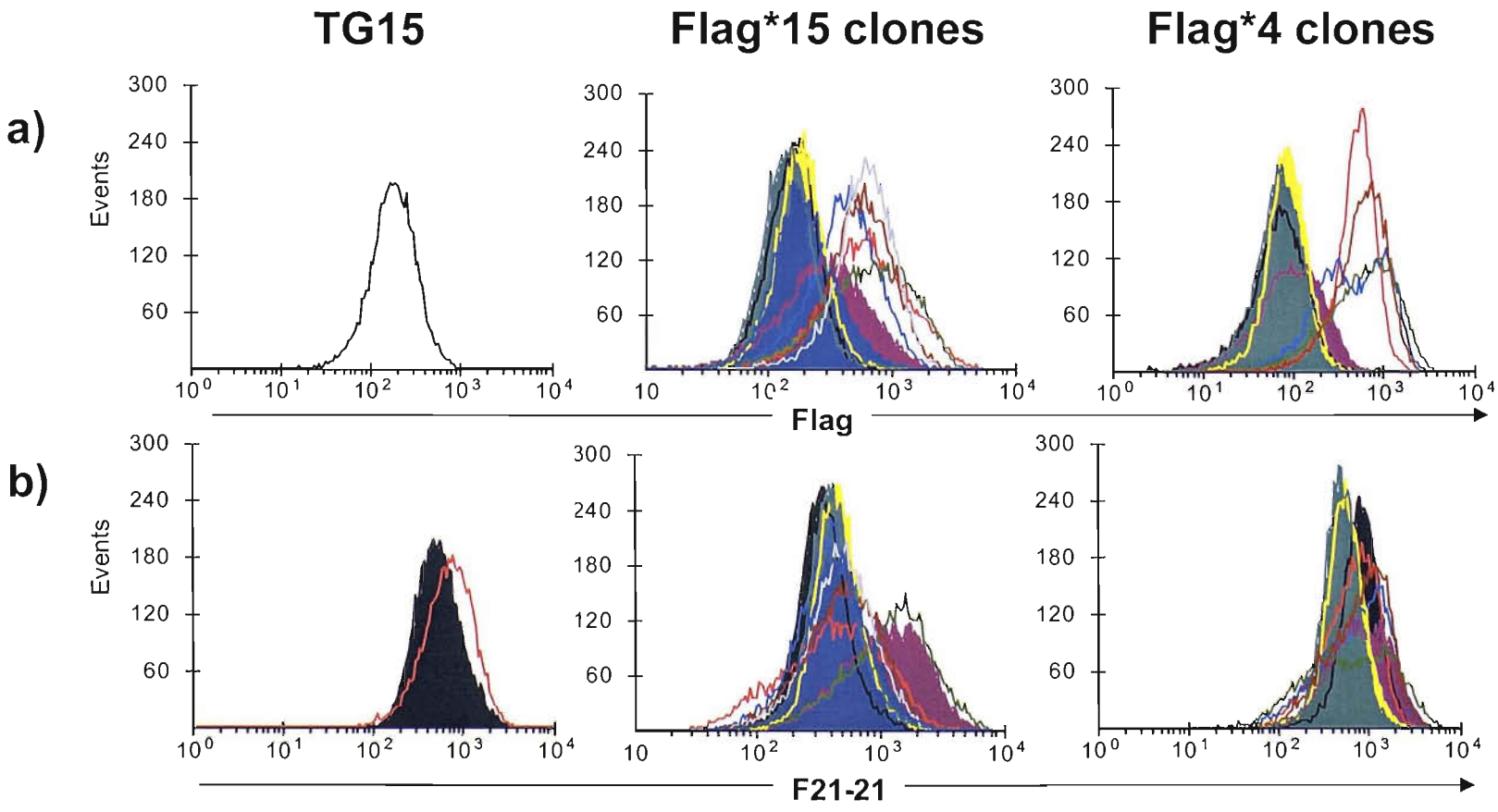


Figure 5.4 Comparison of Flag and chicken β_2m (F21-21) expression levels following surface staining or after permeabilisation with saponin.

a) Histograms depict Flag staining. Filled histograms show surface staining, open histograms show staining following permeabilisation.

b) As above, but representing F21-21 staining.

The finding that the Flag staining was increased to a greater extent than the β_2m staining following saponin permeabilisation could suggest that there is a larger intracellular pool of Flag-BF2 molecules than β_2m molecules. However, in a separate experiment (experiment 186, table 5.1) the Flag staining observed for TG15 cells, which express no Flag epitope-tagged molecules, was 3.6 times greater following saponin permeabilisation. This suggests that there is greater non-specific Flag staining than non-specific β_2m staining following saponin permeabilisation. Therefore the increase in Flag staining evident in the Flag-BF2*15 transfectants is likely to represent non-specific staining.

5.2.3 Exogenous peptide stabilisation assays increase Flag-BF2*15 surface expression levels to a greater extent than for Flag-BF2*4 at physiological temperatures

To further characterise the differences in expression of Flag-BF2 molecules in TG15 cells, exogenous peptide class I stabilisation assays were performed. The peptides chosen, KRLIGKRY for B15, and ADVEEYEE for B4, were identified as individual peptides eluted in high amounts in several peptide elution studies from B15 and B4 blood cells (table 4.1 (Wallny et al., 2006)), and would therefore be expected to be high affinity peptides. The addition of exogenous peptide may be expected to stabilise surface expressed class I molecules that are loaded with sub-optimal peptides, and may reach the ER to drive otherwise ER-retained molecules to the cell surface (Day et al., 1997; Chun et al., 2001; Lybarger et al., 2001).

In light of the likely scarcity of peptides suitable for the Flag-BF2*4 molecules that are supplied by the endogenous TAP proteins, it was expected that the surface expressed Flag-BF2*4 molecules would be more susceptible to exogenous peptide stabilisation than the Flag-BF2*15 molecules, which are more likely to be loaded with optimal peptides. Numerous assays were performed using a variety of different experimental conditions, which ultimately produced similar results. The results of three experiments are discussed.

When KRLIGKRY peptide was added to Flag-BF2*15 transfectants and cultured overnight at either 41°C or 28°C, the expression level of Flag-BF2*15 molecules was increased 31-38% at 41°C, and 1-36% at 28°C, in comparison to cells which did not receive peptide (table 5.2 and figure 5.5). In addition, the expression level of β_2m was increased by KRLIGKRY peptide addition, both for TG15 cells and Flag-BF2*15 transfectants, at either temperatures, confirming the peptide also stabilised endogenous BF2*15 class I molecules in addition to Flag-BF2*15 molecules.

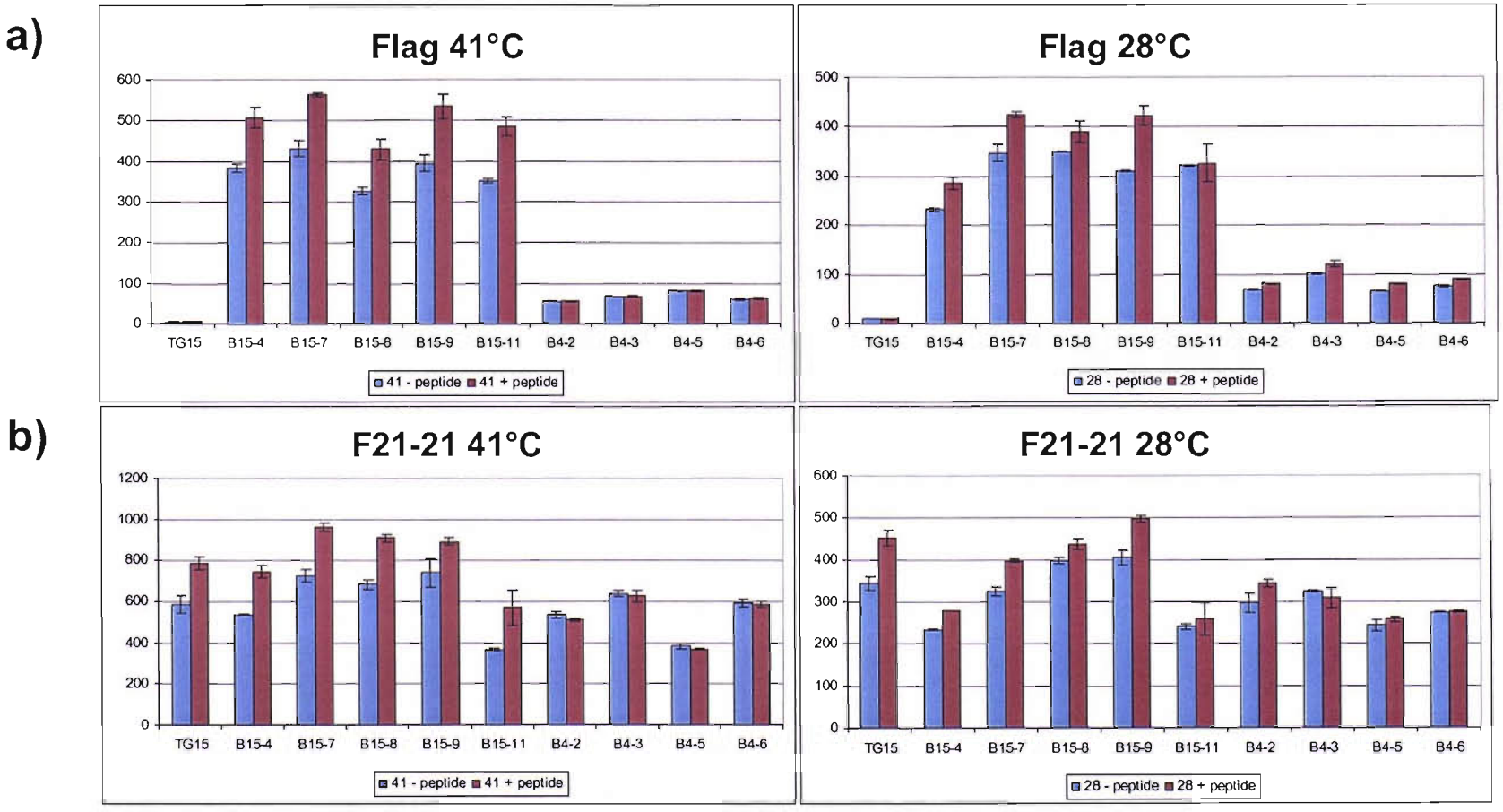


Figure 5.5 Flag and β_2m (F21-2) surface expression levels following culture with or without specific peptide at 41°C or 28°C.
a) Graphs of Flag MFI for each clone when cultured in the presence or absence of specific peptide (KRLIGKRY for B15, ADVEEYEE for B4) overnight at 41°C or 28°C. y axis represents mean MFI. Error bars depict the standard deviation of triplicate cultures.
b) As above, but data represents F21-21 staining.

Clone	Flag 41 Fold Increase KRLIGKRY	Flag 41 Fold Increase ADVEEYEE	Flag 28 Fold Increase KRLIGKRY	Flag 28 Fold Increase ADVEEYEE	F21-21 41 Fold Increase No peptide	F21-21 41 Fold Increase ADVEEYEE	F21-21 28 Fold Increase KRLIGKRY	F21-21 28 Fold Increase ADVEEYEE
TG15	1.32		1.04		1.34		1.32	
B15-4	1.32		1.23		1.40		1.20	
B15-7	1.31		1.22		1.33		1.23	
B15-8	1.31		1.12		1.33		1.10	
B15-9	1.35		1.36		1.21		1.23	
B15-11	1.38		1.01		1.56		1.08	
B4-2		1.02		1.19		0.96		1.16
B4-3		1.00		1.19		0.98		0.95
B4-5		1.03		1.25		0.96		1.06
B4-6		1.05		1.20		0.99		1.00

Table 5.2 The fold increase in Flag-BF2 (Flag) or β_2m (F21-21) cell surface expression levels following overnight culture in the presence or absence of specific peptide at 41°C or 28°C. Fold increase = (MFI+peptide)/(MFI–peptide).

In contrast, ADVEEYEE peptide addition led to a modest increase in Flag staining of up to 5% when the Flag-BF2*4 transfectants were cultured at 41°C. Peptide stabilisation assays performed over shorter intervals of time failed to stabilise a larger proportion of Flag-BF2*4 molecules, confirming that the overnight peptide incubation used in this experiment was not masking a short-term stabilisation of Flag-BF2*4 molecules. However, when ADVEEYEE peptide was added during culture overnight at 28°C, there was a larger proportion (between 19-25%) of molecules that could be stabilised, suggesting that when Flag-BF2*4 cells are cultured at temperatures which should permit the egress of sub-optimally peptide-loaded molecules to the cell surface, that these molecules can be stabilised by the exchange of their presumably low affinity peptides for the exogenous peptide. As expected from their mutually exclusive peptide-binding motifs, ADVEEYEE peptide addition did not increase the expression level of β_2m at 41°C, suggesting this peptide did not bind to endogenous BF2*15 molecules.

To confirm the apparent lack of stabilisation of Flag-BF2*4 molecules (at 41°C) by exogenous peptide addition was not related to the choice of peptide, two other peptides identified in peptide elution studies: LDVEDVKF and IDWFDGKE, were compared to ADVEEYEE and KRLIGKRY in peptide stabilisation assays performed at 41°C. All three B4 peptides performed reasonably similarly (table 5.3 and figure 5.6); stabilising a smaller proportion of Flag-BF2*4 molecules than the addition of KRLIGKRY stabilised Flag-BF2*15 molecules.

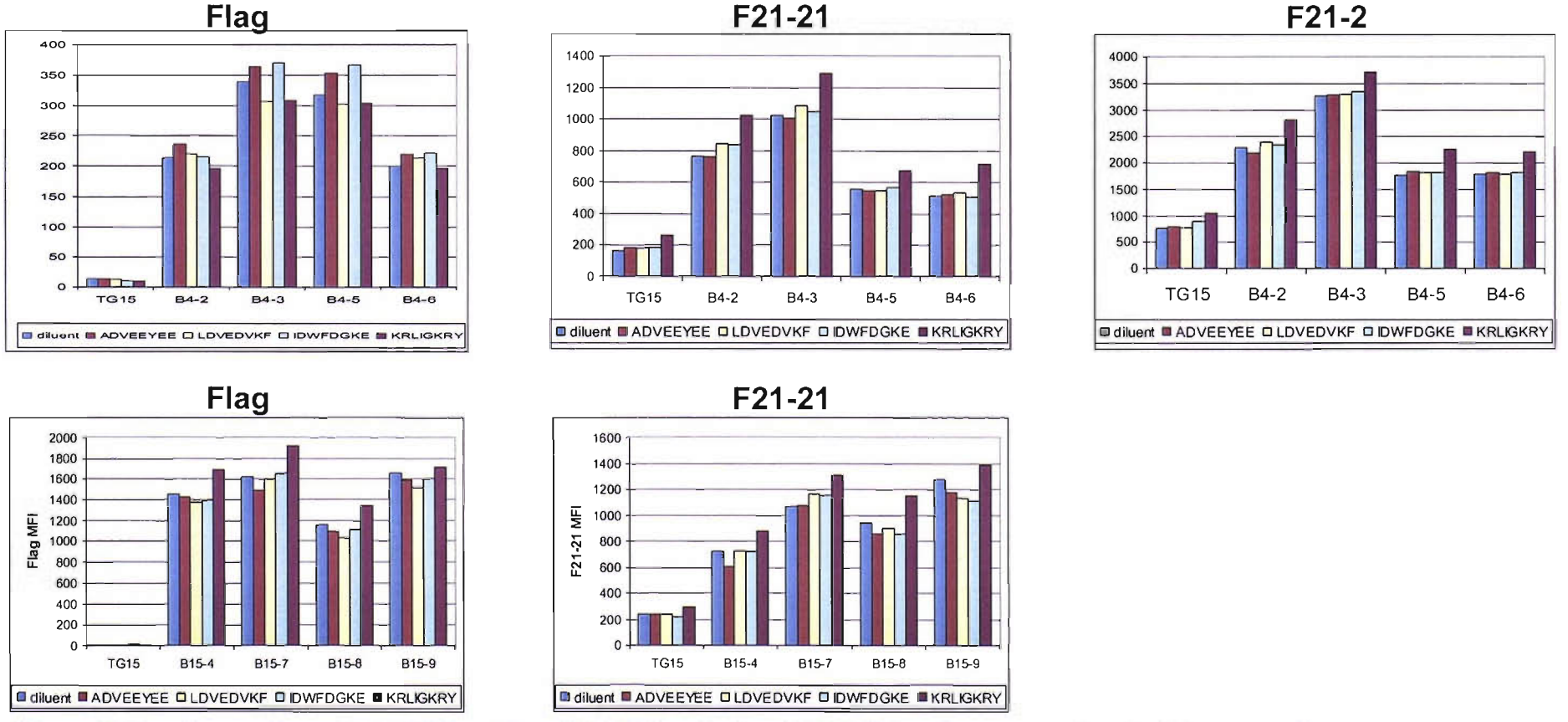


Figure 5.6 Comparison of total Flag, β_2m (F21-21), and class I (F21-2) surface expression levels upon culture with or without various peptides overnight at 41°C overnight.

Graphs depicting Flag, F21-21 and F21-2 MFI values for each Flag-BF2*4 clone (top) and for each Flag-BF2*15 clone (bottom) when cultured in the presence or absence of the indicated peptide overnight at 41°C.

Expt 219			
Clone	Peptide	F21-21 Fold Increase	Flag Fold Increase
TG15	diluent	1.00	1.00
	ADVEEYEE	0.99	0.93
	LDVEDVKF	0.99	1.01
	IDWFDGKE	0.92	1.09
	KRLIGKRY	1.24	0.95
B15-4	diluent	1.00	1.00
	ADVEEYEE	0.84	0.98
	LDVEDVKF	1.01	0.95
	IDWFDGKE	1.00	0.96
	KRLIGKRY	1.22	1.17
B15-7	diluent	1.00	1.00
	ADVEEYEE	1.01	0.92
	LDVEDVKF	1.09	0.99
	IDWFDGKE	1.08	1.02
	KRLIGKRY	1.23	1.18
B15-8	diluent	1.00	1.00
	ADVEEYEE	0.91	0.95
	LDVEDVKF	0.95	0.89
	IDWFDGKE	0.90	0.97
	KRLIGKRY	1.22	1.17
B15-9	diluent	1.00	1.00
	ADVEEYEE	0.92	0.96
	LDVEDVKF	0.89	0.92
	IDWFDGKE	0.88	0.97
	KRLIGKRY	1.09	1.03

Expt 216				
Clone	Peptide	F21-2 Fold Increase	F21-21 Fold Increase	Flag Fold Increase
TG15	diluent	1.00	1.00	1.00
	ADVEEYEE	1.04	1.13	1.02
	LDVEDVKF	1.02	1.11	0.97
	IDWFDGKE	1.17	1.12	0.74
	KRLIGKRY	1.36	1.61	0.76
Flag4-2	diluent	1.00	1.00	1.00
	ADVEEYEE	0.96	0.99	1.10
	LDVEDVKF	1.04	1.10	1.03
	IDWFDGKE	1.02	1.09	1.01
	KRLIGKRY	1.23	1.33	0.92
Flag4-3	diluent	1.00	1.00	1.00
	ADVEEYEE	1.00	0.99	1.07
	LDVEDVKF	1.01	1.06	0.90
	IDWFDGKE	1.03	1.02	1.09
	KRLIGKRY	1.14	1.26	0.91
Flag4-5	diluent	1.00	1.00	1.00
	ADVEEYEE	1.04	0.98	1.11
	LDVEDVKF	1.04	0.98	0.95
	IDWFDGKE	1.04	1.02	1.15
	KRLIGKRY	1.28	1.21	0.96
Flag4-6	diluent	1.00	1.00	1.00
	ADVEEYEE	1.03	1.02	1.10
	LDVEDVKF	1.00	1.05	1.06
	IDWFDGKE	1.02	0.99	1.11
	KRLIGKRY	1.24	1.40	0.98

Table 5.3 The fold increase in Flag-BF2 (Flag), β_2m (F21-21) or total class I (F21-2) cell surface expression levels following overnight culture in the presence or absence of various peptides at 41°C. Fold increase = (MFI+peptide)/(MFI–peptide).

The limited extent of peptide-induced stabilisation of Flag-BF2*4 molecules at 41°C suggests that the smaller number of Flag-BF2*4 molecules that reach the cell surface represent one of two extremes: either the surface expressed Flag-BF2*4 molecules are loaded with optimal repertoires of peptides that are hard to displace with exogenous peptides; or that the surface expressed Flag-BF2*4 molecules were loaded with sub-optimal peptides, that were rapidly lost from the peptide-binding grooves, leading to dissociation of Flag-BF2*4 heterodimers, leaving a small population of free Flag-BF2*4 heavy chains, unable to bind peptide in the absence of β_2m.

5.2.4 Tapasin*4 transfection does not improve Flag-BF2*4 expression

Numerous attempts were made to transfect the Flag-BF2*4-expressing transfectants with V5 epitope tagged Tapasin*4 (step 3 in figure 5.1), which would determine how the expression of the appropriate tapasin allele (in addition to the endogenous Tapasin*15 allele) affected the expression of the haplotype-mismatched Flag-BF2*4 molecules in TG15 cells. However, as experienced with the human .220 cell line, transient transfection or attempts to generate stable transfectants were all unsuccessful, apart from one stable Tapasin*4-expressing transfectant derived from the Flag-BF2*4 transfectant 4-2 (figure 5.7c).

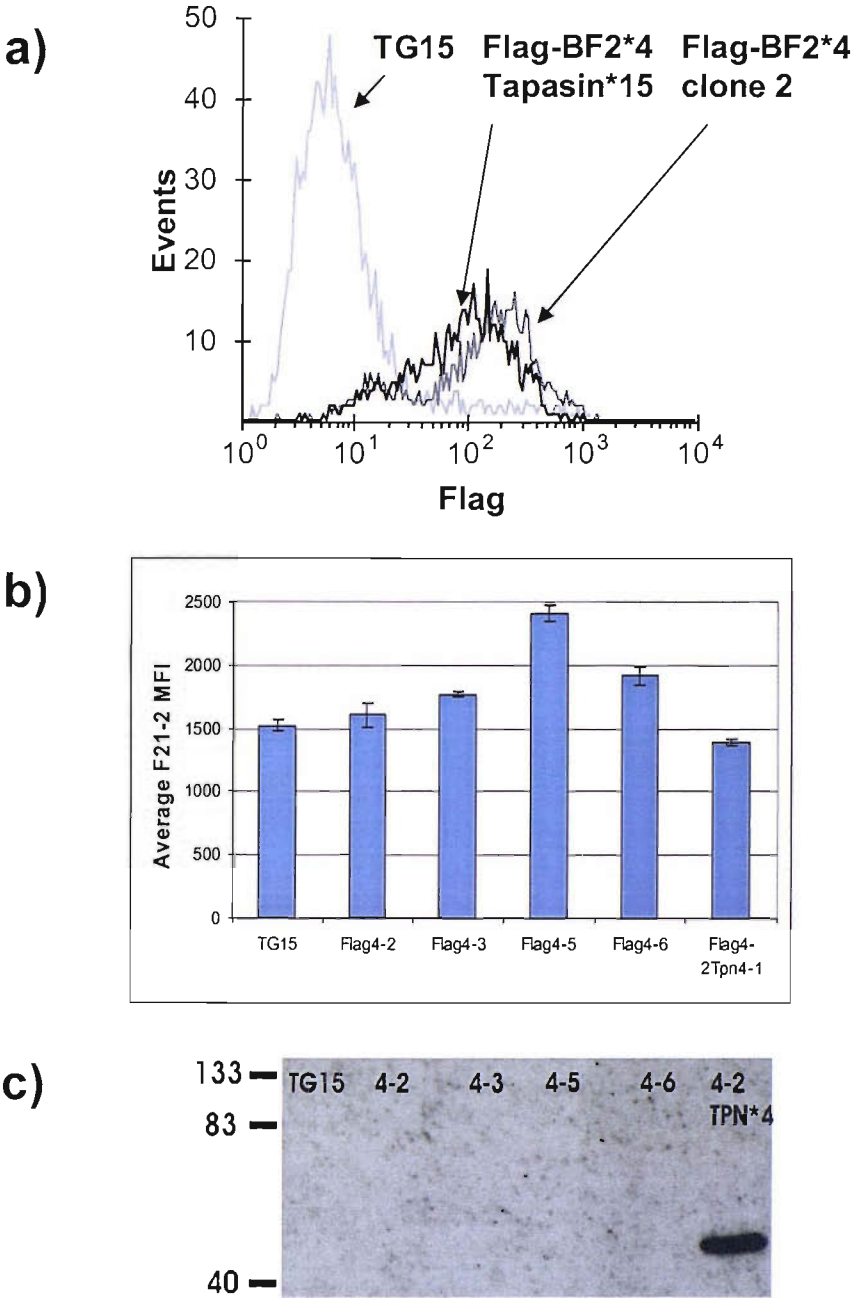


Figure 5.7 Expression of Tapasin*4 does not increase the steady state surface expression levels of Flag-BF2*4 or total class I molecules

a) Histogram representing steady state Flag surface expression on clones 4-2Tpn*4 (black line), 4-2 (grey line) or TG15 (grey filled). Note a proportion of cells in clone 4-2 have partially lost expression of Flag-BF2*4.
b) Graph depicting steady state total class I (F21-2) surface expression levels upon TG15 and Flag-BF2*4 clones. Error bars depict standard deviation of triplicate cultures.
c) Anti-V5 western blot confirming expression of Tapasin*4 protein in clone 4-2Tpn*4. Molecular weight markers shown on left hand side in kDa.

In comparison to the parental Flag-BF2*4 transfectant 4-2, Tapasin*4 expression did not increase the surface expression level of either Flag-BF2*4 or total class I molecules (table 5.4 and figures 5.7a-b). In addition, Tapasin*4 expression did not increase the proportion of peptide-receptive Flag-BF2*4 molecules as assessed by exogenous peptide stabilisation assays (table 5.4). These findings suggest that Flag-BF2*4 expression is likely to be compromised primarily by an absence of suitable peptides.

Clone	Peptide	Average Flag MFI	St dev	Fold Increase Flag
TG15	none	5.57	1.65	
	ADVEEYEE	4.89	0.01	0.88
	KRLIGKRY	4.74	0.02	0.85
Flag4-2	none	109.48	7.38	
	ADVEEYEE	105.57	6.91	0.96
	KRLIGKRY	75.58	7.42	0.69
Flag4-3	none	196.15	10.70	
	ADVEEYEE	203.17	8.55	1.04
	KRLIGKRY	131.61	5.76	0.67
Flag4-5	none	446.53	17.71	
	ADVEEYEE	441.37	41.87	0.99
	KRLIGKRY	317.18	12.07	0.71
Flag4-6	none	170.59	62.04	
	ADVEEYEE	208.74	2.48	1.22
	KRLIGKRY	140.90		0.83
Flag4-2 Tpn*4	none	84.62	0.77	
	ADVEEYEE	79.08	3.00	0.93
	KRLIGKRY	66.78	0.28	0.79

Table 5.4 Cell surface expression levels of Flag-BF2*4 molecules following culture in the absence of peptide, or presence of specific peptide (ADVEEYEE), or irrelevant peptide (KRLIGKRY). Fold increase = (MFI+peptide)/(MFI–peptide).

5.2.5 Culture with brefeldin A suggests Flag-BF2*4 molecules decay faster than Flag-BF2*15 molecules

To further characterise the expression of Flag-BF2 molecules, the five Flag-BF2*15, four Flag-BF2*4 and Flag-BF2*4 Tpn*4 transfectants were cultured for various lengths of time with brefeldin A, a drug that prevents protein traffic between the ER and Golgi. Table 5.5 shows both the Flag expression levels recorded for each transfectant at each time point, as well as a summarised version, with graphical representations of the results provided in figure 5.8a-c.

Clone	Time	#Events	Flag MFI	%Flag 0hr	#Events	TRT1 MFI
TG15	0hr	1230	7.31	100.00	229	2.60
	2hr	4244	3.60	49.25		
	4hr	6075	3.40	46.51		
	6hr	10633	3.25	44.46		
	8hr	10789	3.22	44.05		
15-4	0hr	740	463.45	100.00	351	4.76
	2hr	2190	339.21	73.19		
	4hr	7779	221.00	47.69		
	6hr	9998	189.61	40.91		
	8hr	11623	162.03	34.96		
15-7	0hr	703	522.37	100.00	279	14.96
	2hr	577	474.25	90.79		
	4hr	4646	300.91	57.60		
	6hr	9132	231.28	44.28		
	8hr	10717	181.29	34.71		
15-8	0hr	3838	1099.78	100.00	1549	7.45
	2hr	4996	1020.25	92.77		
	4hr	11687	715.06	65.02		
	6hr	11812	589.38	53.59		
	8hr	12164	508.35	46.22		
15-9	0hr	2549	586.16	100.00	786	14.94
	2hr	3276	533.66	91.04		
	4hr	9569	370.28	63.17		
	6hr	11397	300.66	51.29		
	8hr	12243	251.59	42.92		
15-11	0hr	2225	581.37	100.00	685	10.42
	2hr	2106	489.56	84.21		
	4hr	5479	367.56	63.22		
	6hr	7628	322.64	55.50		
	8hr	8555	267.08	45.94		
4-2	0hr	5067	46.42	100.00	1422	8.35
	2hr	2042	27.33	58.88		
	4hr	8903	16.43	35.39		
	6hr	10994	13.55	29.19		
	8hr	9003	9.78	21.07		
4-3	0hr	586	112.91	100.00	227	10.20
	2hr	962	38.63	34.21		
	4hr	7608	41.03	36.34		
	6hr	8692	34.24	30.33		
	8hr	10331	22.99	20.36		
4-5	0hr	1306	175.61	100.00	387	6.30
	2hr	2352	102.01	58.09		
	4hr	5990	65.06	37.05		
	6hr	8668	49.26	28.05		
	8hr	8833	41.92	23.87		
4-6	0hr	5093	122.27	100.00	1464	5.78
	2hr	8453	70.26	57.46		
	4hr	10196	47.51	38.86		
	6hr	10547	38.39	31.40		
	8hr	10161	33.25	27.19		
4-2 Tpn*4	0hr	6000	66.93	100.00	1251	6.76
	2hr	7177	38.15	57.00		
	4hr	8070	28.00	41.83		
	6hr	7902	22.63	33.81		
	8hr	7599	20.19	30.17		

Clone	Time	% Flag 0hr
TG15	0	100.00
	2	49.25
	4	46.51
	6	44.46
	8	44.05
Average Flag*15	0	100.00
	2	86.40
	4	59.34
	6	49.11
	8	40.95
Average Flag*4	0	100.00
	2	52.16
	4	36.91
	6	29.74
	8	23.12
4-2Tpn*4	0	100.00
	2	57.00
	4	41.83
	6	33.81
	8	30.17

Table 5.5 Flag-BF2 surface expression levels following culture with brefeldin A for the indicated times.

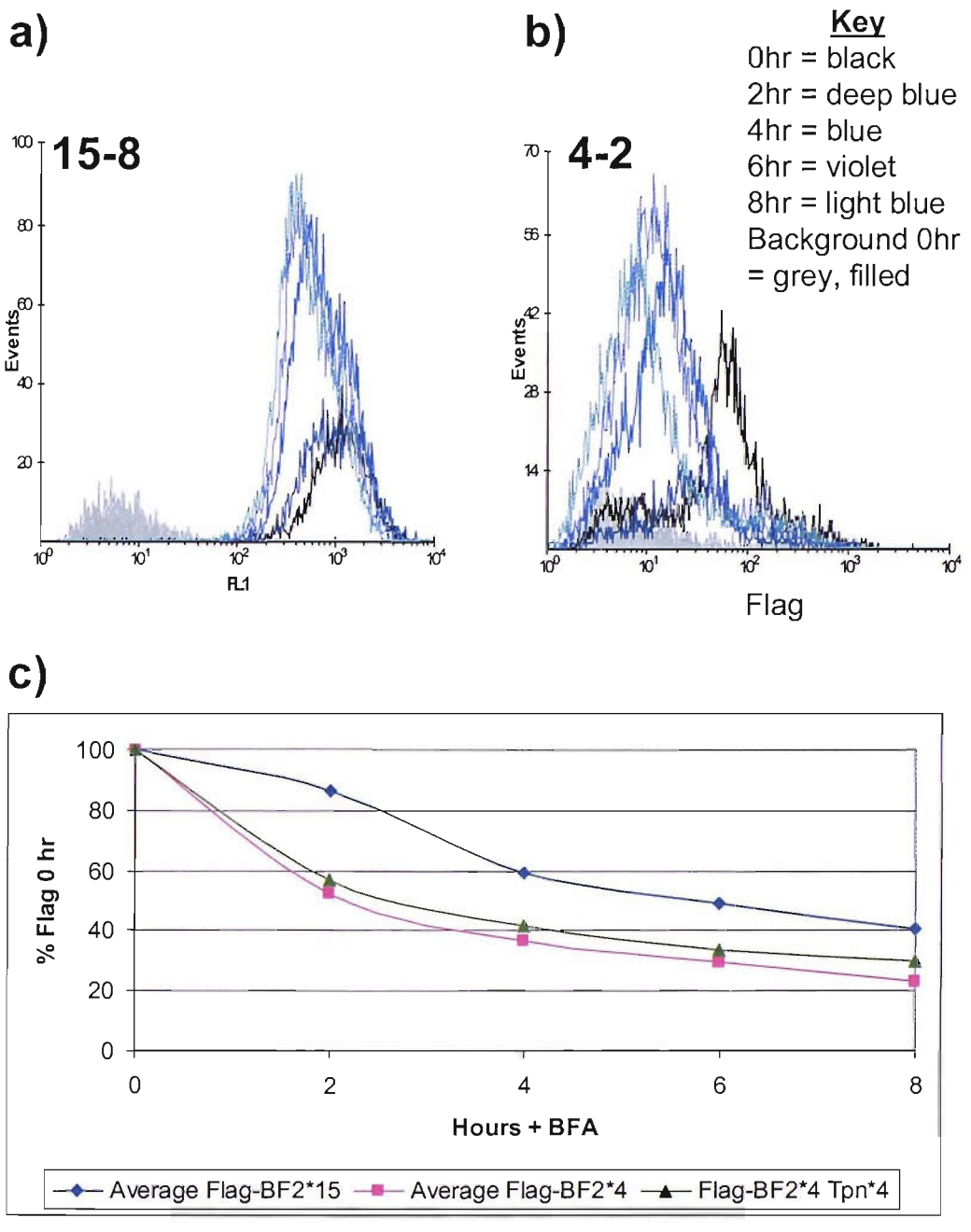


Figure 5.8 Flag staining levels following culture with brefeldin A (BFA).

a. Histogram depicting Flag expression levels for clone 15-8 following culture with BFA for 0 hr (black line), 2 hr (deep blue), 4 hr (blue), 6 hr (violet) or 8 hr (light blue). Background MFI represented by TRT1 staining of cells at 0 hr (grey, filled).

b. As for above (b) but representing clone 4-2. Note a proportion of cells in clone 4-2 have lost expression of Flag-BF2*4

c. Graph depicting the Flag expression levels relative to 0 hr for 1) an average of the five Flag-BF2*15 ("average Flag-BF2*15"), 2) an average of the four Flag-BF2*4 clones ("average Flag-BF2*4"), 3) the Flag-BF2*4 Tpn*4 clone.

The Flag expression level decayed to a lesser extent following eight hours of culture with BFA for the Flag-BF2*15 transfectants (reaching an average of 41% of the initial expression level) than the Flag-BF2*4 transfectants (reaching an average of 23%). This suggests that the Flag-BF2*4 molecules were loaded with a lower quality repertoire of peptides than were loaded into the Flag-BF2*15 molecules. Although the Tapasin*4-expressing Flag-BF2*4 transfectant exhibited higher Flag expression than the parental transfectant in this experiment, its Flag expression level was still below the other three Flag-BF2*4 clones. In addition, expression of Tapasin*4 did not slow BFA decay, reaching 30% of the initial expression level following eight hours of culture. This is in agreement with the previous findings confirming Tapasin*4 expression did not substantially improve Flag-BF2*4 expression.

5.2.6 Failure to introduce an alternative source of peptides for Flag-BF2*4 molecules

The experimental strategy that was attempted relied upon comparing the expression of Flag-BF2*4 molecules in the transfectants already produced (steps 1 and 3, figure 5.1) with additional transfectants that should have an abundant source of peptides within the ER (steps 2 and 4, figure 5.1), which should determine the full extent of the impairment caused by the endogenous TAP transport specificity.

To achieve this, two approaches were taken. Firstly, attempts were made to transfect the TAP1*4 and TAP2*4 genes into the Flag-BF2*4 transfectants. Unfortunately these transfections proved unsuccessful, either by transient transfection or by attempting to make stable transfectants. The second approach involved placing nucleotides encoding the desired peptide sequence behind the BF2 signal sequence to make peptide minigenes that would be delivered to the ER lumen via the sec61 protein translocator. The plasmid vector that was chosen for the peptide minigene construct encoded a truncated version of the human P75 nerve growth receptor (referred to as ΔNGFR) that is expressed at the cell surface of the transfected cell. The ΔNGFR was linked by an internal ribosome entry site to the peptide minigene, and therefore the surface ΔNGFR expression level is likely to be an accurate reflection of the peptide minigene expression level. Two expression constructs were made encoding either the KRLIGKRY or ADVEEYEE peptides.

The Flag-BF2*4-expressing transfectants were transiently transfected with the peptide minigene expression constructs on numerous occasions (two examples are shown in figure 5.9). In comparison to mock transfected cells, the peptide minigene transfected cells exhibited a slightly greater, broader distribution of cells following staining with the anti-P75 antibody (figure 5.9b and table 5.6), suggesting that few cells expressed large amounts of the peptide constructs (the transfection efficiency did not ever exceed 2.5%).

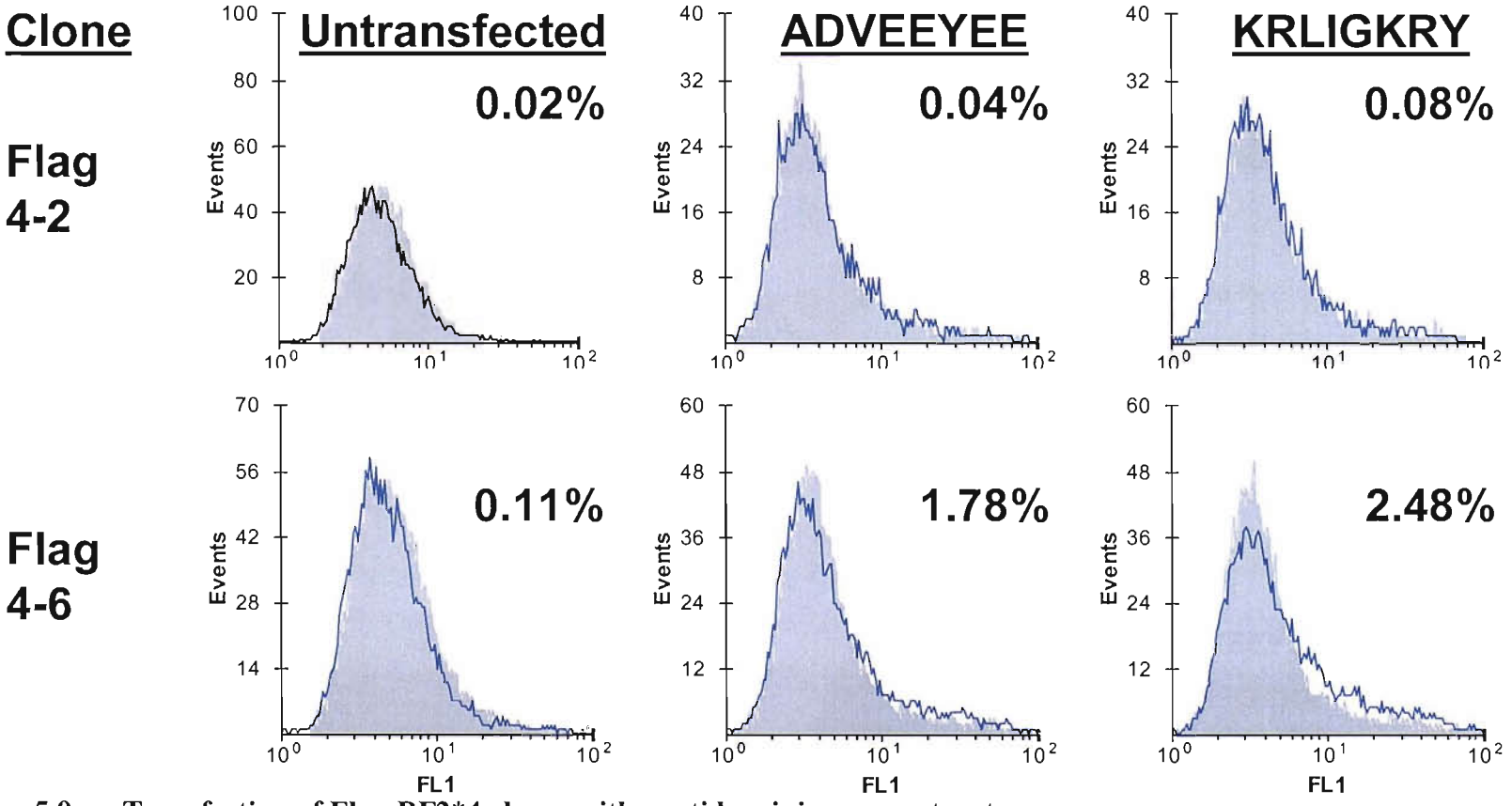


Figure 5.9 Transfection of Flag-BF2*4 clones with peptide minigene constructs

a) The transfectants were transiently transfected with a peptide minigene expression construct (comprising the BF2 signal sequence preceding the indicated peptide) or left untransfected. Two days later the cells were stained with either anti-P75 antibody (blue line) or an isotype control (TRT1) antibody (grey, filled). The transfection efficiency is indicated as the percentage of cells that were classified as live (by exclusion of propidium iodide), and gave fluorescence greater than background in the FL1 channel.

Continued

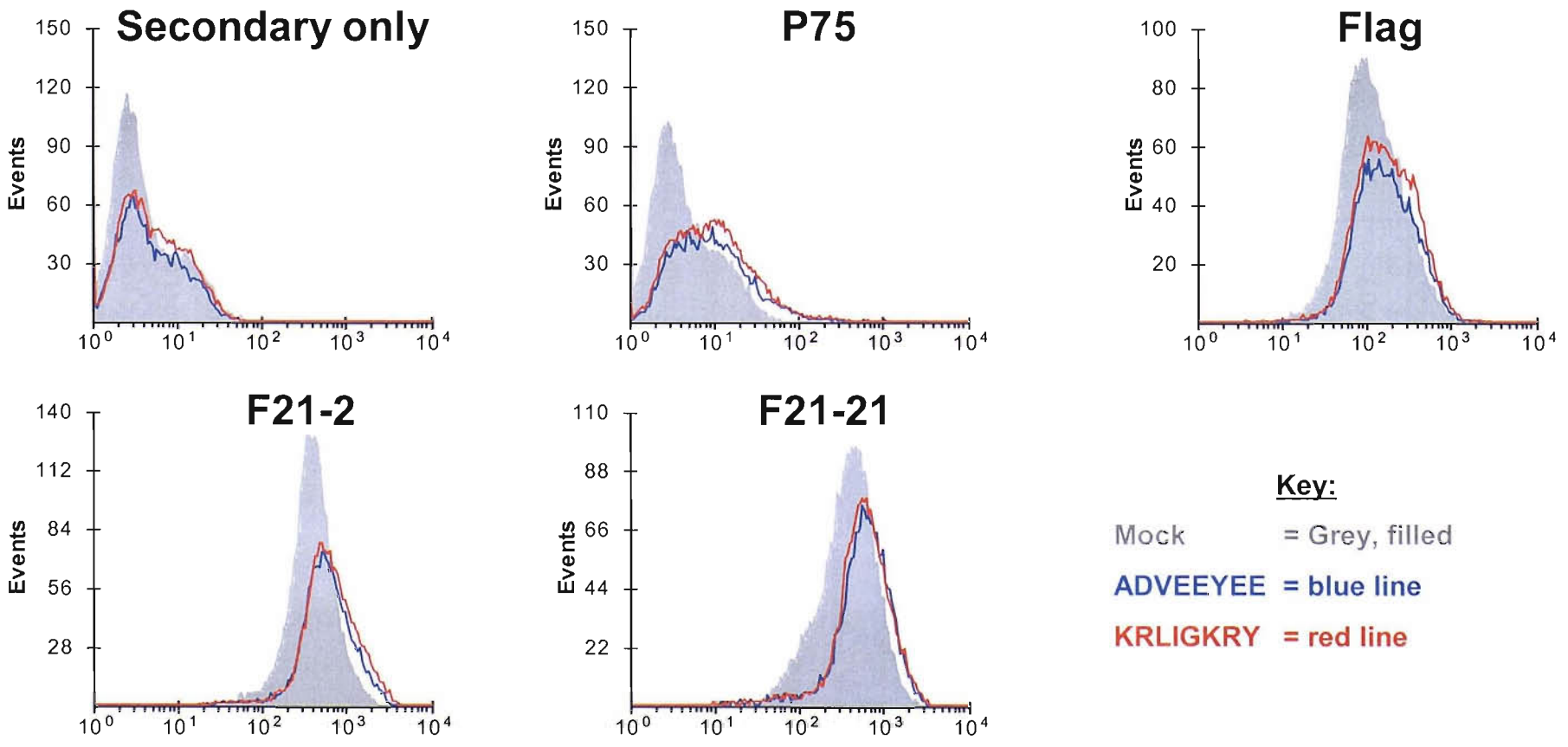


Figure 5.9 Transfection of Flag-BF2*4 clone 6 with peptide minigene constructs

b) Flag-BF2*4 clone 4-6 was transfected with either ADVEEYEE or KRLIGKRY peptide minigene constructs, or was transfected with no DNA (mock). Three days later the transfectants were compared by flow cytometry following staining with the indicated antibodies.

Flag-BF2*4 clone 6 Transfected with	mAb	MFI
Mock	2ary only	4.14
ADVEEYEE		4.78
KRLIGKRY		5.25
Mock	P75	4.43
ADVEEYEE		9.31
KRLIGKRY		9.42
Mock	Flag	116.61
ADVEEYEE		166.48
KRLIGKRY		170.57
Mock	F21-2	371.86
ADVEEYEE		550.7
KRLIGKRY		602.71
Mock	F21-21	347.43
ADVEEYEE		558.69
KRLIGKRY		554.71

Table 5.6 Expression levels of ΔNGFR (P75), Flag-BF2*4 (Flag), total class I (F21-2) and β_2m (F21-21) at the cell surface of Flag-BF2*4 clone 6, three days after transient transfection with the indicated construct

In addition to anti-P75 staining, the transfectants were stained with anti-Flag, anti-total class I and anti-β_2m antibodies. Following staining with any of the three antibodies, the peptide minigene transfectants exhibited slightly greater expression levels in comparison to the mock transfected cells. However, there was no difference between the transfectants expressing either peptide minigene. This suggests that the presence of transfected DNA led to an alteration in surface expression levels that was not apparent in mock transfected cells, but that the potential presence of either peptide in the ER did not affect Flag-BF2 expression levels.

Attempts that were made to isolate the transfected cells either by cell-sorting the peptide minigene: ΔNGFR-expressing cells, or by transferring the peptide minigene constructs into an expression vector conferring selectable antibiotic resistance were also unsuccessful, a likely consequence of the low transfection efficiency. The failure to supply an adequate source of peptides suitable for Flag-BF2*4 molecules severely compromised the success of the experimental strategy; as it was not possible to determine the consequence of a potential incompatibility with the endogenous Tapasin*15 molecules in the absence of suitable peptides.

5.2.7 Comparison of maturation rates of Flag-BF2 molecules suggests sub-optimal peptide-loading is occurring for both BF2 alleles

Whilst the experiments described above were being conducted, the Flag-BF2*15 and Flag-BF2*4 transfectants were further compared by assessing the rates of egress of newly synthesized Flag-BF2 molecules, which may reveal differences in the speed or proportions of Flag-BF2 molecules that are permitted to leave the ER to populate the cell surface. To achieve this one Flag-BF2*15 transfectant and one Flag-BF2*4 transfectant were ^{35}S pulse/chased before the Flag-BF2 molecules

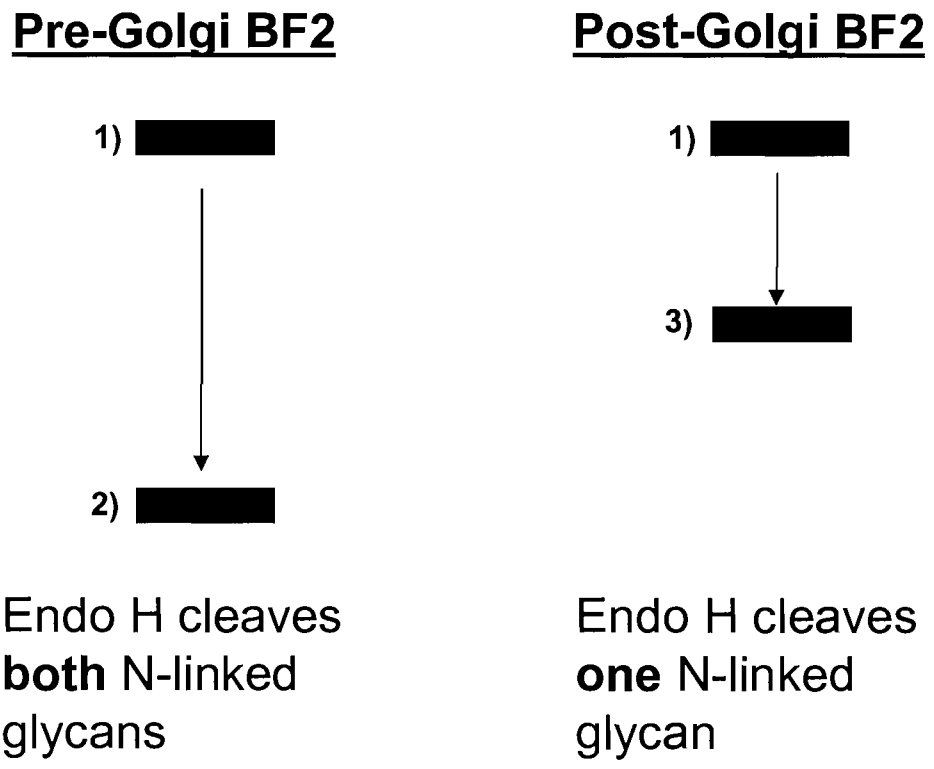
were immunoprecipitated. In addition, the parental TG15 cell line was also pulse/chased before total class I molecules were immunoprecipitated in order to compare the maturation rates of the endogenous BF2*15 molecules to the Flag-BF2 molecules.

The rate of maturation of BF2 molecules was assessed by digestion with endoglycosidase H (endo H) enzyme. Newly made class I molecules are initially glycosylated within the ER with a preformed oligosaccharide ($GlcNAc_2$, Man_9, Glc_3). Subsequently, starting within the ER and progressing in the Golgi glucosidases and mannosidases progressively trim the oligosaccharide. In the medial Golgi the high mannose oligosaccharide (comprising at least the core structure $GlcNAc_2$, Man_5) can be further trimmed and modified to become a complex oligosaccharide (which has a final core of three mannoses) at which point the oligosaccharide becomes resistant to digestion by endo H. Therefore the rate at which class I molecules gain resistance to endo H digestion is a marker for progression into and beyond the medial Golgi.

Chicken class I molecules are N-glycosylated at two sites (N37 and N85), and mature class I molecules (that have reached the cell surface) bear one high mannose and one complex N-linked glycan (Moller et al., 1991). Therefore mature class I molecules are digested to some extent by endo H, whereas immature ER-resident class I molecules are digested to a greater extent by endo H (figure 5.10). Given the differences in Flag-BF2 surface expression levels and cellular distribution, it may be expected that the rate of maturation of Flag-BF2*15 molecules is faster than Flag-BF2*4, or that a larger proportion of Flag-BF2*15 molecules mature than Flag-BF2*4.

The vast majority of the endogenous class I molecules in the TG15 cell line gained resistance to endo H after two hours of chase (figure 5.11a). This suggests that within the first two hours of the chase all of the newly made radiolabelled class I heavy chains progressed to and beyond the medial Golgi. All of the class I heavy chain bands were observed to be slightly blurry, which may reflect the heterogeneous nature of class I molecules caused by post-transcriptional modifications including alternative splicing of exon seven, or post-translational modifications including glycosylation and phosphorylation.

In contrast, the vast majority of the radiolabelled Flag-BF2*4 molecules remained sensitive to endo H digestion throughout the chase (Figure 5.11b), with only a minute proportion gaining resistance to endo H digestion after 22 hours of chase. This agrees with the decreased surface expression level and increased intracellular expression levels observed previously, and fits with the expectation that Flag-BF2*4 molecules are severely compromised in their ability to be expressed upon the cell surface, which is likely to be primarily a consequence of an absence of appropriate peptides.



BF2 molecular weight:

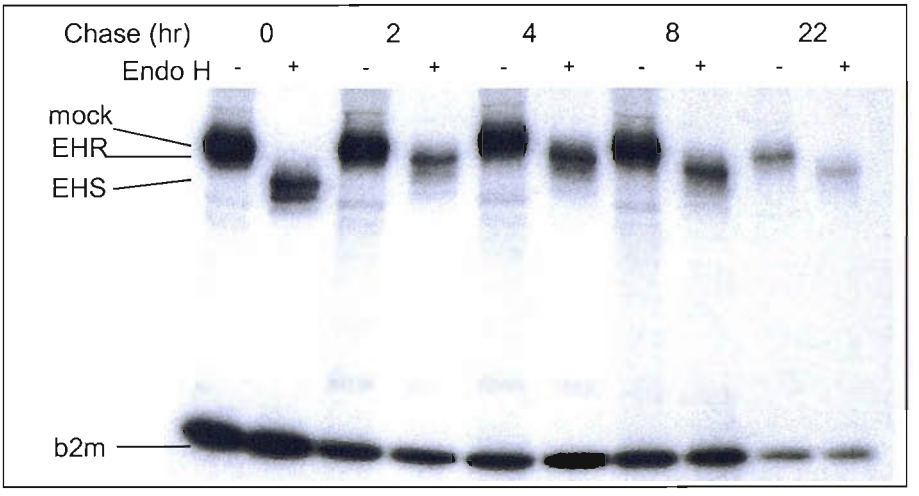
1) Mock-digested with Endo H

2) Immature, pre-medial Golgi BF2
   Both glycans: high mannose.

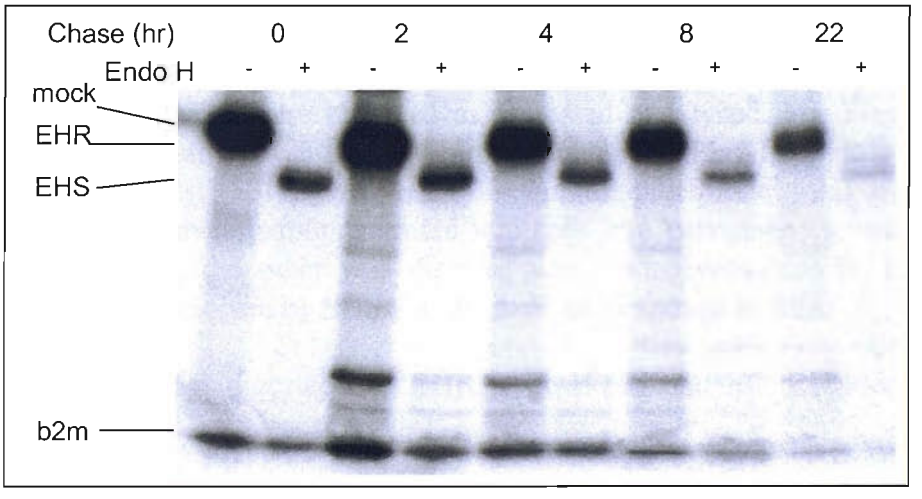
3) Mature, post-medial Golgi BF2
   One glycan: high mannose
   One glycan: complex

Figure 5.10 Cartoon depicting the changes in molecular weight when BF2 molecules are digested with endoglycosidase H (endo H)

a) TG15 total class I (F21-2) IP



b) TG15 Flag-BF2*4 myc IP (clone 3)



c) TG15 Flag-BF2*15 myc IP (clone 8)

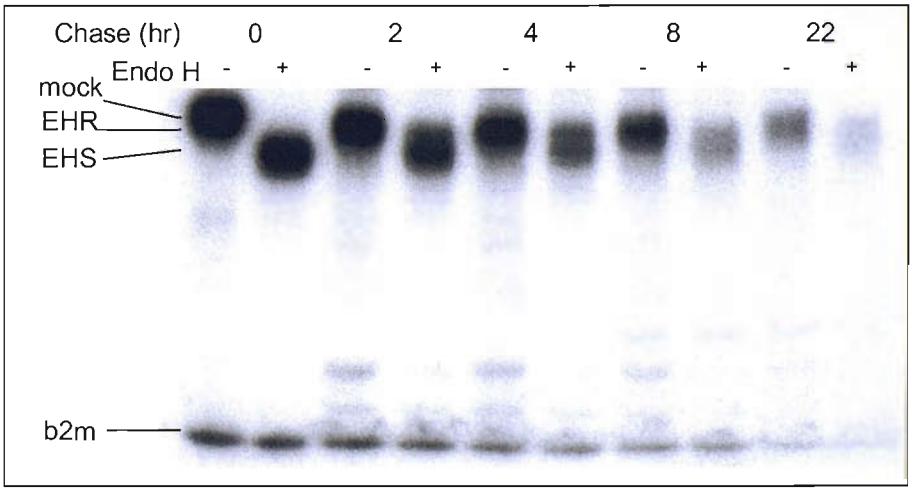


Figure 5.11 Comparison of endogenous BF2*15, Flag-BF2*4 and Flag-BF2*15 maturation rates in TG15 cells (Experiment 197)

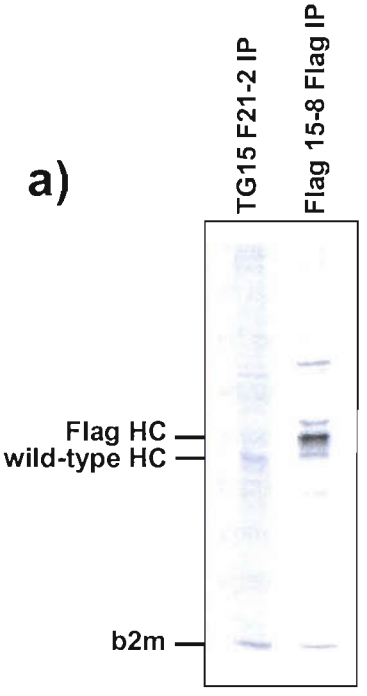
a-c) The indicated clones were pulse labelled for 15 minutes and chased for 22 hours. Aliquots were removed as indicated and lysed. For Flag-BF2 clones: myc agarose was used to immunoprecipitate Flag-BF2 molecules; for TG15 cells: F21-2 was used to immunoprecipitate total class I molecules. Each immunoprecipitate was split into two samples: one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and exposed to film.

Key:Mock: mock digested HC, **EHR**: endo H resistant HC, **EHS**: endo H sensitive HC

When the Flag-BF2*15 transfectant was pulse/chased (figure 5.11c), the Flag-BF2*15 bands resolved at slightly different molecular weights causing blurriness which hampered accurate analysis. However, there was clearly a band comprising endo H resistant Flag-BF2*15 molecules that became evident after two hours of the chase. Unlike the endogenous class I molecules immunoprecipitated from TG15 cells, there was persistence of approximately half the immunoprecipitated Flag-BF2*15 molecules as endo H sensitive molecules throughout the chase. This suggests that whilst some Flag-BF2*15 molecules became loaded with peptides and progressed to the cell surface within the same timeframe as the endogenous class I molecules, an equal sized proportion of Flag-BF2*15 molecules did not reach the medial Golgi, and most likely were retained within the ER. As the Flag-BF2*15 molecules are derived from the same MHC haplotype as the TG15 cells, it was anticipated that the Flag-BF2*15 molecules would be able to interact efficiently with the endogenous chaperones and other molecules, and should be loaded with an optimal repertoire of peptides permitting all of the Flag-BF2*15 molecules to progress to the cell surface. However, the maturation of only approximately half the Flag-BF2*15 molecules suggests an impairment in peptide-loading or cellular trafficking occurred.

The pulse/chase experiments also suggest that fewer Flag-BF2 molecules associate with β_2m than endogenous heavy chain molecules associate with β_2m. This was confirmed in a separate experiment (figure 5.12), where similar amounts of radiolabelled β_2m were co-precipitated from a Flag-BF2*15 transfectant and the parental TG15 cell line, despite nearly nine times more Flag-BF2*15 molecules being immunoprecipitated than endogenous class I molecules. This suggests that the N-terminal Flag epitope tag may compromise the ability of the molecule to associate with β_2m.

a)



b)

PhosphorImager		
Cell, IP	**HC median**	**b2m Median**
TG15 F21-2 IP	336.28	458.85
15-8 Flag IP	2945.87	475.24

Figure 5.12 Comparison of the amount of β2m co-precipitated when endogenous class I molecules are immunoprecipitated in TG15 cells to the amount co-precipitated in Flag-BF2*15 clone 15-8

a) TG15 and Flag-BF2*15 clone 15-8 were pulse labelled for 30 min. The cells were then lysed. For TG15 cells: total class I was immunoprecipitated by F21-2 antibody; for Flag clone 15-8: myc-tagged BF2 molecules were immunoprecipitated.
b) After exposure to phosphor screens, the median band intensity was determined for heavy chain and β_2m.

5.3 Discussion

The strategy of the research presented in this chapter involved the transfection of a B15 homozygous chicken cell line with combinations of genes involved in class I peptide-loading. The starting point for this strategy was the comparison of the expression of two epitope-tagged class I alleles. One of these class I alleles was derived from the same haplotype as that of the transfected cell line (Flag-BF2*15), and so should serve as an example of a class I molecule that is supplied with suitable peptides and can interact with the endogenous chaperones in an optimal fashion. In contrast, the second class I allele (Flag-BF2*4) has a peptide-binding motif that is disparate from the endogenous B15 haplotype, and is likely to be severely compromised by the restrictive TAP transport specificity of the B15 TAP proteins. However, the expression of Flag-BF2*4 molecules may be further compromised by an incompatibility with the endogenous Tapasin*15 molecule. Therefore, for this strategy to succeed, it was necessary to reconstitute the Flag-BF2*4 transfectants with a source of appropriate peptides, Tapasin*4, or both. Unfortunately, it proved impossible to provide a source of peptides suitable for Flag-BF2*4 molecules, which prevented the extent of the potential tapasin incompatibility from being determined.

5.3.1 Impairments in the expression of Flag-BF2*4

Numerous lines of evidence suggest that the Flag-BF2*4 molecules were compromised in their ability to be expressed at the cell surface of TG15 cells, which included the decreased level at which they were expressed at the cell surface and the increased amount of intracellular Flag-BF2*4 molecules in comparison to the Flag-BF2*15 molecules. Pulse/chase experiments also demonstrated that only a small proportion of Flag-BF2*4 molecules progress beyond the medial Golgi.

Two sources of evidence suggest that the small amount of surface-expressed Flag-BF2*4 molecules are not likely to be class I heterodimers optimally loaded with peptides, and instead most probably represent predominantly free heavy chains that may have initially escaped ER retention and eventual degradation, perhaps by being loaded with sub-optimal peptides that were quickly lost, leading to dissociation of the heavy chain from β_2m. Firstly, the extent to which Flag-BF2*4 molecules were lost from the cell surface when cultured with brefeldin A was greater than for Flag-BF2*15 molecules, which is likely to reflect the lower quality and sub-optimal nature of the peptide repertoire of Flag-BF2*4 molecules. Secondly, the addition of specific exogenous peptide stabilised a much smaller proportion of Flag-BF2*4 molecules in comparison to Flag-BF2*15 molecules. The greater stabilisation of Flag-BF2*15 molecules may be a consequence of the greater number of surface expressed Flag-BF2 molecules; similar to the greater number of class I molecules that can be stabilised by exogenous peptides at the surface of wild-type RMA cells in comparison to TAP-deficient RMA-S cells (Day et al., 1995).

5.3.2 Transfection of Tapasin*4 did not rectify the impairments in Flag-BF2*4 expression

The failure for the majority of Flag-BF2*4 molecules to reach the cell surface is most likely a reflection of a substantial impairment in the supply of peptides suitable for Flag-BF2*4 molecules. However, it was possible that the expression of Tapasin*4 may have provided ER retention or peptide-loading facilities that were not provided by the endogenous Tapasin*15 molecules. For example Tapasin*4 may have permitted the loading of suitable peptides that reached the ER through TAP-independent routes (for example peptides derived from leader sequences, or from internalised proteins that are returned to the ER).

However, in the single Tapasin*4-expressing clone, Flag-BF2*4 surface expression levels were unaltered in comparison to non-Tapasin*4 transfectants, which suggests the impairment in Flag-BF2*4 surface expression was not primarily caused by a failure of the endogenous Tapasin*15 molecules in loading suitable peptides into Flag-BF2*4 molecules.

Certain class Ib molecules, such as H-2M3, bear some similarities to the expression characteristics of Flag-BF2*4 molecules in TG15 cells: under normal conditions few H-2M3 molecules are expressed at the cell surface, with the majority being retained intracellularly (Chun et al., 2001; Lybarger et al., 2001). However, the provision of exogenous high affinity peptides permits surface expression of H-2M3 molecules. Whilst tapasin does not mediate the retention of H-2M3 molecules, tapasin facilitates exogenous peptide-loading, which may occur through the endocytosis and subsequent transfer to the ER of exogenous peptide. However, Tapasin*4 expression did not increase the poor ability for Flag-BF2*4 molecules to be stabilised by exogenous peptide. The expression of Tapasin*4 also did not preserve Flag-BF2*4 molecules from degradation when the cells were cultured with brefeldin A. Cumulatively, these findings suggest that Flag-BF2*4 expression was not limited solely by the absence of the correct tapasin allele.

However, it is possible that the transfected Tapasin*4 molecules were not able to function as efficiently as the endogenous Tapasin*15 molecules. For example, the Tapasin*4 molecules may be unable to associate with the endogenous TAP proteins as a result of sequence polymorphisms that prohibit such an association. However, this appears unlikely as there are no sequence polymorphisms between Tapasin*4 and Tapasin*15 alleles within the transmembrane and cytoplasmic domains that are most likely to associate with the TAP proteins. Alternatively, the C-terminal V5 epitope tag incorporated into the Tapasin*4 molecule may have decreased the ability for tapasin to interact with the endogenous TAP proteins. It is also possible that too few Tapasin*4 molecules were expressed to competitively displace the endogenous tapasin molecules from peptide-loading complexes.

5.3.3 Absence of suitable peptides is most likely to account for impairments in Flag-BF2*4 expression

Given the above results it appears that the expression of Flag-BF2*4 molecules was significantly limited by the absence of appropriate peptides. However, in addition to a limited supply of suitable peptides, it remains possible that an impaired ability to load peptides may have existed, which would not become apparent until suitable peptides were available for Flag-BF2*4 molecules.

Attempts to provide sources of appropriate peptides by transfection of TAP1*4 and TAP2*4, or by peptide minigenes into the Flag-BF2*4 cells were severely compromised by low transfection efficiencies. Without appropriate peptides, most Flag-BF2*4 molecules would be expected to either be rapidly degraded, or retained in, or returned back to the ER. The pulse/chase experiment suggests most Flag-BF2*4 molecules remained sensitive to endo H digestion and persisted for at least 22 hours of chase. Furthermore the rate of their degradation appeared no different to that of ER-resident Flag-BF2*15 molecules. This suggests that one or more chaperone molecule was responsible for preventing their rapid degradation. It is possible that Flag-BF2*4 molecules are retained in or returned back to the ER independently of the PLC, or they may be retained in the ER in association with the PLC. However the molecule responsible for such retention/retrieval was not determined in this study.

The molecules that may be involved in the ER-retention of Flag-BF2*4 molecules independently of the PLC may be the chaperones calnexin and/or calreticulin. With calreticulin being a soluble protein, it is possible that Flag-BF2*4 molecules loaded with sub-optimal peptides can be returned to the ER from the Golgi by progressive rounds of retrieval in the characterised quality control cycle. As chicken class I molecules bear two N-linked glycans, it is also possible that both chaperones may bind to the same molecule simultaneously to prevent egress to the cell surface.

Alternatively, it is possible that Flag-BF2*4 molecules were ER-retained in association with the PLC. This would suggest that the endogenous Tapasin*15 molecules associate with the haplotype-mismatched Flag-BF2*4 molecules. However, this potential association does not imply whether the endogenous Tapasin*15 molecules can facilitate the loading of Flag-BF2*4 molecules with suitable peptides if they were present.

A scenario with some similarities to the research presented in this chapter has been described previously using transfectants of rat cells that recreate the *cim* effect. Similar to the low surface expression level of Flag-BF2*4 molecules in TG15 cells, when class I alleles such as RT1.A^a that bind peptides with C-terminal basic residues are transfected into cell lines with a restrictive TAP2B allele, the RT1.A^a molecules are not expressed at a high level at the cell surface (Knittler et al., 1998; Ford et al., 2004). However, perhaps because of the limited polymorphism of the rat tapasin molecule, rat class I molecules can associate with the PLC irrespective of the transport specificity

of the TAP2 allele. In the rat transfectants described above, the peptide-starved RT1.A^a molecules accumulate in the ER, and competitively displace endogenous RT1.A^u molecules from the PLC. Despite the absence of, or at best transient association with the PLC, RT1.A^u molecules do nevertheless gain suitable peptide cargoes and are expressed efficiently on the cell surface. It would therefore be of interest to determine in future experiments whether the peptide-starved Flag-BF2*4 molecules can associate with the haplotype-mismatched PLC in TG15 cells, and whether this leads to the complete or partial exclusion of the endogenous BF2*15 molecules from the PLC. However, such research is limited at the moment by the availability of antibodies specific for components of the chicken PLC.

5.3.4 The N-terminal Flag tag appears to disturb class I expression

A problem with the class I expression constructs used in this chapter became apparent in the pulse/chase experiment. Whilst all of the endogenous class I molecules immunoprecipitated from the TG15 cell line gained resistance to endo H digestion within two hours of the chase; even after 22 hours of the chase only approximately half of the immunoprecipitated Flag-BF2*15 molecules gained resistance to endo H digestion, with an equal sized proportion of the Flag-BF2*15 molecules remaining susceptible to endo H digestion. This suggests that relative to the wild-type class I molecules, the Flag-BF2*15 molecules were compromised in their ability to be expressed at the cell surface.

The only differences between Flag-BF2*15 molecules and the endogenous BF2*15 molecules that may explain such a phenomenon are the N–terminal Flag and C-terminal myc epitope tags. As various mammalian class I molecules have been tagged at their C-terminus by epitope tags or large moieties such as GFP without affecting their ability to be expressed comparably to wild-type molecules, the cause of the impairment is likely to be the N-terminal Flag tag. Clearly the N-terminal Flag tag does not completely prohibit surface expression of Flag-BF2 molecules or inhibit their ability to present antigen, which was apparent both from this study and others (Fulton et al., 1995). However, it should be noted that CTL assays generally do not require many specific pMHC complexes in order to trigger target cell lysis.

Whilst the Flag-BF2 molecules did associate with β_2m, they associated with substantially fewer radiolabelled β_2m molecules in comparison to the endogenous class I molecules immunoprecipitated from the TG15 cell line. This suggests the placement of the Flag epitope tag at the N-terminus of the α1 domain may have led to an impairment occurring either in the early folding stages leading to less heavy chain in a conformation suitable for binding to β_2m; or that the Flag-BF2 molecules had a decreased ability to remain in association with β_2m. However, the Flag epitope tag may also have impaired an association with the PLC, which may lead to decreased ability to load peptides; or it may have impaired an association with molecules that may be responsible for protein transport within the cell.

5.3.5 Heavy chain competition may also contribute to the impaired expression of Flag-BF2 molecules

It is possible that the Flag-BF2 molecules are also compromised by the presence of the endogenous class I molecules. Competition between mouse class I alleles has been described in F1 heterozygous mice, transgenic mice and in transfected cell lines, and can lead to functional differences in antiviral responses (Tourdot and Gould, 2002; Tourdot et al., 2005). The mechanism responsible for such deficiencies is likely to represent competition between heavy chain proteins at an early stage of folding prior to their association with β_2m (as out-competed class I heavy chains do not associate with normal amounts of β_2m, and heavy chain-β_2m single chain constructs can still dominate other class I molecules, and competition can occur in TAP-deficient cell lines). Therefore it is possible that competition between the endogenous class I molecules and Flag-BF2 molecules may contribute to the impairments in the expression of Flag-BF2 molecules that are described above, and may explain why fewer Flag-BF2 molecules associated with β_2m than endogenous class I molecules. Expression of the Flag-BF2 molecules in a class I-deficient chicken cell line would be the ideal host cell to determine the contribution heavy chain competition has on the deficiencies described for Flag-BF2 expression.

5.3.6 Summary

The experimental strategy adopted in this chapter did not achieve the goal of characterising the effect of a potential tapasin incompatibility on the expression of Flag-BF2*4 molecules. Clearly, Flag-BF2*4 molecules are compromised in their ability to be expressed in TG15 cells. However, whether this is simply due to an unsuitable peptide supply, or whether there is an additional tapasin incompatibility (or incompatibility with other MHC-encoded chaperone that is yet to be described) was not determined. In addition, the contribution that heavy chain competition may have had on the impairment was not apparent.

Despite the problems that were encountered pursuing this experimental strategy, two important findings were made that bear significance for the research strategy that is described in the next chapter. Firstly, it became apparent during the course of these experiments that for such MHC-mismatching experiments to succeed, class I alleles must be compared that are both likely to be supplied with suitable peptides by the endogenous TAP proteins, limiting the amount of reconstitution that is needed. Secondly, it became clear that in contrast to the reports in the literature, N-terminal Flag epitope tagged class I molecules are not completely interchangeable substitutes for wild-type class I molecules.

6. Assessing the role of tapasin by recreating the origin and evolution of a natural recombinant MHC haplotype

6.1 Introduction

6.1.1. The origin and evolution of the B19v1 haplotype

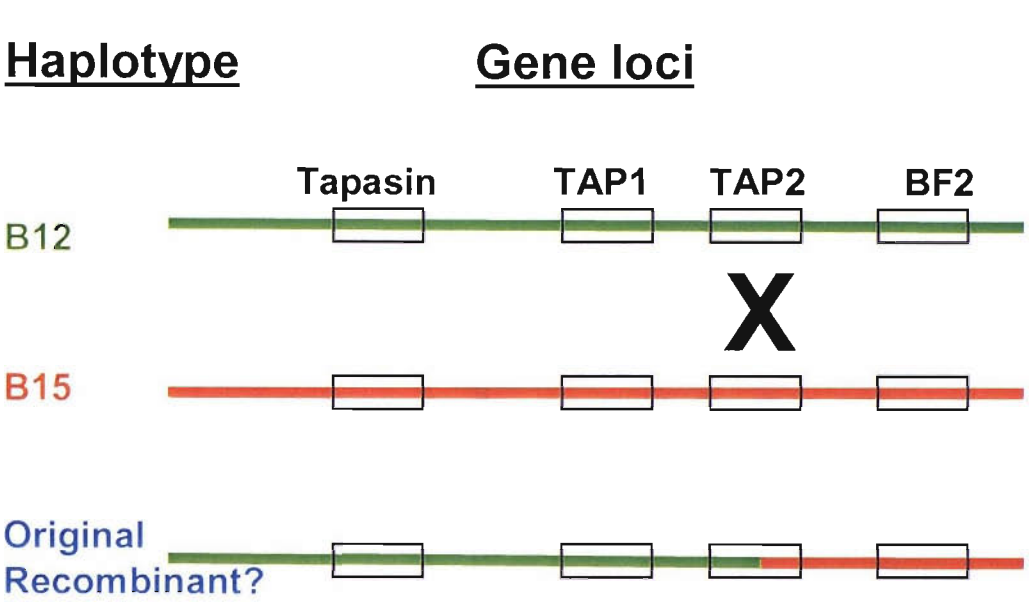
The strategy described in this chapter again involved the comparison of the expression characteristics of two class I alleles expressed separately in a MHC homozygous chicken cell line. One of these class I alleles was derived from the same haplotype as the host cell line, and so should serve as an example of an optimally peptide-loaded class I molecule. However, a potential tapasin incompatibility may have existed for the other class I allele. The design of this experimental strategy took advantage of the knowledge of a naturally occurring recombination event between the tapasin and BF2 loci in the B19v1 haplotype.

The B19v1 haplotype is the product of a rare recombination event which occurred between the B12 and B15 haplotypes (or at least a haplotype very similar to the modern B15 haplotype, J. Kaufman and B. Walker, personal communication, figure 6.1a). The recombination event occurred towards one end of the MHC within the TAP2 locus, resulting in the B19v1 haplotype being virtually identical to the B12 haplotype throughout the MHC region, with the B12 and B19v1 haplotypes sharing an identical tapasin protein (Tapasin*12), and encoding serologically indistinguishable BL antigens (Simonsen et al., 1982). However, the B19v1 haplotype differs from the B12 haplotype at the TAP2 and BF2 loci, where the alleles have derived from the B15 haplotype (figure 6.1b).

The TAP2 and BF2 alleles that are present in the modern B19v1 haplotype are distinct from the B15 alleles that are thought to have inserted into the B12 haplotype to form the new recombinant. This may suggest that immediately after the recombination event, the new combination of alleles that were present in the original recombinant haplotype (Tapasin*12, TAP1*12, TAP2*12/15 and BF2*15) were capable of working together to some extent; otherwise the recombinant haplotype was likely to have been lost from the population. However, the TAP2 and BF2 alleles appear to have subsequently evolved into the alleles that are now present in the modern B19v1 haplotype. It is possible that some degree of incompatibility between the BF2*15 and Tapasin*12 molecules that were present in the original recombinant may have driven the evolutionary changes in the BF2*15 allele to form the BF2*19v1 allele, which may be more compatible for the Tapasin*12 allele.

The class I peptide-binding motif determined from B19v1 homozygous blood cells is highly similar to the B15 class I motif (and is dissimilar to the B12 class I motif, table 6.1) (Kaufman et al., 1995), suggesting that the subsequent evolution within the B19v1 haplotype of the original BF2*15 allele into the BF2*19v1 allele has maintained essentially the same peptide-binding motif. This also suggests that the transport specificity of the TAP proteins encoded by the B19v1 haplotype is similar to that of the TAP proteins encoded by the B15 haplotype. This has since been confirmed in peptide transport assays of iodinated index peptides bearing glycosylation sequence motifs (figure 6.2, B. Walker, J. Kaufman personal communication).

a)



b)

Haplotype	Tapasin	TAP1	TAP2	BF2
B12	12	12	12	12
B15	15	15	15	15
Original?	12	12	12/15	15
B19v1	12	12	19v1 function = 15	19v1 motif = 15

Figure 6.1 Schematic depicting the likely origin of the B19v1 haplotype

a) Cartoons depicting the recombination event occurring within the TAP2 locus, between the B12 and B15 haplotypes, giving rise to the original recombinant haplotype, which may have evolved into the B19v1 haplotype.

b) Table depicting the alleles of tapasin, TAP and BF2 loci present in the B12, B15 and B19v1 haplotypes, and in the original recombinant haplotype (see text for details).

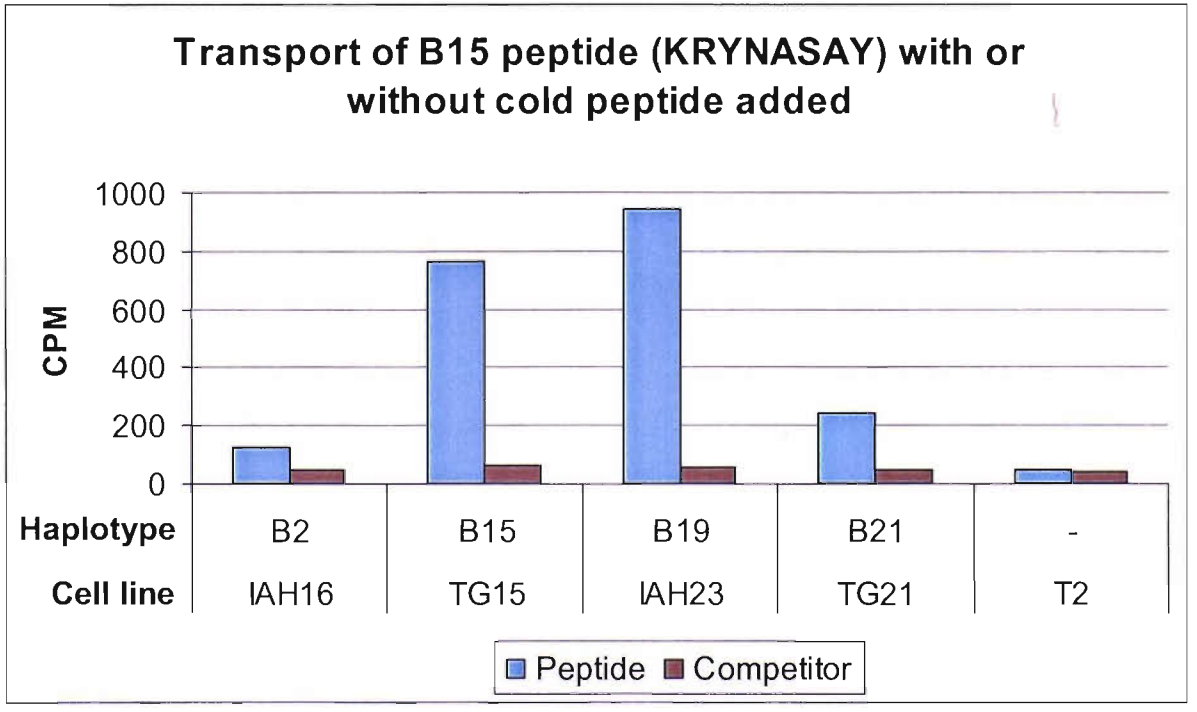


Figure 6.2 The restrictive TAP transport specificity evident in different MHC haplotypes

A peptide (KRYNASAY) that matches the peptide-binding motif of the B15 haplotype was synthesised and iodinated. The iodinated peptide was then cultured, with or without non-radioactive KRYNASAY peptide (competitor), in the presence of ATP with microsomes prepared from each of the cell lines, before being cooled and lysed by the addition of N-P40 lysis buffer. Glycosylated peptides were then recovered by the addition of Con-A sepharose and the amount of radioactivity recovered determined by γ-counting.

Data kindly provided by Drs. B. Walker and J. Kaufman (unpublished).

The reasons for the B19v1 haplotype having a TAP transport specificity similar to that of the B15 haplotype are not immediately apparent; as the B19v1 haplotype encodes a TAP1 protein identical to that found in the B12 haplotype, and the TAP2*19v1 protein differs from the TAP2*12 protein by only four amino acids, which are all located within the nucleotide-binding domain. It is possible that the B12 TAP transport specificity may be sufficiently broad to transport appropriate peptides for both the BF2*12 and BF2*15 molecules.

	B19									B15									B12							
position	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8
Anchor		R									R						Y	y					V			V
																							I			
strong	K		F	P	G	M		P		K		F	P								A	N	K		G	R
			L		A	K		Y				L	I								D	F			A	K
			I					F				H									P	Q			N	Q
			Y					L				N										E				E
												Y														
Weak	D		D	E	F	D	V			R		A	E	G	I	E	F	f	V		I	L		W		L
	V		A	V	N	R						Q	V	R	Q		W	w	K		M	P				I
	T		H	Q		Y						V	A						L			W				
			M	S									L													
			N	K																						
			V																							

Table 6.1 Class I peptide motifs for three MHC haplotypes

Sequences of peptides bound to class I molecules isolated from three chicken haplotypes determined from peptide pools showing anchor, strong and weak signals. Basic residues in blue, acidic residues in red, all others in black. B15 motif taken with permission from (Wallny et al., 2006).

The BF2*15 and BF2*19v1 proteins differ by eight amino acids, which are shown along with the equivalent residue in BF2*12 in table 6.2. Five of the first six polymorphisms are located within the peptide-binding groove (69, 79, 95, 111 and 113) and are presumed to be in contact with the peptide. The last two polymorphisms (126 and 220) are located outside of the peptide-binding groove in loops in the α2 and α3 domains which, when aligned with mammalian class I molecules, occupy positions that are implicated in associating with the mammalian tapasin molecule (appendix seven). Of particular note is the fact that the B12 and B19v1 haplotypes, which have identical Tapasin*12 proteins, also have identical residues at these positions within their BF2 alleles (five of the eight polymorphic positions are shared between the BF2*19v1 and BF2*12 alleles: 22, 79, 95, 126 and 220). If these two BF2 alleles are specifically dependent on the Tapasin*12 protein that is encoded from their haplotypes, then the residues at 126 and 220 may influence the specificity of the interaction with the tapasin molecule; G126 and R220 may define an interaction with Tapasin*12, whilst D126 and Q220 may define an interaction with Tapasin*15.

Amino acid position	Domain	Structural position and orientation	Total number of BF2 alleles	BF2*15	BF2*19v1	BF2*12
22	**α1**	β2 strand, upward	**2**	**Y**	**F**	**F**
69	**α1**	α helix, into groove, peptide contact, polymorphic	**3**	**T**	**S**	**N**
79	**α1**	α helix, into groove, peptide contact, polymorphic	**2**	**T**	**I**	**I**
95	**α2**	β1 strand, upward, peptide contact, polymorphic	**2**	**L**	**W**	**W**
111	**α2**	β2 strand, upward, peptide contact, polymorphic	**4**	**S**	**R**	**Y**
113	**α2**	β2 strand, upward, peptide contact, polymorphic, β_2m contact	**5**	**D**	**Y**	**M**
126	**α2**		**2**	**D**	**G**	**G**
220	**α3**	CD8 contact	**2**	**Q**	**R**	**R**

Table 6.2 Amino acid polymorphisms in BF2*15 and BF2*19v1 alleles

Amino acid polymorphisms between BF2*15 and BF2*19v1 alleles are tabulated together with the amino acid that is present at these positions in the BF2*12 allele. The polymorphic residues are numbered from the first residue of the mature protein and the domain in which they are located. The structural position and orientation is taken from (Kaufman et al., 1992), and takes into account the mainchain and first sidechain bonds, without implying the location of the end of the sidechain. Details are provided for the location, direction into which the sidechain points, and whether the residue is presumed to be in contact with the peptide, β_2m, or CD8, based upon the structural knowledge of HLA-A*0201, and whether the amino acid is polymorphic in human or murine alleles. The number of BF2 alleles with polymorphisms at the indicated position is tabulated (BF2 sequences are provided in appendix).

6.1.2 Potential strategies

There are a number of possible MHC mismatching scenarios that could be established to test such hypotheses, which are discussed below:

1. The presumed original recombinant haplotype could be partly recreated by transfection of the BF2*15 allele into a B12 cell line.

B12 cell line	Tapasin*12	TAP1*12	TAP2*12	Wild-type BF*12	**Transfected** BF2*15

Whilst such transfectants would recreate the mismatching of BF2*15 and Tapasin*12 alleles that is likely to have existed in the original recombinant haplotype, the endogenous TAP proteins may not transport peptides suitable for BF2*15 molecules. In addition, there are no B12 cell lines available for such experiments.

2. The reciprocal scenario to that described above could be created by transfection of the BF2*12 allele into a B15 cell line.

B15 cell line	Tapasin*15	TAP1*15	TAP2*15	Wild-type BF2*15	**Transfected** BF2*12

Such transfectants would again create a mismatching between the endogenous Tapasin*15 and transfected BF2*12, but would also suffer the problem that the endogenous TAP proteins may not transport peptides suitable for BF2*12 molecules.

3. A scenario with similarities to the presumed original recombinant haplotype could however be created by transfection of the BF2*19v1 allele into a B15 cell line.

B15 cell line	Tapasin*15	TAP1*15	TAP2*15	Wild-type BF2*15	**Transfected** BF2*19v1

In this scenario, the endogenous TAP proteins would be expected to transport peptides suitable for BF2*19v1 molecules; based upon peptide transport assay data, and similar peptide motifs eluted from B15 or B19v1 homozygous blood cells. There would also be a potential incompatibility between the BF2*19v1 and Tapasin*15 molecules; as the B19v1 allele may have evolved (from the BF2*15 allele) in order to work more efficiently with the Tapasin*12 allele that is encoded in the B19v1 haplotype.

4. A similar experimental scenario to that described could be established by the transfection of the BF2*15 allele into a B19v1 cell line.

B19v1 cell line	Tapasin*12	TAP1*12	TAP2*19v1	Wild-type BF*19v1	**Transfected** BF2*15

In such transfectants, the peptides supplied by the endogenous TAP molecules are likely to be appropriate for BF2*15 molecules, and there may be a potential incompatibility between the endogenous Tapasin*12 and the BF2*15 molecules.

6.1.3 Strategy: Part one

The experimental strategy that was adopted is based upon the third situation described above (figure 6.3 a-b); whereby C-terminally myc epitope tagged BF2*15 or BF2*19v1 molecules (referred to as BF2myc*15 or BF2myc*19v1 hereafter) were transfected into the B15 homozygous TG15 cell line. BF2myc*15 molecules were included as they should be supplied with appropriate peptides and able to interact with the endogenous Tapasin*15 and other co-factor molecules in an optimal fashion, and should serve as an example of a class I molecule that is loaded with optimal peptides. In contrast, although BF2myc*19v1 molecules are likely to be supplied with appropriate peptides by the endogenous TAP proteins, there may be an incompatibility with the endogenous Tapasin*15 molecules.

It is likely that a potential incompatibility between Tapasin*15 and BF2myc*19v1 molecules has only a moderate phenotype; as it would appear that in the original B19v1 recombinant haplotype (before the presumed period of molecular evolution that led to the B19v1 alleles of TAP2 and BF2) the combination of alleles of tapasin, TAP and BF2 were able to work in a fashion which at the least maintained the recombinant haplotype in the population.

Therefore, once the stable transfectants were obtained, the expression characteristics of the BF2myc molecules were compared in detail to assess whether there was any evidence of a potential incompatibility between the BF2myc*19v1 and Tapasin*15 molecules. Two different experimental techniques were employed in this characterisation: a comparison of thermal-stability of the BF2myc molecules; and a comparison of the rates of maturation of BF2myc molecules.

Thermal-stability assays were performed to compare the BF2myc molecules in the transfectants, as mammalian class I molecules loaded in the absence of tapasin have been demonstrated to have decreased thermal-stability in comparison to those loaded in the presence of tapasin (Williams et al., 2002). In such assays, the ability for class I heterodimers to remain in association, when a detergent lysate is heated to various temperatures, correlates with the affinity of their bound peptide. An incompatibility with Tapasin*15 may lead to BF2myc*19v1 molecules being loaded with less optimal peptides than were loaded in the BF2myc*15 molecules. Therefore, BF2myc*19v1 heterodimers would be expected to dissociate into free heavy chain and β_2m proteins at lower temperatures than would occur for the BF2myc*15 molecules.

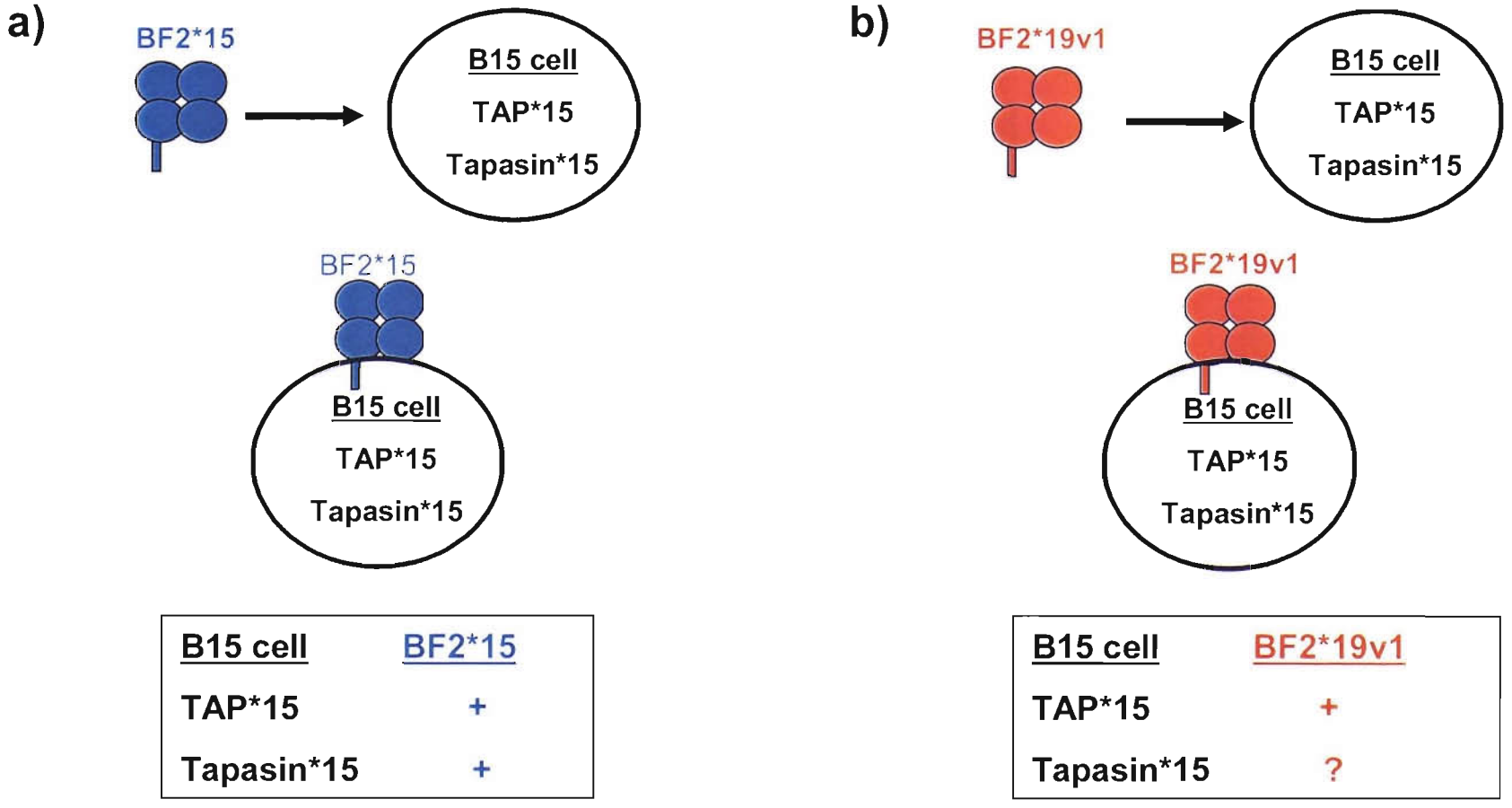
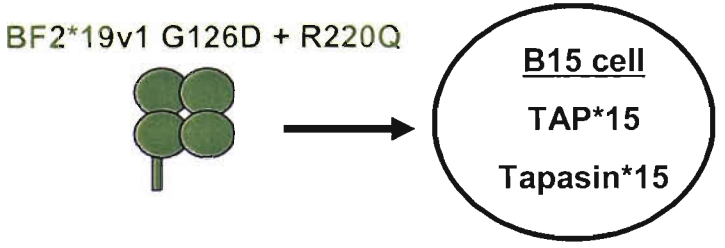
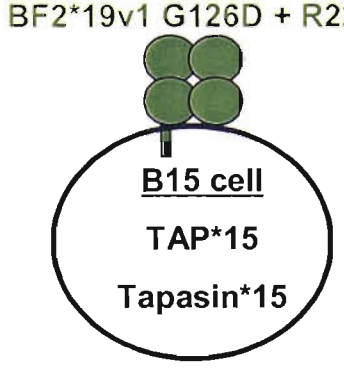


Figure 6.3 **The experimental strategy that is described in this chapter.**
a-d) The TG15 cell line was transfected with the indicated molecules. Tables provide details of suitable interactions (shown as +) and potential incompatibilities (shown as ?) for the transfected BF2 molecules. **Continued on next page**

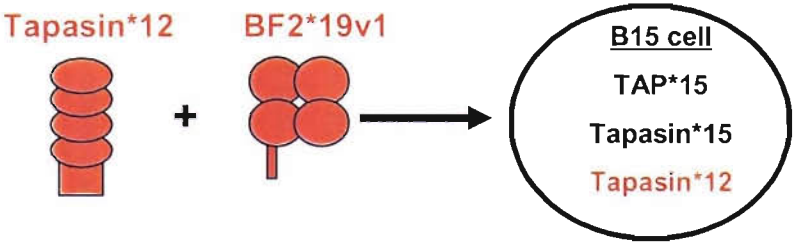
c)

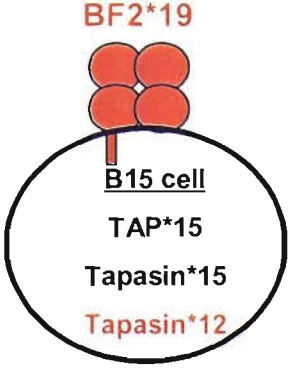




B15 cell	BF2*19v1 G126D + R220Q
TAP*15	+
Tapasin*15	?

d)





B15 cell	BF2*19v1
TAP*15	+
Tapasin*15	?
Tapasin*12	+

Pulse/chase experiments were performed to compare the BF2myc molecules in the transfectants, as an incompatibility between BF2myc*19v1 and Tapasin*15 molecules may lead to a slower rate of peptide-loading, or to a larger proportion of molecules that fail to become loaded with appropriate peptides that permit their egress to the cell surface, than would be apparent for the BF2myc*15 molecules. However, the rate of maturation of class I molecules is likely to be influenced by a number of factors besides the ability of tapasin (and/or other co-factor chaperone molecules) to facilitate optimal peptide-loading, or to retain the class I molecules within the ER. For example, the rate of maturation may be influenced by: the abundance of suitable peptides within the ER, the rate at which ERp57 and/or tapasin may oxidise the disulphide bond within the peptide-binding domain of the class I molecules (or at the N-terminus of tapasin); or perhaps by the extent of competition occurring between different class I molecules prior to their association with β_2m or other components of the PLC.

Although murine class I molecules mature at the same rate in tapasin-deficient and wild-type tapasin-expressing murine cells (Grandea et al., 2000), tapasin expression has been observed to alter the rate and extent of maturation of mammalian class I molecules in other studies. Firstly, in the human tapasin-deficient .220 cell line the products of many class I alleles do not progress beyond the ER in the absence of tapasin-mediated peptide-loading; secondly, tapasin transfection decreased the rate of maturation of H-2K^b molecules expressed in .220 cells (Barnden et al., 2000); thirdly, mutant HLA-A*0201 T134K molecules, which do not associate with the PLC, demonstrate a faster rate of maturation than wild-type HLA-A*0201 molecules (Lewis et al., 1996).

6.1.4 Strategy: Part two

The hypothesis that positions 126 and 220 within the class I molecule influence the specificity of the interaction with tapasin was tested by replacing the residues present at these positions in the BF2myc*19v1 molecule (glycine 126 and arginine 220) with those residues present in the BF2myc*15 molecule (aspartic acid 126 and glutamine 220) (figure 6.3c). The expression characteristics of this BF2myc*19v1 G126D & R220Q mutant and of the reciprocal BF2myc*15 D126G & Q220R mutant were compared to those of wild-type BF2myc*15 and BF2myc*19v1 molecules by pulse/chase and thermal-stability assays to determine whether the expression characteristics of the BF2myc molecules was affected by the mutation of these residues.

6.1.5 Strategy: Part three

It was of interest to determine the effect that transfection of Tapasin*12 has on the expression characteristics of the BF2myc*19v1 molecules (figure 6.3d). It is possible that BF2myc*19 is specifically dependent upon Tapasin*12 in order to achieve optimal peptide-loading. In addition, it is also possible that Tapasin*12 has a superior ability to facilitate peptide-loading compared with Tapasin*15.

6.2 Results

In this chapter BF2myc expression constructs derived from different haplotypes (B15 or B19v1) were transfected into the B15 homozygous TG15 cell line, and stable BF2myc expressing clones were generated. The expression characteristics of BF2myc*19v1 molecules were compared to BF2myc*15 molecules by pulse/chase analysis, and by thermal-stability assays.

6.2.1 BF2myc*19v1 molecules are impaired in their maturation in comparison to BF2myc*15 molecules

In the first experiment (experiment 231), the maturation rates of BF2myc*15, BF2myc*19v1 and BF2myc*4 molecules was determined by a five minute pulse radiolabelling, and two hour chase period. BF2myc*15 molecules were included as they should be able to associate with endogenous co-factor molecules in an optimal fashion and are a control for influences arising from transfection or the addition of a C-terminal myc epitope tag. In the previous chapter, similar pulse/chase experiments suggested that N-terminally tagged Flag-BF2*15 molecules were compromised in their ability to be loaded with peptides and to progress beyond the medial Golgi. Consequently, only the C-terminal myc tag was employed in this study.

BF2myc*4 molecules were included in the experiment as they should serve as an example of a peptide-starved class I molecule; in the previous chapter the vast majority of N-terminally tagged Flag-BF2*4 molecules did not progress beyond the medial Golgi, which is likely to reflect impairments resulting from both the N-terminal Flag tag and the scarcity of suitable peptides supplied by the endogenous TAP molecules.

In a B15 haplotype cell there should be an abundance of peptides within the ER that are suitable for BF2myc*19v1 molecules. However, the peptide-loading of these molecules may be compromised by an incompatibility with the endogenous Tapasin*15 molecules, which may result in altered maturation kinetics of newly made class I heterodimers.

Figure 6.4 shows the extent and rate of maturation of BF2myc molecules. Approximately half of the labelled BF2myc*15 cohort gained resistance to endo H digestion after one hour of the chase, after two hours of the chase virtually the entire cohort were resistant. The almost complete maturation of BF2myc*15 molecules in this experiment suggests that these molecules were not compromised in their ability to load with peptides. This confirmed the N-terminal Flag epitope tag used in the previous chapter was the cause of the impaired maturation evident for such class I molecules.

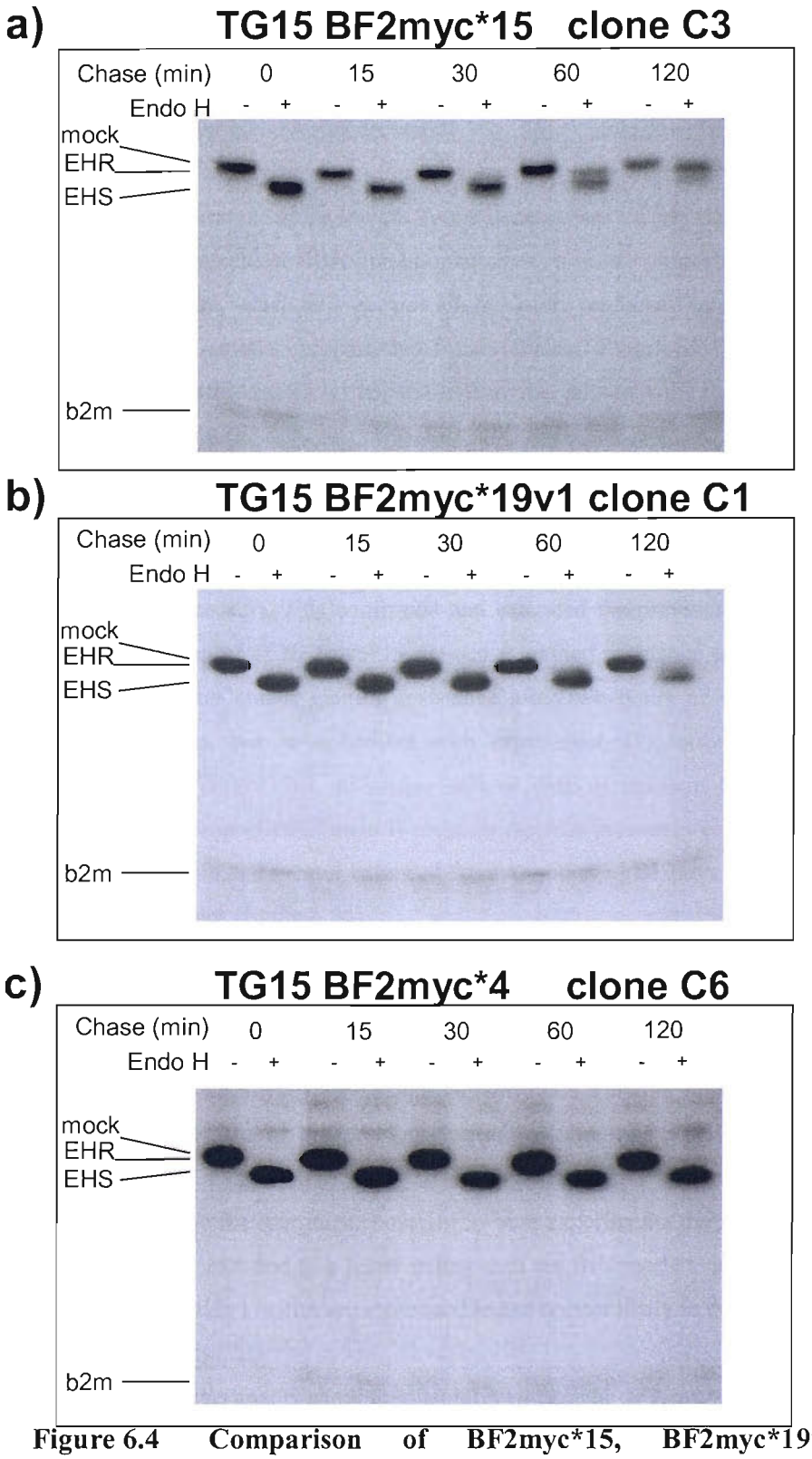


Figure 6.4 Comparison of BF2myc*15, BF2myc*19v1 and BF2myc*4 maturation rates in TG15 cells (Experiment 231)

a-c) The indicated clones were pulse labelled for five minutes and chased for 120 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and exposed to film.

In contrast to BF2myc*15 molecules, no BF2myc*4 molecules become endo H resistant after two hours of chase, which confirms the majority, if not all of these class I molecules do not become loaded with peptide repertoires that permit their egress to the cell surface, and that in effect, the *cim* effect has been recreated in these transfectants.

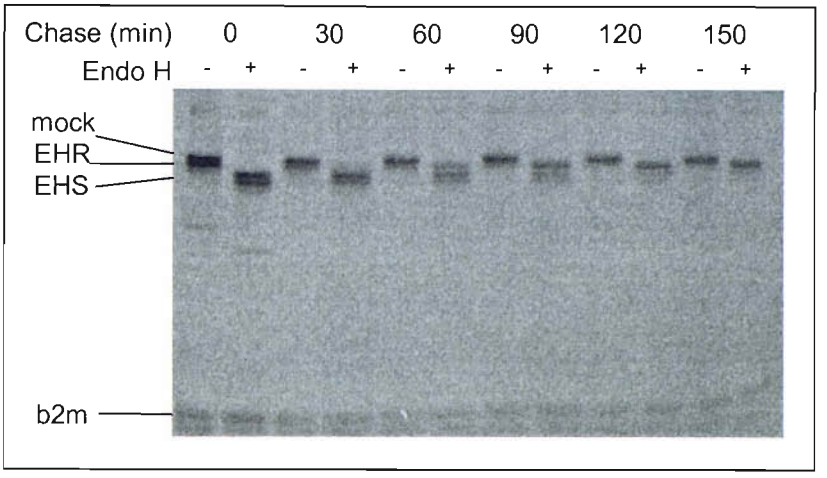
The extent of the maturation of BF2myc*19v1 molecules was clearly decreased in comparison to that of BF2myc*15 molecules. After one hour of chase, a minor proportion of the labelled cohort became endo H resistant, which only became slightly more prominent after a further hour of chase. This suggests that a smaller proportion of the labelled cohort of BF2myc*19v1 molecules progressed beyond the medial Golgi apparatus than was apparent for the BF2myc*15 molecules during the two hour chase.

To confirm these findings, the experiment was repeated with the same BF2myc*15 and BF2myc*19v1 clones being radiolabelled for five minutes and then chased over 210 minutes (experiment 233, figure 6.5). This confirmed and extended the previous findings; approximately half of the labelled cohort of BF2myc*15 molecules gained resistance to endo H after one hour, with virtually the entire cohort gaining resistance after two hours of the chase. In contrast to BF2myc*15 molecules, but in agreement with experiment 231, only approximately half the labelled cohort of BF2myc*19v1 molecules became endo H resistant by the end of the chase. Quantification of the proportion of endo H resistant material present at each timepoint of the chase confirmed the faster rate and greater extent of the maturation of BF2myc*15 molecules relative to BF2myc*19v1 molecules (figure 6.5d).

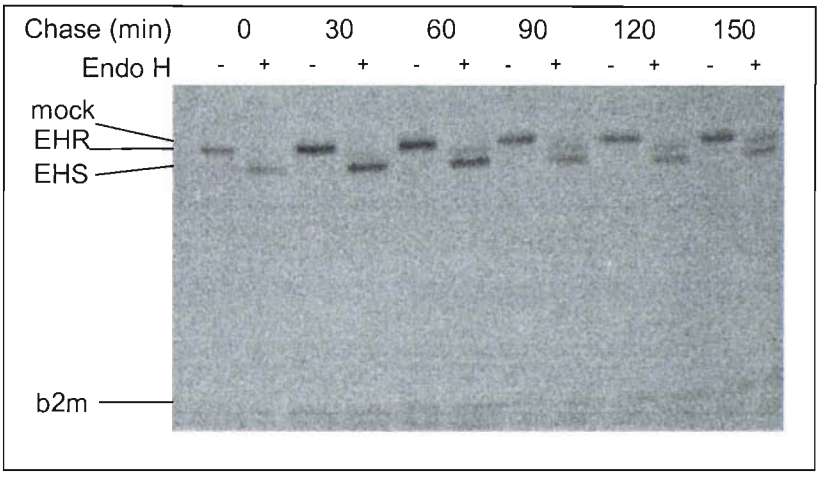
To ensure the differential maturation of BF2myc molecules was not an artefact unique to these specific clones, the experiment was repeated with two different clones (experiment 236, figure 6.6). In this experiment the BF2myc*15 molecules demonstrated a similar rate and extent of maturation to that which was evident in the previous experiment. In this experiment the BF2myc*19v1 molecules matured at a faster rate, and to a greater extent, than was evident for BF2myc*19v1 molecules in the previous experiment, however in both experiments the BF2myc*19v1 molecules still matured at slower rate and to a lesser extent than the BF2myc*15 molecules, confirming that the differential maturation kinetics are reproducible and are not likely to be clone-specific artefacts.

Cumulatively these experiments show that haplotype-matched BF2myc*15 molecules mature at a slightly faster rate and to a much greater extent than haplotype-mismatched BF2myc*19v1 molecules, when expressed in TG15 cells.

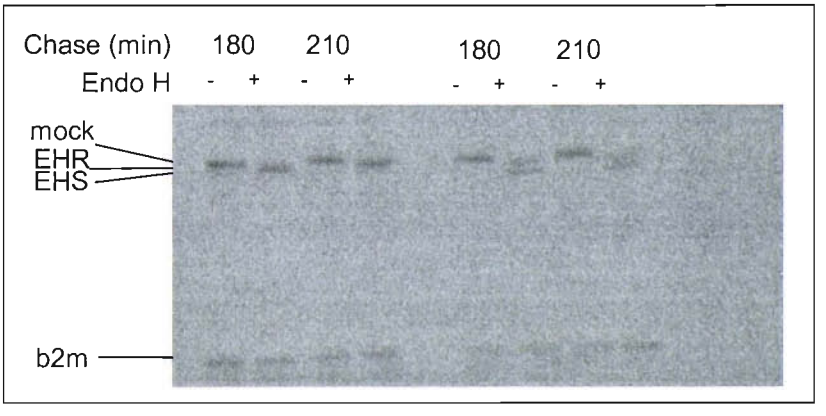
a) TG15 BF2myc*15 clone C3



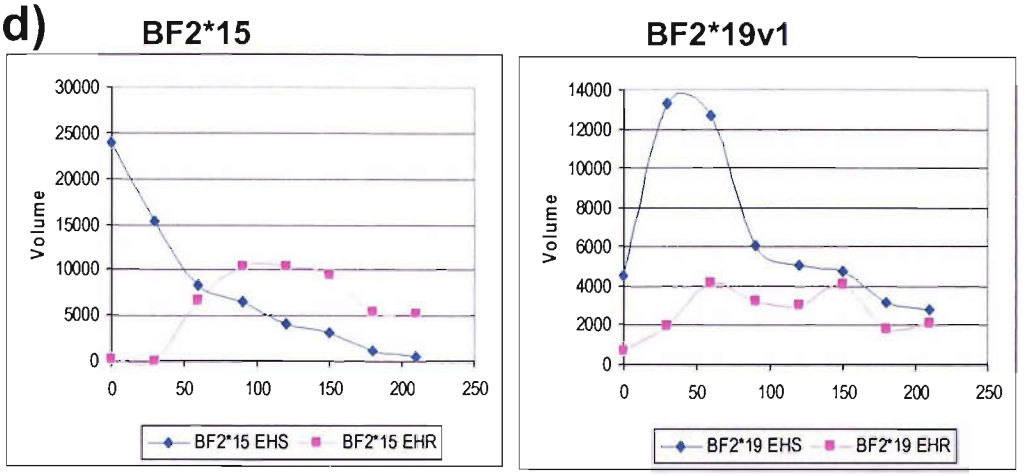
b) TG15 BF2myc*19v1 clone C1



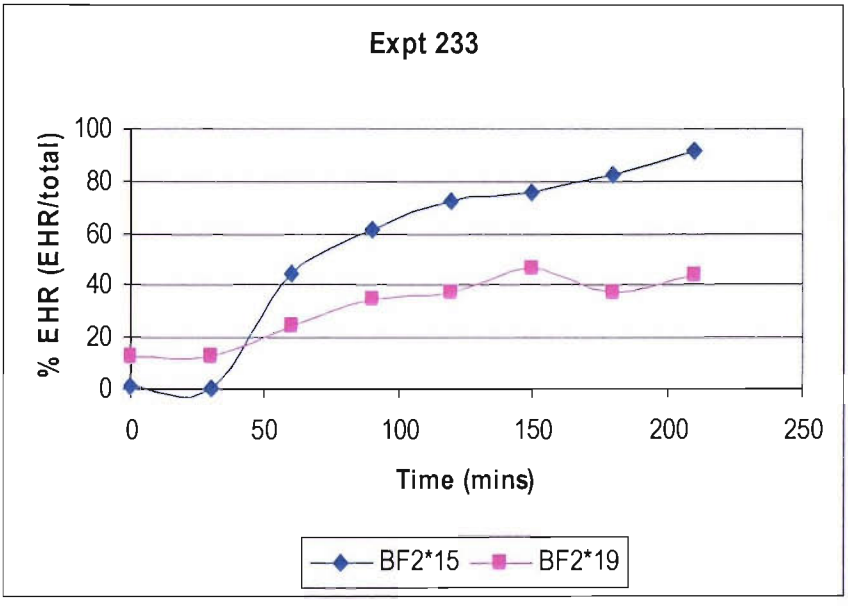
c) TG15 BF2myc*15 TG15 BF2myc*19v1



d)


BF2*15
BF2*19v1
Volume
30000
25000
20000
15000
10000
5000
0
14000
12000
10000
8000
6000
4000
2000
0
50
100
150
200
250
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR

e)


Expt 233
% EHR (EHR/total)
100
80
60
40
20
0
0
50
100
150
200
250
Time (mins)
BF2*15
BF2*19

f)

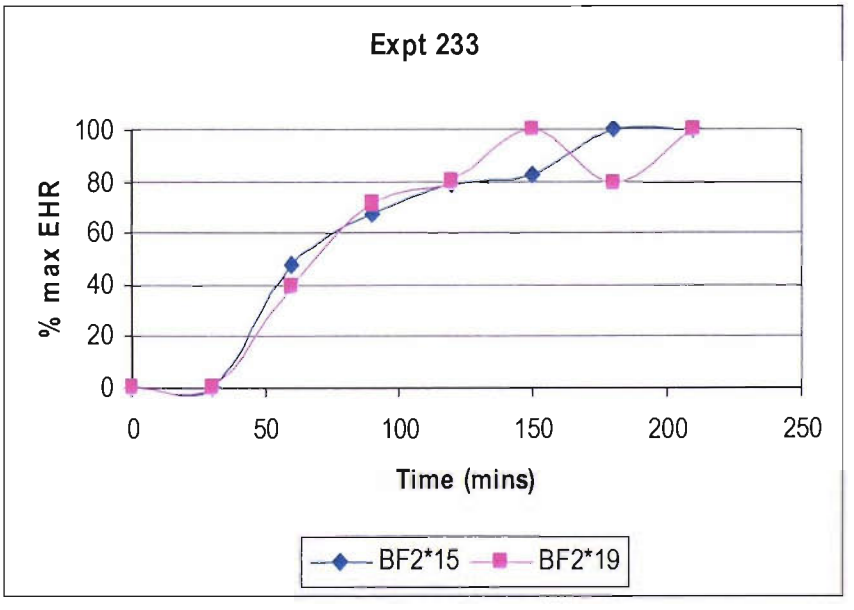

Expt 233
% max EHR
100
80
60
40
20
0
0
50
100
150
200
250
Time (mins)
BF2*15
BF2*19

Figure 6.5 Comparison of BF2myc*15 and BF2myc*19v1 maturation rates in TG15 cells (Experiment 233)

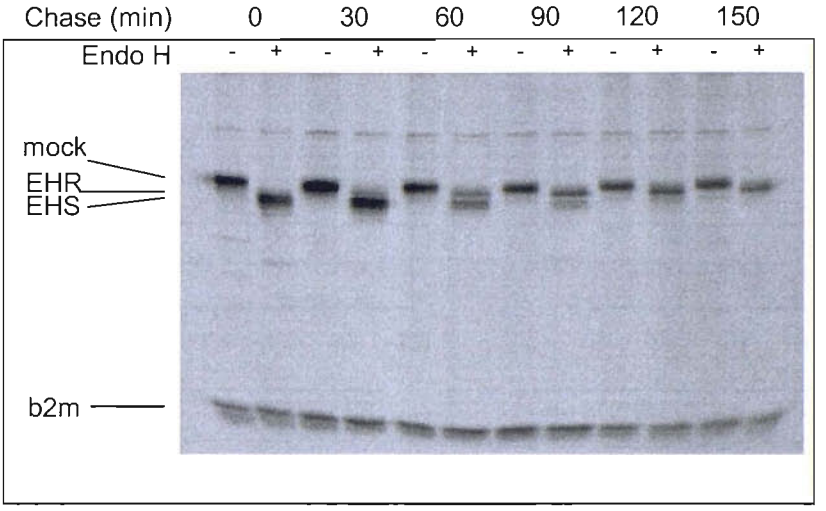
a-c) The indicated clones were pulse labelled for five minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis.

d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

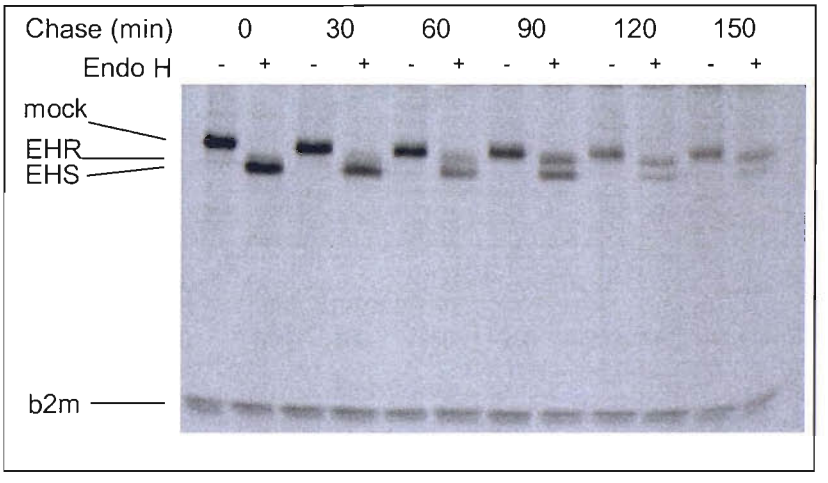
e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

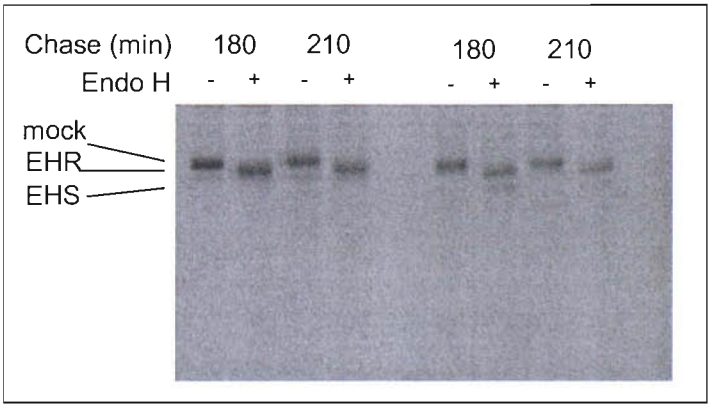
a) TG15 BF2myc*15 clone C7



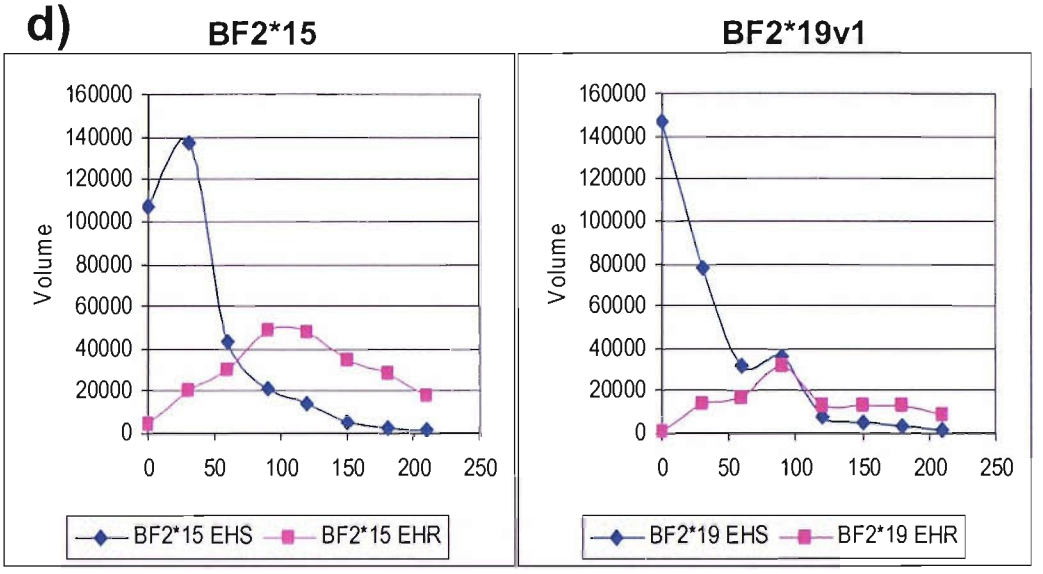
b) TG15 BF2myc*19v1 clone C9



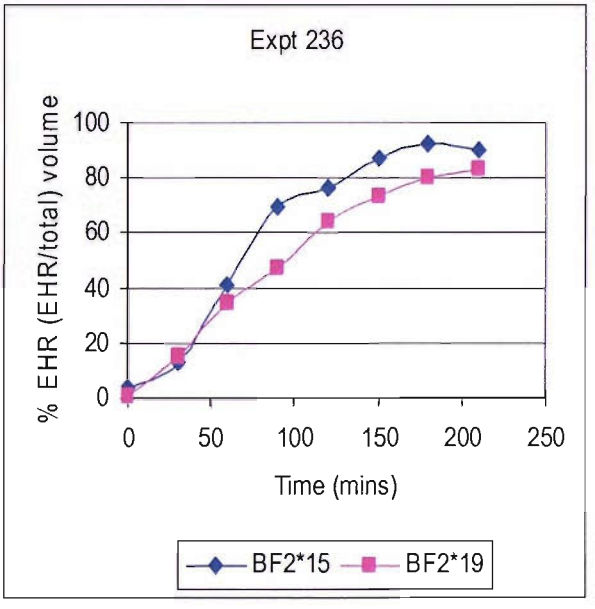
c) TG15 BF2myc*15 TG15 BF2myc*19v1



d)


BF2*15
BF2*19v1
Volume
160000
140000
120000
100000
80000
60000
40000
20000
0
0
50
100
150
200
250
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR

e)


Expt 236
% EHR (EHR/total) volume
100
80
60
40
20
0
0
50
100
150
200
250
Time (mins)
BF2*15
BF2*19

f)

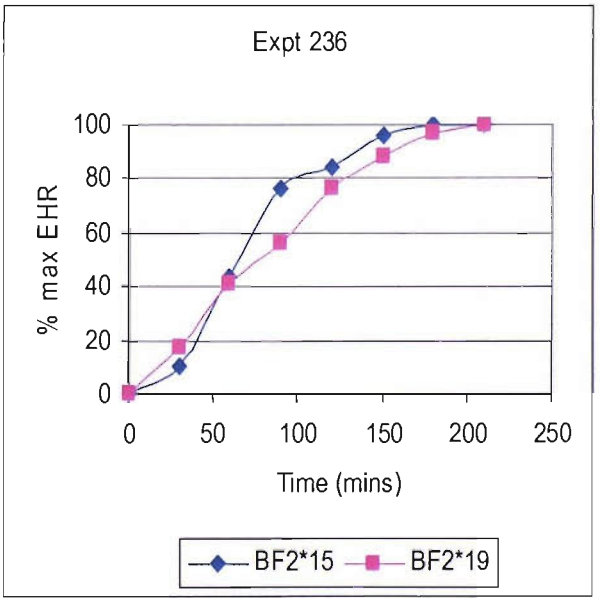

Expt 236
% max EHR
100
80
60
40
20
0
0
50
100
150
200
250
Time (mins)
BF2*15
BF2*19

Figure 6.6 Comparison of BF2myc*15 and BF2myc*19v1 maturation rates in TG15 cells (Experiment 236)

a-c) The indicated clones were pulse labelled for five minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis.

d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

6.2.2 Comparison of BF2myc alleles by thermal-stability assays

To determine whether the mismatching of Tapasin*15 and BF2myc*19v1 alleles that may have caused the decreased maturation characteristics was a result of a sub-optimal peptide cargo, as is observed for mammalian class I molecules loaded in the absence of tapasin, thermal-stability assays were performed. Two different experimental protocols were performed to compare the thermal-stability of the BF2myc molecules.

The first protocol involved a western blot analysis of the thermal-stability of the total pool of BF2myc molecules. Numerous experiments were performed using BF2myc*15 clone C3 and BF2myc*19v1 clone C1 with slight variations in the experimental conditions, with a representative example (experiment 281) being shown in figure 6.7. In this experiment, lysates were prepared and aliquots were either kept on ice or heated for one hour at various temperatures before being cooled. The BF2myc molecules were then immunoprecipitated following 90 minute incubation with an anti-myc antibody coupled to agarose (anti-myc agarose), before the immunoprecipitated proteins were washed, and then eluted and resolved by SDS PAGE. Following western blotting, the blots were divided and incubated with anti-myc or anti-β_2m antibodies.

For each clone roughly similar amounts of BF2myc molecules were immunoprecipitated and loaded into each lane of the gel (figure 6.7a). However, more BF2myc*19v1 molecules were immunoprecipitated than BF2myc*15 molecules from the lysates, which was observed in every experiment. Despite this difference in the amount of immunoprecipitated BF2myc molecules, a similar amount of co-precipitating β_2m molecules was detected in each clone when the lysates were not heated and kept on ice instead (figure 6.7b). Additionally, as far as can be determined, the heterodimers that were immunoprecipitated from each clone demonstrated equal thermal-stability; with the amount of co-precipitated β_2m molecules progressively decreasing when the lysates were heated at temperatures greater than 48°C, until no β_2m molecules were detected following heating at 56°C. This suggests firstly, that the BF2myc: β_2m heterodimers that were recovered from either clone were loaded with repertoires of peptides that conferred equal stability, and secondly that a substantially smaller proportion of BF2myc*19v1 molecules were recovered in association β_2m than was evident for BF2myc*15 molecules.

To confirm these findings a second thermal-stability protocol was performed using two different clones. This involved a pulse/chase radio-labelling of the transfectants before a similar thermal-stability assay to that described above was performed, which offers the advantage of examining the thermal-stability of the BF2myc molecules as they mature. However in this assay the immunoprecipitation step involved overnight incubation of the lysate with anti-myc agarose. This experiment (figure 6.8, table 6.3) confirmed and extended the findings of the "western blot" thermal-stability assay, by allowing three points to be made.

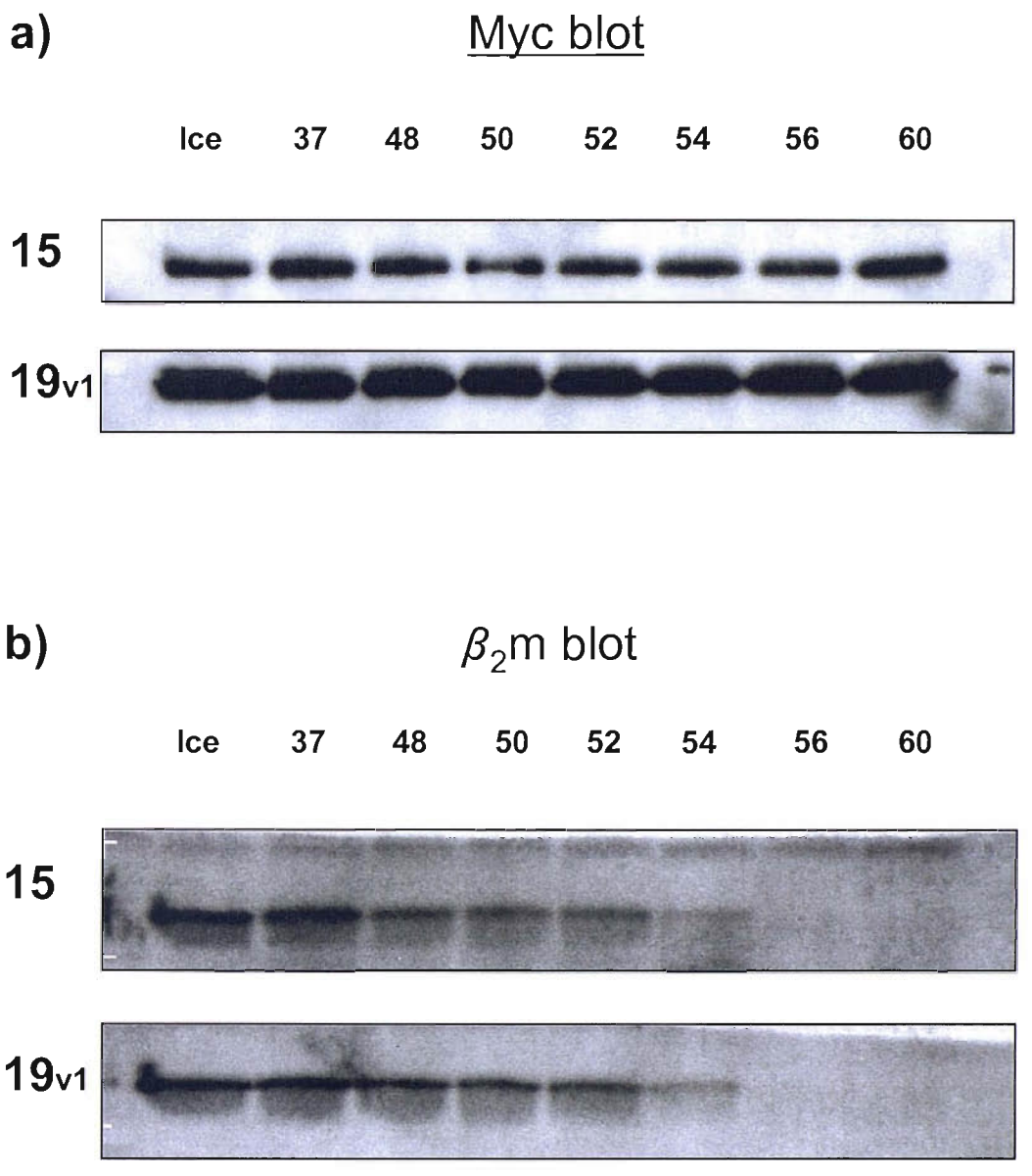
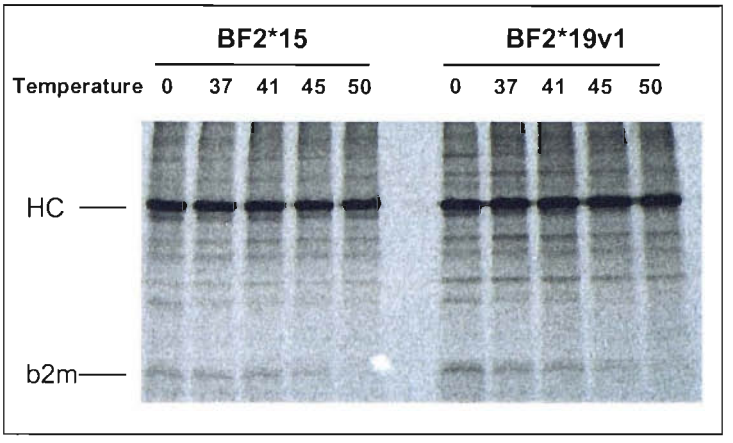


Figure 6.7 Thermal-stability of BF2myc*15 and BF2myc*19v1 molecules

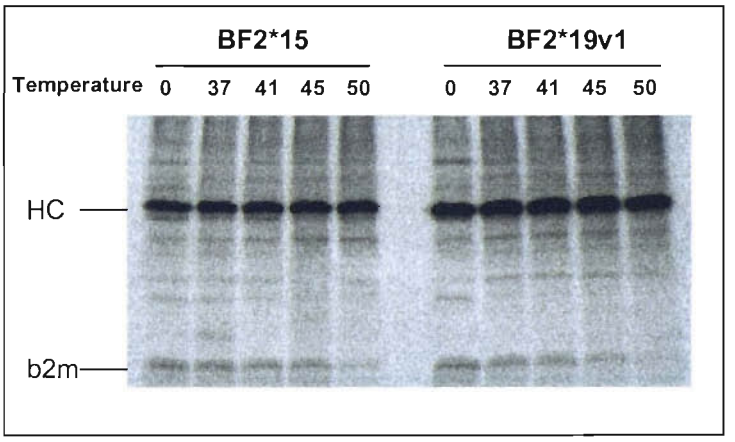
Lysates were made from TG15 BF2myc*15 clone C3 (labelled 15) and TG15 BF2myc*19v1 clone C1 cells (labelled 19). Volumes equivalent to $5x10^6$ cells were heated at the indicated temperatures for one hour, before myc agarose used to immunoprecipitate BF2myc molecules. Following SDS PAGE and western blotting, the blots were divided into sections according to the location of molecular weight markers and incubated with the indicated antibodies. Only the bands corresponding to either BF2myc or β_2m are shown for clarity.

a) Myc western blot
b) Rabbit anti-β_2m western blot

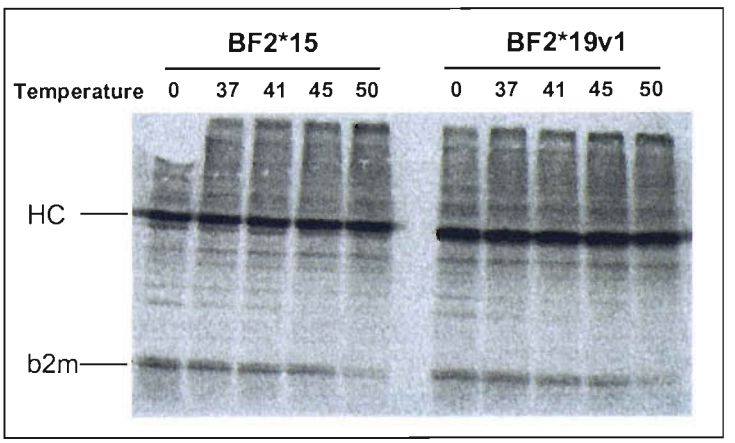
a)


0 hr chase
BF2*15
BF2*19v1
Temperature
0 37 41 45 50
0 37 41 45 50
HC
b2m

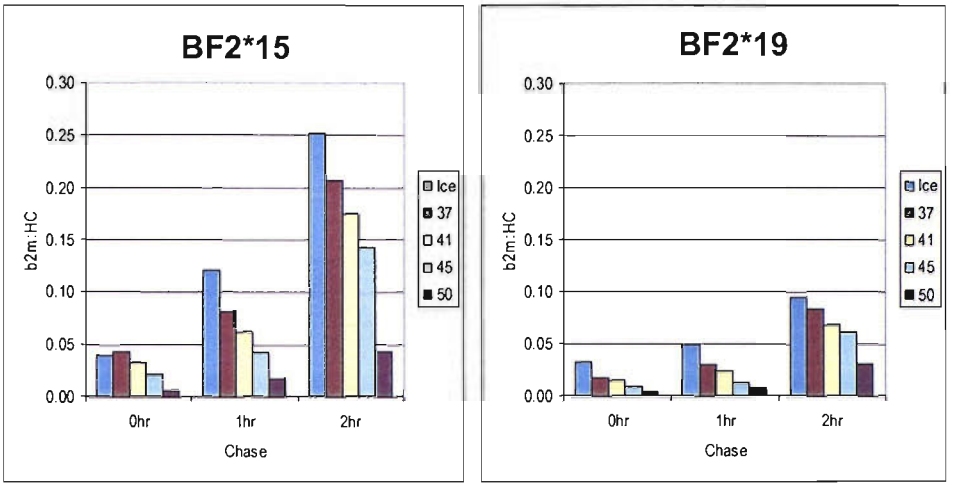
b)


1 hr chase
BF2*15
BF2*19v1
Temperature
0 37 41 45 50
0 37 41 45 50
HC
b2m

c)


2 hr chase
BF2*15
BF2*19v1
Temperature
0 37 41 45 50
0 37 41 45 50
HC
b2m

d)



e)

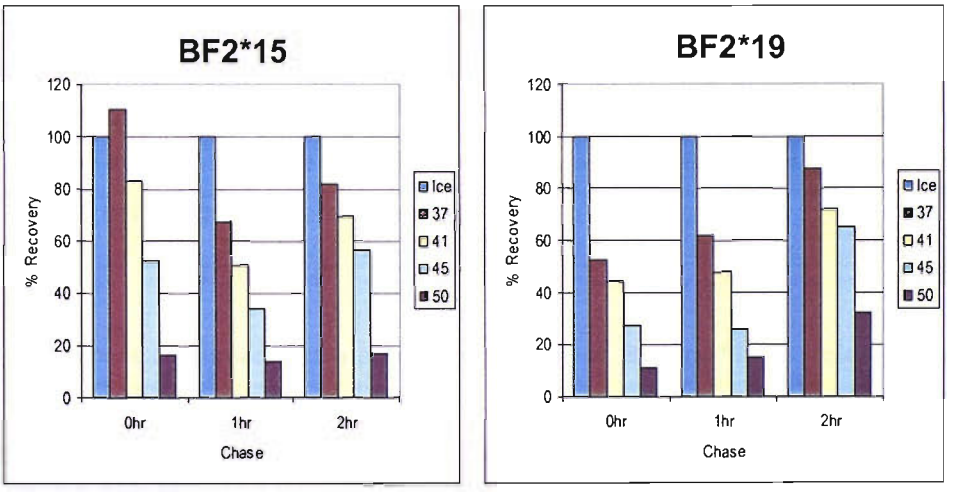


Figure 6.8 Comparison of BF2myc*15 and BF2myc*19v1 molecules by thermal stability assays (experiment 292).

a-c) Clones BF2myc*15 C7 and BF2myc*19v1 C9 were radio-labelled for 20 minutes and then chased for 120 minutes. Aliquots were removed as indicated and lysed, before being heated for one hour at the indicated temperatures. Myc tagged molecules were then immunoprecipitated and separated by SDS PAGE gels, fixed and subjected to phosphorimager analysis.

d) Graphs depicting the ratio of β2m molecules that were co-precipitated with BF2myc molecules (ratio = β2m band volume: myc HC band volume).

e) Graphs depicting the stability of the immunoprecipitated BF2myc:β2m heterodimers. Shown as the "Recovery %" for each clone, which is the percentage of heterodimers that were recovered following heating for each point of the chase relative to those that were recovered after incubation on ice at the same point of the chase.

BF2*15
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.039	0.122	0.252
37	0.043	0.082	0.207
41	0.033	0.062	0.175
45	0.021	0.042	0.143
50	0.006	0.017	0.043

BF2*15
Recovery %

	0hr	1hr	2hr
Ice	100.0	100.0	100.0
37	110.4	67.5	82.0
41	82.9	50.8	69.5
45	52.3	34.1	56.5
50	16.4	13.9	17.1

BF2*19
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.032	0.048	0.095
37	0.017	0.030	0.083
41	0.014	0.023	0.068
45	0.009	0.012	0.062
50	0.004	0.007	0.031

BF2*19
Recovery %

	0hr	1hr	2hr
Ice	100.0	100.0	100.0
37	53.1	62.1	87.6
41	44.5	47.9	72.2
45	27.1	25.7	65.1
50	11.0	15.0	32.3

Table 6.3 Pulse/chase thermal-stability assays comparing BF2myc*15 and BF2myc*19v1 molecules. Left column: the ratios of β_2m: heavy chain (HC) band volumes were calculated for each immunoprecipitate to normalise for differential recovery of HC ("b2m: HC ratio"). Right column: the b2m: HC ratio calculated at each time-point after incubation on ice was set to 100%, and the b2m:HC ratio that was calculated following heat treatment at that same time-point was calculated as a percentage of this value ("Recovery %").

Firstly, this experiment confirmed that a smaller proportion of BF2myc*19v1 molecules were recovered in association with β_2m than was evident for BF2myc*15 molecules (figure 6.8d). Secondly, this experiment confirmed that the heterodimers that were recovered at all time-points from either clone demonstrated comparable thermal-stability (figure 6.8e); as the percentage of heterodimers that were recovered following heating at each time-point decreased in similar amounts, in relation to the proportion recovered following incubation on ice. Thirdly, the amount of co-precipitating β_2m molecules increased throughout the chase, which is likely to reflect that for both class I alleles the peptide repertoire was optimised over time. Therefore after two hours of chase, the peptide repertoire had been optimised, permitting a larger proportion of heavy chains to associate stably with β_2m, allowing an increased recovery of heterodimers.

However, it is possible that in these experiments a substantial proportion of sub-optimally peptide-loaded heterodimers may have rapidly dissociated during the immunoprecipitation procedure, which would leave only those heterodimers loaded with optimal peptides (i.e. those peptides of an appropriate length and specificity to bind stably to the class I molecule) to be recovered. If so, this may be a consequence of using an anti-epitope tag antibody rather than a conformationally-specific antibody in the immunoprecipitation step. An antibody that is specific for class I heterodimers (if one were available) would have not only specifically immunoprecipitated β_2m-associated

heterodimers rather than capturing the total pool of BF2myc molecules, but would have stabilised the association between proteins.

Therefore to determine the extent to which β_2m-associated heterodimers dissociate during the immunoprecipitation step, clones expressing either BF2myc*15 or BF2myc*19v1 were radiolabelled, before being lysed in the presence or absence of a high concentration of specific peptide. Without an intervening preclear step, the lysates were incubated with anti-myc agarose for increasing periods of time, before being quickly washed and resolved by electrophoresis (figure 6.9, experiment 296).

The autoradiograms suggest that under these conditions both BF2myc alleles co-precipitated broadly similar amounts of β_2m molecules, a finding which contrasts with previous thermal-stability assays. The autoradiograms also showed that the efficiency of the immunoprecipitation procedure improved with increasing time, despite some proteolysis of the heavy chain proteins occurring at later time-points.

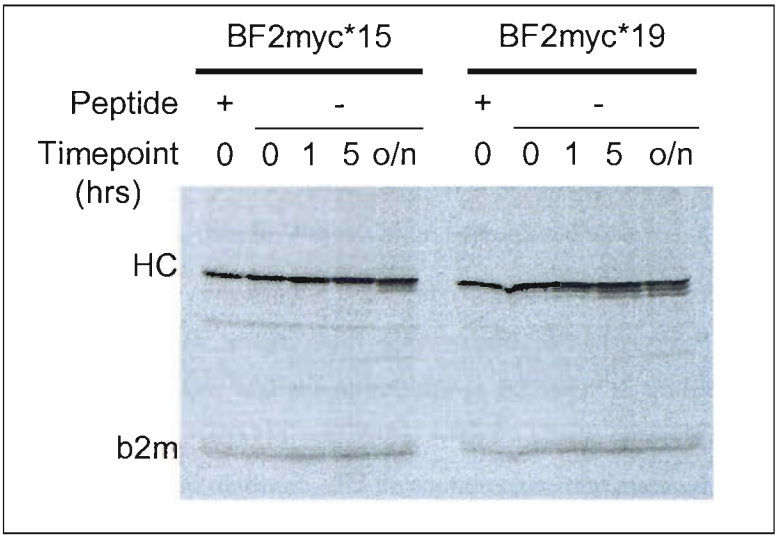
In order to normalise for the differential recovery of BF2myc molecules, the β_2m: BF2myc HC ratio was calculated. Analysis of these ratios revealed that when a high concentration of exogenous peptide was added during lysis, a virtually identical amount of β_2m molecules was co-precipitated with either BF2myc allele. However, in the absence of exogenous peptide, the amount of β_2m molecules that co-precipitated with BF2myc*19v1 molecules progressively decreased to an extent that was slightly greater than was apparent for the BF2myc*15 molecules. It is possible that with a preclear step and more comprehensive washing steps (such as used in the previously described thermal-stability assays) the β_2m: BF2myc*19v1 heavy ratio would have decreased further.

Therefore, this suggests that in both thermal-stability assays discussed above (experiments 281 and 292), a larger proportion of BF2myc*19v1 heterodimers dissociated during the immunoprecipitation procedure than was apparent for the BF2myc*15 molecules. The increased dissociation of BF2myc*19v1 heterodimers is likely to reflect a lower affinity peptide repertoire.

6.2.3 The maturation characteristics of BF2myc*19v1 molecules can be altered by mutation of specific residues

To test the prediction that the amino acids at positions 126 and 220 of the BF2myc alleles influence the interaction with tapasin, the residues present at these positions in the BF2myc*19v1 molecule (glycine 126 and arginine 220) were mutated to those residues present in the BF2myc*15 molecule (aspartic acid 126 and glutamine 220), with the mutant molecule being labelled BF2myc*19v1 G126D&R220Q. Transfectants of TG15 cells expressing such constructs were generated, and were compared to transfectants expressing BF2myc*15 and BF2myc*19v1 molecules by pulse/chase analysis.

a)



b)

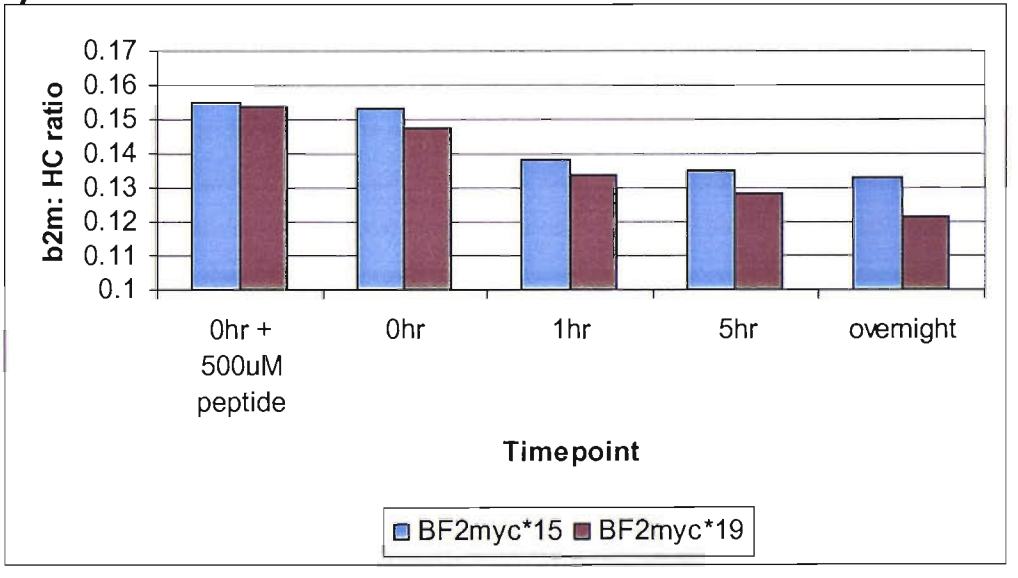


Figure 6.9 Dissociation of pMHC heterodimers over time (experiment 296)

a) TG15 BF2myc*15 clone C3 and TG15 BF2myc*19v1 clone C9 were radiolabelled for 30 minutes at 41°C, and then lysed in the presence or absence of 500μM KRLIGKRY peptide. Myc agarose was then added for a minimum of 20 minutes (0hr), and aliquots removed after a further hour, 5 hours or after overnight incubation. The immunoprecipitates were rapidly washed and frozen, before being resolved by electrophoresis and exposed to phosphor screens.

b) Graph depicting the calculated b2m: HC ratio (to account for differences in amount of immunoprecipitated heavy chains) for either clone.

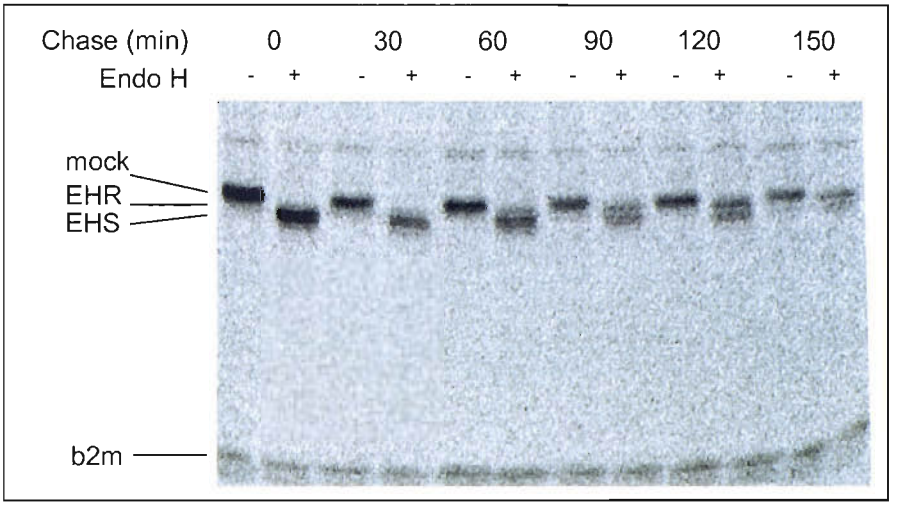
The extent and the rate of maturation of BF2myc*19v1 G126D&R220Q molecules was very similar to that of BF2myc*15 molecules (figure 6.10, experiment 245); being faster and proceeding to a greater extent than was evident for BF2myc*19v1 molecules. It must be noted that in this experiment the rates and the extents of maturation of BF2myc*15 and BF2myc*19v1 molecules were decreased relative to those observed in previous experiments, which may be a consequence of experimental variation, or of the comparison of a greater number of clones with this experimental protocol.

To confirm these findings the experiment was repeated with slight modifications to the experimental protocol (table 6.4, experiment 277). Again, BF2myc*19v1 G126D&R220Q molecules matured to similar extents and at similar rates as BF2myc*15 molecules (figure 6.11), which was greater than for BF2myc*19v1 molecules, where a smaller proportion of the labelled cohort gained resistance to endo H digestion, and those molecules that matured did so at a slower rate.

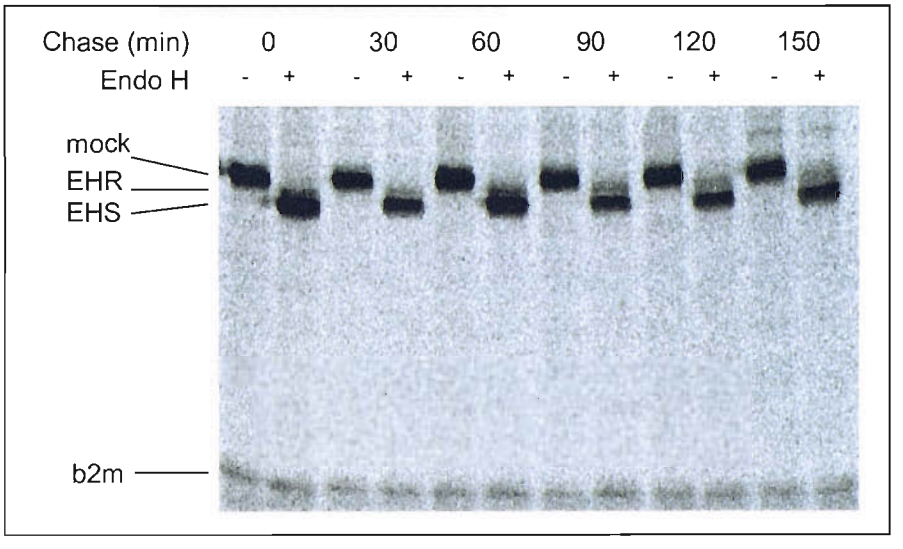
To determine whether one or both of the residues that occupy positions 126 and 220 are responsible for the alterations in maturation of BF2myc*19v1 molecules, single position mutants of BF2myc*19v1 were produced, which were labelled BF2myc*19v1 G126D or BF2myc*19v1 R220Q, and transfectants of TG15 cells were generated, which were compared with each other by pulse/chase analysis (figure 6.12, experiment 240). Both single position mutants demonstrated similar rates of maturation, which proceeded to equal extents. The result of this pulse/chase experiment is similar to that observed previously for BF2myc*15 molecules (experiments 231, 233 and 236) suggesting that both positions contribute to the improved maturation of BF2myc*19v1 mutant molecules.

To determine whether making the reciprocal mutations to the BF2myc*15 molecule also affected the rate and extent of maturation: the residues present at positions 126 and 220 in the BF2myc*15 molecule (aspartic acid 126 and glutamine 220) were mutated to those residues present in the BF2myc*19v1 molecule (glycine 126 and arginine 220); with the resultant molecule being labelled BF2myc*15 D126G&Q220R. Transfectants of TG15 cells expressing BF2myc*15 D126G&Q220R molecules were generated and were compared to transfectants expressing either BF2myc*15, BF2myc*19v1 or BF2myc*19v1 G126D&R220Q molecules by pulse/chase analysis.

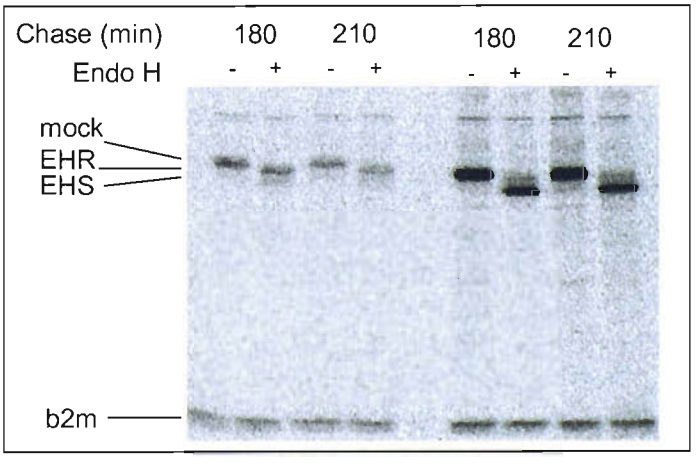
a) TG15 BF2myc*15 clone C7



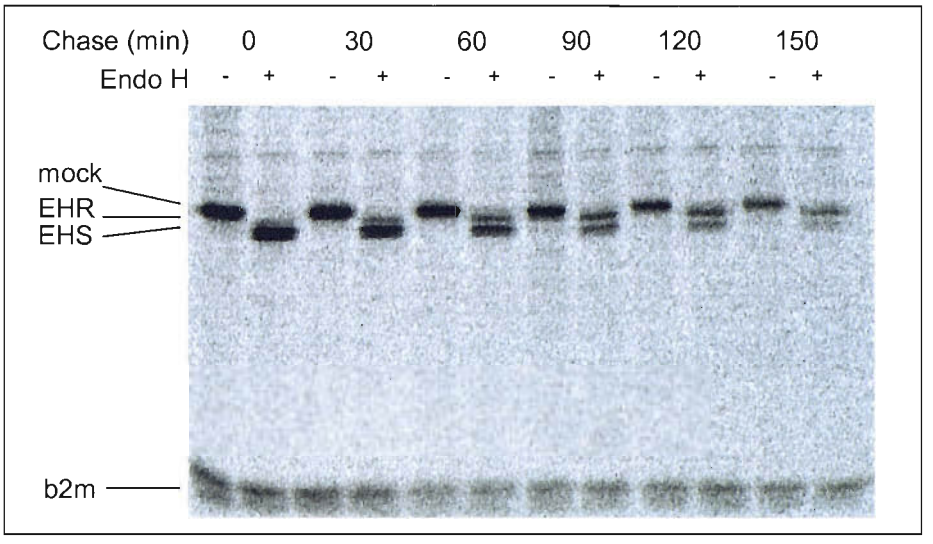
b) TG15 BF2myc*19v1 clone C9



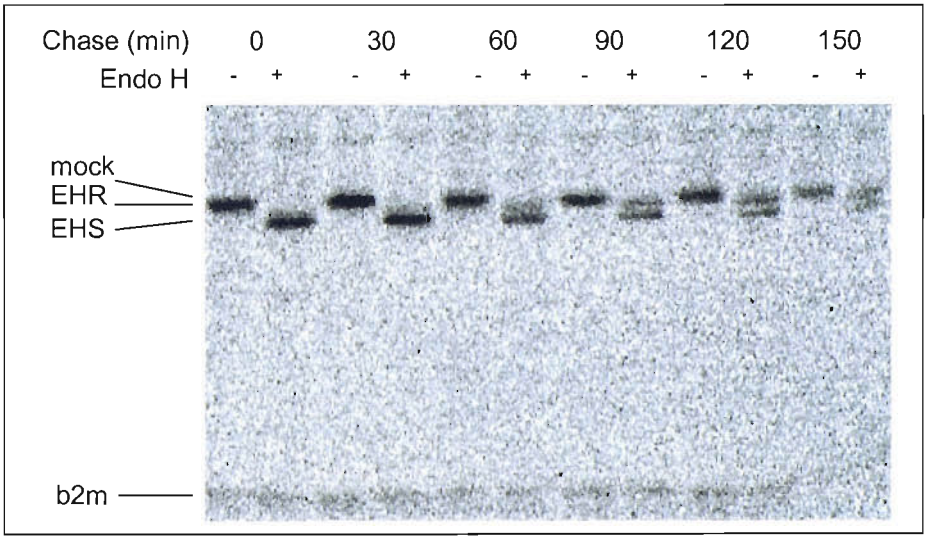
c) TG15 BF2myc*15 TG15 BF2myc*19v1



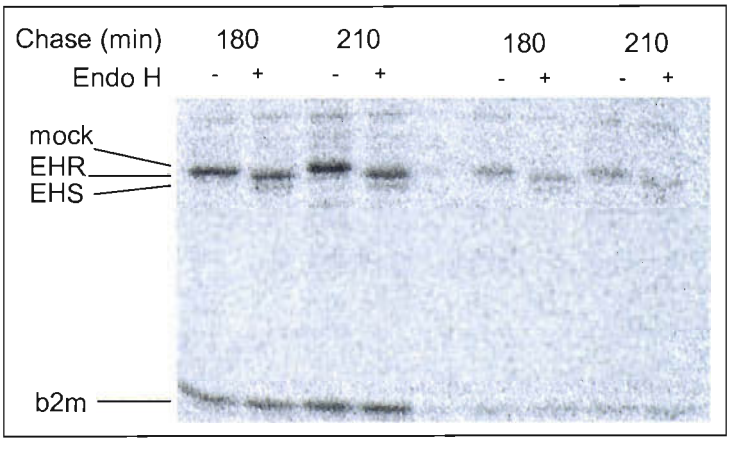
d) TG15 BF2myc*19v1 G126D + R220Q clone C3



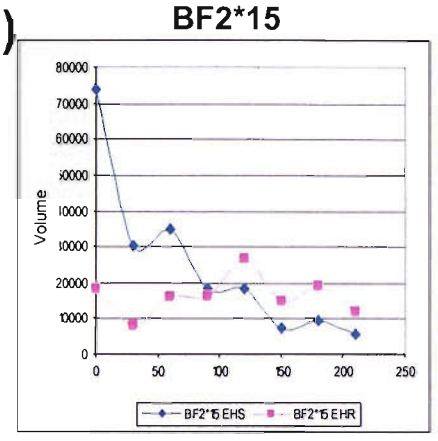
e) TG15 BF2myc*19v1 R111S + Y113D clone C1

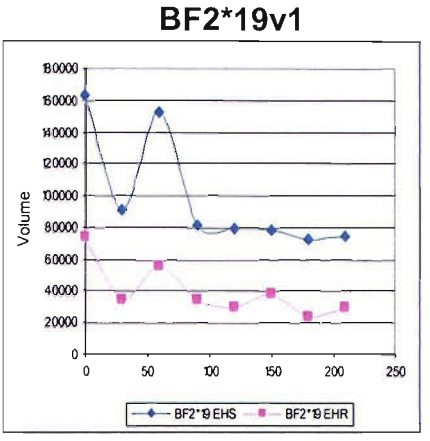


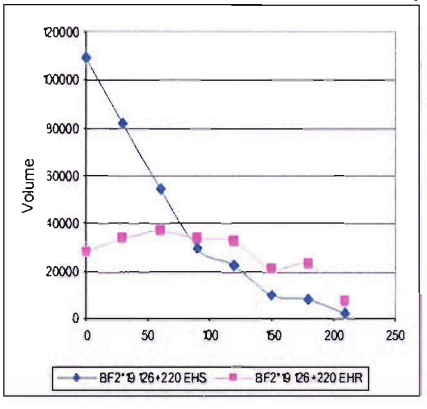
f)

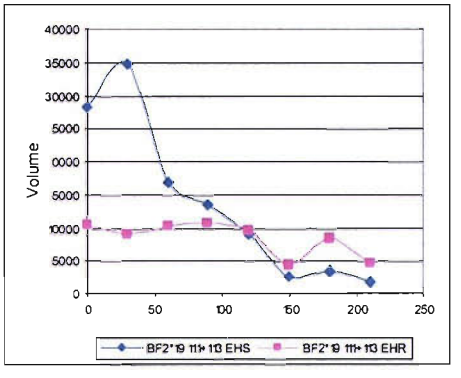


g)

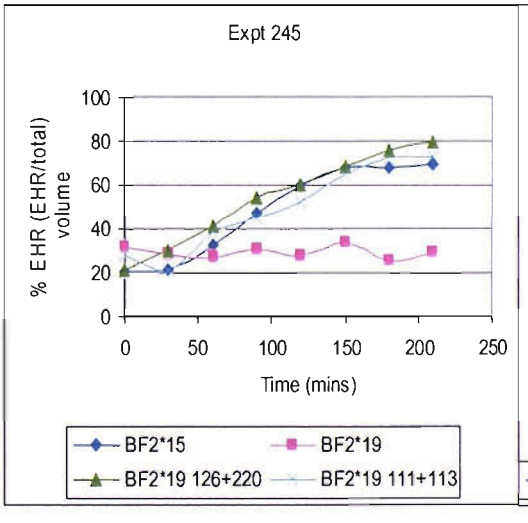

BF2*15
Volume
BF2*15 EHS
BF2*15 EHR


BF2*19v1
Volume
BF2*19 EHS
BF2*19 EHR


BF2*19v1 G126D+R220Q
Volume
BF2*19 126+220 EHS
BF2*19 126+220 EHR


BF2*19v1 R111S+Y113D
Volume
BF2*19 111+113 EHS
BF2*19 111+113 EHR

h)


Expt 245
% EHR (EHR/total) volume
Time (mins)
BF2*15
BF2*19
BF2*19 126+220
BF2*19 111+113

i)

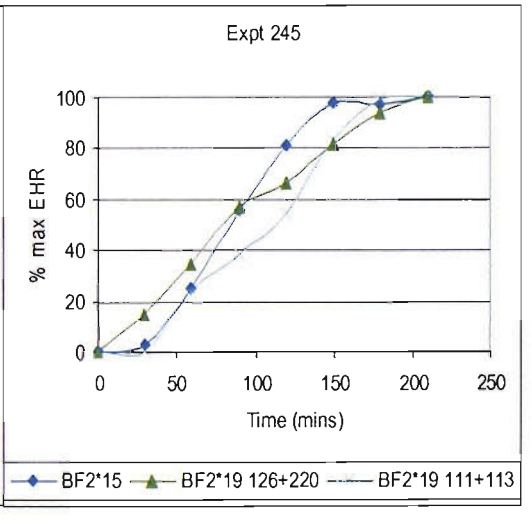

Expt 245
% max EHR
Time (mins)
BF2*15
BF2*19 126+220
BF2*19 111+113

Figure 6.10 Comparison of BF2myc*15, BF2myc*19v1, BF2myc*19v1 G126D+R220Q and BF2myc*19v1 R111S+Y113D maturation rates in TG15 cells (Experiment 245)

a-f) The indicated clones were pulse labelled for five minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis.

g) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

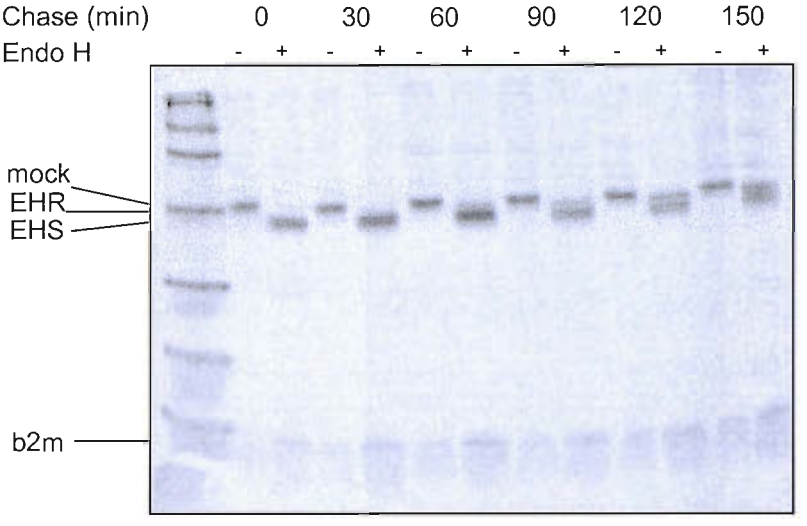
h) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

i) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion). Note that the BF2*19v1 data was omitted.

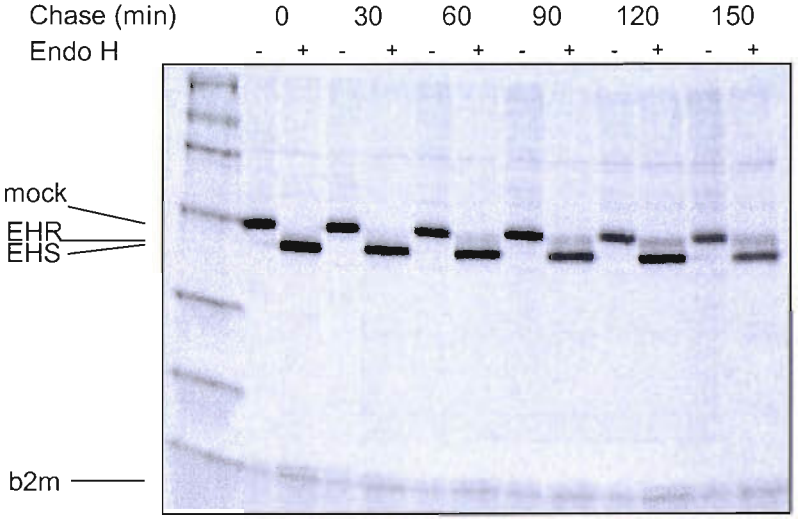
Experiment number, Figure	BF2myc clones	Length of pulse	Initiation of chase	Length of chase
Expt 231 Fig 6.4	BF2myc*15 C3 BF2myc*19 C1	5 minutes	380 g, one minute	120 minutes
Expt 233 Fig 6.5	BF2myc*15 C3 BF2myc*19 C1	5 minutes	380 g, one minute	210 minutes
Expt 236 Fig 6.6	BF2myc*15 C7 BF2myc*19 C9	5 minutes	380 g, one minute	210 minutes
Expt 240 Fig 6.12	BF2myc*19 G126D C3 BF2myc*19 R220Q C6	5 minutes	380 g, one minute	210 minutes
Expt 241 Fig 6.14	BF2myc*19 G126D+R220Q C3 BF2myc*19 R111S+Y113D C1	5 minutes	380 g, one minute	210 minutes
Expt 245 Fig 6.10	BF2myc*15 C7 BF2myc*19 C9 BF2myc*19 G126D+R220Q C3 BF2myc*19 R111S+Y113D C1	5 minutes	380 g, one minute	210 minutes
Expt 277 Fig 6.11	BF2myc*15 C3 BF2myc*19 C1 BF2myc*19 G126D+R220Q C1	30 minutes	10 fold dilution in non-radioactive media	240 minutes
Expt 282 Fig 6.18	BF2myc*15 C3 BF2myc*19 C1 BF2myc*19 Tapasin*12 A9	20 minutes	320g, three minutes	210 minutes
Expt 285 Fig 6.13	BF2myc*15 C3 BF2myc*19 C1 BF2myc*15 D126G+Q220R A1 BF2myc*19 G126D+R220Q C1	20 minutes	320g, three minutes	210 minutes
Expt 290 Fig 6.15	BF2myc*15 C3 BF2myc*19 C1 BF2myc*15 D126G+Q220R A1 BF2myc*15 S111R+D113Y C1 BF2myc*15 4x (D126G+Q220R S111R+D113Y) C1	20 minutes	320g, three minutes	210 minutes
Expt 291 Fig 6.19	BF2myc*15 C7 BF2myc*19 C9 BF2myc*19 Tapasin*12 A10	20 minutes	320g, three minutes	210 minutes
Expt 292 Fig 6.8	BF2myc*15 C7 BF2myc*19 C9	20 minutes	320g, three minutes	120 minutes
Expt 293 Fig 6.16	BF2myc*15 C7 BF2myc*19 C9 BF2myc*15 D126G+Q220R A1 BF2myc*19 G126D+R220Q C1	30 minutes	320g, three minutes	120 minutes

Table 6.4 Details of experimental conditions

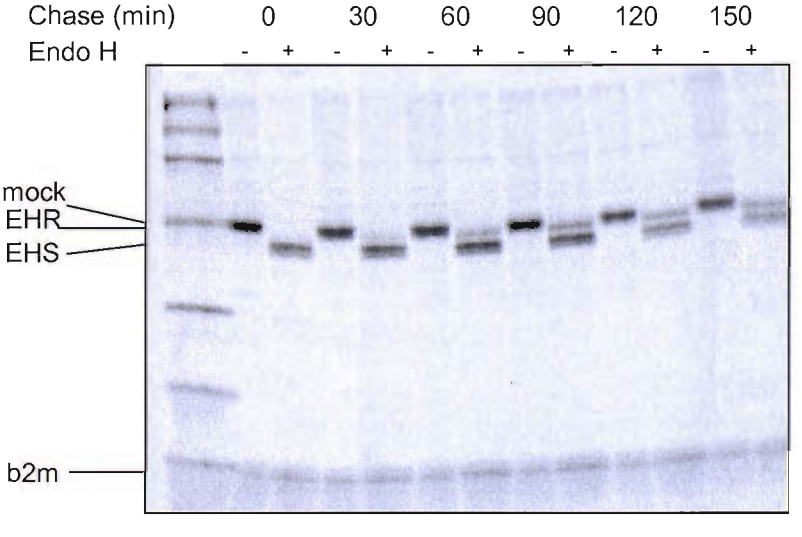
a) TG15 BF2myc*15 clone C3



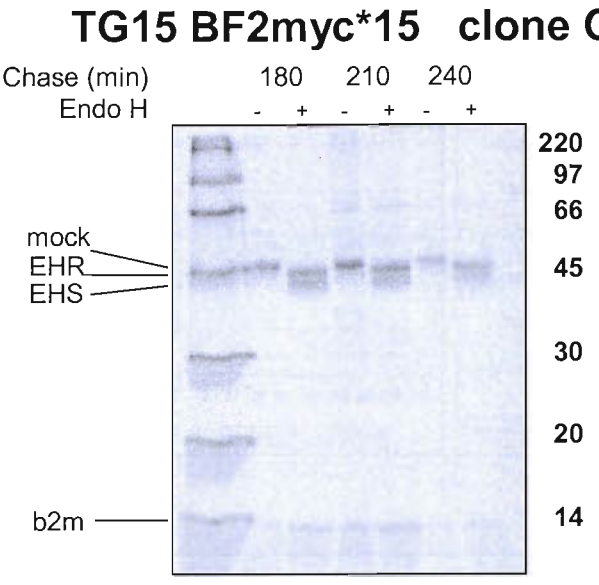
b) TG15 BF2myc*19v1 clone C1



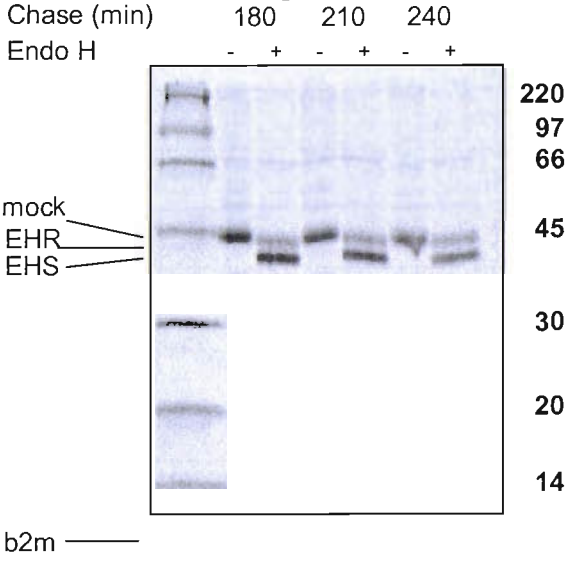
c) TG15 BF2myc*19v1 G126D+R220Q clone C1



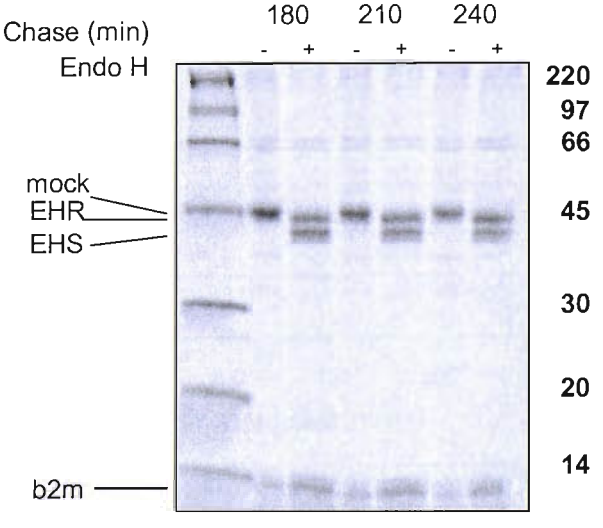
d)


TG15 BF2myc*15 clone C3
Chase (min)
180 210 240
Endo H
- + - + - +
mock
EHR
EHS
b2m
220
97
66
45
30
20
14

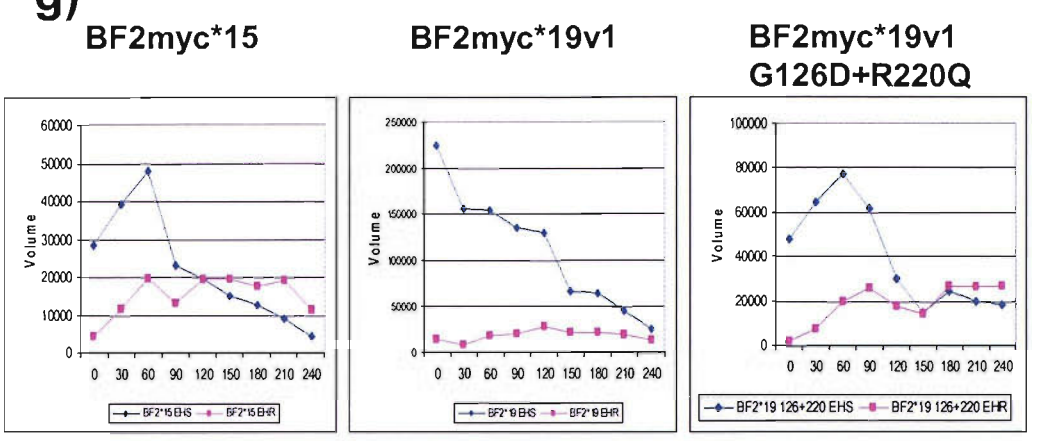
e)


TG15 BF2myc*19v1 clone C1
Chase (min)
180 210 240
Endo H
- + - + - +
mock
EHR
EHS
b2m
220
97
66
45
30
20
14

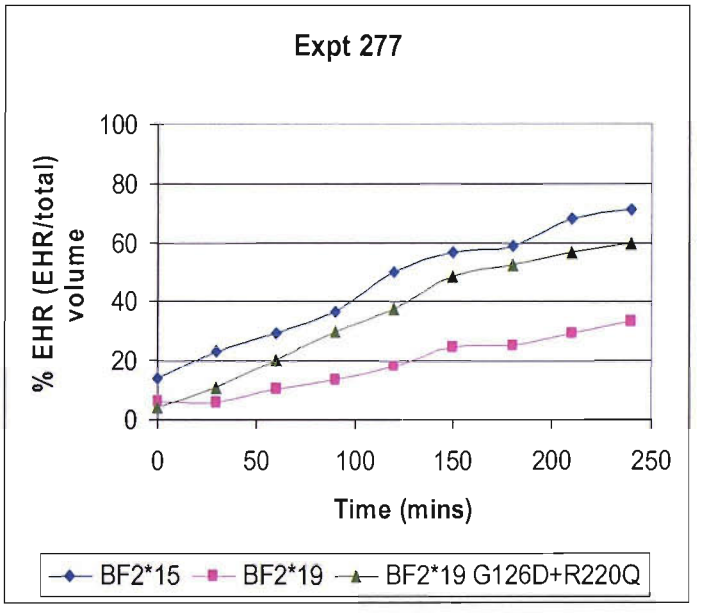
f)


TG15 BF2myc*19v1 G126D+R220Q clone C1
Chase (min)
180 210 240
Endo H
- + - + - +
mock
EHR
EHS
b2m
220
97
66
45
30
20
14

g)


BF2myc*15
BF2myc*19v1
BF2myc*19v1 G126D+R220Q
Volume
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR
BF2*19 126+220 EHS
BF2*19 126+220 EHR

h)


Expt 277
% EHR (EHR/total) volume
Time (mins)
BF2*15
BF2*19
BF2*19 G126D+R220Q

i)

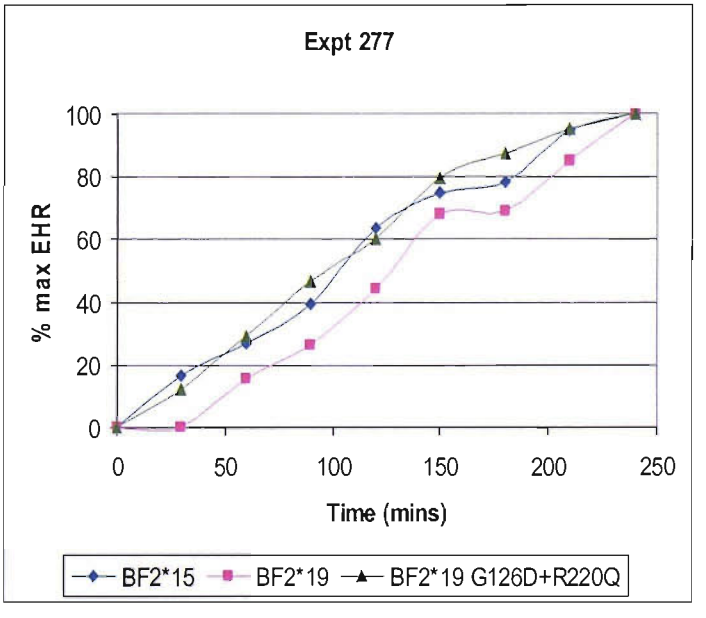

Expt 277
% max EHR
Time (mins)
BF2*15
BF2*19
BF2*19 G126D+R220Q

Figure 6.11 Comparison of BF2myc*15, BF2myc*19v1 and BF2myc*19v1 G126D+R200Q maturation rates in TG15 cells (Experiment 277)

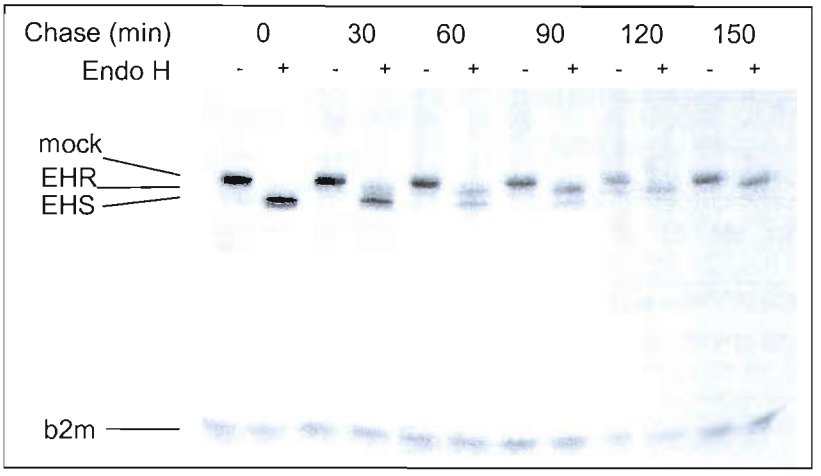
a-f) The indicated clones were pulse labelled for 20 minutes and chased for 240 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was digested with endoglycosidase H, apart from one aliquot which was mock digested, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis. In the left hand lane on each gel is a ^{14}C molecular weight marker. The molecular weights are indicated in kDa.

g) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

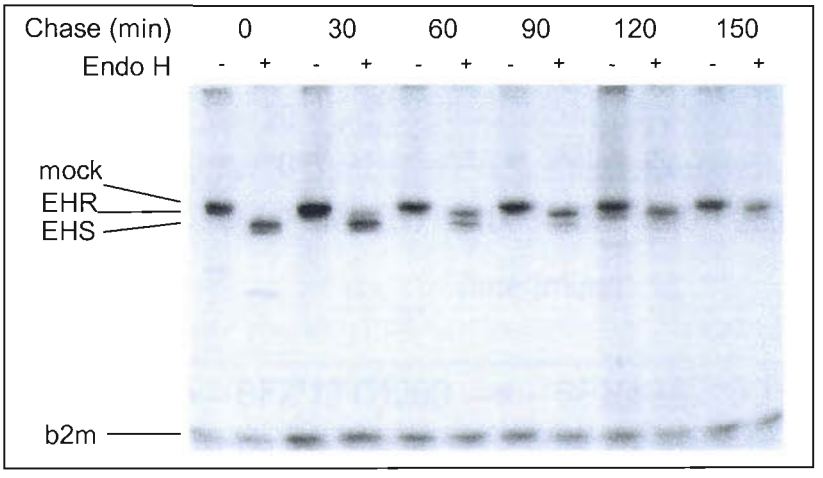
h) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

i) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

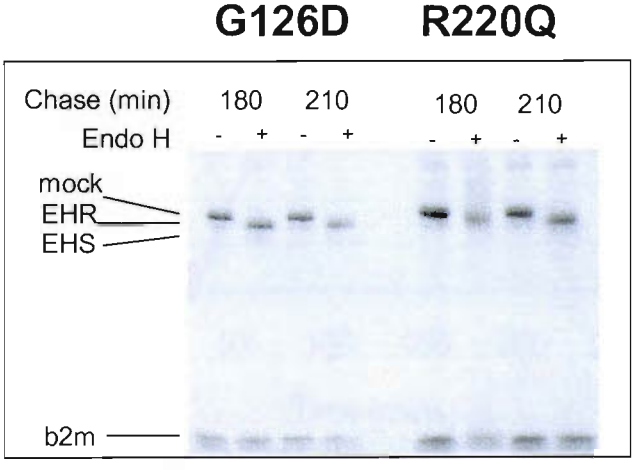
a) TG15 BF2myc*19v1 G126D clone C3



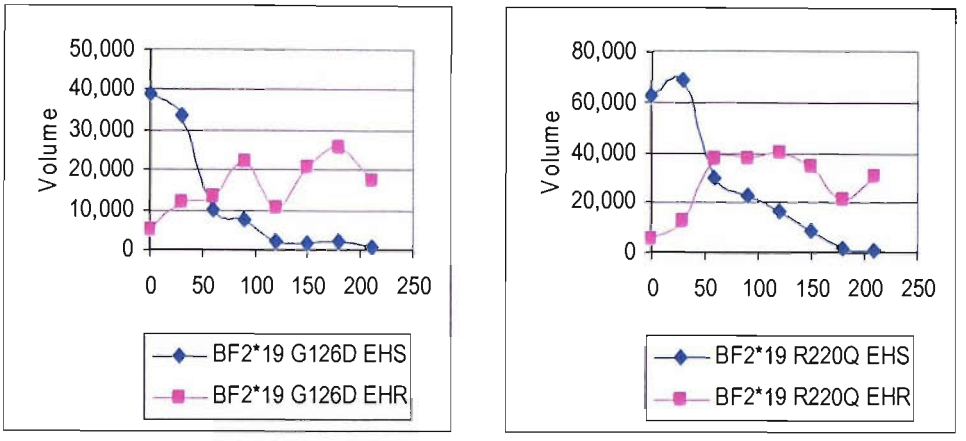
b) TG15 BF2myc*19v1 R220Q clone C6



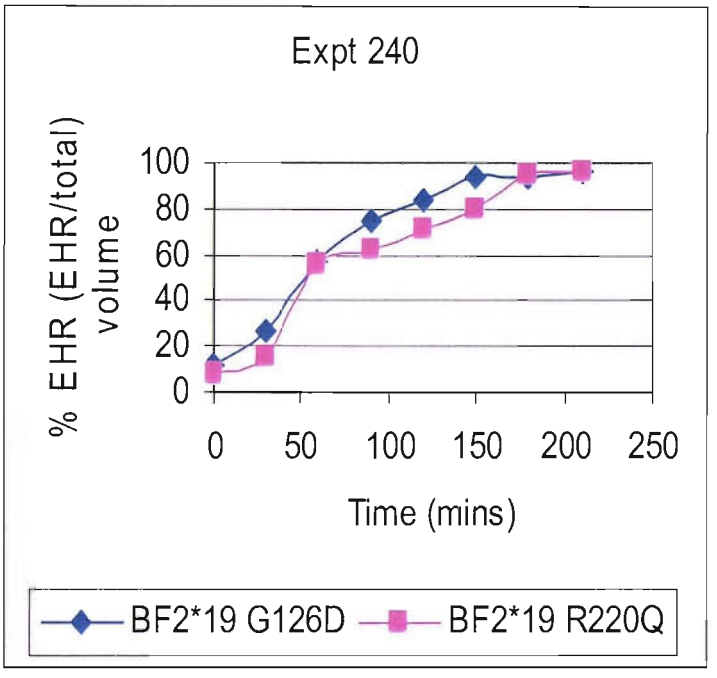
c)



d)


BF2*19v1 G126D
BF2*19v1 R220Q
Volume
50,000
40,000
30,000
20,000
10,000
0
0 50 100 150 200 250
BF2*19 G126D EHS
BF2*19 G126D EHR
Volume
80,000
60,000
40,000
20,000
0
0 50 100 150 200 250
BF2*19 R220Q EHS
BF2*19 R220Q EHR

e)


Expt 240
% EHR (EHR/total) volume
100
80
60
40
20
0
0 50 100 150 200 250
Time (mins)
BF2*19 G126D
BF2*19 R220Q

f)

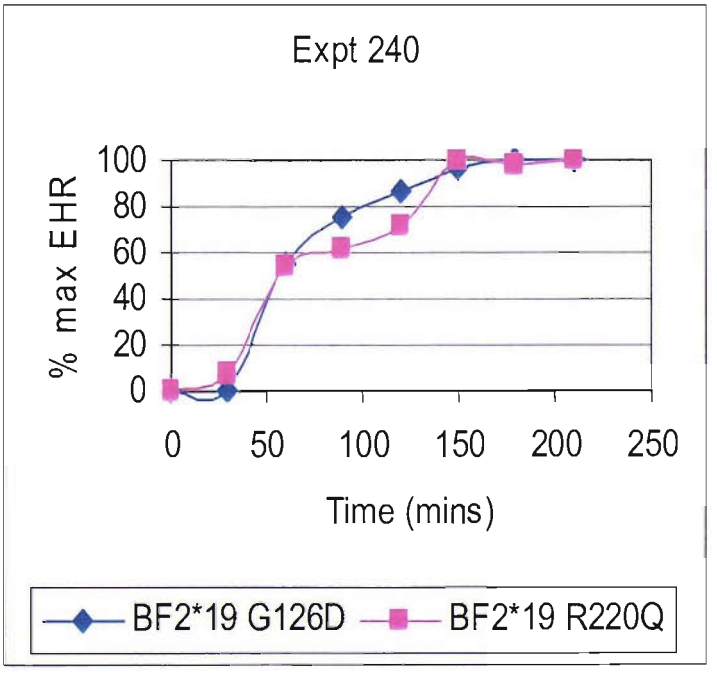

Expt 240
% max EHR
100
80
60
40
20
0
0 50 100 150 200 250
Time (mins)
BF2*19 G126D
BF2*19 R220Q

Figure 6.12 Comparison of BF2myc*19v1 G126D and BF2myc*19v1 R220Q maturation rates in TG15 cells (Experiment 240)

a-c) The indicated clones were pulse labelled for five minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with endoglycosidase H, before samples were resolved by SDS PAGE, fixed and exposed to film analysis.

d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

In this experiment, quantification was performed by densitometric analysis using an electronic image of a film exposure.

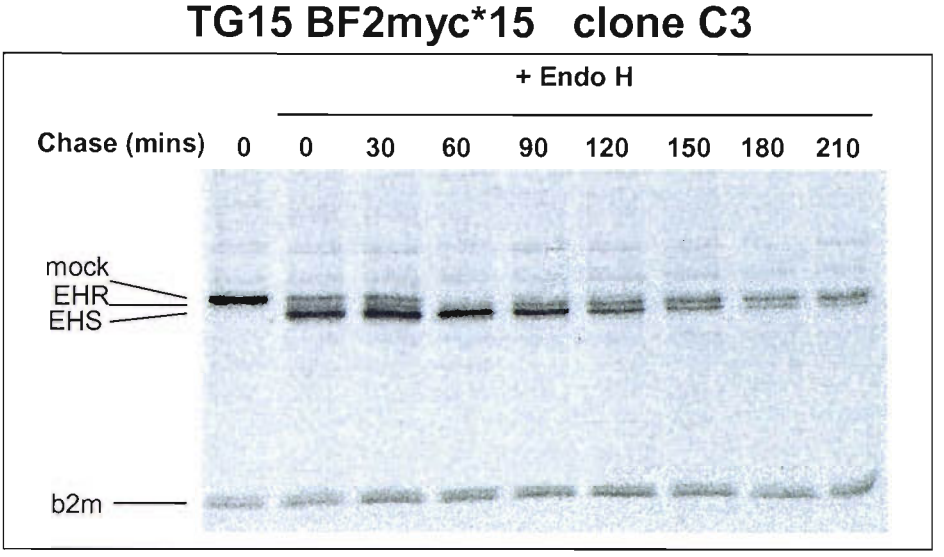
In this experiment (figure 6.13, experiment 285), BF2myc*15, BF2myc*19v1 G126D&R220Q and BF2myc*15 D126G&Q220R molecules all matured at similar rates and to similar extents, which was greater than for BF2myc*19v1 molecules. This experiment: firstly confirmed the differential maturation of BF2myc*15 and BF2myc*19v1 molecules; secondly, confirmed that BF2myc*19v1 G126D&R220Q molecules matured in a similar fashion to that of BF2myc*15 molecules; and thirdly showed that when the reciprocal mutations were made to generate the BF2myc*15 D126G&Q220R mutant, this did not adversely affect the maturation of the BF2myc molecules.

It was of interest to examine the influence that residues 111 and 113, which are located within the peptide-binding groove, have on BF2myc maturation, as the equivalent residues (residues 114 and 116) in mammalian class 1 molecules appear to determine tapasin-dependency (Williams et al., 2002; Park et al., 2003). Therefore the residues present at these positions in the BF2myc*19v1 molecule (arginine 111 and tyrosine 113) were replaced with those residues present in the BF2myc*15 molecule (serine 111 and aspartic acid 113), with the mutant molecule being labelled BF2myc*19v1 R111S&Y113D. Transfectants of TG15 cells expressing such constructs were generated, and were compared to a transfectant expressing BF2myc*19v1 G126D&R220Q molecules by pulse/chase analysis (figure 6.14, experiment 241).

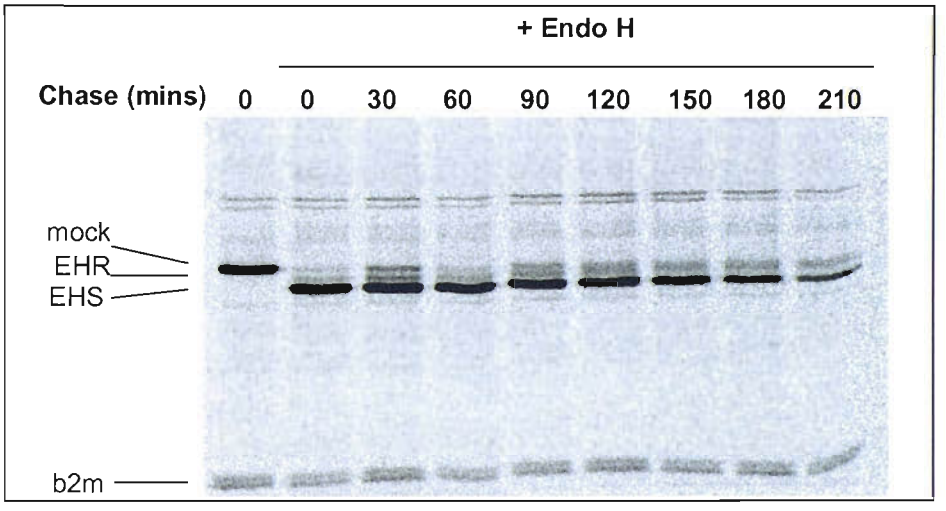
Both BF2myc*19v1 mutants displayed virtually identical maturation characteristics, which are similar to the characteristics expected of BF2myc*15 molecules (experiments 231, 233 and 236). Similar results were also observed in experiment 245 (figure 6.10), where the BF2myc*19v1 R111S&Y113D demonstrated a similar rate and extent of maturation as BF2myc*15 molecules that were included in the same experiment. These two experiments suggest that substituting residues from the BF2*15 molecule into the BF2myc*19v1 molecule at positions 111 and 113 improved BF2myc*19v1 maturation in a similar fashion to the mutations to positions 126 and 220.

It was also of interest to determine whether making the reciprocal mutations to the BF2myc*15 molecule at positions 111 and 113 affected their maturation: therefore the residues present at positions 111 and 113 in the BF2myc*15 molecule (serine 111 and aspartic acid 113) were mutated to those residues present in the BF2myc*19v1 molecule (arginine 111 and tyrosine 113); with the resultant expression construct being labelled BF2myc*15 S111R&D113Y. In light of the unaltered maturation of the BF2myc*15 D126G&Q220R mutant, an additional BF2myc*15 mutant was produced, which combined the mutations to positions 126 and 220 with the mutations to positions 111 and 113, with the resultant expression construct being labelled BF2myc*15 S111R&D113Y & D126G&Q220R, or "4x" for simplicity.

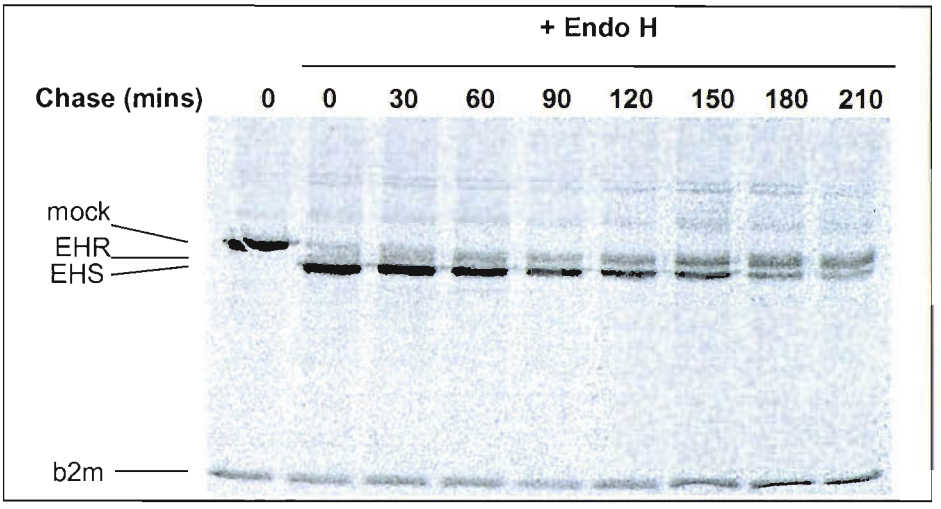
a)


TG15 BF2myc*15 clone C3
+ Endo H
Chase (mins)
0
0
30
60
90
120
150
180
210
mock
EHR
EHS
b2m

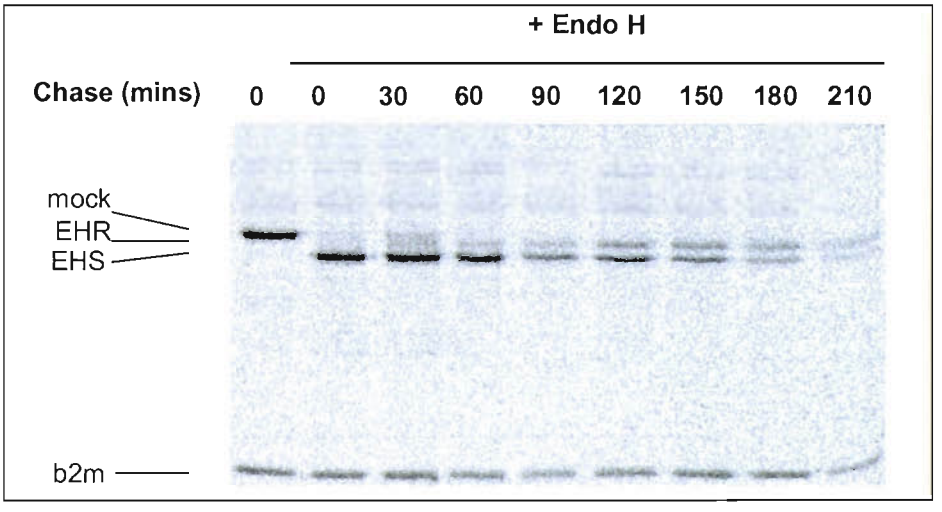
b)


TG15 BF2myc*19v1 clone C1
+ Endo H
Chase (mins)
0
0
30
60
90
120
150
180
210
mock
EHR
EHS
b2m

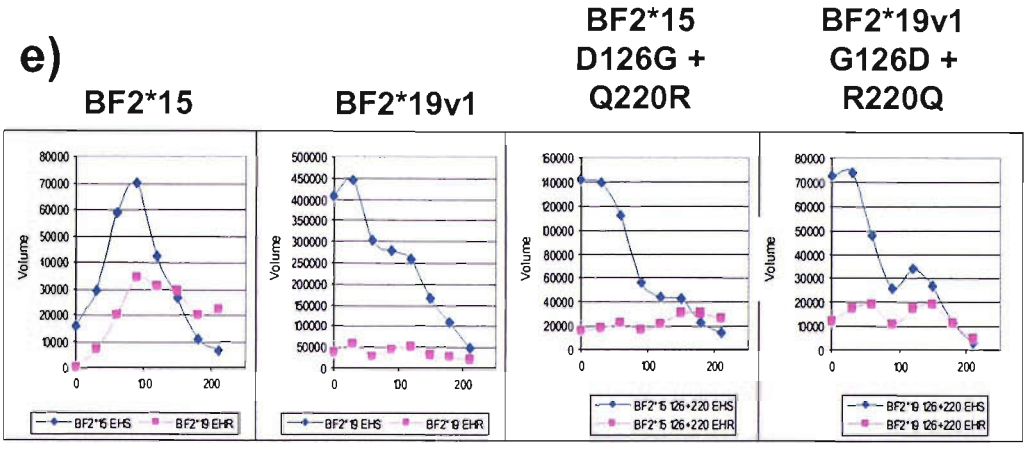
c) TG15 BF2myc*15 D126G + Q220R clone A1



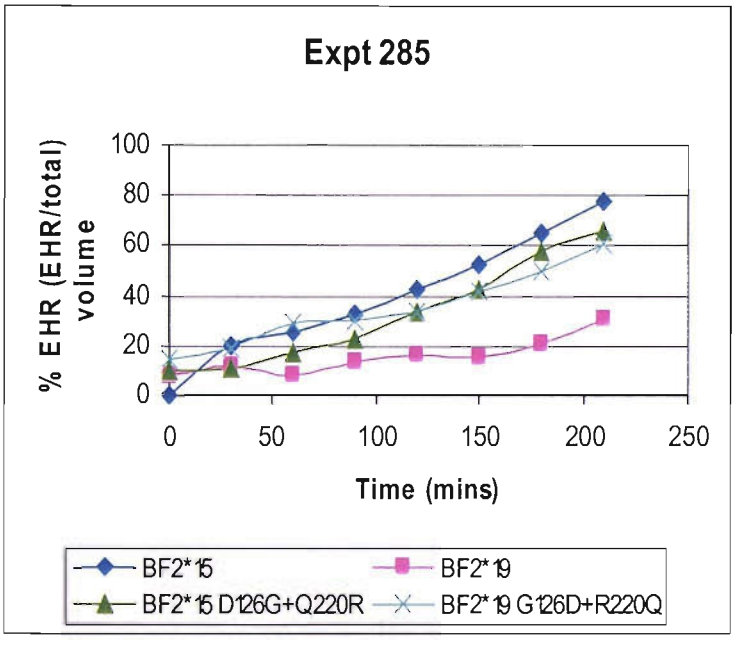
d) TG15 BF2myc*19v1 G126D + R220Q clone C1



e)


BF2*15
BF2*19v1
BF2*15 D126G + Q220R
BF2*19v1 G126D + R220Q
Volume
BF2*15 EHS
BF2*19 EHR
BF2*15 126+220 EHS
BF2*15 126+220 EHR
BF2*19 126+220 EHS
BF2*19 126+220 EHR

f)


Expt 285
% EHR (EHR/total) volume
Time (mins)
BF2*15
BF2*19
BF2*15 D126G+Q220R
BF2*19 G126D+R220Q

g)

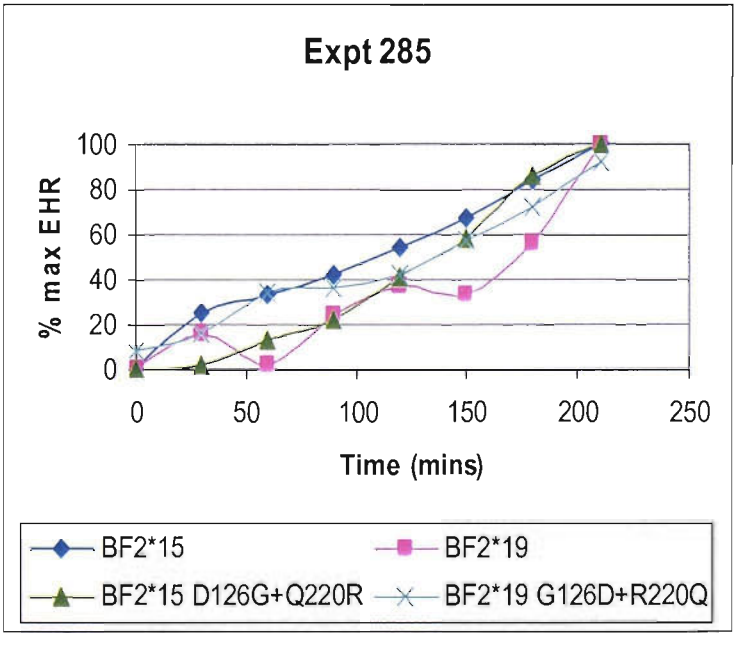

Expt 285
% max EHR
Time (mins)
BF2*15
BF2*19
BF2*15 D126G+Q220R
BF2*19 G126D+R220Q

Figure 6.13 Comparison of BF2myc*15, BF2myc*19v1, BF2myc*15 D126G+Q220R and BF2myc*19v1 G126D+R200Q maturation rates in TG15 cells (Experiment 285)

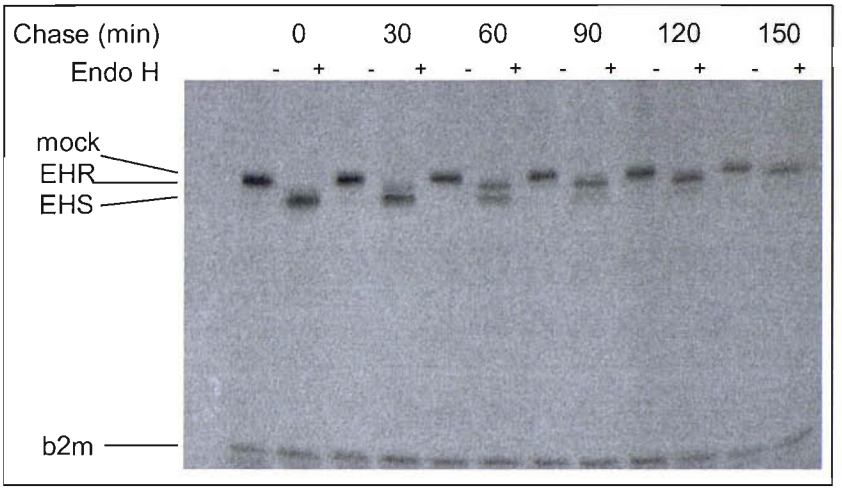
a-d) The indicated clones were pulse labelled for 20 minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc tagged molecules were immunoprecipitated and endoglycosidase H digested, before being run on SDS PAGE gels, fixed and subjected to phosphorimager analysis.

e) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

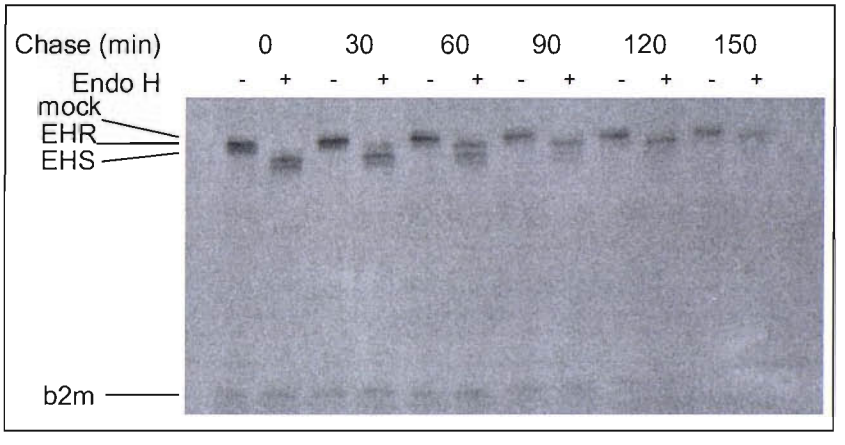
f) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

g) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

a) TG15 BF2myc*19v1 R111S + Y113D clone C1

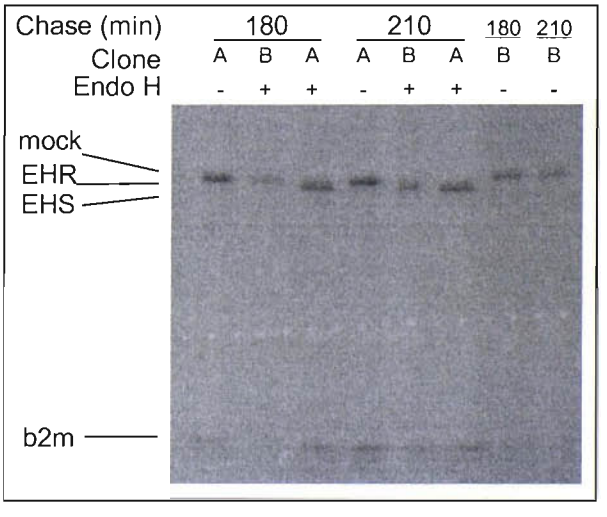


b) TG15 BF2myc*19v1 G126D + R220Q clone C3

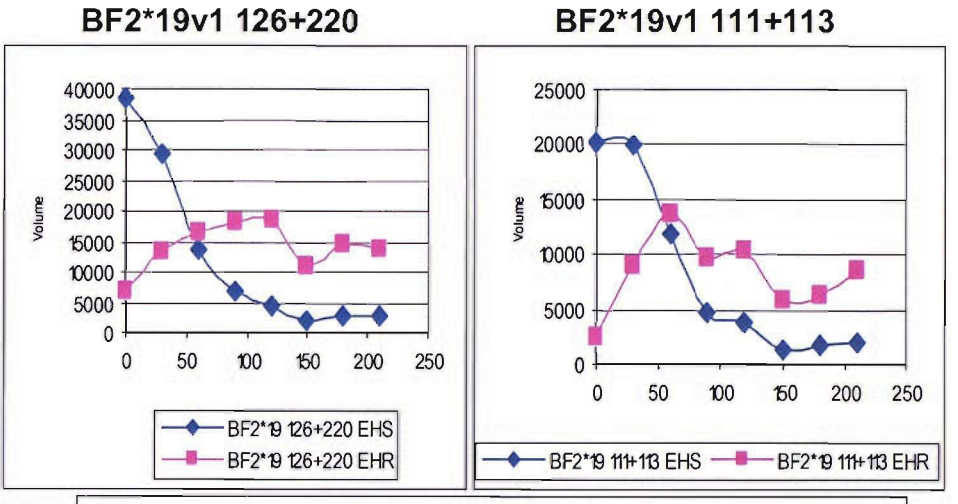


c)

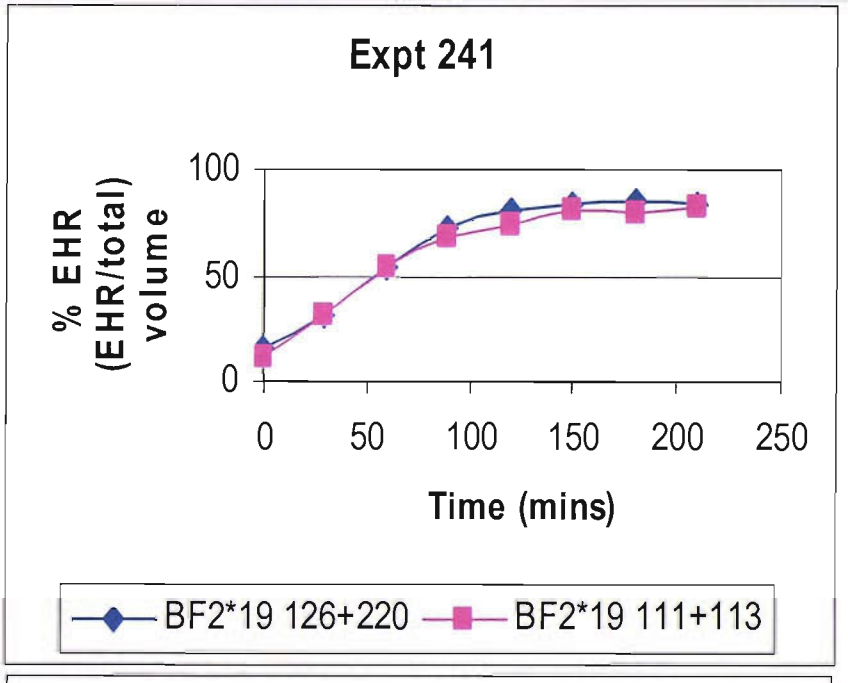
Clone A = BF2myc*19v1 R111S + Y113D clone C1
Clone B = BF2myc*19 v1 G126D + R220Q clone C3



d)


BF2*19v1 126+220
BF2*19v1 111+113
Volume
40000
35000
30000
25000
20000
15000
10000
5000
0
0 50 100 150 200 250
BF2*19 126+220 EHS
BF2*19 126+220 EHR
25000
20000
15000
10000
5000
0
BF2*19 111+113 EHS
BF2*19 111+113 EHR

e)


Expt 241
% EHR (EHR/total) volume
100
50
0
0 50 100 150 200 250
Time (mins)
BF2*19 126+220
BF2*19 111+113

f)

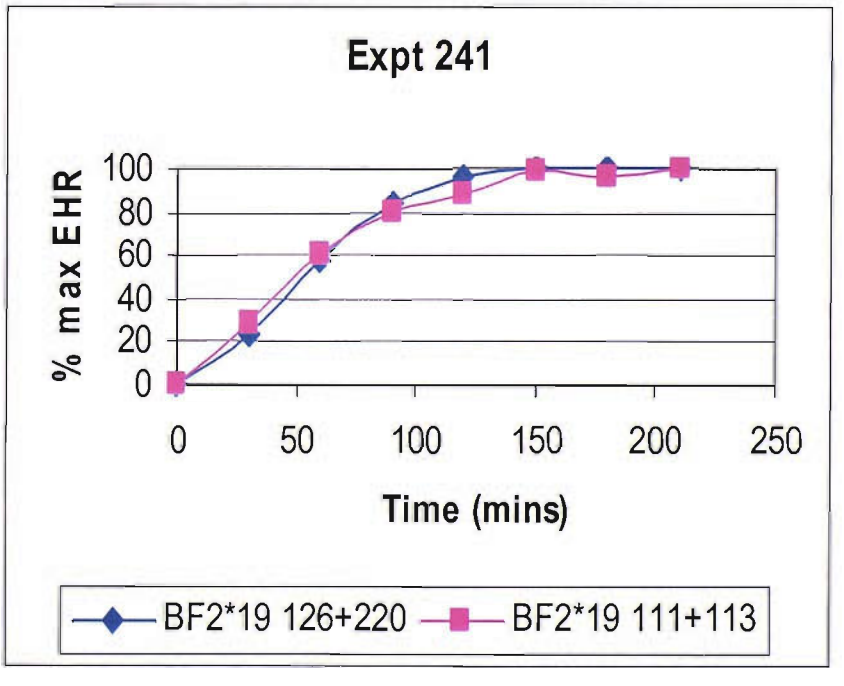

Expt 241
% max EHR
100
80
60
40
20
0
0 50 100 150 200 250
Time (mins)
BF2*19 126+220
BF2*19 111+113

Figure 6.14 Comparison of BF2myc*19v1 mutants maturation rates in TG15 cells (Experiment 241)

a-c) The indicated clones were pulse labelled for five minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was split into two samples, one was mock digested, the other was digested with Endoglycosidase H, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis.

d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

Transfectants expressing either BF2myc*15, BF2myc*19v1, BF2myc*15 D126G&Q220R, BF2myc*15 S111R&D113Y or BF2myc*15 4x were compared by pulse/chase analysis (figure 6.15, experiment 290). In this experiment all three of the BF2myc*15 mutants demonstrated similar rates and extents of maturation as wild-type BF2myc*15 molecules, which was greater than was evident for BF2myc*19v1 molecules. This suggests that mutations made to either positions 126 and 220, or positions 111 and 113, or the combination of these mutations, did not affect the maturation of the BF2myc*15 molecule.

Having demonstrated that the extent and rate of maturation of BF2myc*19v1 molecules can be increased by replacing the amino acids at positions 126 and 220 (or 111 and 113) with the equivalent residues from the BF2myc*15 molecule, it was important to determine that this was not a consequence of a mutation similar to that described for the HLA-A*0201 T134K mutant (Lewis et al., 1996; Lewis and Elliott, 1998). T134K mutants do not associate with the PLC and are rapidly released from the ER loaded with sub-optimal peptides. Therefore the thermal-stability of BF2myc*19v1 G126D&R220Q molecules was compared with that of BF2myc*15, BF2myc*19v1 and BF2myc*15 D126G&Q220R molecules using the pulse/chase thermal-stability assay and overnight immunoprecipitation step that was described previously (figure 6.16, table 6.5, experiment 293).

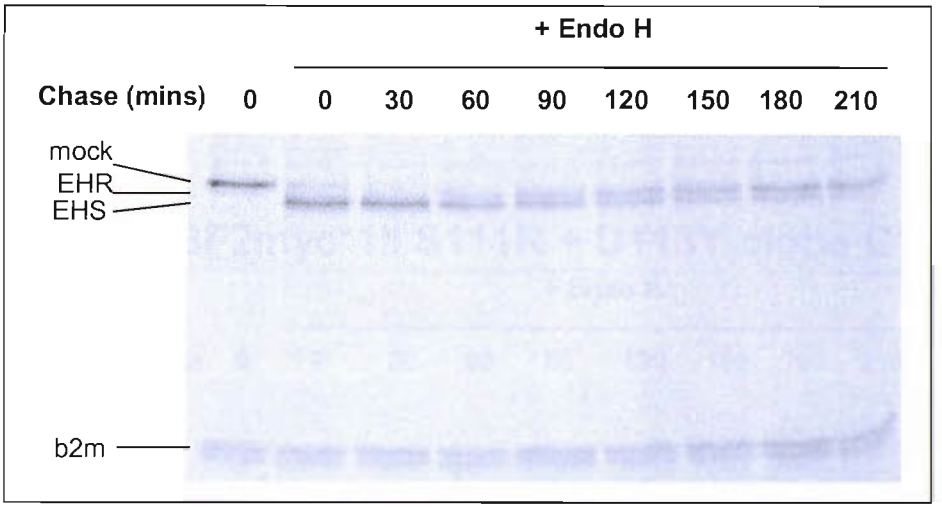
This experiment allows the following points to be made. Firstly, for all class I alleles there was a time-dependent increase in the amount of co-precipitating β_2m molecules, suggesting that the mutations did not prevent the class I molecules from optimising their peptide repertoires over time.

Additionally, the experiment confirmed that a smaller proportion of BF2myc*19v1 molecules co-precipitated β_2m than was evident for BF2myc*15 molecules, which is likely to reflect dissociation of sub-optimally loaded BF2myc*19v1 heterodimers during the immunoprecipitation procedure.

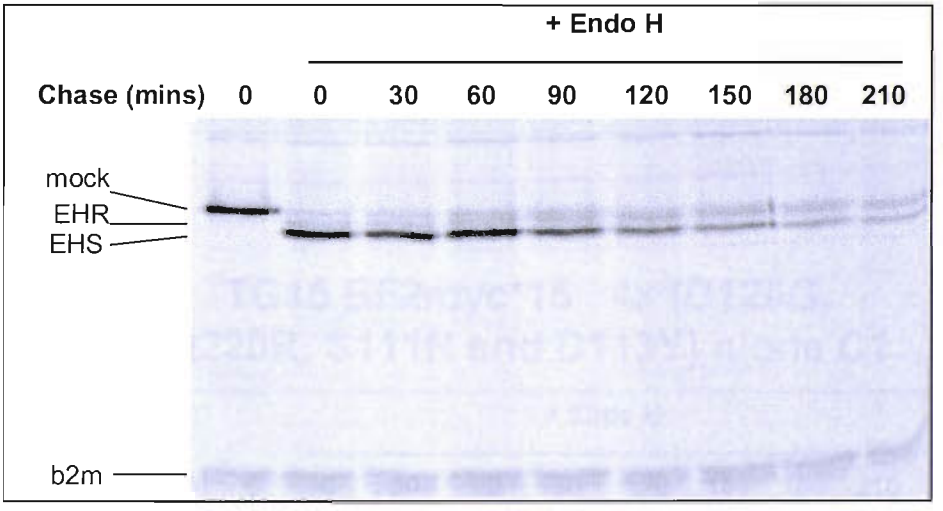
The mutations made to the BF2myc*15 allele at positions 126 and 220, which did not affect maturation of the molecule, appeared to modestly decrease the amount of β_2m molecules that were co-precipitated with BF2myc*15 molecules. This suggests that their peptide repertoire was slightly inferior in quality than that of BF2myc*15 molecules.

The BF2myc*19v1 G126D&R220Q mutant, which demonstrated improved maturation, co-precipitated a slightly larger amount of β_2m molecules than the wild-type BF2myc*19v1 molecules. This suggests the peptide repertoire that was loaded was modestly improved as a consequence of the mutation. However, the HC: β_2m ratio remained much lower than for either BF2myc*15 or BF2myc*15 D126G&Q220R molecules, suggesting a substantial proportion of BF2myc*19v1 G126D&R220Q molecules were loaded with sub-optimal peptides despite the mutation.

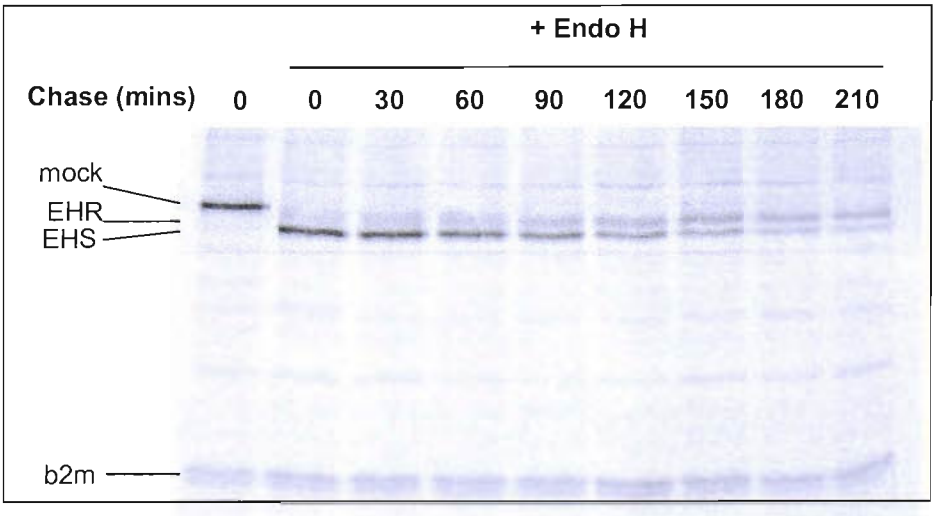
a) TG15 BF2myc*15 clone C3



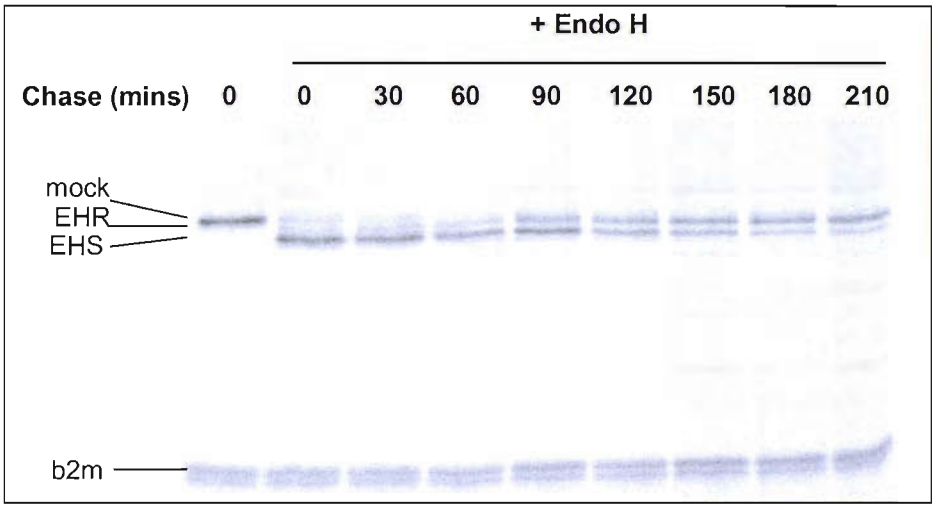
b) TG15 BF2myc*19v1 clone C1



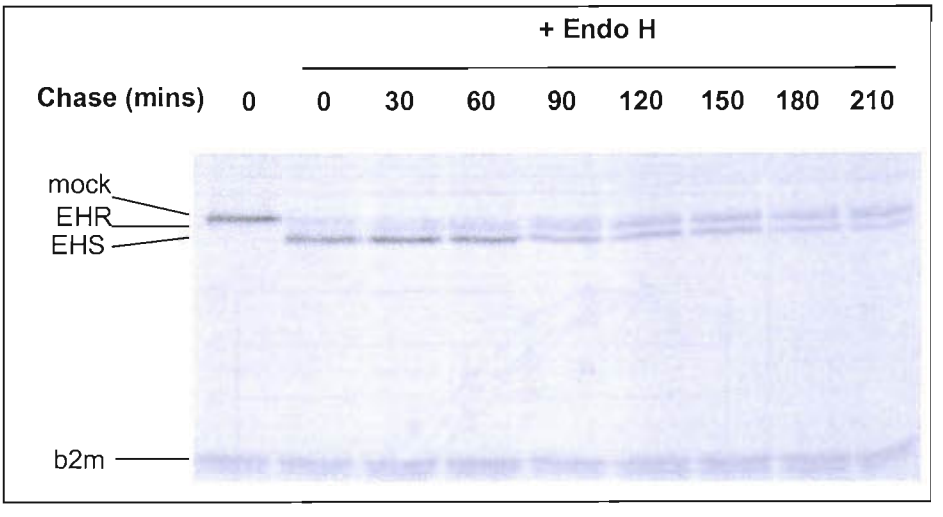
c) TG15 BF2myc*15 D126G + Q220R clone A1



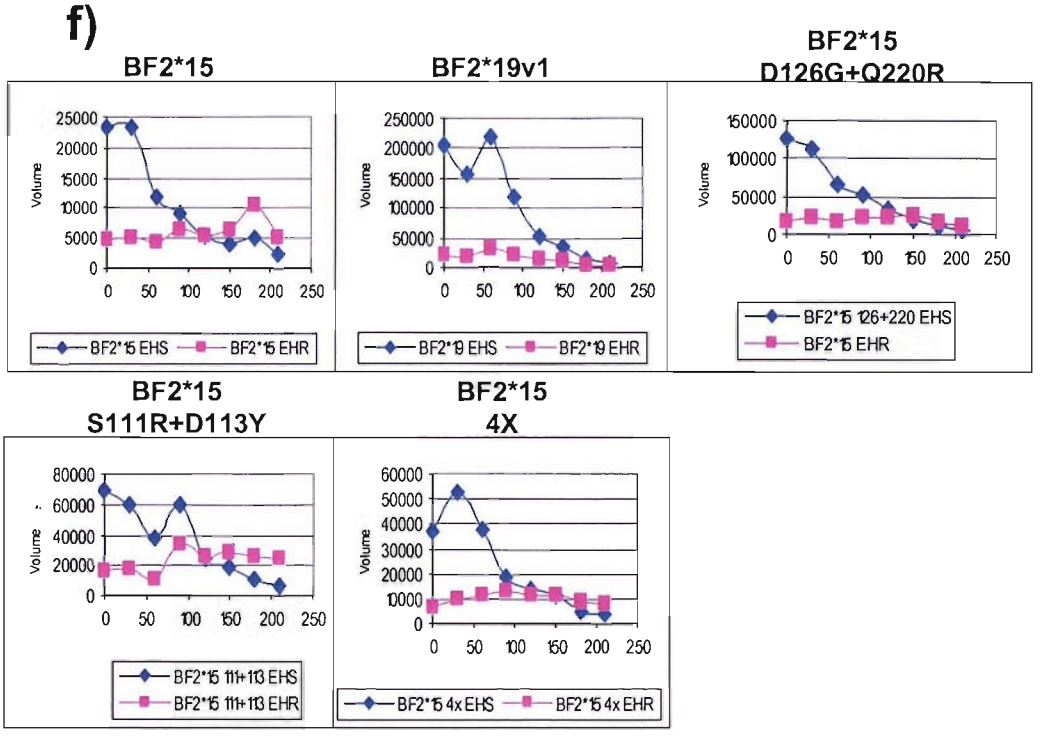
d) TG15 BF2myc*15 S111R + D113Y clone C1



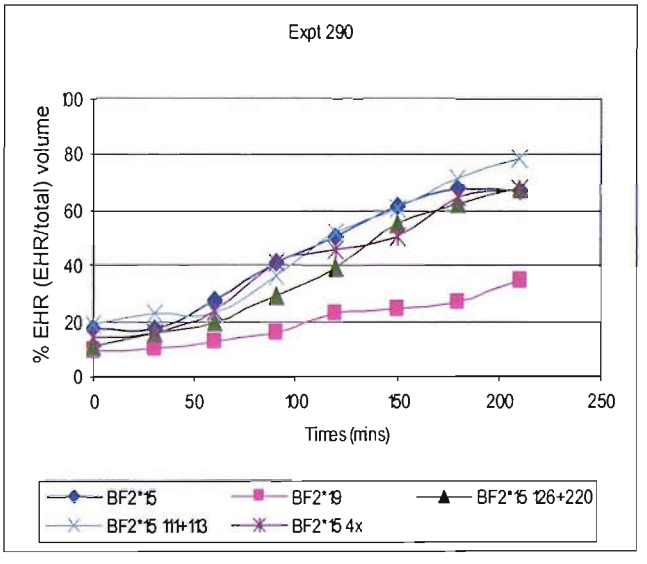
e) TG15 BF2myc*15 4x (D126G, Q220R, S111R and D113Y) clone C1



f)


BF2*15
BF2*19v1
BF2*15 D126G+Q220R
Volume
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR
BF2*15 126+220 EHS
BF2*15 EHR
BF2*15 S111R+D113Y
BF2*15 4X
BF2*15 111+113 EHS
BF2*15 111+113 EHR
BF2*15 4x EHS
BF2*15 4x EHR

g)


Expt 290
% EHR (EHR/total) volume
Times (mins)
BF2*15
BF2*19
BF2*15 126+220
BF2*15 111+113
BF2*15 4x

h)

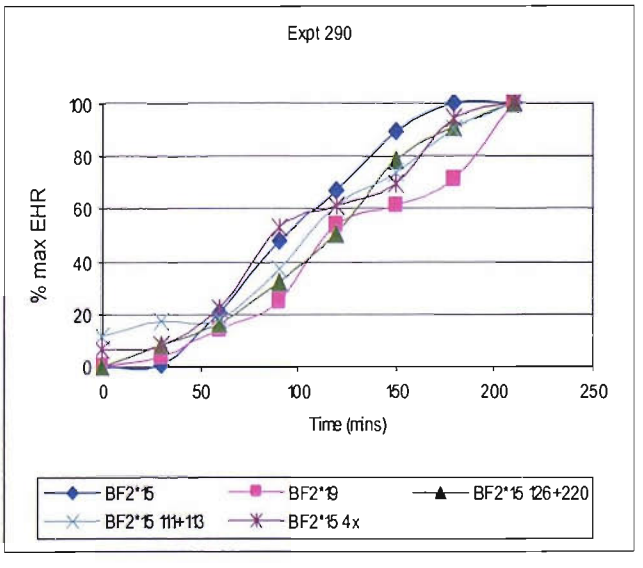

Expt 290
% max EHR
Time (mins)
BF2*15
BF2*19
BF2*15 126+220
BF2*15 111+113
BF2*15 4x

Figure 6.15 Comparison of BF2myc*15, BF2myc*19v1, BF2myc*15 D126G+Q220R, BF2myc*15 S111R+D113Y and BF2myc*15 4x (G126D, R200Q, S111R and D113Y) maturation rates in TG15 cells (Experiment 290)

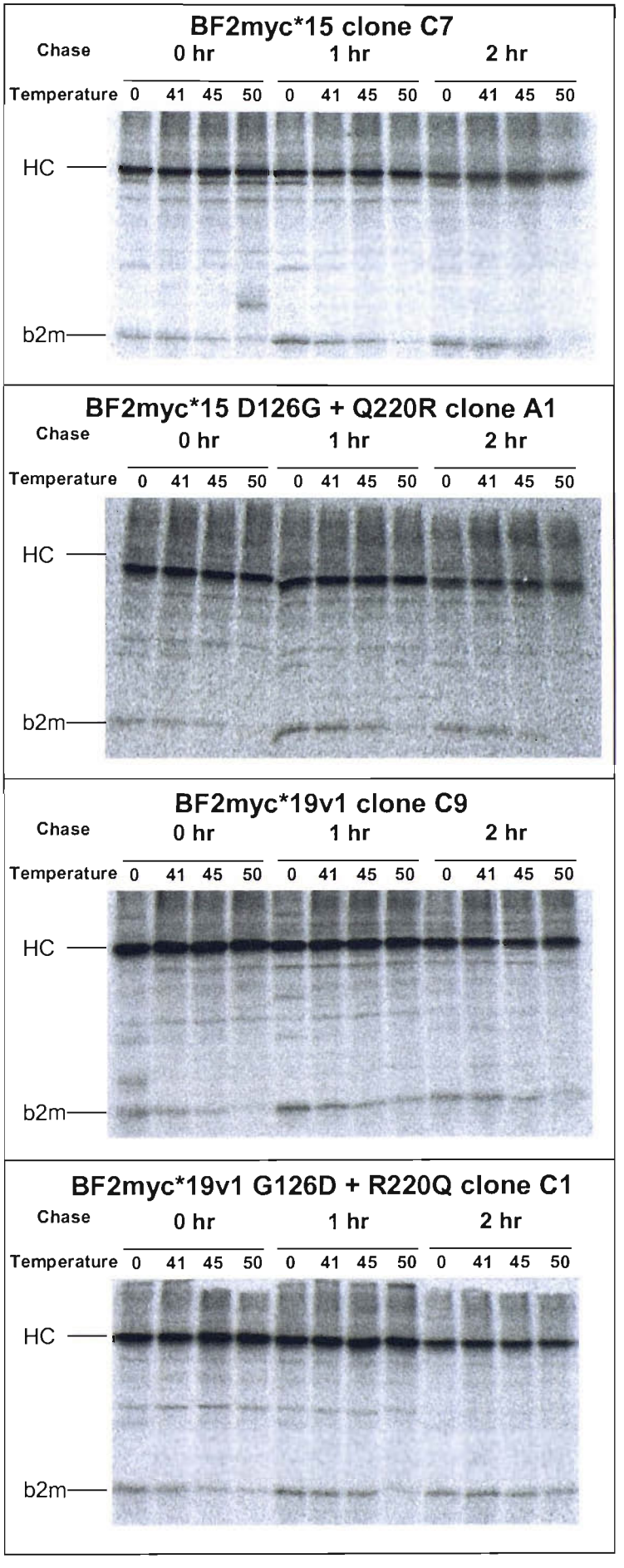
a-e) The indicated clones were pulse labelled for 20 minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc agarose was used to precipitate BF2myc molecules. Each immunoprecipitate was digested with Endoglycosidase H, apart from one aliquot which was mock digested, before samples were resolved by SDS PAGE, fixed and subjected to phosphorimager analysis.

f) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

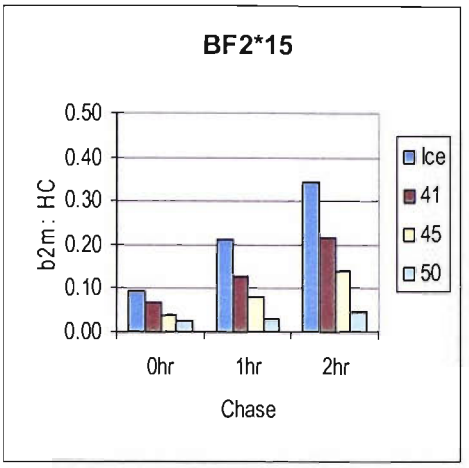
g) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

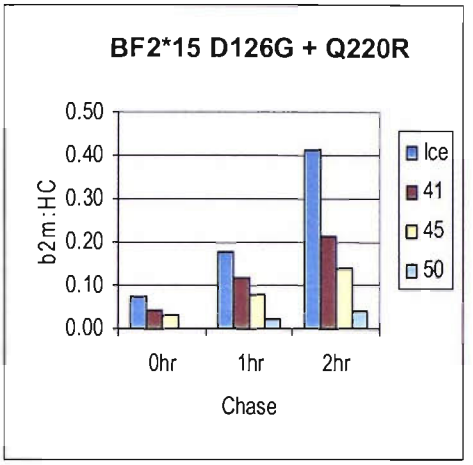
h) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

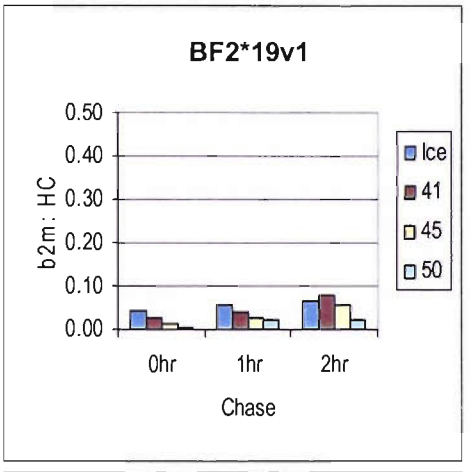
a)

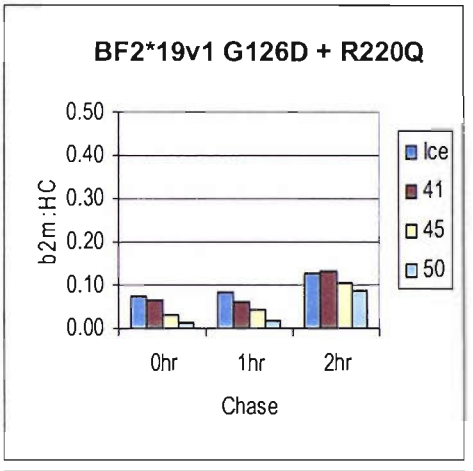

BF2myc*15 clone C7
Chase
0 hr
1 hr
2 hr
Temperature
0 41 45 50 0 41 45 50 0 41 45 50
HC
b2m
BF2myc*15 D126G + Q220R clone A1
Chase
0 hr
1 hr
2 hr
Temperature
0 41 45 50 0 41 45 50 0 41 45 50
HC
b2m
BF2myc*19v1 clone C9
Chase
0 hr
1 hr
2 hr
Temperature
0 41 45 50 0 41 45 50 0 41 45 50
HC
b2m
BF2myc*19v1 G126D + R220Q clone C1
Chase
0 hr
1 hr
2 hr
Temperature
0 41 45 50 0 41 45 50 0 41 45 50
HC
b2m

b)

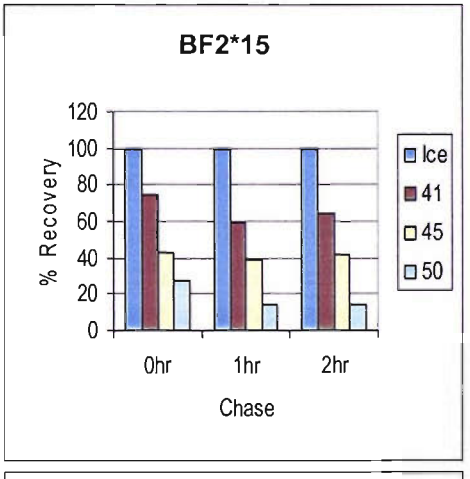

BF2*15
b2m: HC
0.50
0.40
0.30
0.20
0.10
0.00
0hr
1hr
2hr
Chase
Ice
41
45
50

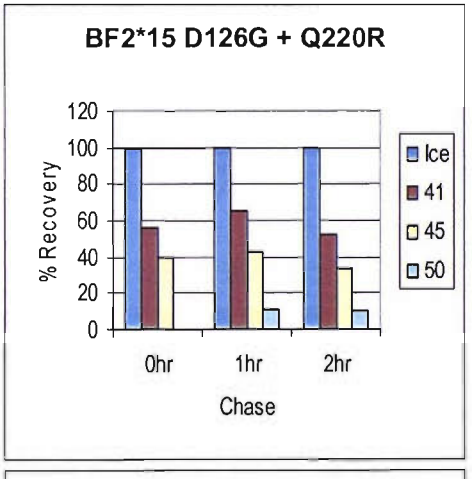

BF2*15 D126G + Q220R
b2m:HC
0.50
0.40
0.30
0.20
0.10
0.00
0hr
1hr
2hr
Chase
Ice
41
45
50

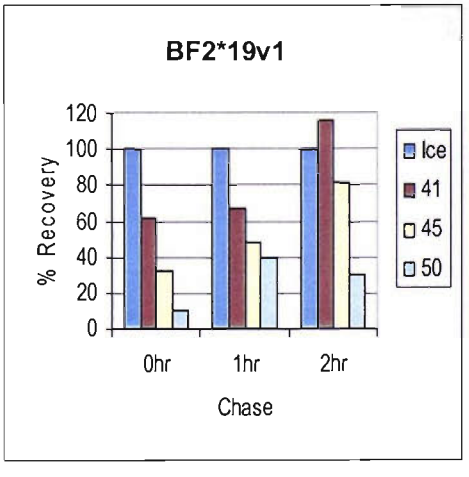

BF2*19v1
b2m: HC
0.50
0.40
0.30
0.20
0.10
0.00
0hr
1hr
2hr
Chase
Ice
41
45
50


BF2*19v1 G126D + R220Q
b2m:HC
0.50
0.40
0.30
0.20
0.10
0.00
0hr
1hr
2hr
Chase
Ice
41
45
50

c)


BF2*15
% Recovery
120
100
80
60
40
20
0
0hr
1hr
2hr
Chase
Ice
41
45
50


BF2*15 D126G + Q220R
% Recovery
120
100
80
60
40
20
0
0hr
1hr
2hr
Chase
Ice
41
45
50


BF2*19v1
% Recovery
120
100
80
60
40
20
0
0hr
1hr
2hr
Chase
Ice
41
45
50

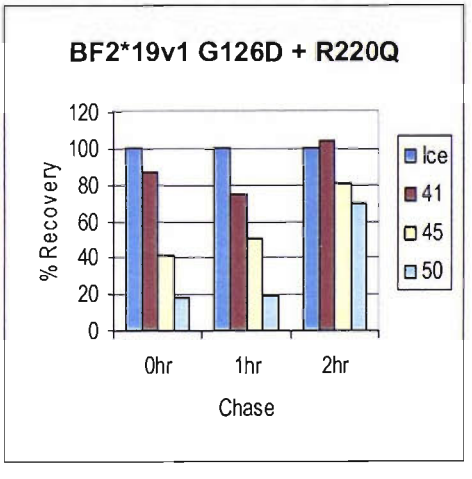

BF2*19v1 G126D + R220Q
% Recovery
120
100
80
60
40
20
0
0hr
1hr
2hr
Chase
Ice
41
45
50

Figure 6.16 Comparison of BF2myc*15 and BF2myc*19v1 molecules by thermal stability assays (experiment 293).

a) Clones BF2myc*15 C7, BF2myc*15 D126G + Q220R A1, BF2myc*19v1 C9 and BF2myc*19v1 G126D + R220Q C1 were pulse labelled for 30 minutes and chased for 120 minutes. Aliquots were removed as indicated and lysed, before being heated for one hour at the indicated temperature. Myc tagged molecules were then immunoprecipitated and separated by SDS PAGE gels, fixed and subjected to phosphorimager analysis.

b) Graphs depicting the ratio of β_2m molecules that were co-precipitated with BF2myc molecules (ratio = β_2m band volume: myc HC band volume)

c) Graphs depicting the stability of the immunoprecipitated BF2myc:β_2m heterodimers. Shown as the "Recovery %" for each clone, which is the percentage of heterodimers that were recovered following heating for each point of the chase relative to those that were recovered after incubation on ice at the same point of the chase.

BF2*15
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.094	0.212	0.342
41	0.070	0.126	0.218
45	0.040	0.082	0.142
50	0.026	0.031	0.048

BF2*15
Recovery %

	0hr	1hr	2hr
Ice	100.0	100.0	100.0
41	73.9	59.3	63.6
45	42.6	38.7	41.4
50	27.4	14.5	13.9

BF2*19
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.042	0.056	0.067
41	0.026	0.037	0.078
45	0.013	0.027	0.055
50	0.005	0.022	0.020

BF2*19
Recovery %

	0hr	1hr	2hr
Ice	100.0	100.0	100.0
41	61.1	66.3	115.8
45	31.9	48.0	81.4
50	10.8	39.2	30.3

BF2*15 126+220 mutant
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.076	0.179	0.411
41	0.042	0.117	0.215
45	0.030	0.077	0.139
50	0.000	0.021	0.040

BF2*15 126+220 mutant
Recovery %

	0hr	1hr	2hr
Ice	100.00	100.00	100.00
41	55.8	65.1	52.2
45	39.7	42.9	33.8
50	0.0	11.6	9.8

BF2*19 126+220 mutant
b2m:HC ratio

	0hr	1hr	2hr
Ice	0.075	0.082	0.127
41	0.066	0.061	0.131
45	0.031	0.041	0.103
50	0.014	0.015	0.088

BF2*19 126+220 mutant
Recovery %

	0hr	1hr	2hr
Ice	100.0	100.0	100.0
41	86.9	74.5	103.8
45	41.2	50.7	81.2
50	18.5	19.0	69.5

Table 6.5 Pulse/chase thermal-stability assays comparing wild-type and mutant BF2myc*15 and BF2myc*19v1 molecules. Left column: the ratios of β_2m: heavy chain (HC) band volumes were calculated for each immunoprecipitate to normalise for differential recovery of HC ("b2m: HC ratio"). Right column: the b2m: HC ratio calculated at each time-point after incubation on ice was set to 100%, and the b2m:HC ratio that was calculated following heat treatment at that same time-point was calculated as a percentage of this value ("Recovery %").

A peptide stabilisation assay was performed to confirm that a substantial proportion of BF2myc*19v1 and BF2myc*19v1 G126D&R220Q molecules were loaded with sub-optimal peptides (figure 6.17, experiment 295, note this experiment was conducted prior to experiment 296, which is shown in figure 6.9). This experiment confirmed that a larger amount of β_2m molecules were co-precipitated with BF2myc*19v1 G126D&R220Q molecules than were co-precipitated with BF2myc*19v1 molecules, with BF2myc*15 molecules co-precipitating the most β_2m molecules. Surprisingly however, peptide addition did not substantially increase the amount of co-precipitated β_2m molecules, which may be a consequence of the lower concentration of peptide, and the conditions that were used in this experiment.

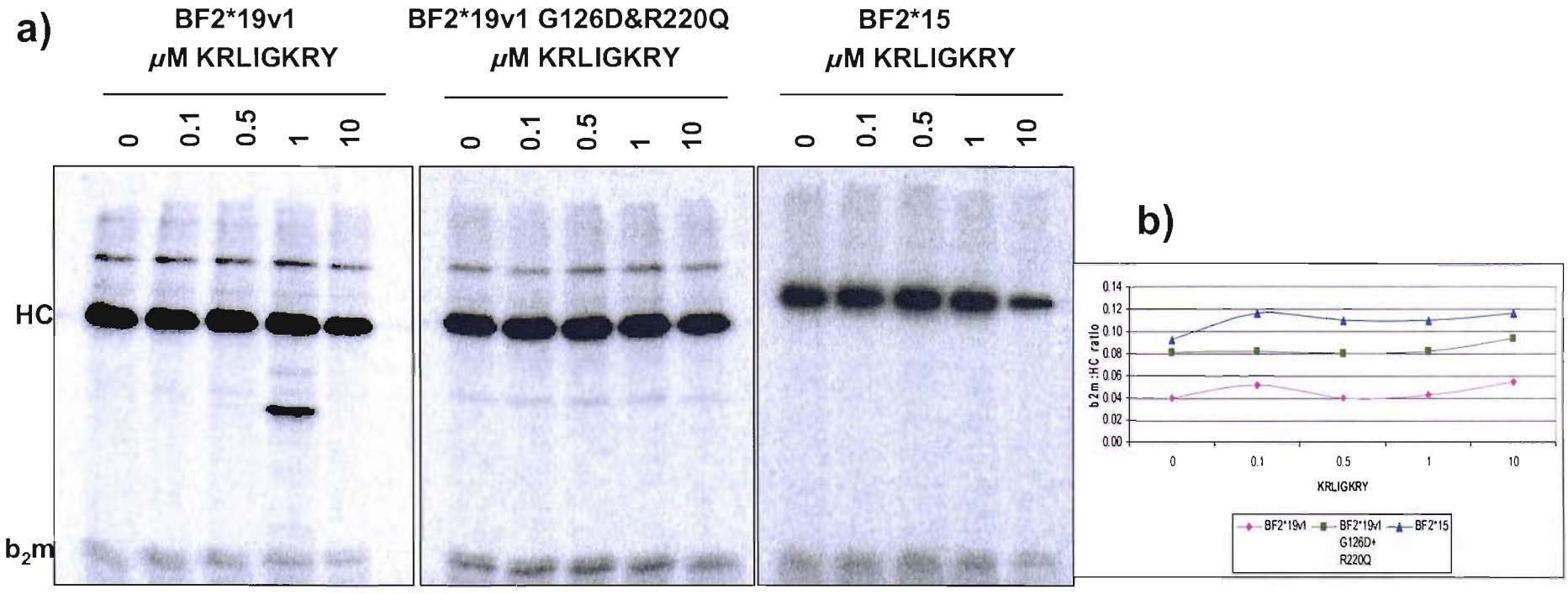


Figure 6.17 Peptide stabilisation assays (experiment 295).

a) TG15 B72myc*19v1 clone C1, TG15 BF2myc*19v1 G126D&R220Q clone A6 and TG15 BF2myc*15 clone C7 were radiolabelled before being lysed in lysis buffer supplemented with the indicated concentration of peptide, before being digested with endoglycosidase H and resolved by SDS PAGE.

b) Graph depicting the b2m: Heavy Chain (HC) ratios that were calculated following phosphorimager analysis.

6.2.4 Transfection of Tapasin*12 does not alter the maturation characteristics of BF2myc*19v1 molecules

The decreased maturation characteristics of wild-type BF2myc*19v1 molecules in TG15 cells may be a consequence of an impaired ability for Tapasin*15 to facilitate peptide-loading of BF2myc*19v1 molecules. Therefore BF2myc*19v1 and V5 epitope-tagged Tapasin*12 molecules were expressed in the same cell in order to determine whether the maturation characteristics of BF2myc*19v1 molecules was altered when the correct allele of tapasin was present (figure 6.3d).

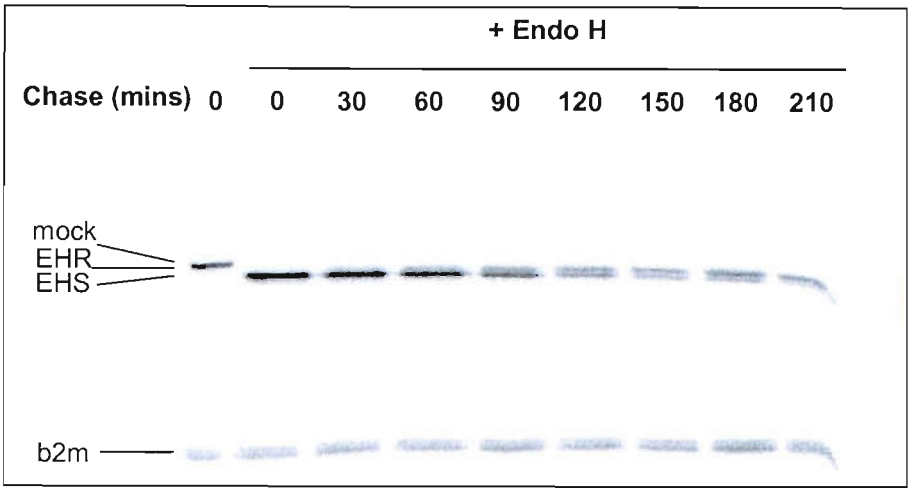
BF2myc*19v1 and Tapasin*12 double transfectants were generated by the simultaneous transfection of both expression constructs into TG15 cells, and application of both selection antibiotics. Interestingly, sequential transfection of these expression constructs was consistently unsuccessful. Stable transfectants expressing both proteins were generated and the maturation of BF2myc*19v1 molecules was compared with single transfectants expressing either BF2myc*15 or BF2myc*19v1 molecules by pulse/chase analysis (figure 6.18, experiment 282). Quantification of the proportion of endo H resistant material present throughout the chase (figure 6.18d) demonstrated that the presence of transfected Tapasin*12 molecules did not alter the maturation of BF2myc*19v1 molecules, which was confirmed using different clones in a second experiment (figure 6.19, experiment 291).

6.2.5 Mutation of five amino acids in human tapasin does not disturb HLA-B*4402 surface expression

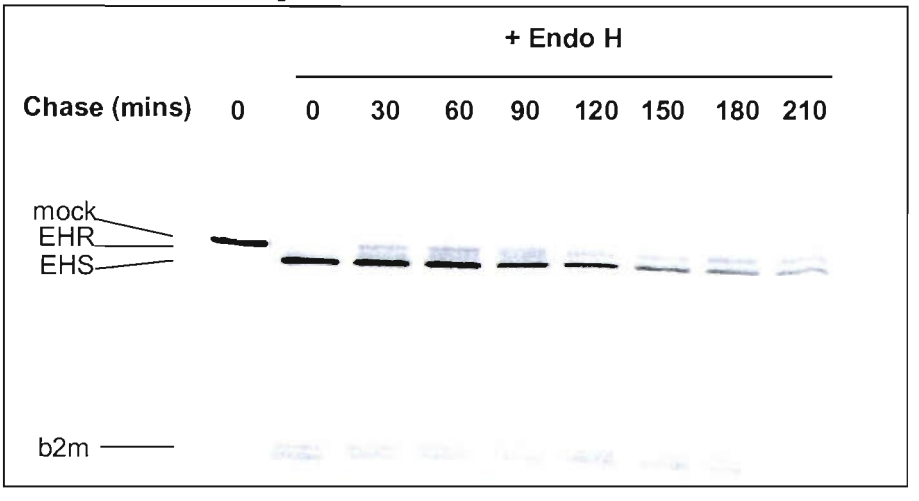
The experiments described above demonstrate that the nature of amino acids at certain positions that vary between BF2myc*15 and BF2myc*19v1 alleles can influence the maturation characteristics of BF2myc*19v1 molecules. The tapasin alleles present in the B15 and B19 haplotypes also vary, and it is possible that these polymorphic amino acids interact specifically with polymorphic amino acids present in the BF2 allele; such that BF2 and tapasin alleles participate in a haplotype-specific interaction, leading to optimal peptide-loading when the combination of tapasin and class I alleles are derived from the same haplotype. However, the experiments described above do not demonstrate directly what, if any, influence the polymorphic amino acids in tapasin have on such an interaction.

Without tapasin-deficient or class I-deficient chicken cell lines it is difficult to determine what the influence of polymorphic residues in chicken tapasin molecules have on the expression of BF2 molecules, as all chicken cells and cell lines endogenously express chicken tapasin and class I proteins. Therefore, to determine whether polymorphic residues in chicken tapasin molecules influence class I peptide-loading, the human tapasin-deficient .220B*4402 cell line was transfected with a mutant human tapasin protein, and the surface expression characteristics of the tapasin-dependent HLA-B*4402 allele was compared to that of tapasin-deficient and wild-type human tapasin-expressing cell lines.

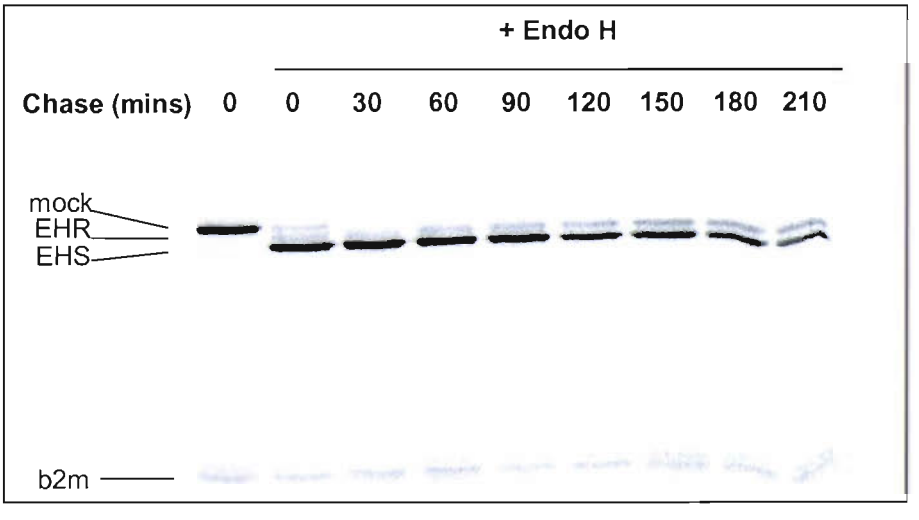
a) TG15 BF2myc*15 clone C3



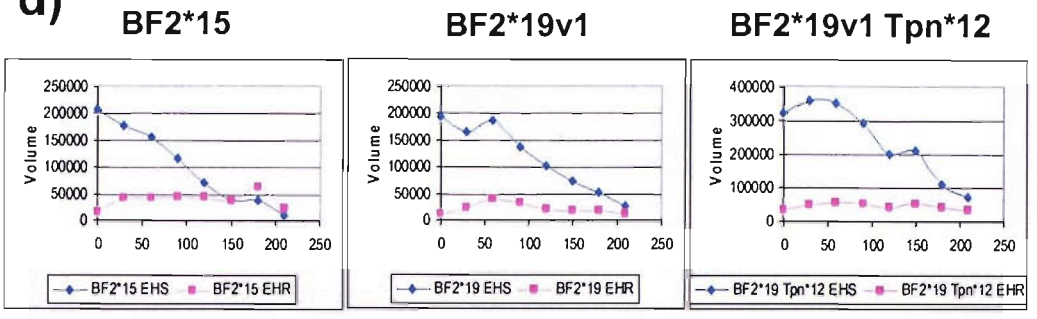
b) TG15 BF2myc*19v1 clone C1



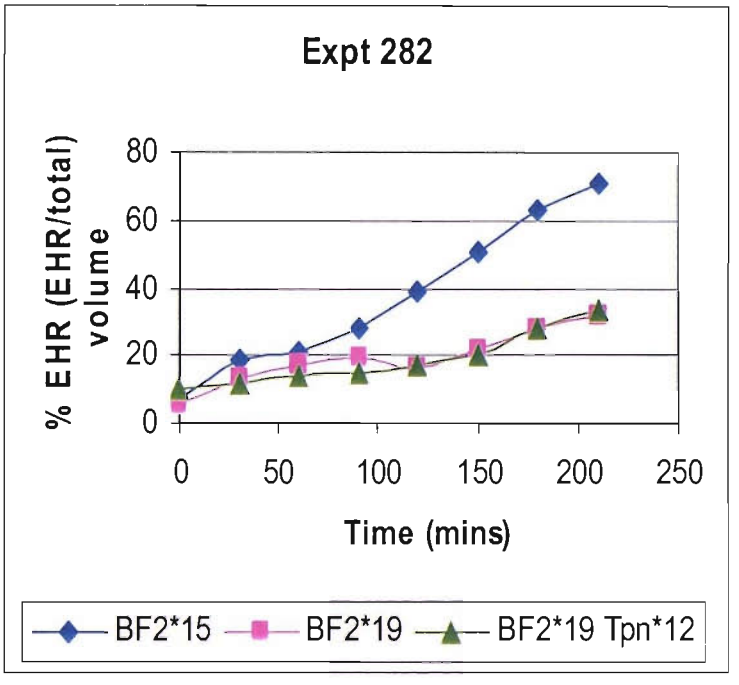
c) TG15 BF2myc*19v1 Tpn*12 clone A9



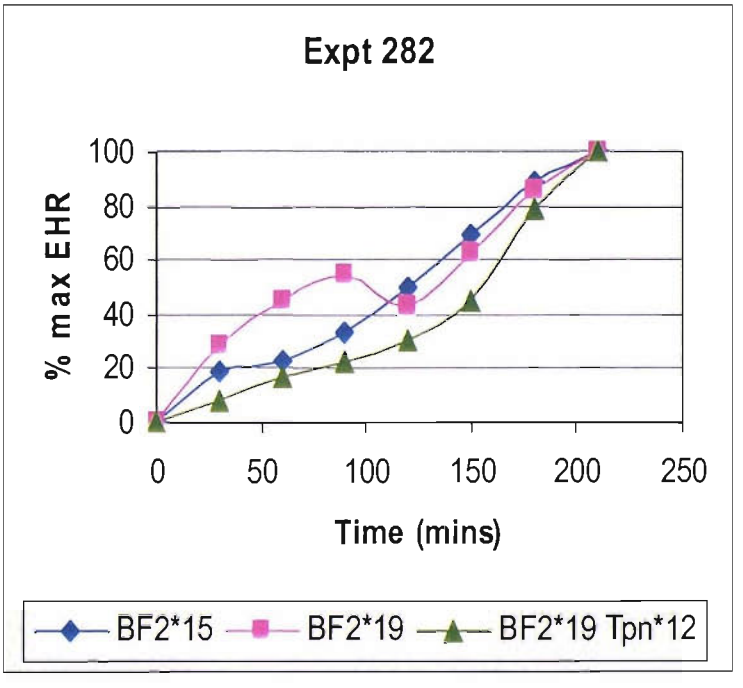
d)


BF2*15
BF2*19v1
BF2*19v1 Tpn*12
Volume
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR
BF2*19 Tpn*12 EHS
BF2*19 Tpn*12 EHR

e)


Expt 282
% EHR (EHR/total) volume
Time (mins)
BF2*15
BF2*19
BF2*19 Tpn*12

f)


Expt 282
% max EHR
Time (mins)
BF2*15
BF2*19
BF2*19 Tpn*12

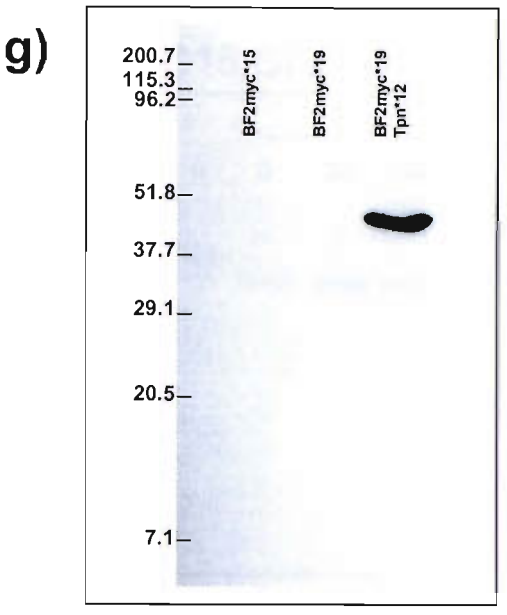


Figure 6.18 Tapasin*12 transfection does not increase BF2*19v1 maturation (experiment 282)

a-c) The indicated clones were pulse labelled for 20 minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc tagged molecules were immunoprecipitated and endoglycosidase H digested, before being run on SDS PAGE gels, fixed and subjected to phosphorimager analysis.

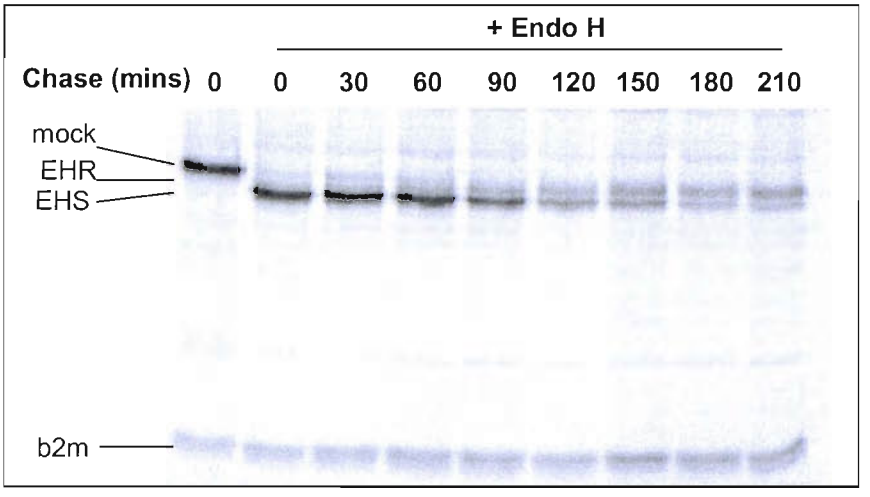
d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

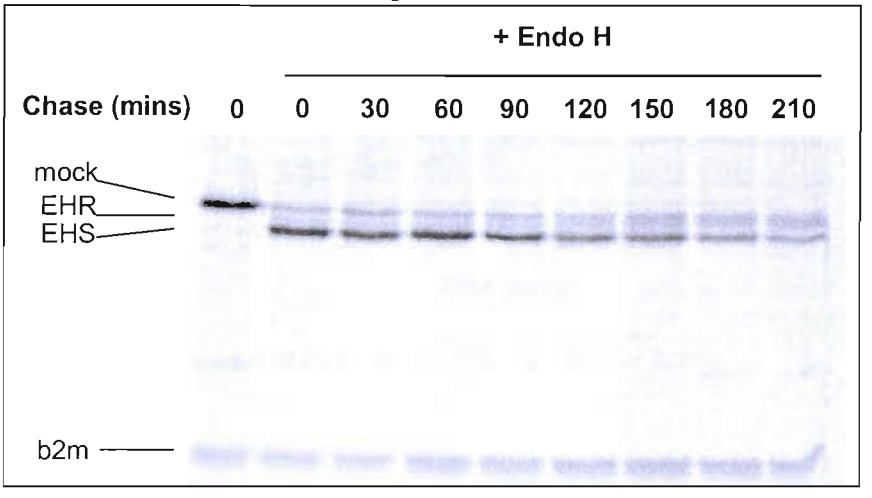
f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

g) Western blot confirming expression of V5-tagged Tapasin*12 molecules. Lysates were made from each of the transfectants, a volume equivalent to $5x10^5$ cells was separated by SDS PAGE and western blotted.

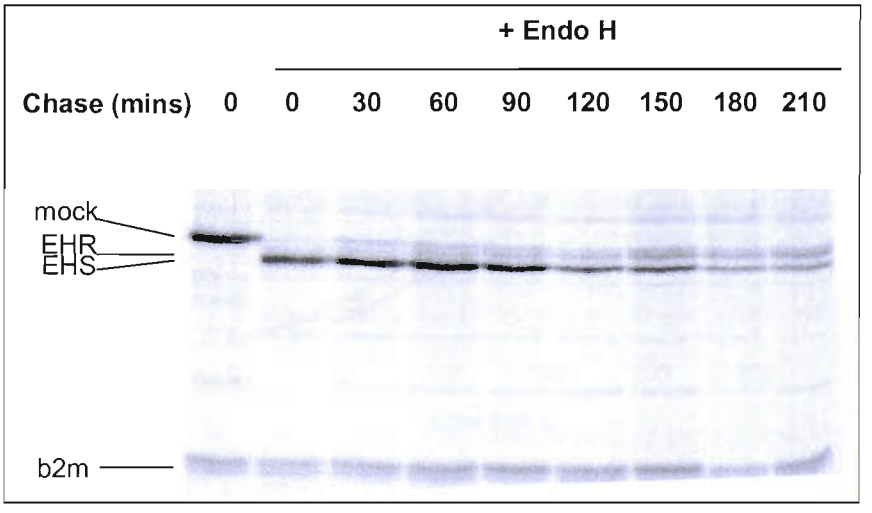
a) TG15 BF2myc*15 clone C7



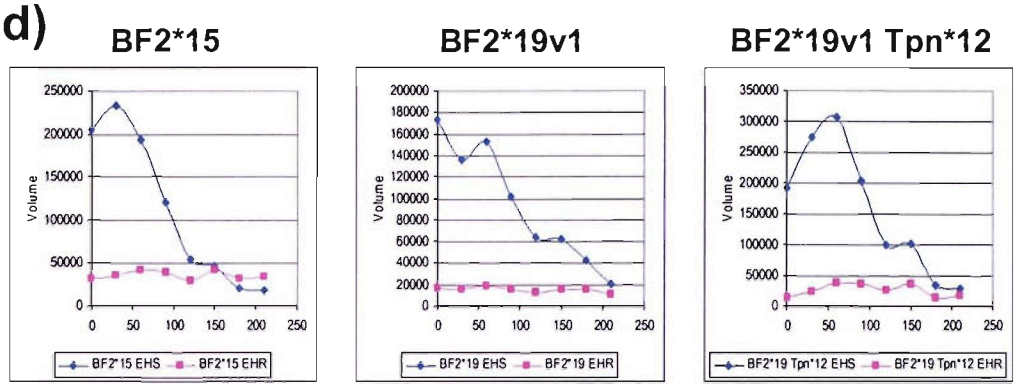
b) TG15 BF2myc*19v1 clone C9



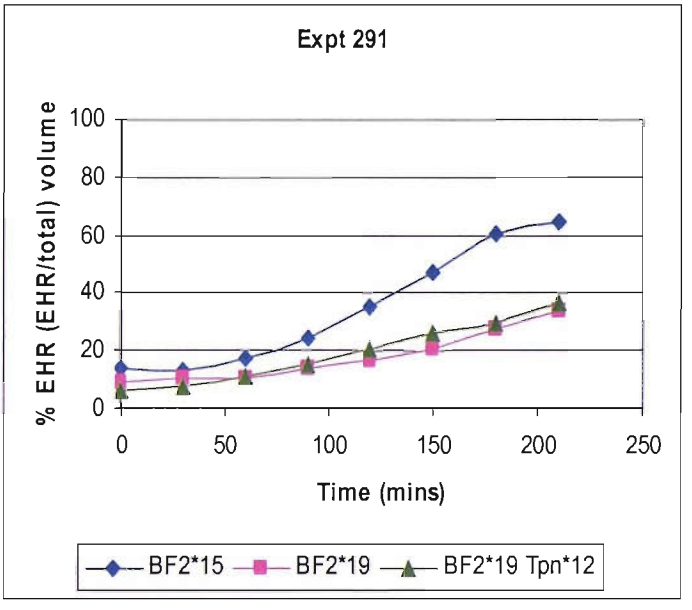
c) TG15 BF2myc*19v1 Tpn*12 clone A10



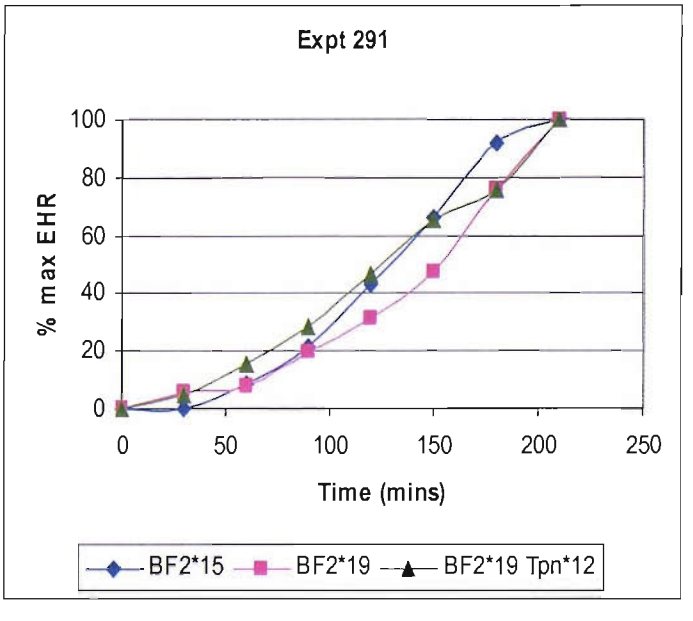
d)


BF2*15
BF2*19v1
BF2*19v1 Tpn*12
Volume
BF2*15 EHS
BF2*15 EHR
BF2*19 EHS
BF2*19 EHR
BF2*19 Tpn*12 EHS
BF2*19 Tpn*12 EHR

e)


Expt 291
% EHR (EHR/total) volume
Time (mins)
BF2*15
BF2*19
BF2*19 Tpn*12

f)


Expt 291
% max EHR
Time (mins)
BF2*15
BF2*19
BF2*19 Tpn*12

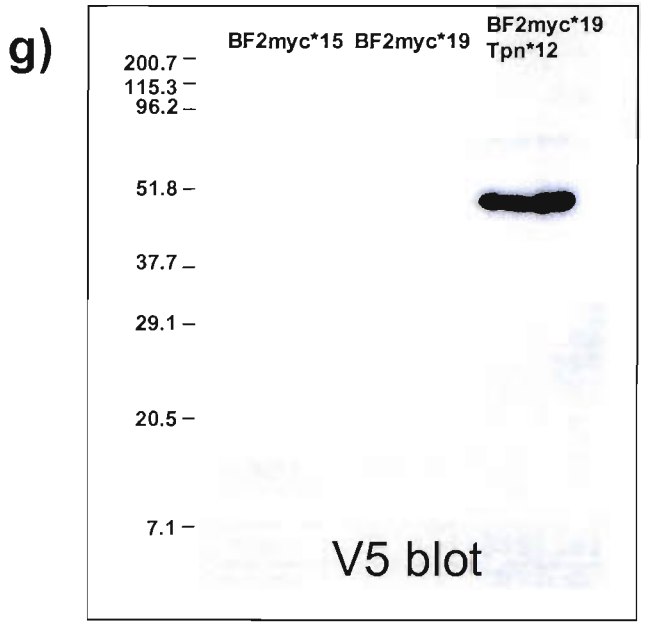


Figure 6.19 Tapasin*12 transfection does not increase BF2*19v1 maturation (experiment 291).

a-c) The indicated clones were pulse labelled for 20 minutes and chased for 210 minutes. Aliquots were removed as indicated and lysed. Myc tagged molecules were immunoprecipitated and endoglycosidase H digested, before being run on SDS PAGE gels, fixed and subjected to phosphorimager analysis.

d) Graphs depicting the measured band volumes of endo H sensitive and endo H resistant BF2myc molecules (displayed in arbitrary units) for each clone.

e) Graph depicting the rate and extent of the acquisition of resistance of the BF2myc molecules to endo H digestion following phosphorimager analysis (the volume of the endo H resistant material is shown as a percentage of total volume of endo H resistant and sensitive material).

f) Graph depicting the rate at which maximal endo H resistance was gained for each clone (note that this does not take into account the proportion of BF2myc molecules that gained resistance to endo H digestion).

g) Western blot confirming expression of V5-tagged Tapasin*12 molecules. Lysates were made from each of the transfectants, a volume equivalent to $5x10^5$ cells was separated by SDS PAGE and western blotted.

The location of the residues in human tapasin that were mutated was guided by an alignment of chicken and human tapasin proteins performed by Frangoulis *et al* (figure 6.20, although it must be noted other alignments are possible, especially concerning the N-terminus). The nature of the mutations was guided by the five polymorphic residues that differ between Tapasin*12 and Tapasin*15 alleles with the resulting human tapasin mutant being labelled "5x", in reference to the five mutated residues (table 6.6).

	Exon				
	4	5			
	amino acid position				
Allele	192	317	319	324	365
Tapasin*12	S	A	G	Q	I
Tapasin*15	P	T	S	R	V

Human	189	334	336	341	379
Wild-type	S	G	P	Q	R
5x mutant	P	T	S	R	I

Table 6.6 Amino acid polymorphisms in chicken Tapasin*12 and Tapasin*15 alleles, and the mutations that were made to create the human tapasin "5x" mutant.

Where an identical amino acid occupied equivalent positions in one chicken tapasin allele and in the human tapasin protein, the human tapasin residue was mutated to the amino acid that is present in the other chicken allele (positions 192 and 324). When a different amino acid residue occupied the equivalent position in human tapasin, this residue was mutated to either of the residues present in the chicken tapasin alleles (positions 317, 319 and 365). The human tapasin "5x" expression construct was transfected into the .220B*4402 cell line and two stable tapasin "5x"-expressing clones were generated.

The human tapasin "5x" mutant protein ran with decreased mobility relative to wild-type human tapasin during electrophoresis (figure 6.21), which may reflect an alteration in the tapasin protein structure or charge. "5x" clone 1 had a very low level of tapasin protein present in the lysate, whilst "5x" clone 2 had a greater amount of tapasin protein present, similar to the amount of tapasin protein expressed by the human tapasin R48A mutant clone that was included as a control, along with the human tapasin R48E mutant clone, which expressed the greatest amount of tapasin protein. Both position 48 mutant tapasin clones and the "5x" mutant clone 2 expressed a greater amount of tapasin protein than the wild-type tapasin transfectant.

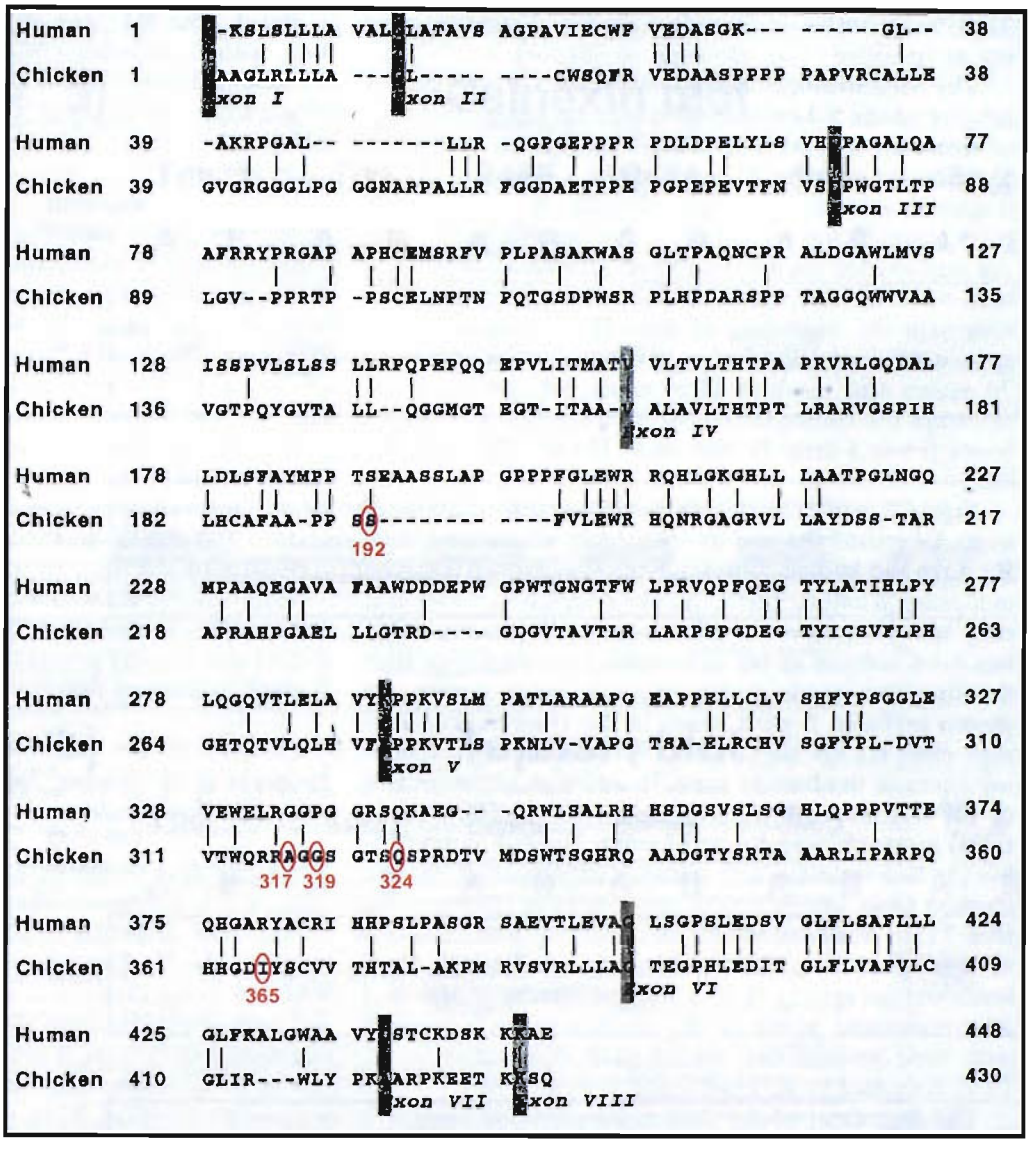


Figure 6.20 Alignment of human and chicken tapasin protein sequences.

The residues that differ between Tapasin*12 and Tapasin*15 alleles are indicated by red circles and are numbered from the initiating methionine. The alignment was produced by (Frangoulis et al., 1999), but other alignments are possible, especially with regard to the N terminus.

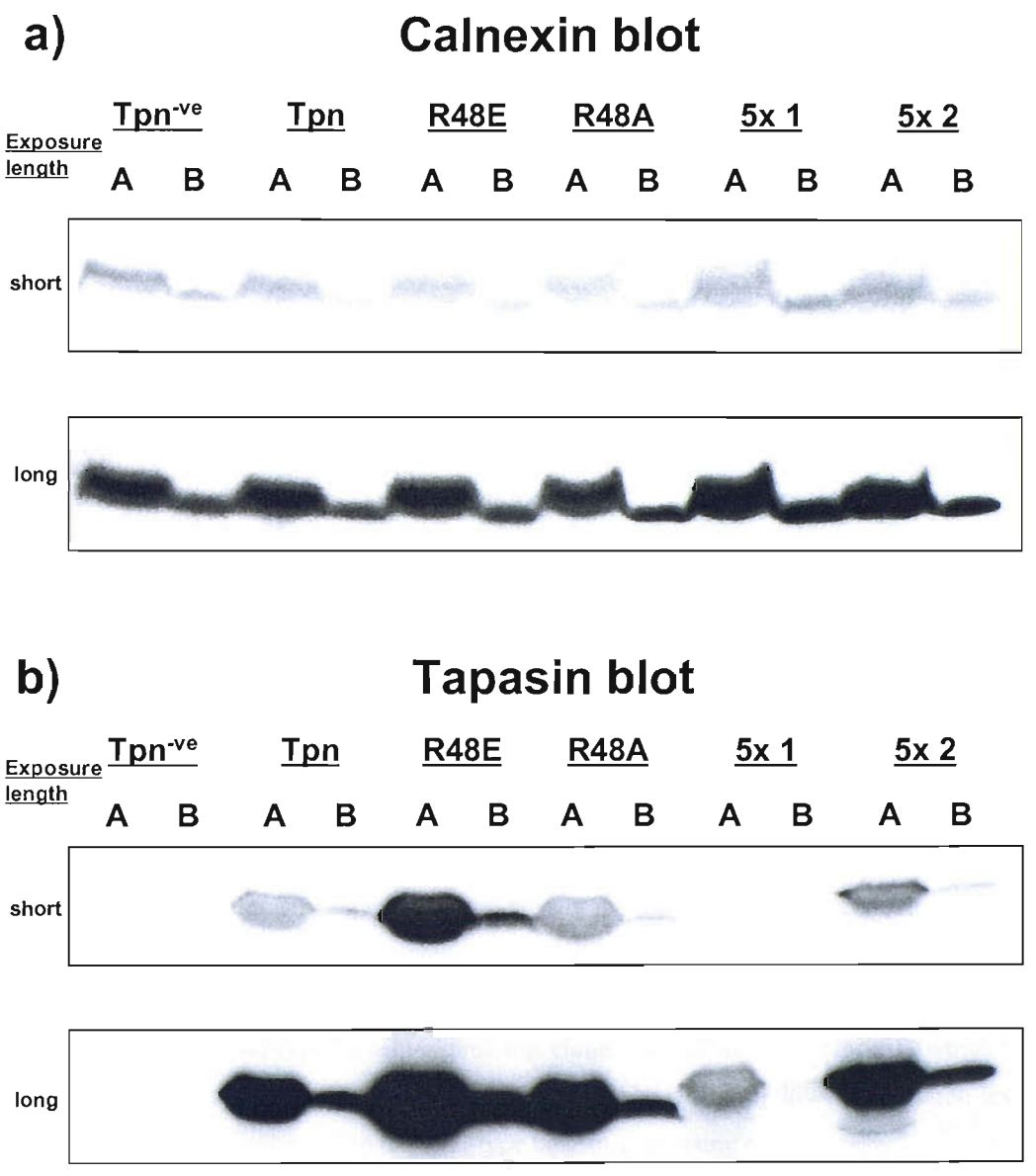


Figure 6.21 Expression levels of tapasin protein in .220B*4402 transfectants

Lysates were made from the indicated transfectants, and volumes derived from equal numbers of cells were loaded into each lane (as detailed below) and separated by SDS PAGE and western blotted. The blots were incubated with the indicated antibodies. Only the bands corresponding to either calnexin or human tapasin proteins are shown (from either a short or longer exposure) for clarity.

Tpn^{-ve} = .220B*4402	Tpn = .220B*4402wild-type tapasin
R48E = .220B*4402TpnR48	R48A = .220B*4402 Tpn R48A
5x 1 = .220B*4402Tpn5x	5x 2 = .220B*4402Tpn5x
A = $3.3x10^5$ cell equivalents	B = $3.3x10^4$ cell equivalents
a) Anti-calnexin antibody	b) Anti-human tapasin antibody

Flow cytometry was performed on the transfectants following staining with one of four antibodies (table 6.7): W6/32 (which reacts with HLA: β_2m heterodimers); an anti-HLA-B antibody referred to as clone Joan-1 (which reacts with a determinant present on HLA-B locus encoded gene products); TT4-A20 (which specifically recognises the HLA-Bw4 public determinant controlled by the polymorphic residues 77 and 80-83 in the α1 domain, in a fashion that is sensitive to the peptide cohort); and HC10 (which preferentially recognises β_2m-free HLA-B and HLA-C heavy chain molecules).

	W6/32		Joan-1		TT4-A20		HC10	
Clone	Average MFI	St Dev	Average MFI	St Dev	Average MFI	St Dev	Average MFI	St Dev
.220B4402	13.59	0.11	3.09	0.02	2.41	0.02	3.33	0.04
.220B4402 hTpn	134.14	1.44	35.70	0.79	93.96	1.46	5.58	0.11
.220B4402 hTpn R48E1	160.45	3.89	53.96	1.32	143.73	3.11	8.48	0.23
.220B4402 hTpn R48A1	149.63	1.21	50.11	0.10	148.94	1.38	8.39	0.04
.220B4402 hTpn 5x 1	116.95	1.64	38.93	0.63	97.71	1.37	7.83	0.12
.220B4402 hTpn 5x 2	147.41	1.31	48.49	1.31	125.66	1.22	8.54	0.19

Table 6.7 Cell surface expression levels of HLA-B*4402 molecules in transfectants expressing different tapasin proteins.

Staining with W6/32 antibody revealed there was a massive increase in the surface expression level of HLA: β_2m heterodimers which occurred when peptide-loading occurred in the presence of either wild-type tapasin or a mutated tapasin protein (table 6.7). This suggests that the five mutations that were made to human tapasin did not grossly affect the ability of tapasin to facilitate peptide-loading of HLA-B*4402 molecules. In keeping with the tapasin protein expression levels evident from the western blots, both position 48 mutant clones and "5x" clone 2 (which appeared to express more tapasin protein than the wild-type tapasin expressing clone) had greater expression levels of class I heterodimers than the wild-type tapasin-expressing clone. In contrast "5x" clone 1 (which had a decreased amount of tapasin protein present in the lysate) had a slightly lower expression level of class I heterodimers than observed in the wild-type tapasin-expressing clone.

In agreement with the W6/32 results and the western blots, the surface expression level of HLA-B molecules detected following Joan-1 antibody staining (table 6.7) was increased above that of the wild-type tapasin-expressing clone for both position 48 mutants and the "5x" clone 2 mutant. However, "5x" clone 1 (which expressed the smallest amount of tapasin protein and had a slight decrease in class I heterodimer expression level relative to the wild-type tapasin-expressing clone as detected by W6/32) gave an equivalent surface expression level of HLA-B molecules as the wild-type tapasin-expressing clone. This suggests that for this clone there may have been a defect specifically relating to HLA-C peptide-loading.

Following TT4-A20 staining (table 6.7), the surface expression level of HLA-B molecules evident in "5x" mutant clone 2 was decreased slightly below that observed for the position 48 mutant clones, which suggests that a slightly different repertoire of peptides was loaded into this clone.

Despite this, the expression level in these clones was greater than either the wild-type tapasin-expressing or "5x" mutant clone 1 clones, which gave similar expression levels.

In all tapasin-expressing clones there was a modest increase in free heavy chain surface expression, detected following HC10 staining (table 6.7), which was slightly greater than seen in the tapasin-deficient clone. This suggests that as the number of class I molecules expressed on the cell surface increases, the amount of free heavy chain also increases. This contrasts with the findings of Sesma *et al* who observed that following mammalian tapasin transfection of .220B*2705 cells, the amount of β_2m-free heavy chains expressed on the cell surface decreased, which may reveal specific differences between the two class I alleles (Sesma et al., 2005).

Cumulatively these experiments show that firstly, the amount of tapasin protein expressed in a cell can influence the amount of surface expressed class I molecules. Secondly, that the surface expression of HLA-B*4402 molecules was not adversely affected when mutations were made at five positions in the human tapasin protein.

6.3 Discussion

In this chapter most of the research that was conducted concerned the comparison of two class I alleles (BF2myc*15 and BF2myc*19v1) that were expressed separately in a B15 homozygous cell line (figure 6.3a-b). These transfectants were produced based upon the presumed sequence of events that led to the generation of the recombinant B19v1 haplotype. In the first part of the strategy, which was designed to reproduce the molecular situation immediately after the recombination event, the haplotype-mismatched BF2myc*19v1 molecules were likely to be supplied with adequate peptides by the endogenous TAP molecules, but their ability to be peptide-loaded may have been compromised by an incompatibility with the endogenous Tapasin*15 molecules. In order to characterise the effect of potential tapasin incompatibility, thermal-stability and pulse/chase assays were performed to compare the two class I alleles.

In the second part of the strategy (figure 6.3c), the effect of mutations that were made to the BF2myc alleles was determined. The nature of the mutations was guided by the presumed evolutionary changes that occurred within the recombinant haplotype after the recombination event. In addition, the maturation rate of BF2myc*19v1 molecules was also determined when reconstituted with the appropriate tapasin allele (figure 6.3d).

6.3.1 The expression of BF2myc*19v1 molecules was compromised in comparison to BF2myc*15 molecules

Pulse/chase experiments suggested that a smaller proportion of newly made haplotype-mismatched BF2myc*19v1 molecules matured than was evident for the haplotype-matched BF2myc*15 molecules, and that those BF2myc*19v1 molecules which did leave the ER, did so at a slightly slower rate.

Thermal-stability assays (and related assays) suggested that a substantially larger proportion of haplotype-mismatched BF2myc*19v1 heterodimers dissociated during immunoprecipitation than was evident for the haplotype-matched BF2myc*15 molecules. This resulted in the recovery of a small pool of BF2myc*19v1 heterodimers that were loaded with peptides that conferred comparable stability to that achieved by a larger proportion of BF2myc*15 molecules.

These findings suggest that the haplotype-mismatched BF2myc*19v1 molecules were loaded with an inferior repertoire of peptides in comparison to haplotype-matched BF2myc*15 molecules. The inferior peptide repertoire of BF2myc*19v1 molecules is likely to be a consequence of an impaired ability for the endogenous Tapasin*15 molecules to facilitate peptide-loading of BF2myc*19v1 molecules.

These transfectants were designed to recreate the molecular situation that was likely to have been present immediately after the recombination event between the B12 and B15 haplotypes occurred, which eventually led to the modern B19v1 haplotype. The impaired expression of haplotype-mismatched BF2myc*19v1 molecules in these transfectants suggests that after the recombination event a similar impairment in the peptide-loading of BF2*15 molecules existed. The fact that a proportion of haplotype-mismatched BF2myc*19v1 molecules succeeded in gaining a peptide repertoire that permitted their egress from the ER and partial recovery following immunoprecipitation suggests that the impairment was not absolute. This may reflect an impaired ability for the endogenous tapasin molecules to facilitate peptide-loading of the haplotype-mismatched BF2myc*19v1 molecules, or may reflect an inherent ability of the BF2myc*19v1 molecules to self-optimise their peptide repertoire. The mild phenotype evident in these transfectants had been anticipated, as a recombinant haplotype in which there is a complete inability to present peptides would be expected to be lost from the population.

6.3.2 The effect of mutations upon the expression of the BF2myc alleles

Analysis of the modern B12, B15 and B19v1 haplotypes suggests that following the recombination event, the BF2*19v1 allele has evolved from the BF2*15 allele that was likely to have been present in the original recombinant haplotype. This evolution may have been driven by the selection of BF2 mutants that associated more productively with the endogenous tapasin allele. The eight amino acid differences between the BF2*15 and BF2*19v1 alleles may reflect this evolution.

Therefore mutations were made to the haplotype-mismatched BF2myc*19v1 allele, where the equivalent residues from the haplotype-matched BF2myc*15 allele replaced the residues present at 126 and 220, which are the only polymorphisms between the BF2*15 and BF2*19 alleles that are located in predicted sites of direct interaction with the tapasin protein. These mutations resulted in an increased proportion of BF2myc*19v1 mutant molecules maturing at a rate that was comparable to that of BF2myc*15 molecules. Both positions 126 and 220 appeared to contribute to this phenomenon. In addition, there was a modest increase in the recovery of BF2myc*19v1 G126D&R220Q mutant heterodimers following immunoprecipitation in comparison to wild-type BF2myc*19v1 molecules. This implies that BF2myc*19v1 G126D&R220Q mutant molecules had a peptide repertoire that was quantitatively and qualitatively improved in comparison to that of wild-type BF2myc*19v1 molecules. This is likely to reflect that BF2myc*19v1 G126D&R220Q mutant molecules had an increased ability to profit from the endogenous Tapasin*15 molecules than was apparent for wild-type BF2myc*19v1 molecules.

However, whilst the maturation characteristics of BF2myc*19v1 G126D&R220Q mutant molecules were directly comparable to that of haplotype-matched BF2myc*15 molecules, the proportion of BF2myc*19v1 G126D&R220Q mutant molecules that were immunoprecipitated as β_2m-associated heterodimers was substantially lower than was evident for BF2myc*15 molecules.

Therefore despite the beneficial effect of the mutations (in comparison to wild-type BF2myc*19v1 molecules), the BF2myc*19v1 G126D&R220Q mutant molecules were loaded with an inferior peptide repertoire that conferred decreased stability to the heterodimer in comparison to BF2myc*15 molecules in thermal-stability assays.

The replacement of the residues at positions 111 and 113 in the BF2myc*19v1 molecule with the equivalent residues from the BF2myc*15 molecule also produced similar improvements in maturation characteristics to the BF2myc*19v1 G126D&R220Q mutant. Positions 111 and 113 are located within the peptide-binding groove, and mutation of the equivalent residues in mammalian class I molecules (residues 114 and 116) affects both the ability of the class I molecules to associate with the PLC, and alters their dependency on tapasin for optimal peptide-loading (Neisig et al., 1996; Turnquist et al., 2002; Park et al., 2003). The molecular mechanism behind such observations is likely to be that the nature of the amino acids in these positions around the F pocket, in addition to defining peptide specificity, may affect the conformation or the amount of flexibility that exists in the portions of the α helices that surround the C-terminus of the bound peptide. It has been postulated that the differential tapasin binding properties of mammalian class I alleles may be affected by the conformation or the amount of flexibility that exists in this region, or in the nearby loop within the α2 domain (chicken residues 125-133) to which tapasin presumably binds (Wright et al., 2004; Elliott and Williams, 2005). It can therefore be anticipated that the replacement of the BF2myc*19v1 residues at these positions with the equivalent residues from the BF2myc*15 allele altered the conformation or the amount of flexibility of the peptide-binding domain, or the nearby loop within the α2 domain (chicken residues 125-133), which resulted in BF2myc*19v1 mutant molecules profiting from the endogenous Tapasin*15 molecules to a greater extent than was evident for wild-type BF2myc*19v1 molecules, leading to faster and more comprehensive peptide-loading. However, it remains possible that alternate explanations apply; the mutation of these positions may have increased the ability of BF2myc*19v1 molecules to self-optimise their peptide repertoire, or changed the peptide-binding specificity to be more compatible for the peptides supplied by the TAP proteins endogenous to the B15 cell line.

In addition to the BF2myc*19v1mutants that are discussed above, the effect of making the reciprocal mutations to the haplotype-matched BF2*15 molecule was also determined; where positions 126 and 220, or 111 and 113, or a combination of all four positions were exchanged for the equivalent residues from the BF2myc*19v1 allele. These mutations were made because it was anticipated that they would decrease the ability of the BF2*15 molecule to associate productively with tapasin. However, the effect of these mutations was modest; the maturation characteristics of the BF2myc*15 mutants were directly comparable to wild-type BF2myc*15 molecules. Thermal-stability assays suggested that BF2myc*15 D126G&Q220R mutant molecules had a slightly inferior peptide repertoire than wild-type BF2myc*15 molecules, which resulted in a slightly decreased recovery of heterodimers following immunoprecipitation. This implies that this mutant

had a decreased ability to associate productively with endogenous Tapasin*15 molecules. It is possible that other assays may reveal a more pronounced phenotype for these mutants.

Comparison of the two BF2myc 126&220 mutants revealed that both the BF2myc*15 126&220 mutant and the BF2myc*19v1 126&220 mutant trafficked at the same rate (which was equivalent to wild-type BF2myc*15 molecules), but differed dramatically in stability, with the BF2myc*19v1 position 126&220 mutant demonstrating decreased stability in comparison to the BF2myc*15 position 126&220 mutant. The greater stability of BF2myc*15 molecules may be a reflection of a more suitable peptide repertoire supplied by the endogenous TAP molecules. The comparable rates of maturation suggest that BF2myc heterodimers which are loaded predominantly with peptides that confer decreased levels of stability (in thermal-stability assays) are able to egress beyond the ER as efficiently as peptides that confer greater stability to the heterodimers. Assessing the maturation characteristics of BF2myc molecules therefore appears to be particularly sensitive to modest improvements to the peptide repertoire of poorly loaded BF2myc molecules, but generally is a less comprehensive indication of "optimal" peptide-loading than thermal-stability assays ("optimal" peptide-loading may best be defined in this system as the rate of production of stable complexes at the cell surface).

These transfectants suggest that following the recombination event between the B12 and B15 haplotypes which eventually led to the B19v1 haplotype, the mutation of residues at positions 126 and 220 of the BF2*15 allele may have allowed the new BF2 molecule to gain a slightly superior peptide repertoire, which is likely to be a consequence of an increased ability to associate productively with the endogenous tapasin molecules. Whilst this may have only slightly improved the thermal-stability of BF2myc heterodimers, the most significant result of this mutation may have been to substantially increase the proportion of BF2 molecules that gain a peptide repertoire that permits egress from the ER to eventually be expressed at the cell surface.

6.3.3 Failure of Tapasin*12 to alter the expression characteristics of BF2myc*19v1 molecules

In light of the beneficial effect that the mutations had on the expression characteristics of BF2myc*19v1 molecules, it was anticipated that the transfection and expression of Tapasin*12 molecules would produce similar improvements in the expression of wild-type BF2myc*19v1 molecules. However, the expression of Tapasin*12 proteins did not alter the maturation characteristics of BF2myc*19v1 molecules.

As discussed in previous chapters, it is possible that in these transfectants the transfected V5 epitope-tagged Tapasin*12 molecules were prevented from performing functions that Tapasin*12 molecules normally provide in MHC homozygous wild-type cells. For example, the epitope tag may have prevented the transfected tapasin protein from associating with the endogenous TAP molecules; in addition, it is possible too few transfected molecules were expressed to competitively

displace the endogenous Tapasin*15 molecules from PLCs; alternatively the polymorphic amino acids in Tapasin*12 may have prohibited an association with the endogenous TAP molecules. However, it is also possible that the transfected BF2myc*19v1 molecules did associate with the endogenous Tapasin*15 molecules but were largely unable to become loaded with peptides, similar to the non-productive association of murine tapasin and HLA-B*4402 molecules (Peh et al., 1998). Therefore the transfected Tapasin*12 molecules may have been unable to displace sufficient amounts of the BF2myc*19v1 molecules from the endogenous Tapasin*15 molecules to alter the expression characteristics of BF2myc*19v1 molecules.

Alternatively it is possible that the Tapasin*12 molecules provided no additional functions that were not provided by endogenous Tapasin*15 molecules, and that the BF2myc*19v1 molecules associated comparably with either tapasin allele. The beneficial effect that the mutations had on the expression of BF2myc*19v1 molecules may have been the result of increasing the ability of BF2myc*19v1 molecules to associate with, and profit from, the action of any tapasin molecules.

6.3.4 Mutation of human tapasin did not alter HLA-B*4402 surface expression levels

In addition to the experiments described above that were conducted using chicken cells, an attempt was made to assess whether the residues that differ between chicken Tapasin*12 and Tapasin*15 alleles directly affect the expression of class I molecules. Due to the absence of chicken tapasin-deficient cell lines, the human tapasin-deficient .220B*4402 cell line was transfected with a mutant human tapasin protein. However, the five residues that were mutated in human tapasin did not affect the expression of HLA-B*4402 molecules at steady state; with the mutant human tapasin molecule functioning comparably to the wild-type human tapasin. This suggests that the mutated tapasin residues were not involved in specific interactions with the human class I molecule. However, it must be noted that the mutations to the human tapasin protein were not always non-conservative, and a more pronounced effect may have become apparent with such mutations. In addition other assays (such as culture with BFA) may have revealed a more prominent phenotype.

The increased chance of a recombination event separating specific tapasin and class I alleles in the human MHC, in comparison to the chicken MHC, means it is likely that human class I and tapasin proteins have not co-evolved as polymorphic molecules. In contrast, the potential co-evolution of chicken tapasin and class I molecules may mean the interaction between these proteins is specifically dependent on a combination of polymorphic residues.

6.3.5 Summary

This study identified naturally polymorphic residues of BF2 molecules that altered their maturation and thermal-stability characteristics, when expressed in a MHC homozygous cell line. This is likely to reflect alterations in the ability of the class I molecules to profit from the action of the endogenous tapasin molecules. Similar to mammalian class I molecules, tapasin-dependent

peptide-loading appeared to be influenced by amino acid residues at positions 126 and 220, within loops in the α2 and α3 domains that are likely to associate with tapasin directly, and indirectly influenced by amino acids at positions 111 and 113, within the peptide-binding groove, which may affect the conformational signature that tapasin may recognise (Wright et al., 2004; Elliott and Williams, 2005).

The results that were obtained are consistent with haplotype-mismatched BF2myc*19v1 molecules benefiting to a lesser extent from an interaction with endogenous Tapasin*15 molecules than haplotype-matched BF2myc*15 molecules. This led to sub-optimal peptide-loading, which resulted in decreased stability of β_2m-associated heterodimers, and a smaller proportion of BF2myc*19v1 molecules gaining a peptide cargo that permitted egress from the ER. The replacement of residues 126 and 220, or 111 and 113, with the equivalent residues from the BF2myc*15 allele appeared to increase the ability of the BF2myc*19v1 mutant molecule to profit from the action of Tapasin*15 molecules. This led to the peptide repertoire conferring slightly improved stability to the molecules, with substantially more molecules egressing beyond the ER. The exact molecular explanation for such an improvement in peptide-loading was not determined in this study, but may reflect an increased or restored ability of the mutant class I molecule to associate with, or become loaded with peptides by Tapasin*15 molecules. In future experiments it would be of interest to determine whether wild-type and mutant BF2myc*19v1 molecules associate comparably with Tapasin*15 molecules; such research is currently limited by the absence of appropriate serological reagents.

If we now consider the findings of these experiments in the context of the presumed evolution of the B19v1 haplotype; there is circumstantial evidence that in the original recombinant haplotype, BF2*15 molecules were compromised to some extent by impairment in the ability of endogenous Tapasin*12 molecules to facilitate their peptide-loading. In the original recombinant haplotype, BF2*15 molecules may have exhibited expression characteristics similar to the haplotype-mismatched BF2myc*19v1 molecules in this study. If the ability to present peptides was affected more severely, then the recombinant haplotype would have been likely to have been rapidly lost from the population. Any new tapasin or class I mutants that subsequently appeared in the recombinant haplotype that led to more optimal and efficient peptide-loading are likely to have been positively selected. The replacement of aspartic acid 126 and glutamine 220 with glycine 126 and arginine 220 in the class I molecule expressed by the recombinant haplotype may reflect the fine-tuning of the class I antigen presentation process driven by an evolutionary advantage of matching tapasin and class I alleles.

6.3.6 Future research aims

It would be of interest to compare the same two BF2myc alleles when they are expressed in a B19v1 cell line. This would determine whether BF2myc*15 molecules would show the defects that were apparent for BF2myc*19v1 molecules when expressed in the B15 cell line. If such

observations were made, this would imply that the BF2myc*15 allele was specifically dependent on the Tapasin*15 allele in order to become optimally loaded with peptides, and was less able to function with the Tapasin*12 allele in the B19v1 cell line. In contrast, if BF2myc*15 molecules demonstrated superior expression characteristics to BF2myc*19v1 molecules when expressed in a B19v1 cell line, this would suggest that BF2myc*15 molecules profited from the action of whichever tapasin molecules are present to a greater extent than BF2myc*19v1 molecules. In addition it would be of interest to determine whether the transfection of Tapasin*15 had an effect on expression characteristics of BF2myc*15 molecules in these cells. Unfortunately attempts to produce stable or transient transfectants of B19v1 homozygous cell lines have so far proven unsuccessful.

7. Final discussion

7.1 Aims of the project

The chicken MHC has been the focus of intensive research for many years; whilst much of this research has followed and extended the dogma that may have been anticipated from the study of typical mammals, other features of the chicken MHC are significantly different to what is often assumed to be the "norm". For example, the presence of CD1 and NK-receptor genes in the chicken MHC would not have been expected, and may represent the primordial MHC organisation. In addition, the striking association of particular chicken MHC haplotypes with disease resistance or susceptibility is in contrast to the generalised pathogen protection that is conferred by the multigene family of MHC molecules expressed by typical mammals.

The chicken MHC is perhaps the smallest and simplest MHC to be described in detail to date, and exhibits an extremely low rate of recombination. In light of this, and of the close proximity of the chicken tapasin gene to the polymorphic TAP and class I genes, it is likely that these genes have co-evolved. Therefore in order to investigate the nature of such a possibility, the first aim of this project was to determine the extent to which chicken tapasin varies in different haplotypes.

In the event that tapasin were polymorphic, the second aim of the project was to characterise whether the polymorphic amino acid residues in tapasin influenced the association with the class I and TAP proteins in a haplotype-specific fashion. The limited opportunity for recombination to disrupt alleles of tapasin, TAP and class I may allow these genes to co-evolve, perhaps to the point that these proteins are only able to interact when all are encoded by alleles from the same haplotype. Such a possibility would prevent class I molecules from associating with TAP molecules with inappropriate peptide-translocating specificities.

The third aim of the project was to detail whether polymorphic amino acid residues in tapasin otherwise affect the ability of tapasin to function. For example it is possible that particular tapasin alleles facilitate peptide-loading to a greater or lesser extent than other alleles. This may mean that "optimal" peptide-loading (i.e. those peptides of the appropriate length and specificity to bind to the class I molecule with high affinity and perhaps more crucially a slow dissociation rate) proceeds to differing extents in different haplotypes.

7.2 Chicken tapasin is highly polymorphic and moderately divergent in sequence

To determine the extent of allelic polymorphism and sequence diversity in chicken tapasin, the gene was sequenced from each of the seven MHC haplotypes that are available at IAH Compton.

In total 31 polymorphic nucleotide positions were found in the protein-coding region of the tapasin gene in the seven MHC haplotypes. Whilst 16 of these polymorphic nucleotides positions did not change the selection of amino acid in chicken tapasin, 15 polymorphic nucleotide positions led to an amino acid substitution. The chicken tapasin proteins that are expressed in these seven

haplotypes are distinct from each other (apart from one haplotype which is the product of a recombination event), and differ from a consensus protein sequence by up to seven amino acids. It is likely that chicken tapasin is as polymorphic as the chicken class I genes, in that each haplotype is likely to express distinct tapasin and class I alleles. Therefore the first two major findings of this project are that chicken tapasin is highly polymorphic, and that the products of these alleles are moderately divergent in sequence. This contrasts with the situation in typical mammals such as humans, where there are two alleles of human tapasin, which differ by only a single amino acid.

Given the low rate of recombination within the chicken MHC, any new tapasin allele that rises in gene frequency and is maintained in the population through any mechanism (including natural selection, genetic drift or hitch-hiking) is unlikely to be exchanged by recombination and potentially lost from the population. Therefore the presence of polymorphic amino acid residues does not automatically imply that they have been positively selected because they conferred an improved ability to function. However, it remains possible that the products of different tapasin alleles possess differential abilities to function, and that there is a selective pressure to maintain a high number of alleles in the population. In support of the possibility that polymorphic residues in tapasin can lead to functional changes in the ability of the protein to function is a murine tapasin mutant (H334F & H335Y), which has an enhanced ability to facilitate the peptide-loading of H-2L^{d} molecules (Turnquist et al., 2001).

If the polymorphisms in tapasin do alter the ability of the protein to function, it would be expected that any polymorphism that leads to a deleterious amino acid replacement would be under selective pressure to be lost from the population, although this may be a slow process (due to MHC heterozygosity, and the low rate of recombination). In addition, it is possible that amongst the products of different chicken tapasin alleles, there is one (or more) "superior" protein that enhances peptide-loading to a greater extent than others, that may be capable of associating with TAP and class I molecules irrespective of haplotypes. Such potential variance in the ability of tapasin to facilitate peptide-loading between different haplotypes may have been selected by unknown pressures (such as pathogens) experienced by different haplotypes. Whilst it may be expected that any such tapasin allele that has a superior ability to function would eventually become fixed throughout the population, this may have yet to occur due to the limited opportunity for recombination to allow a superior tapasin allele to replace other tapasin alleles.

However, it is possible that despite the low rate of recombination within the chicken MHC that gene conversion can operate. If gene conversion were to occur between the tapasin alleles present in an MHC heterozygous individual, then if the tapasin alleles were functionally neutral, this may be expected to lead to one allele (or a superior tapasin allele, if there is one) eventually fixing throughout the population. With the caveat that the occurrence of gene conversion within the

chicken MHC is unknown, the presence of multiple tapasin alleles suggests the polymorphisms in tapasin alleles are not (all) functionally neutral.

In order to assess the potential significance of the polymorphisms in chicken tapasin, the tapasin sequences were analysed by assessing the ratio of nonsynonymous to synonymous nucleotide substitutions and by phylogenetic analysis. The ratio of nonsynonymous to synonymous nucleotide substitutions suggests that tapasin is not under strong selective pressure to be divergent, and may be under pressure for the majority of the protein to remain invariant. This may be necessary for forming the framework of the protein, or for forming (invariant) interaction sites with other proteins. However, a phylogenetic analysis of the tapasin protein sequences suggests that the polymorphic residues are unlikely to have accumulated simply through evolutionary drift, which is thought to be the case for most genes encoded within the chicken MHC that are not under strong selective pressure (J. Kaufman, personal communication). Additionally the polymorphic residues in tapasin do not appear to be simply the result of genetic hitch-hiking driven by the neighbouring class II major BLB2 gene that is likely to be under strong selective pressure. Therefore these analyses suggest the polymorphisms in tapasin are not easily assigned to either natural selection or to neutrality (where polymorphisms may accumulate through mechanisms such as evolutionary drift and genetic hitch-hiking). It is therefore likely that some neutral polymorphisms are maintained in the population by a variety of mechanisms, whilst other polymorphisms are maintained because they confer a selectable advantage to individuals carrying that tapasin allele.

7.3 Analysis of the nature and distribution of polymorphic amino acid residues in chicken tapasin

All of the 15 amino acid substitutions in chicken tapasin are dimorphic. The first polymorphic amino acid is located within the putative signal sequence (serine/phenylalanine at position 15 in the primary protein sequence) which may affect signal sequence cleavage. Additionally, there is a single polymorphic amino acid within the transmembrane and cytoplasmic domains (arginine/glutamine at residue 421 in the primary protein sequence), which may indicate that the interaction between tapasin and TAP is mediated largely through invariant residues. The remaining 13 polymorphic amino acids are located in the ER luminal domains.

One of the polymorphic amino acids in the ER luminal domains of chicken tapasin (arginine/glutamine at residue 246 in the primary protein sequence) occupies an equivalent position to the single amino acid dimorphism in human tapasin (arginine/threonine at residue 260 in the primary protein sequence). The coincidence of the location and nature (basic/polar) of this substitution may imply a functional significance for this polymorphism in both proteins; however no functional difference has been ascribed to either human tapasin allele to date.

Additionally the absence of an unpaired cysteine equivalent to cysteine 95 of the mature mammalian tapasin proteins was found not to be a polymorphism specific for the Tapasin*12 protein that had previously been sequenced. This suggests that ERp57 may be recruited to the avian PLC through other mechanisms.

The other polymorphic amino acids in chicken tapasin are largely scattered throughout the ER luminal domains. However, there is a prominent clustering of polymorphic amino acids at the base of the membrane proximal Ig domain that is encoded by exon 5, which may influence the interaction between tapasin and either the class I or TAP molecules. Importantly, these clustered polymorphic residues are located in a similar location to the residues that were mutated in murine tapasin that led to enhanced peptide-loading of H-2L^{d} molecules (Turnquist et al., 2001).

Exactly how tapasin facilitates peptide-loading is unknown. However, the absence of clustered polymorphic amino acids at the N-terminus of chicken tapasin suggests that this domain of tapasin is unlikely to directly transfer peptides into the peptide-binding groove, or physically occupy the peptide-binding groove of class I molecules; because a greater extent of polymorphism would have been expected to be necessary to allow this domain of tapasin to bind peptides suitable for, or physically occupy, peptide binding grooves which have distinct peptide-binding requirements.

7.4 Tapasin may be involved in a haplotype-specific interaction with the class I, but not the TAP molecules

Having defined the extent of allelic polymorphism and sequence diversity in the chicken tapasin gene, efforts were made to determine whether the polymorphic amino acids in tapasin alter the ability of the protein to function, via tapasin participating in a haplotype-specific interaction with the class I and TAP proteins. The first approach taken involved a comparison of the topologies of phylogenetic trees constructed for MHC-encoded proteins. The second approach involved a comparison of their polymorphic amino acid residues in order to identify patterns of complementary polymorphic amino acids, which may mediate a haplotype-specific interaction.

The comparison of phylogenetic tree topologies suggested that the major class I (BF2) and TAP2 proteins may be co-evolving with tapasin and participating in a haplotype-specific interaction. Although in each case (tapasin and BF2, and tapasin and TAP2) the topologies did not correlate perfectly, which is likely to indicate that not all of the polymorphic residues are involved in mediating such an interaction.

However, the suggestion that tapasin and TAP2 are involved in a haplotype-specific interaction is not directly supported by an analysis of the distribution of polymorphic amino acid residues in tapasin and TAP2, assuming that chicken tapasin and TAP2 interact through the same sites as the mammalian proteins (i.e. the transmembrane and cytoplasmic domains of tapasin interact with the

N-terminal domains of the TAP proteins). In chicken tapasin there is only a single polymorphic residue in the transmembrane and cytoplasmic domains, whilst in TAP2 there are only three polymorphic residues located in the predicted tapasin-binding N-domain. Whilst the clustered polymorphic amino acid residues at the base of the Ig domain encoded by exon 5 of tapasin may be ideally placed to influence an interaction with the TAP heterodimer, no correlations were observed between these residues (or any others in tapasin) and residues located in the transmembrane or ER luminal domains of either TAP protein. Therefore it appears that chicken tapasin is not involved in a haplotype-specific interaction with the TAP proteins, and that the product of any chicken tapasin allele may interact with the TAP proteins of any haplotype.

However in the case of the tapasin and BF2 proteins, the abundance of polymorphic residues in the ER luminal domains of tapasin may allow tapasin to interact with the BF2 protein in a haplotype-specific fashion, as predicted by the phylogenetic tree analysis, although no simple correlations of polymorphic resides were apparent between the two proteins that were easily interpreted as functionally significant. This suggests that multiple polymorphic amino acids may collectively contribute to mediating a haplotype-specific interaction between these proteins.

Therefore these approaches have led to the prediction that most, but possibly not all, of the polymorphic amino acid residues in chicken tapasin do alter the ability of the protein to function, via tapasin participating in a haplotype-specific interaction with class I molecules, but not with the TAP proteins. As there is a selective advantage in being MHC heterozygous, and the chicken TAP, tapasin and class I genes are highly polymorphic, there are likely to be two alleles of tapasin, TAP and class I expressed by most individuals. If the above predictions are true, and that tapasin is capable of associating with TAP proteins irrespective of haplotype, but can only associate with class I molecules encoded by the same haplotype (or associates most productively with the haplotype-matched class I allele), this means that class I molecules are likely to associate, through tapasin, with TAP molecules which may have an inappropriate transport specificity. A precedent for such an interaction is the non-productive association of particular rat class I (such as RT1.A^{a}) molecules with TAP2B containing transporters that is evident in certain MHC homozygous recombinant rat haplotypes that characterise the "cim" phenomenon. However, such pairing of chicken class I and TAP molecules with disparate peptide-binding and peptide-translocating requirements may be permitted in MHC heterozygous chickens, because there is likely to be appropriate peptides circulating within the ER lumen that are available for tapasin-mediated peptide-loading, that are supplied by the other TAP heterodimer which has an appropriate transport specificity and is encoded from the same haplotype as the class I molecule.

7.5 Attempts to demonstrate that tapasin is involved in a haplotype-specific interaction with class I molecules

Having established the polymorphic nature of chicken tapasin and predicted that some of the polymorphic amino acids in tapasin are involved in mediating a haplotype-specific interaction with the class I protein, the next aim of the project was to functionally demonstrate such an interaction. This was attempted using three different transfection strategies which involved creating a haplotype-mismatching of tapasin and class I alleles. In these strategies the expression characteristics of a haplotype-mismatched pair of alleles was compared with a pair of haplotype-matched alleles. However, these approaches were complicated by the absence of chicken cell lines that are deficient for tapasin or class I molecules, and by the limited number of serological reagents that are available for chicken proteins.

The most successful strategy involved the transfection of a B15 homozygous chicken cell line with the haplotype-mismatched BF2*19v1 class I allele; a combination of haplotypes which was guided by the hypothesized sequence of events that led to the generation and evolution of the B19v1 haplotype. This haplotype is thought to be the product of a recombination event which occurred between the B12 and B15 haplotypes (or at least a haplotype very similar to the modern B15 haplotype), with the recombination event separating the tapasin and BF2 loci. However, the BF2*19v1 allele that is present in the haplotype now is distinct from the BF2*15 allele which is likely to have been present in the original recombinant haplotype. As the B19v1 haplotype expresses an identical allele of tapasin to the B12 haplotype (Tapasin*12), and the BF2*19v1 allele shares certain features with the BF2*12 allele, including residues that are located in the predicted tapasin-binding regions, it was hypothesized that following the recombination event the BF2*15 molecules were compromised to some extent in their ability to be optimally loaded with peptides by the Tapasin*12 molecules, and that subsequently the BF2*15 allele evolved into the BF2*19v1 allele, which maintained essentially the same peptide-binding specificity, but gained an improved ability to associate productively with Tapasin*12 molecules.

Evidence for this hypothesis was sought in two stages: firstly, by establishing a situation closely related to that of the original recombinant haplotype, in which a potential incompatibility between tapasin and BF2 molecules may have existed, which could lead to deficiencies in the loading of the haplotype-mismatched class I molecule with peptides; and secondly, by determining whether reproducing the presumed molecular evolution that led to the BF2*15 allele evolving into the BF2*19v1 allele, rectified any defects that were apparent in the first situation.

Analysis of the transfectants which were created to reproduce a similar molecular situation to that which was likely to be present in the original recombinant haplotype showed that the haplotype-mismatched BF2*19v1 molecules were compromised to some extent in their ability to be loaded with peptides. The cause of which was likely to be an incompatibility with the endogenous

Tapasin*15 molecules. However the phenotype was relatively mild, which may have been anticipated, because if a haplotype-mismatched combination of alleles were totally incapable of working together, and the haplotype-mismatched class I allele could not self-optimise its peptide repertoire to any extent, then it is likely that individuals homozygous for the recombinant haplotype would be lost from the population. However, it is possible that even a mild class I peptide-loading deficiency may still produce a significant immunological deficiency.

In the second part of the experimental strategy, mutations were made to the haplotype-mismatched BF2*19v1 allele, which were designed to reproduce some of the functionally important molecular changes that led to the BF2*19v1 allele evolving from the BF2*15 allele. Two residues of the haplotype-mismatched BF2*19v1 allele (positions 126 and 220) located within the predicted tapasin-binding α2 and α3 loops were exchanged for the equivalent residues of the haplotype-matched BF2*15 allele. The BF2*19v1 126 & 220 mutant demonstrated improvements in its expression characteristics, which is likely to reflect an improved ability for the mutant class I molecules to associate productively with the endogenous Tapasin*15 molecules, however the exact molecular mechanism that brought about this improvement was not determined in this study. In agreement, the reciprocal mutations that were made to the haplotype-matched BF2*15 allele, which were anticipated to decrease the ability to associate productively with tapasin, appeared to modestly decrease the stability of the BF2*15 mutant heterodimer. With the recombinant B19v1 haplotype in mind, these findings suggest that the evolution of the BF2*15 allele into the BF2*19v1 allele led to an improvement in the Tapasin*12-mediated peptide-loading of the class I molecules, and that this evolution represents fine-tuning of the interaction between the class I and tapasin proteins, with individuals carrying the "evolved" B19v1 haplotype likely to have an evolutionary advantage over individuals carrying the original recombinant haplotype.

It is however possible that the residues present at positions 126 and 220 (glycine and arginine) in the haplotype-mismatched BF2*19v1 molecule confer an inherently poor ability to productively associate with <u>any</u> tapasin molecule (rather than just the Tapasin*15 molecule). Whilst it would be expected that if a class I mutant were to appear in which G126 and R220 were replaced by residues that associate with tapasin more productively, that this mutant would be positively selected (based upon there being an evolutionary advantage to tapasin-mediated peptide-loading, as opposed to self-optimisation in the absence of tapasin); this may have not yet occurred because of the low rate of recombination within the chicken MHC. Alternatively, it is possible that in the B19v1 haplotype, there is a selective advantage to individuals carrying the "evolved" B19v1 haplotype over the original recombinant haplotype, which is based upon the BF2*19v1 allele being less dependent upon tapasin to facilitate peptide-loading. In support of this possibility are certain mammalian class I alleles (such as HLA-B*2705 and HLA-B*4405) that are less dependent upon tapasin to facilitate peptide-loading. Typically, mammalian tapasin-"independent" alleles can self-optimise their peptide repertoires, but benefit from tapasin-mediated peptide-loading to a lesser extent than

tapasin-"dependent" alleles, which are however unable to self-optimise their peptide repertoires (Williams et al., 2002). Such tapasin-independent alleles are thought to have been maintained in the population because they are less susceptible to pathogen subversion of PLC-mediated peptide-loading. However, the occurrence of glycine (position 126) and arginine (position 220) in the BF2 alleles present in both the B12 and B19v1 haplotypes, which share the same tapasin allele, suggests this explanation is unlikely (although not impossible). Instead, it appears more likely that the BF2 allele in the recombinant haplotype has evolved, through positions 126 and 220, to specifically interact with the Tapasin*12 molecules, conferring an advantage to individuals carrying the "evolved" haplotype.

In total there are four different combinations of residues that are present in the tapasin-binding α2 loop, and two different combinations of residues in the α3 loop, which suggests the nature of the interaction between chicken tapasin (which is highly polymorphic and diverse in sequence) and the chicken class I molecules varies in different haplotypes. In contrast in humans, the interaction between the tapasin and class I molecules is likely to be mediated by invariant residues, as there is only a single dimorphism at position 131 (arginine/serine) in the class I molecule, and there are only two tapasin alleles, that differ by a single amino acid at position 240 (arginine/threonine).

α2 tapasin-binding loop

(chicken residues 125-133)

BF2B2	KGTMTFTAA
BF2B4	.DMK.....
BF2B12	....
BF2B14	....
BF2B15	.D.......
BF2B19	....
BF2B21	.....L...
Human-1	EDLRSW...
Human-2	EDLSSW...
Murine	EDLKTW...

(mammalian 128-136)

α3 tapasin-binding loop

(chicken residues 218-223)

BF2B2	AVRGQD
BF2B4	......
BF2B12	......
BF2B14	......
BF2B15	..Q...
BF2B19	......
BF2B21	......
Human	EDQT..
Murine	EDLTE.

(mammalian 222-227)

During the course of these transfection studies, a second BF2*19v1 mutant molecule was produced by replacing residues 111 and 113 with the equivalent residues from the haplotype-matched BF2*15 allele. These residues are located surrounding the F pocket within the peptide-binding groove, and the equivalent residues of mammalian class I molecules (residues 114 and 116) influence the level of dependence that the class I mutants rely upon tapasin to facilitate optimal peptide-loading (Williams et al., 2002; Park et al., 2003). This mutant demonstrated a similar improvement in maturation characteristics to that of the BF2*19v1 126 & 220 mutant. Again, the mechanism of the improvement was not determined in this study, but may involve the induction of a conformation of the class I molecule that is more suitable for productive tapasin association, or may create a structure that benefits to a greater extent from the peptide-editing function of tapasin.

Analysis of both BF2*19v1 mutants (126 & 220, and 111 & 113) suggests the nature of the interaction between chicken tapasin and class I molecules is largely the same as has been determined in mammalian species, and may be influenced directly by some residues (126 & 220), and indirectly by others (111 &113). The complexity of the interaction between tapasin and class I molecules supports the absence of an obvious correlation of polymorphic amino acid residues between the tapasin and BF2 proteins that are likely to mediate a haplotype-specific interaction.

Therefore, the two major findings of this part of the project are: firstly, the identification of naturally polymorphic residues of the chicken class I molecule which influence the ability of the class I molecule to associate productively with tapasin (the interaction between mammalian tapasin and class I proteins has previously been determined using class I mutants); and secondly that there is evidence that chicken tapasin and class I alleles have co-evolved to mediate optimal peptide-loading within a haplotype.

7.6 Discussion of the significance of tapasin participating in a haplotype-specific interaction with the class I molecule

In this study evidence is presented that the polymorphic residues in tapasin are involved in mediating a haplotype-specific interaction with the class I protein. This implies that there is an advantage in peptide-loading from the association of tapasin and class I molecules that are both encoded by genes of the same haplotype. A haplotype-specific interaction may lead to the same objective being achieved in different haplotypes, that of loading of high affinity peptides which ultimately produces stable peptide-loaded class I molecules, with the nature of the interaction between the tapasin and class I proteins varying in different haplotypes. Alternatively, a haplotype-specific interaction may achieve different results in different haplotypes. For example, in some haplotypes a haplotype-specific interaction may result in the loading of a large repertoire of peptides, allowing a high level of surface expression. In other haplotypes, a haplotype-specific interaction may lead to the loading of a restrictively small pool of peptides, which may lead to a low level of surface expression. Therefore this may explain the haplotype-determined hierarchy of

cell surface expression levels of class I molecules, and may explain why some haplotypes differ in resistance to pathogen infections. Such differences in peptide-loading may have been selected by the pressures of pathogen selection experienced by the different haplotypes.

If tapasin is involved in a haplotype-specific interaction with the class I molecules, this may mean that tapasin proteins are only able to associate with class I molecules that are encoded by the same haplotype, which could be referred to as a "restrictive" haplotype-specific interaction. This would lead to tapasin only being able to facilitate the peptide-loading of the class I molecule that is encoded by the same haplotype. Alternatively, it is possible that the specificity of such an interaction is not absolute, and that tapasin can facilitate the peptide-loading of haplotype-mismatched class I alleles, which could perhaps be referred to as a "permissive" haplotype-specific interaction. However, due to the potential for co-evolution to occur between haplotype-matched alleles of tapasin and BF2, peptide-loading of mismatched alleles may occur less efficiently than for haplotype-matched pairs of alleles.

If tapasin is involved in a permissive haplotype-specific interaction with the class I molecules, the potential for tapasin to (partially) facilitate the peptide-loading of mismatched class I alleles may contribute to the generation of new haplotypes. For example, if a recombination event were to separate tapasin and class I alleles, such as is likely to have occurred in the B19vI haplotype, then a complete inability to facilitate the peptide-loading of mismatched class I molecules that are unable to self-optimise their peptide repertoires is likely to act as a recessive lethal mutation. If an individual were to be homozygous for such a recombinant haplotype, this is likely to lead to diminished class I antigen presentation, and to a selective pressure against the recombinant haplotype.

Additionally, if the specificity of tapasin participating in a haplotype-specific interaction with class I molecules is not absolute, and that the products of different tapasin alleles can facilitate the peptide-loading of mismatched alleles of class I molecules to some extent, this may increase the diversity of peptides that are presented in MHC heterozygous individuals. If we consider the following examples:

Situation	Tapasin alleles	BF2 alleles	TCR repertoire in the periphery & comments
I	**A/A**	**X/X**	Recognises BF2*X and peptides presented by Tpn*A
2	**A/B**	**X/Y**	Recognises **both**: BF2*X and peptides presented by Tpn*A, and BF2*Y and peptides presented by Tpn*B. **However**, the repertoire of peptides presented by the BF2*X & Tpn*B combination may be different to that presented by the BF2*X & Tpn*A combination, which may **increase the total diversity** of peptides presented. A similar argument can be made for BF2*Y.

In the first situation, where the individual is homozygous for both the Tapasin*A and BF2*X alleles, the TCR repertoire in the periphery would be specific for the BF2*X molecule and may recognise the peptides that are efficiently presented by the Tapasin*A allele. In the second example, there is the obvious heterozygote advantage in having a TCR repertoire that may be capable of recognising peptides presented by two class I alleles with different peptide-binding specificities. However, there may be an additional advantage in that the diversity of pMHC complexes that are formed for both BF2*X and BF2*Y alleles may be increased if the "mismatched" combination of BF2 and tapasin alleles loaded a slightly different repertoire of peptides to that loaded by "matched" combination of alleles, even if this occurred with lower efficiency. In contrast, in the MHC of typical mammals, a similar level of diversity in pMHC complexes may be maintained despite the distant location and limited polymorphism of tapasin by an increased number of expressed class I genes.

If tapasin is involved in a permissive haplotype-specific interaction this may also help explain why no exact correlations were predicted between the polymorphic residues within individual alleles of chicken tapasin and class I proteins.

7.7 Polymorphisms in chicken tapasin may explain phenomena attributed to the chicken MHC

As briefly discussed above, functional polymorphism in chicken tapasin and other MHC-encoded proteins may explain some attributes of the chicken MHC. For example, the haplotype-determined hierarchy of class I molecules expressed at the cell surface may potentially be explained if the products of different chicken TAP alleles possess differential peptide transport rates as well as different peptide translocation specificities, additionally the products of different chicken tapasin alleles may possess differential abilities to facilitate peptide-loading, as discussed above. Therefore, in haplotypes such as B21 where there is a low level of class I molecules expressed at the cell surface, there may be a scarcity of appropriate peptides that are transported into the ER, or the tapasin-mediated loading of the B21 class I molecules with optimal peptides may proceed more slowly, or to a lesser extent, than is evident in other haplotypes.

In addition, the polymorphic nature of chicken tapasin may explain the haplotype-determined resistance to some pathogen infections. For example, a pathogen protein may bind to chicken tapasin in order to inhibit class I antigen presentation. It is possible that any new tapasin allele that appears which can escape from such subversion, perhaps through mutation of the residues that constitute the site of interaction with the pathogen protein, would be positively selected, and would rise in gene frequency.

7.8 Future studies

In this study much has become apparent about the polymorphic nature of chicken tapasin, and of the naturally polymorphic residues of the chicken class I molecule that influence the ability to benefit from tapasin-mediated optimisation of the peptide repertoire. However, much more remains to be determined. For example the third aim of this project, to characterize other influences that polymorphic residues in chicken tapasin have on class I peptide-loading, was not fulfilled in this study. One approach to address this could be to compare the thermal-stability of class I molecules expressed by cells taken directly from chickens of different haplotypes. This could reveal whether class I alleles differ in their extent of "optimal" peptide-loading, which may be a reflection of tapasin-mediated peptide-loading proceeding to different extents in these haplotypes. An additional, complementary approach that could be taken is to compare the survival of surface expressed class I molecules when these cells are cultured with brefeldin A. Additionally the full extent to which different class I alleles rely upon tapasin to achieve their usual level of cell surface expression remains unknown. To address this issue, it would be of interest to determine the ability with which different haplotype-matched (as well as haplotype-mismatched) combinations of tapasin and class I molecules can associate. A complementary approach that could be conducted in future studies would be to determine the effect that depletion of tapasin (by RNA interference) has on the peptide-loading of different class I alleles.

Whilst the chicken class I, TAP and tapasin genes have been the focus of intensive research, much remains unknown about the biochemistry of the peptide-loading process of chickens. For example the absence of the N-terminal predicted tapasin-binding domain of the chicken TAP1 protein makes it likely that the stoichiometry of the chicken PLC is different to that of typical mammals, but is consistent with the significant role in producing stable pMHC complexes that has been reported for the N-terminal tapasin-binding domain of the mammalian TAP2, but not the TAP1 protein (Leonhardt et al., 2005). In future studies it would be of interest to determine the exact constitution of the chicken PLC (chicken orthologues of calreticulin, calnexin, ERp57 or ER luminal aminopeptidases have yet to be described), and it would be of interest to determine whether there is a functional reason for the loss of the N-terminal domain of TAP1.

Over the last few years it has become increasingly clear that ERp57 may play a role in shaping the class I peptide repertoire. It would be of interest to determine whether the absence of a lone, unpaired cysteine in chicken tapasin (equivalent to cysteine 95 of the mature human tapasin protein) affects the peptide-loading process, which may reveal how ERp57 helps shape the peptide repertoire. For example it appears that a disulphide bridge between tapasin and ERp57 may not be necessary for chicken tapasin to function correctly, unless a disulphide bridge is formed with another cysteine of chicken tapasin. In addition, it is possible that another component of the chicken PLC may recruit ERp57.

7.9 Summary

This study has shown chicken tapasin to be highly polymorphic and moderately divergent in sequence, which contrasts with the situation in typical mammals, such as humans, where there are two alleles of human tapasin, which differ by only a single amino acid. However, the low rate of recombination that characterises the chicken MHC may allow polymorphisms in tapasin to be maintained in the population through a variety of mechanisms, including evolutionary drift, genetic hitch-hiking, as well as the positive selection of alleles bearing beneficial polymorphic amino acid residues.

Therefore several approaches were taken in order to assess the potential significance of the polymorphic residues in chicken tapasin. The first approach involved an analysis of the ratios of nonsynonymous and synonymous nucleotide substitutions, which suggested that chicken tapasin is not under strong selective pressure to be divergent. Despite this, a phylogenetic analysis of chicken MHC-encoded protein sequences suggested firstly that the majority of the polymorphic residues in chicken tapasin are not likely to have accumulated simply through evolutionary drift, as it thought to have occurred for most chicken MHC-encoded proteins that are not under strong pressures of selection; and secondly that the majority of the polymorphic residues in chicken tapasin are not likely to be a result of tapasin being tightly linked to the polymorphic class II major BLB2 gene, which is under strong selective pressure. Therefore, the polymorphic amino acids of chicken tapasin are not easily assigned to either being functionally relevant (and maintained in the population by the pressures of selection), or functionally neutral (and maintained in the population by evolutionary drift, hitch-hiking and other pressures).

However, there is supportive evidence that at least some of the polymorphisms in tapasin may result in alterations to the ability of the protein to function. Firstly, there is the theoretical possibility that gene conversion may operate within the chicken MHC, despite the low rate of recombination. If the tapasin alleles were functionally neutral, gene conversion may be expected to lead to one tapasin allele eventually becoming fixed throughout the population. If gene conversion does occur within the chicken MHC, the presence of multiple tapasin alleles suggests that the polymorphisms in chicken tapasin are likely to be functionally relevant, and maintained for a reason that is beneficial to the haplotype. Secondly, some of the polymorphic residues in chicken tapasin are located in similar locations to mutations made to murine tapasin, which resulted in an enhanced ability to facilitate peptide-loading. Further evidence to support the possibility that the polymorphisms in tapasin result in alterations to the ability of the protein to function came from phylogenetic analysis and an analysis of the polymorphic amino acid residues of chicken tapasin and other MHC-encoded proteins, as well as a functional approach that was taken based upon the findings of these analyses, which are surmised below.

Comparison of the topologies of phylogenetic trees constructed for various chicken MHC-encoded proteins allowed the prediction that the majority of the polymorphic amino acid residues in tapasin mediate a haplotype-specific interaction with the major class I (BF2) and TAP2 proteins. However, the virtual absence of polymorphic residues in the regions of the tapasin and TAP2 proteins that are likely to interact does not support the possibility that tapasin and TAP2 are involved in a haplotype-specific interaction, and therefore suggests that the tapasin and TAP2 proteins can associate irrespective of haplotype. However, the majority of the polymorphic amino acids in tapasin are located within the ER luminal domains, which may allow tapasin to participate in a haplotype-specific interaction with the class I molecule, as predicted by the phylogenetic tree analysis.

In order to functionally demonstrate a haplotype-specific interaction between the tapasin and BF2 proteins a series of transfection strategies were employed. The most successful transfection strategy was based upon recreating the presumed sequence of events involved in the generation and evolution of a natural recombinant haplotype.

The first part of this transfection strategy was designed to replicate the molecular situation that was likely to have been present immediately after the recombination event. In these transfectants, deficiencies in the expression characteristics of haplotype-mismatched BF2*19v1 class I molecules suggested that the haplotype-mismatched class I molecules benefited poorly from endogenous tapasin-mediated peptide-loading. This suggests that following the recombination event that generated the recombinant haplotype, peptide-loading was likely to have been similarly impaired.

The second part of this transfection strategy was designed to reproduce the molecular evolution which is hypothesised to have occurred within the recombinant haplotype, after the recombination event. In these transfectants, the mutations that were made to the haplotype-mismatched BF2*19v1 class I allele improved the expression characteristics, in comparison to wild-type BF2*19v1 molecules. This suggests that in these transfectants the BF2*19v1 mutant molecules benefited to a greater extent from endogenous tapasin-mediated peptide-loading than the wild-type BF2*19v1 molecules. This may suggest that in the recombinant haplotype, the BF2 allele evolved in order to specifically interact with the endogenous tapasin allele. However, an alternate possibility is that in the recombinant haplotype the BF2 allele evolved (into the BF2*19v1 molecule that was used in the first part of the transfection strategy) in order to interact less efficiently with <u>any</u> tapasin allele, which may nevertheless, offer an advantage in times of pathogen subversion of the action of the PLC. However, as the modern B12 and B19v1 haplotypes, which express an identical tapasin protein, also share the same residues at the positions that were mutated in the BF2 alleles in these transfectants; this suggests that both BF2 alleles use these residues to specifically interact with the same tapasin allele. Therefore this part of the transfection strategy functionally demonstrated naturally polymorphic residues of the chicken class I molecule which specifically influenced the interaction with tapasin alleles, and suggested that, following the recombinational event which

generated the recombinant haplotype, tapasin and class I alleles have co-evolved in order to achieve optimal peptide-loading..

In conclusion, based upon the data presented in this thesis, it is unlikely that all of the novel polymorphisms in chicken tapasin are functionally neutral (being maintained in the population because of the low rate of recombination). Instead, it is more likely that some of the polymorphisms in chicken tapasin result in alterations in the ability of the protein to function, leading to a specific interaction with the BF2 alleles that are expressed in the same haplotype.

8. Appendices

8.1 Appendix 1 Tapasin sequencing clones

For each haplotype tapasin was PCR amplified, cloned and sequenced in four steps:

1. The protein-coding region, from cDNA
2. The protein-coding region, from genomic DNA (i.e. introns & exons)
3. The genomic region upstream of exon 1
4. The genomic region downstream of exon 8

For each project, at least two PCR amplifications were performed, with at least one clone fully sequenced per PCR. Where discrepancies existed, additional clones or PCRs were examined.

Clone nomenclature:

Each clone was given a unique description consisting:

expt, line, clone#

e.g. 19A2 = expt19, A = B2 L6, 2 = clone #2

Haplotype	Label
B2 L6	A
B2 L7	B
B4	C
B12	D
B14	E
B15	F
B19 LP	G
B21 L0	H
B21 LN	I

Tapasin cDNA protein-coding sequencing: **Expts 19 & 24.**

- PCR primers to produce clones : c1353 & c469

Haplotype	Expt 19	Expt 24
B2 L6	A2,A3	A1,A4
B2 L7	B1,B3	B1,B2
B4	C1,C2	C2,C3
B12	D1,D2,D4	D2,D3
B14	E2,E3	E1,E3,E4,E5
B15	F1,F3	F2,F4
B19 LP	G1,G2,G3	G1,G2
B21 L0	H1,H3,H4,H5	H1,H2
B21 LN	I1,I3	I3,I4

Tapasin upstream sequencing: **Expts 57 & 63 & 64.**

- PCR primers to produce clones: c1576 & c1578

Haplotype	Expt 57	Expt 63	Expt 64
B2 L6	A1	A1	
B2 L7	B1	B1	B2
B4	C1	C2	
B12	D1	D1	D1
B14	E1	E2	
B15	F1	F1	F3
B19 LP	G1	G1	
B21 L0	H2	H1	
B21 LN	I1	I1	

Tapasin downstream sequencing: **Expts 56 & 70/73 & 71.**

- PCR primers to produce clones: c1571 & c1573

Haplotype	Expt 56	Expt 70/73	Expt 71
B2 L6	A1	70A2	A3
B2 L7	B1	70B2	
B4	C2	70C1	
B12	D1	70D1	
B14	E1	70E2	
B15	F1	73F1	
B19 LP	G2	70G1	G3
B21 L0	H3	70H1	H1
B21 LN	I1	70I1	I3

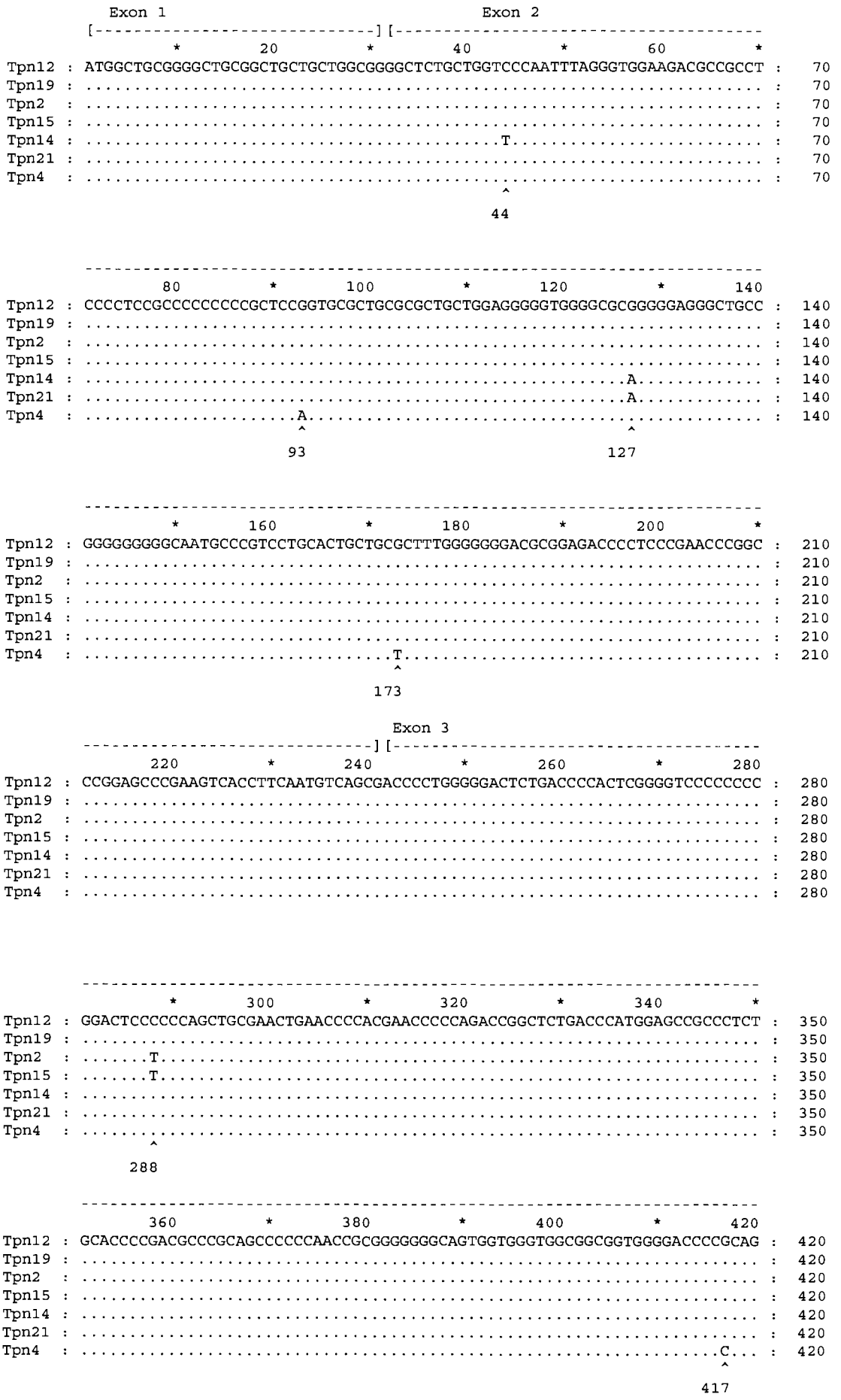
Tapasin genomic protein-coding region sequencing: **Expts 32, 46 mini-preps & maxi-preps, 50, 51.**

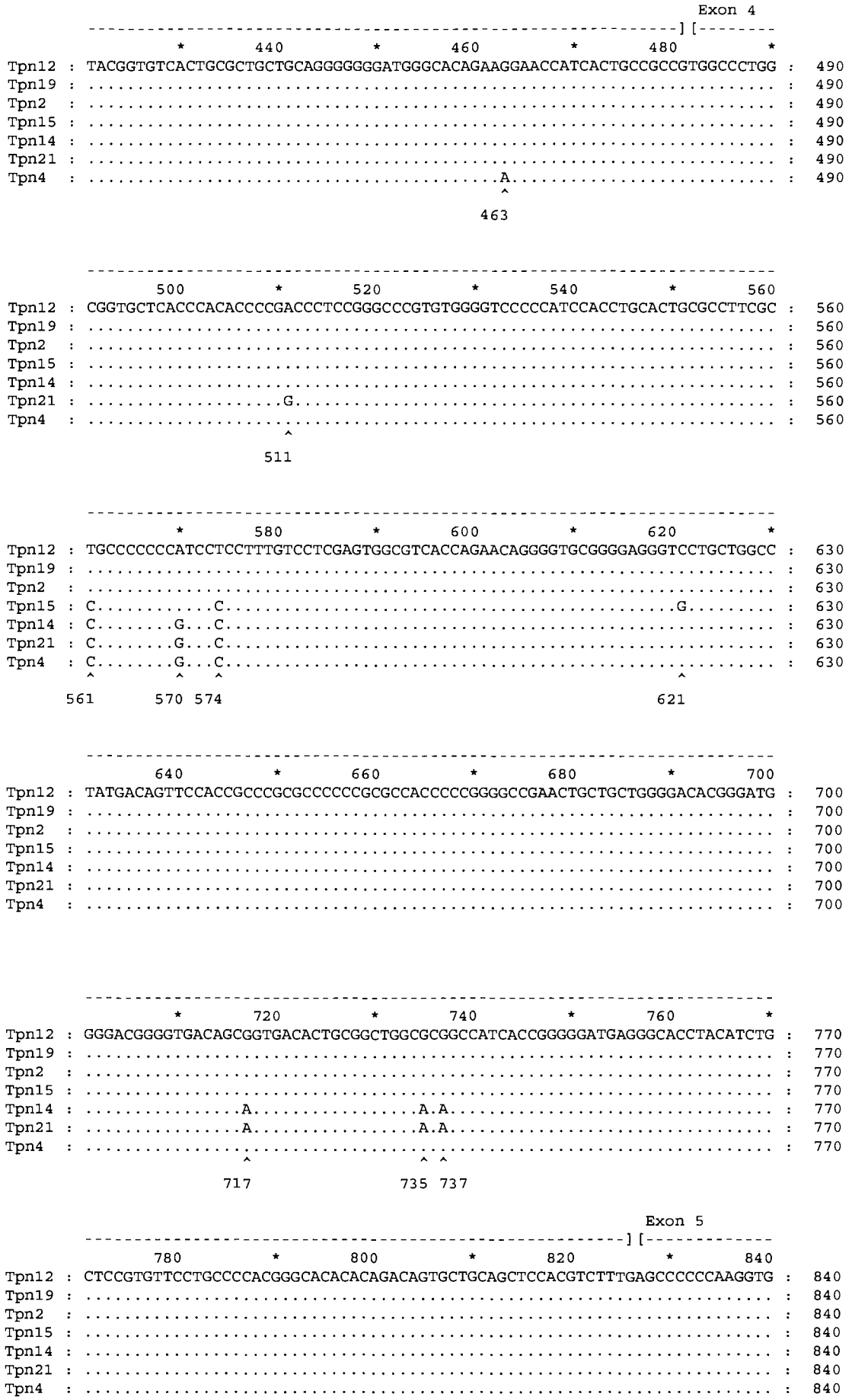
- PCR primers to produce clones: c1353 & c469

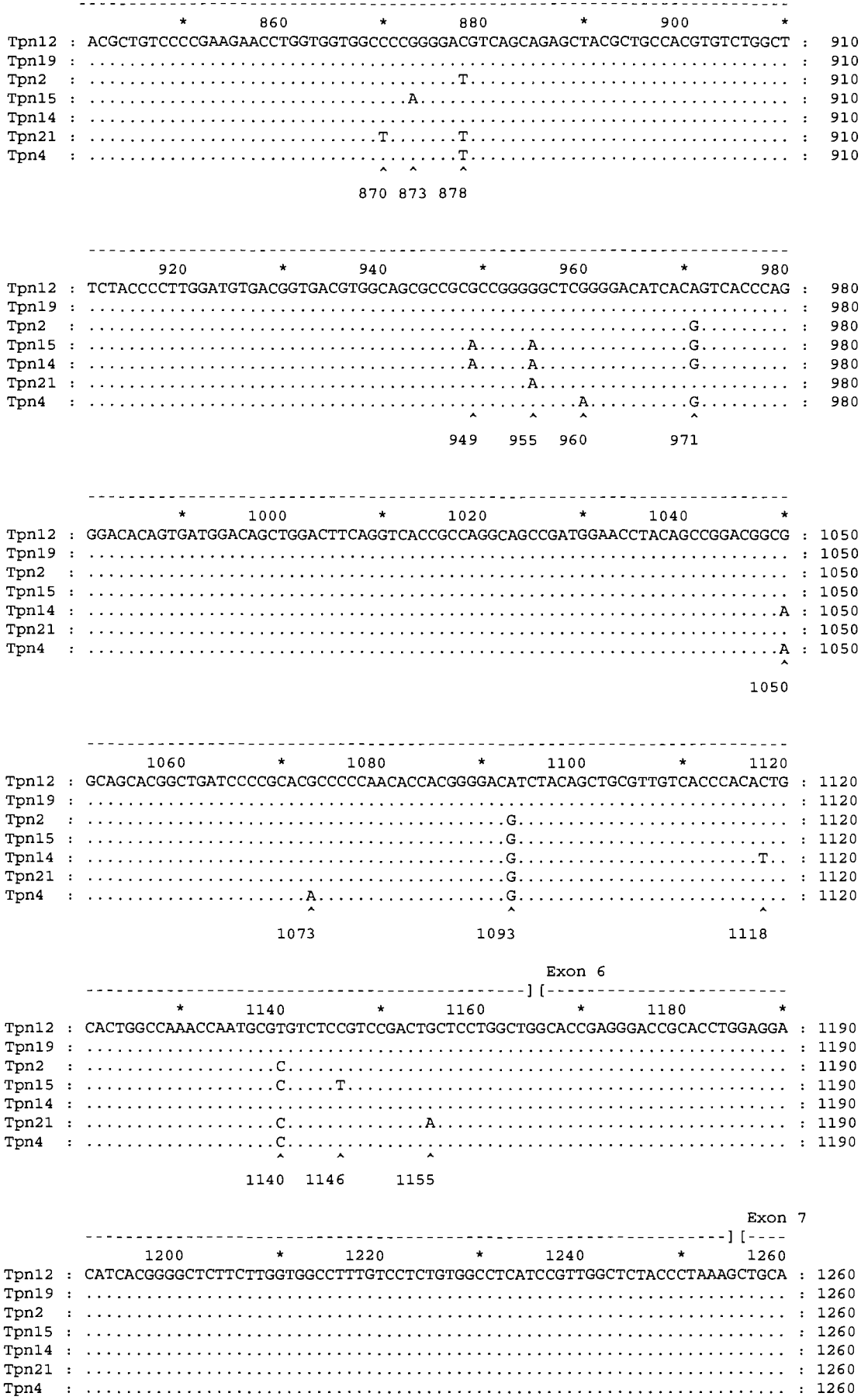
*= partial sequence, † = potentially mixed clone

Haplotype	Expt 32	Expt 46	Expt 50	Expt 51
B2 L6	A1,A2*	A1	A1	A1
B2 L7	B2,B6*	B1*	B6	B2
B4	C2,C10*	C2	C1	C1
B12	D1	D1	D1	D1
B14	E2,E6*	E1	E1	E1
B15	F3,F5*	F4	F1	F2
B19 LP	G1,G2*	G3	G1	G1
B21 L0	H3,H4*	H1	H3	H1*
B21 LN	I1†,I3*	I1	I1*	I1

8.2 Appendix 2 Nucleotide alignment of tapasin coding sequences from seven MHC haplotypes




Exon 4
Exon 5
Tpn12
Tpn19
Tpn2
Tpn15
Tpn14
Tpn21
Tpn4
463
511
561
570 574
621
717
735 737

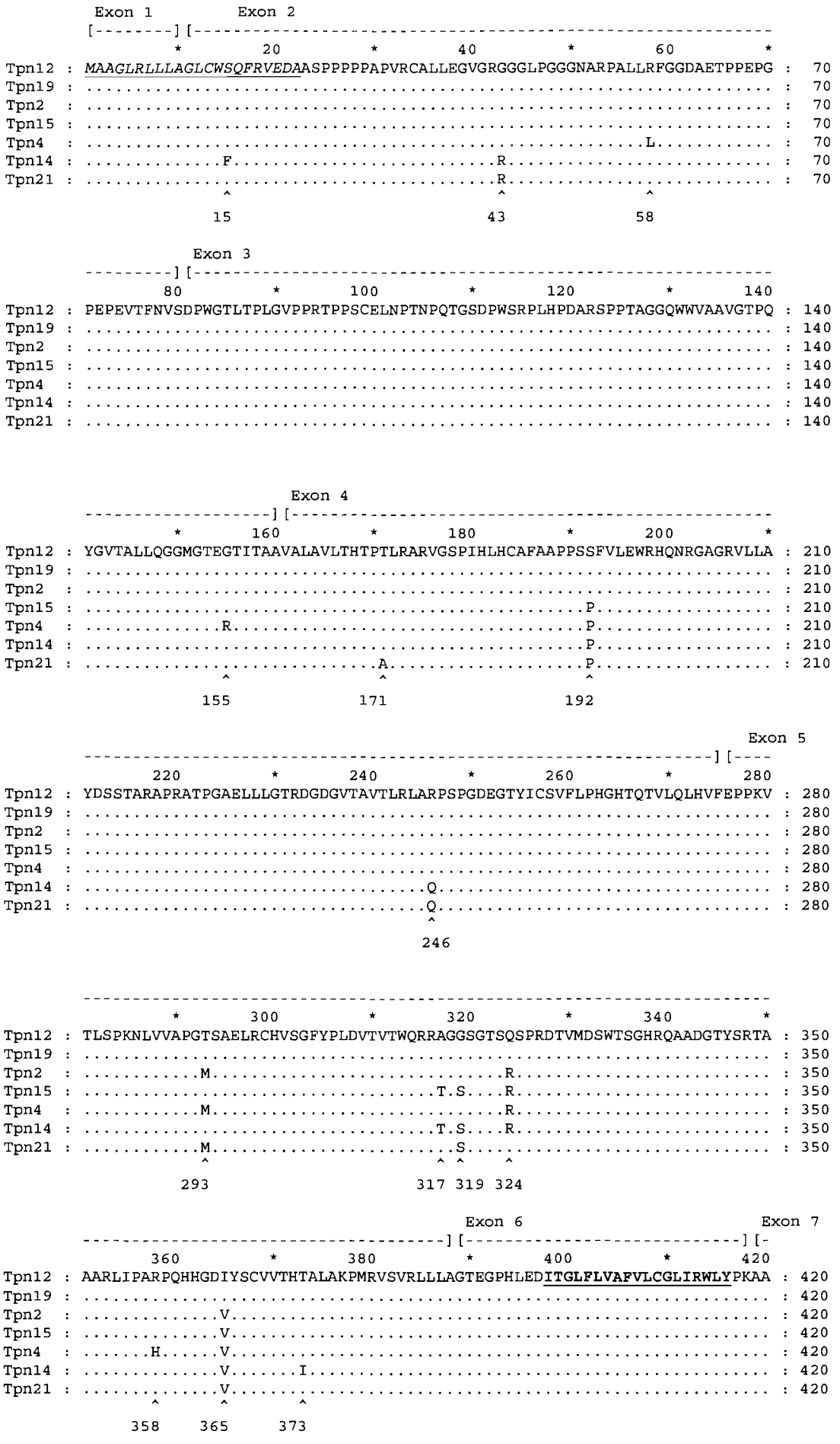

Tpn12 : ACGCTGTCCCCGAAGAACCTGGTGGTGGCCCCGGGGACGTCAGCAGAGCTACGCTGCCACGTGTCTGGCT : 910
Tpn19 : ..... : 910
Tpn2 : .....T..... : 910
Tpn15 : .....A..... : 910
Tpn14 : ..... : 910
Tpn21 : .....T.....T..... : 910
Tpn4 : .....T..... : 910
870 873 878
Tpn12 : TCTACCCCTTGGATGTGACGGTGACGTGGCAGCGCCGCGCCGGGGGCTCGGGGACATCACAGTCACCCAG : 980
Tpn19 : ..... : 980
Tpn2 : .....G..... : 980
Tpn15 : .....A.....A.....G..... : 980
Tpn14 : .....A.....A.....G..... : 980
Tpn21 : .....A..... : 980
Tpn4 : .....A.....G..... : 980
949 955 960 971
Tpn12 : GGACACAGTGATGGACAGCTGGACTTCAGGTCACCGCCAGGCAGCCGATGGAACCTACAGCCGGACGGCG : 1050
Tpn19 : ..... : 1050
Tpn2 : ..... : 1050
Tpn15 : ..... : 1050
Tpn14 : .....A : 1050
Tpn21 : ..... : 1050
Tpn4 : .....A : 1050
1050
Tpn12 : GCAGCACGGCTGATCCCCGCACGCCCCCAACACCACGGGGACATCTACAGCTGCGTTGTCACCCACACTG : 1120
Tpn19 : ..... : 1120
Tpn2 : .....G..... : 1120
Tpn15 : .....G..... : 1120
Tpn14 : .....G.....T.. : 1120
Tpn21 : .....G..... : 1120
Tpn4 : .....A.....G..... : 1120
1073 1093 1118
Exon 6
Tpn12 : CACTGGCCAAACCAATGCGTGTCTCCGTCCGACTGCTCCTGGCTGGCACCGAGGGACCGCACCTGGAGGA : 1190
Tpn19 : ..... : 1190
Tpn2 : .....C..... : 1190
Tpn15 : .....C.....T..... : 1190
Tpn14 : ..... : 1190
Tpn21 : .....C.....A..... : 1190
Tpn4 : .....C..... : 1190
1140 1146 1155
Exon 7
Tpn12 : CATCACGGGGCTCTTCTTGGTGGCCTTTGTCCTCTGTGGCCTCATCCGTTGGCTCTACCCTAAAGCTGCA : 1260
Tpn19 : ..... : 1260
Tpn2 : ..... : 1260
Tpn15 : ..... : 1260
Tpn14 : ..... : 1260
Tpn21 : ..... : 1260
Tpn4 : ..... : 1260

```
                                Exon 8
        ---------------------][----------]
                    *      1280         *
Tpn12  : CGACCCAAAGAGGAAACCAAGAAATCGCAGTGA : 1293
Tpn19  : ................................. : 1293
Tpn2   : ................................. : 1293
Tpn15  : ................................. : 1293
Tpn14  : .A............................... : 1293
Tpn21  : ................................. : 1293
Tpn4   : ........................A........ : 1293
           ^                     ^
       1262                  1287
```

Appendix 2 Nucleotide alignment of tapasin coding sequences from seven MHC haplotypes.

Start to stop cDNA sequences were deduced by analysis of cDNA and genomic clones from seven MHC haplotypes, and aligned using GCG and Genedoc analysis programs. Nucleotide identity with the B12 sequence is represented by dots (.). Residues are numbered according to position in the alignment shown above the B12 sequence. Exons boundaries are indicated above the alignment. The position of residues which vary between haplotypes is shown beneath the alignment.

8.3 Appendix 3 Deduced amino acid sequences of tapasin from seven MHC haplotypes.



```
                Exon 8
        ------][--]
                 *
Tpn12 : RPKEETKKSQ- : 430
Tpn19 : ..........- : 430
Tpn2  : ..........- : 430
Tpn15 : ..........- : 430
Tpn4  : ..........- : 430
Tpn14 : Q.........- : 430
Tpn21 : ..........- : 430
        ^
      421
```

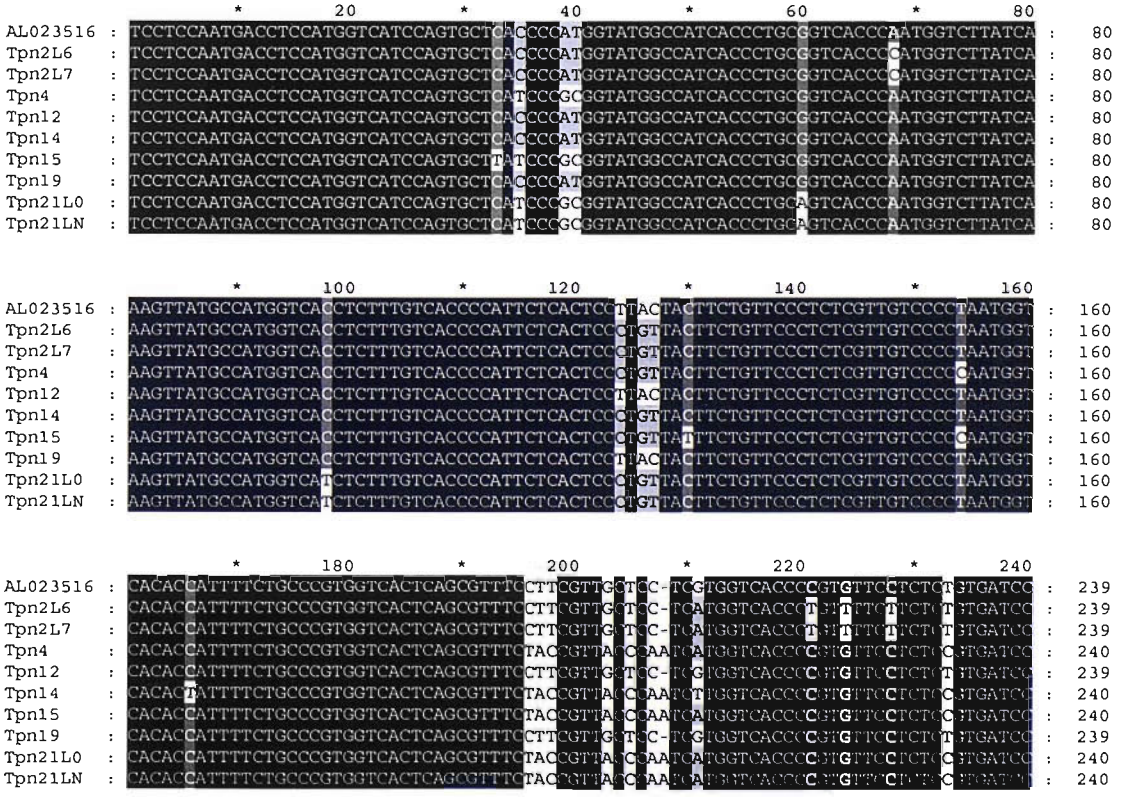
Appendix 3 Deduced amino acid sequences of tapasin from seven MHC haplotypes.

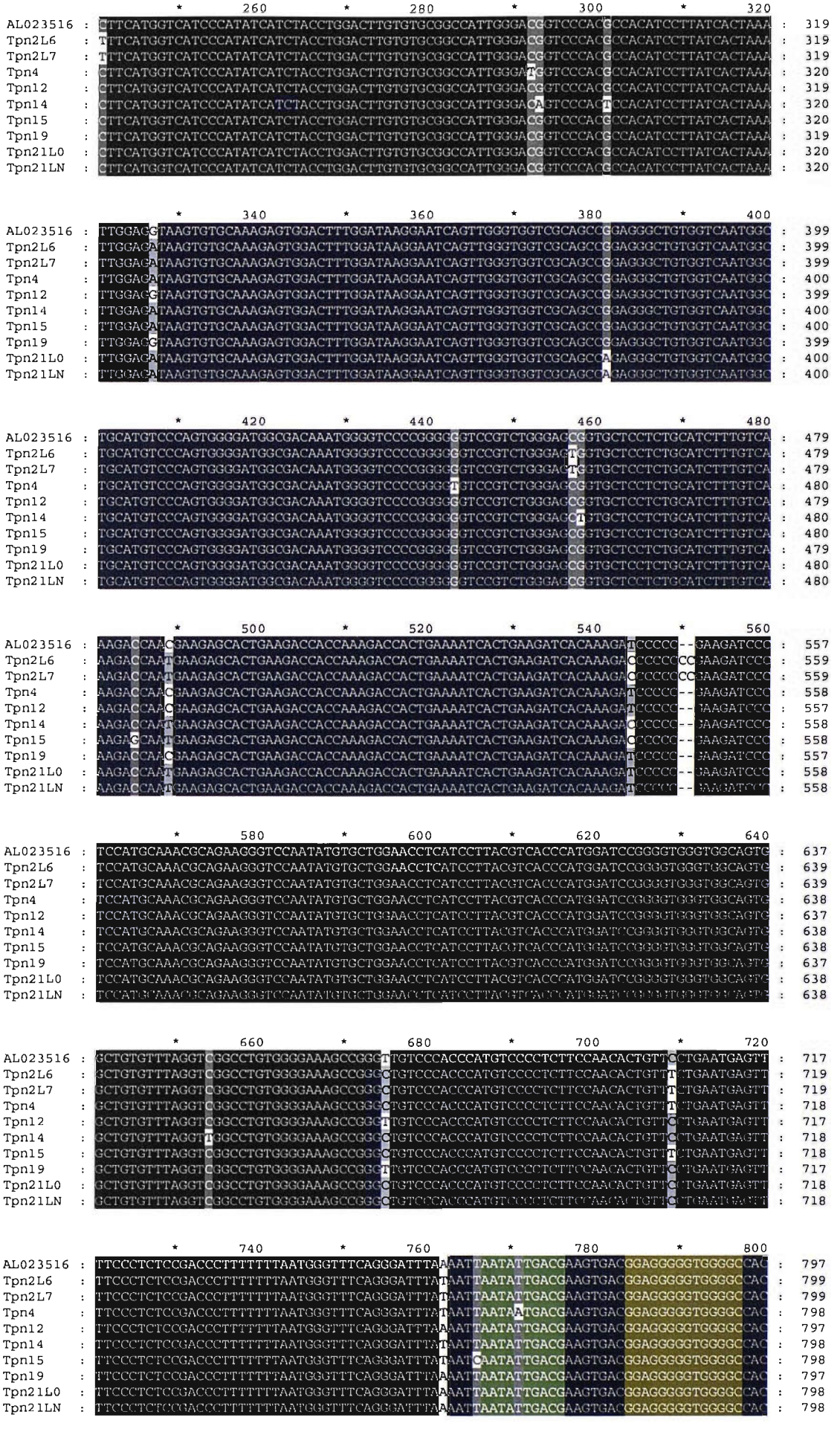
Polypeptide sequences were deduced from the cDNA sequences, and aligned using GCG and Genedoc analysis programs. Amino acid identity with the B12 sequence is represented by dots (.). Exon boundaries are indicated above the alignment. The predicted signal sequence is italicised and underlined and predicted transmembrane domain is underlined and in bold. The position of residues which vary between haplotypes is shown beneath the alignment.

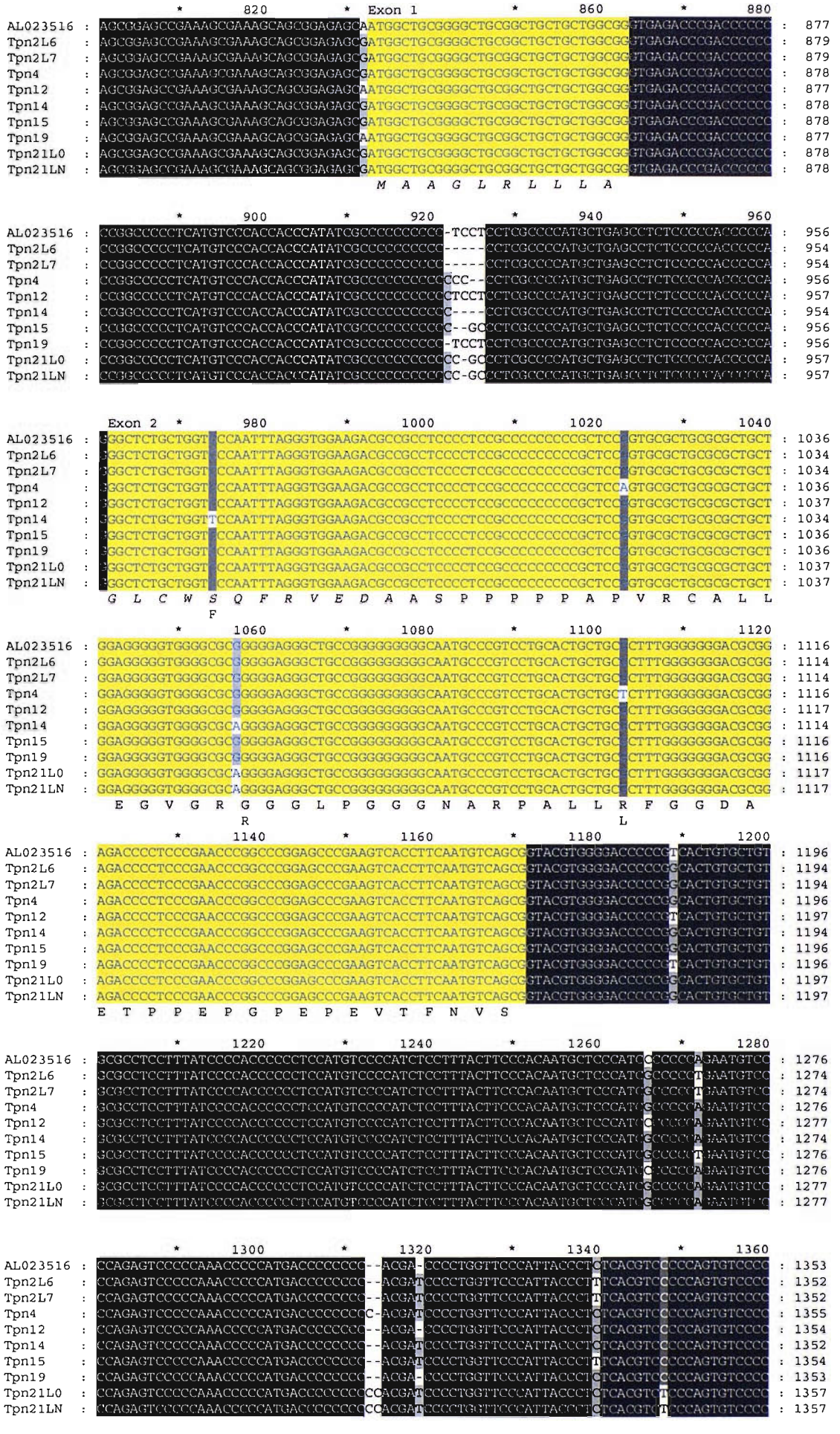
8.4 Appendix 4 Alignment of tapasin genomic sequences.

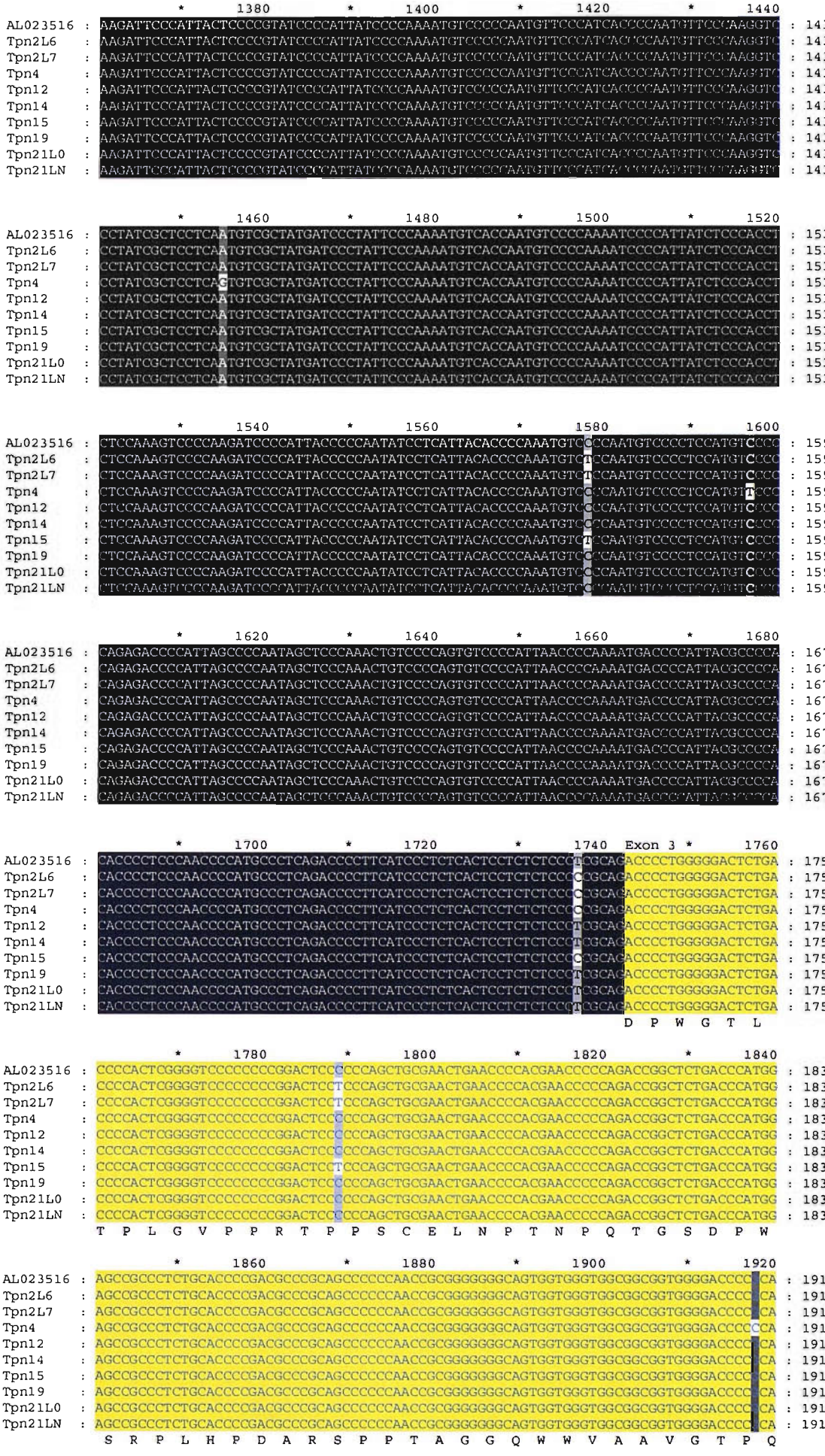
Consensus genomic sequences were produced for each haplotype and aligned using ClustalX and Genedoc analysis programs. The genomic sequences represent a combination of the four regions amplified and sequenced (5' region, tapasin exons and introns, 3' region, and tapasin cDNA). The top line of the alignment represents the corresponding region of the AL023516 accession.

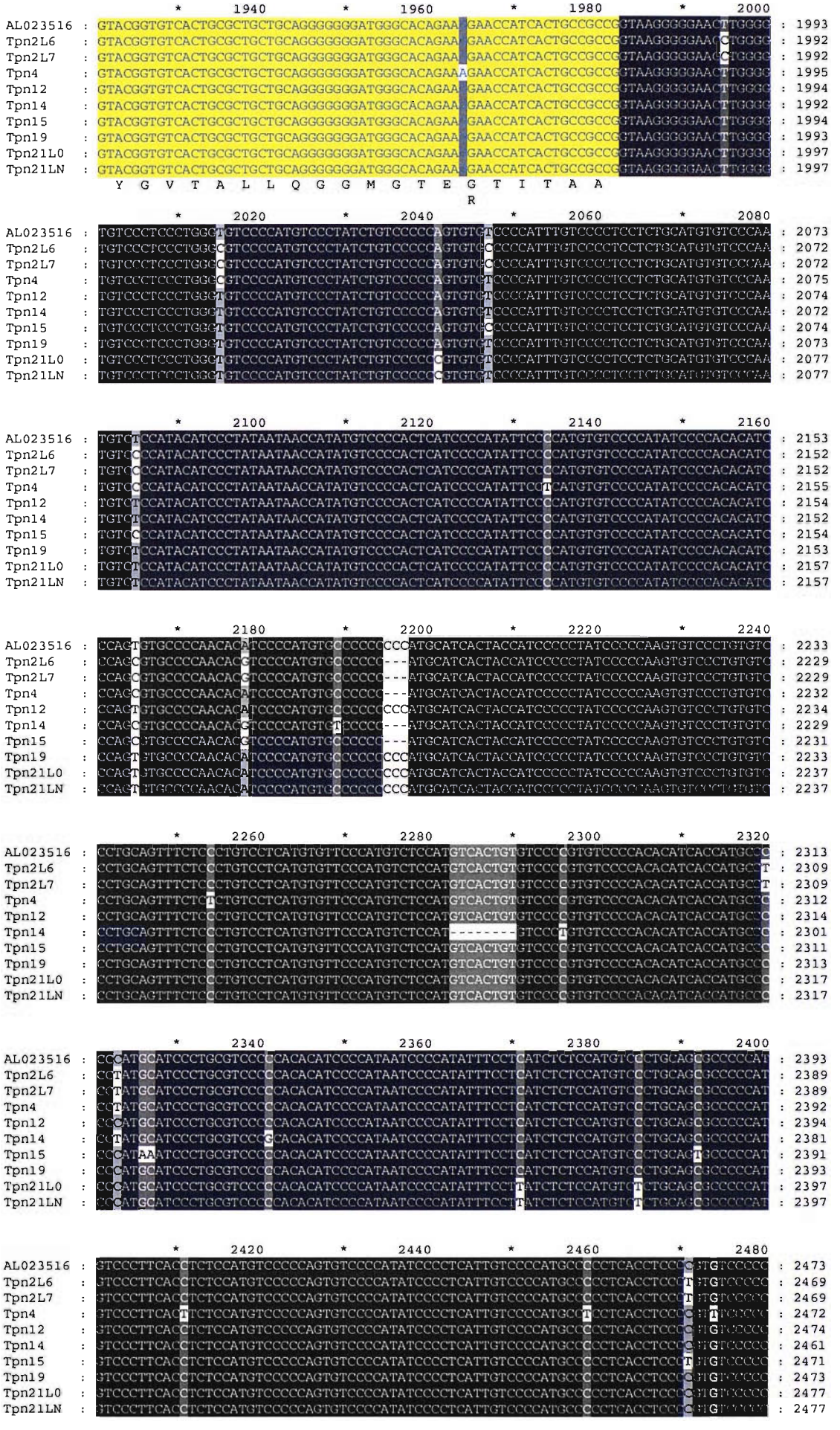
- **Exons** are indicated by blue text with predominantly yellow background (except for polymorphism), and labelled at start of each exon.
- **Introns, 5'UTR and 3'UTR** are indicated by white text with predominantly black background (except for polymorphism).
- **Polymorphisms** are indicated by grey/white background and black/blue text.
- **Amino acid translation** is shown below the second nucleotide of the codon, and refers to the B12 sequence, with alternate amino acids shown on the line below.
- Potential TATAA box (predicted by *Frangoulis et al*) indicated by green background. Putative Sp1 binding sites (predicted by *Frangoulis et al*) indicated by dark yellow background.
- Canonical polyadenylation site (predicted by *Frangoulis et al*) indicated by violet background. Poly A addition site (predicted by *Frangoulis et al*) indicated by red background.
- Transmembrane domain (predicted by *Frangoulis et al*) indicated by underlined amino acid (residues 398-416).
- Predicted signal sequence (SignalP 3.0) indicated by italicised amino acids (residues 1-22).

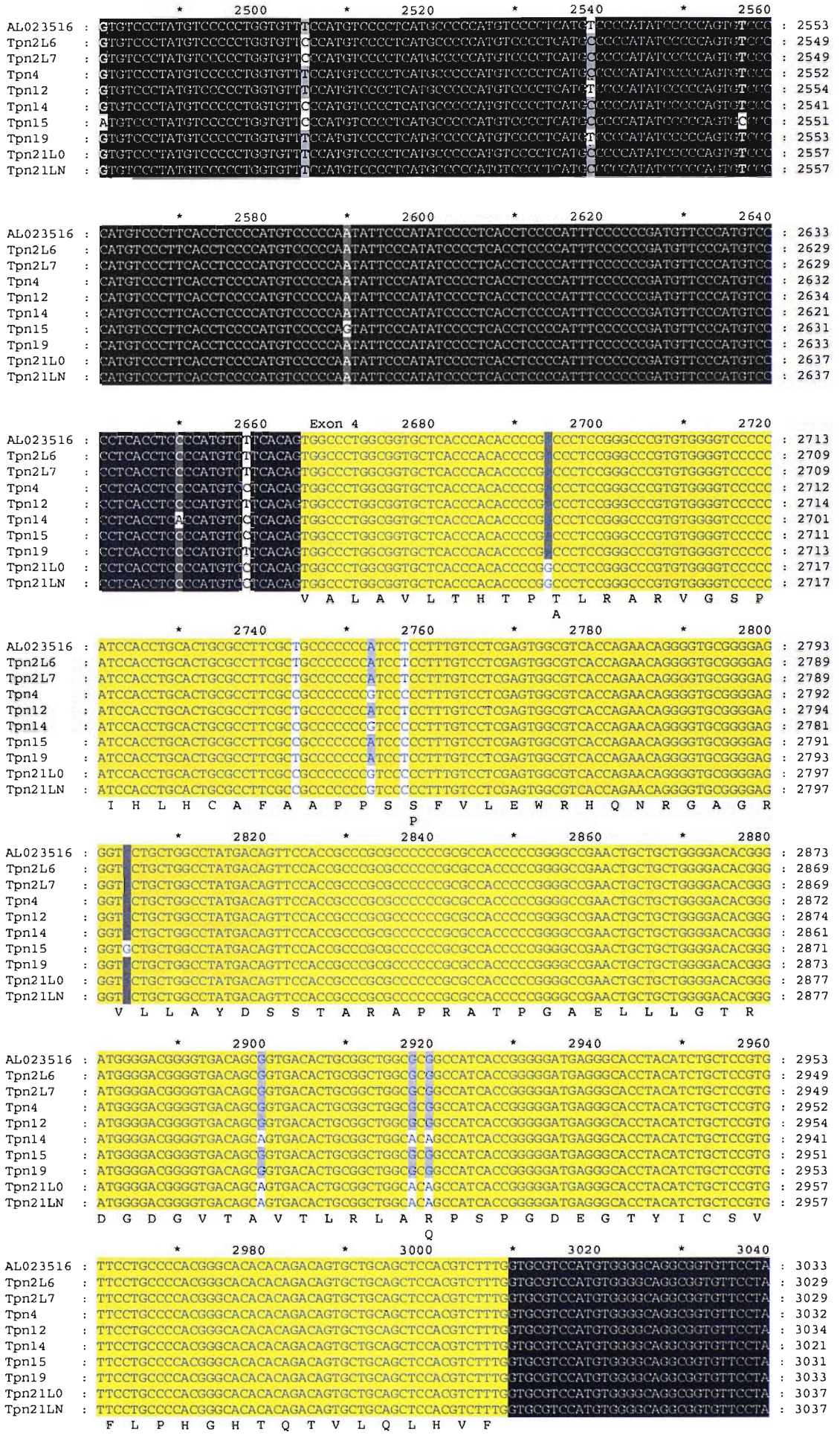


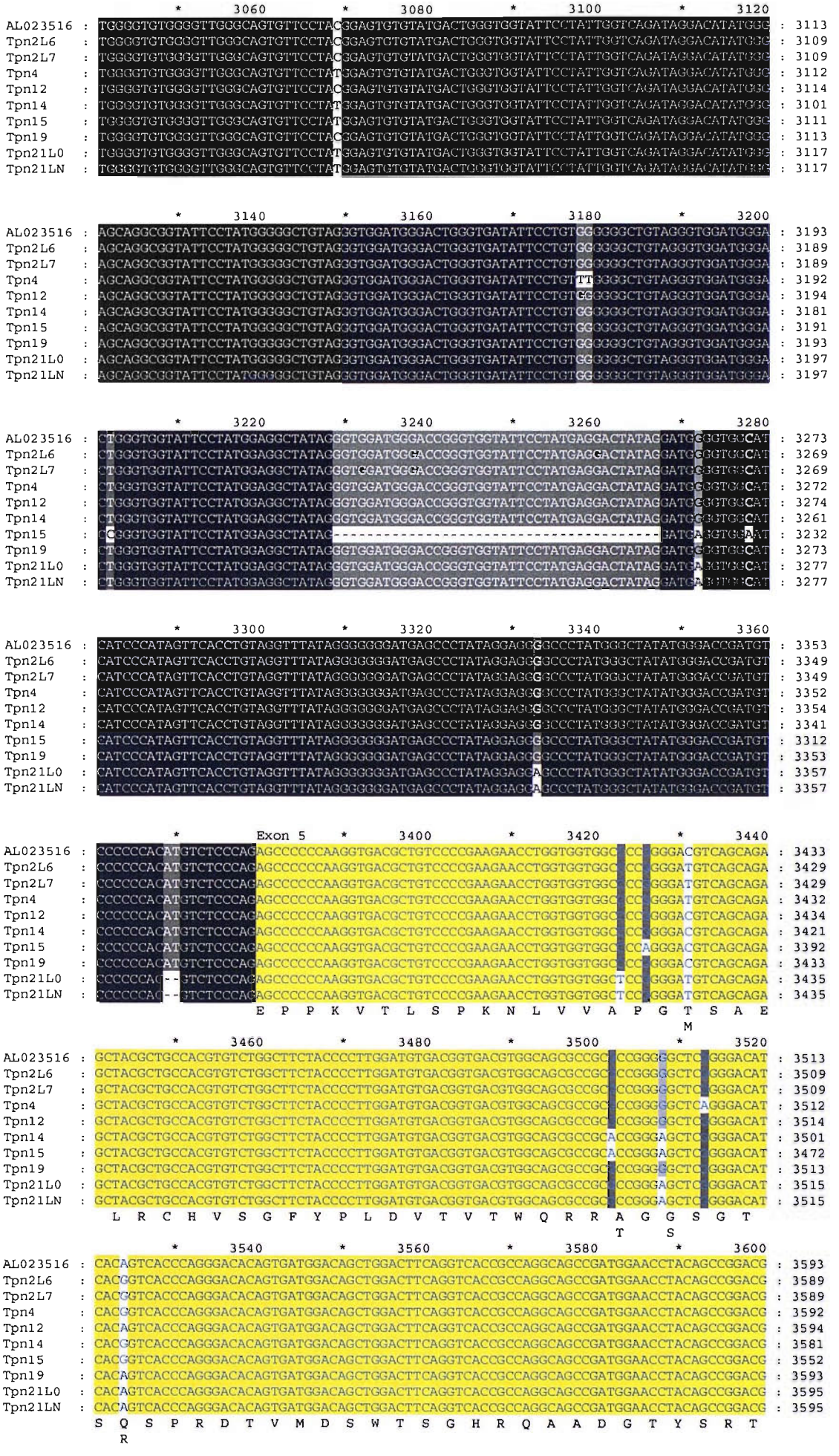

AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN

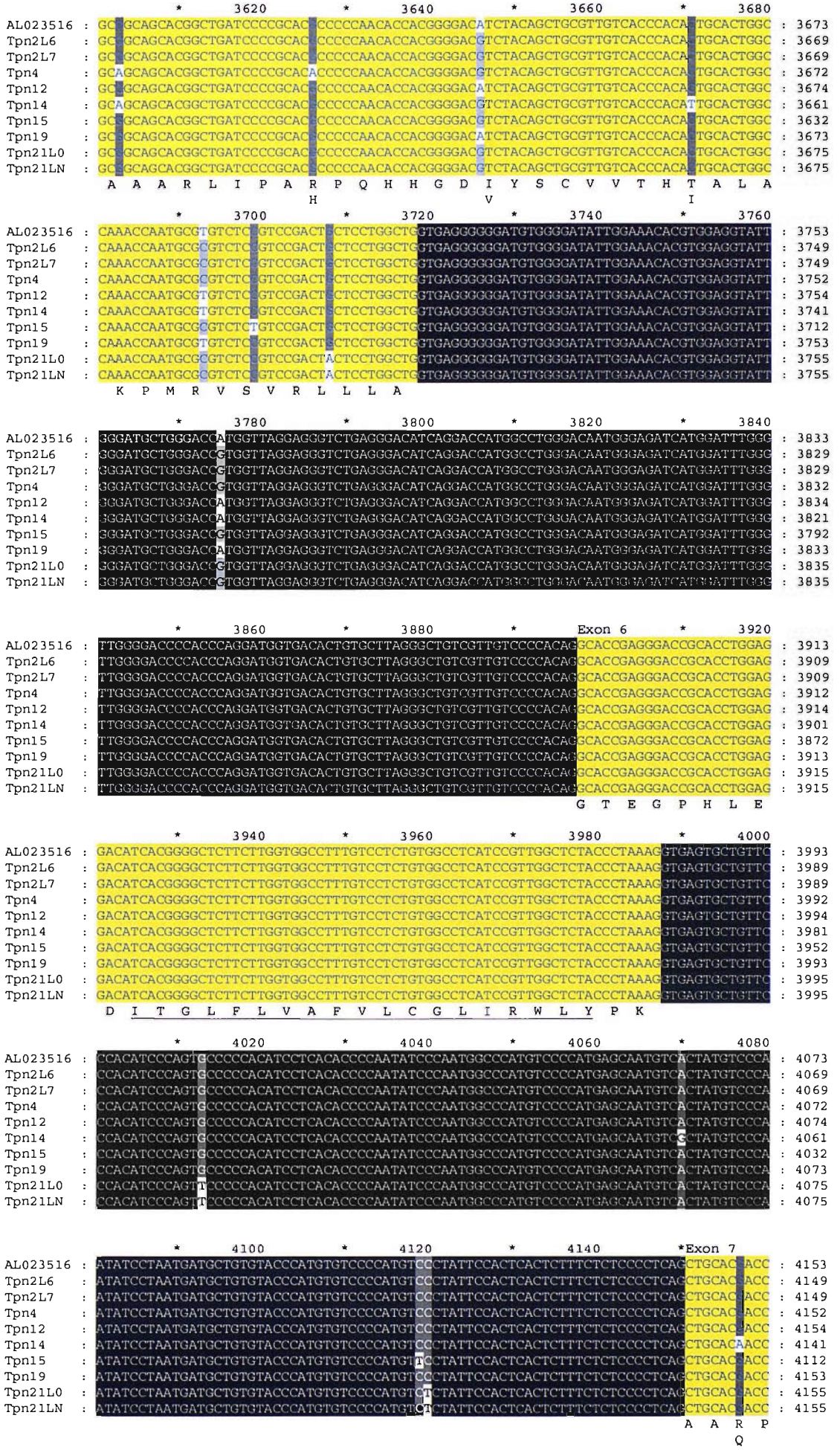

Exon 1
Exon 2
AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN

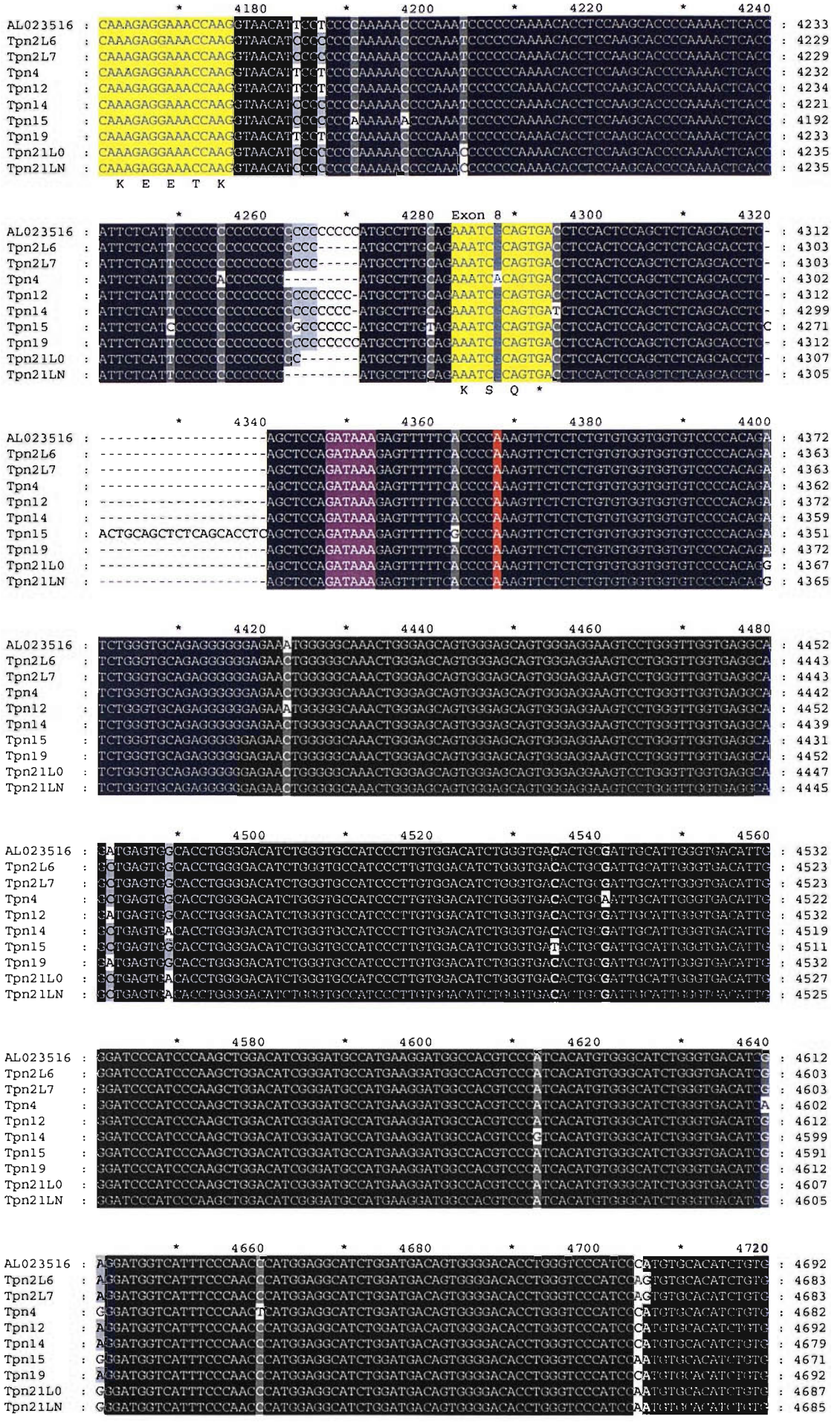

AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Exon 3
D P W G T L
T P L G V P P R T P P S C E L N P T N P Q T G S D P W
S R P L H P D A R S P P T A G G Q W W V A A V G T P Q

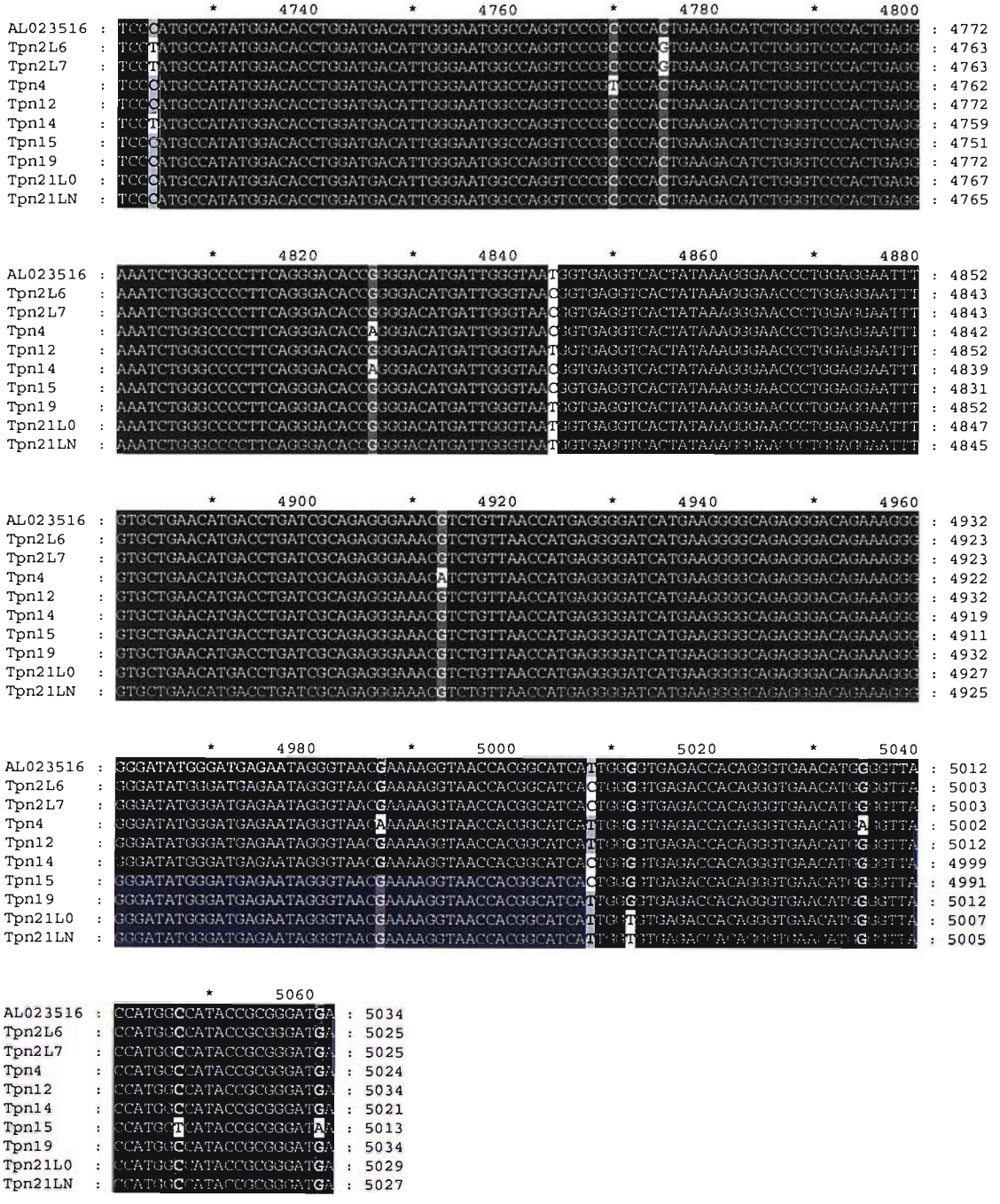

AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Y G V T A L L Q G G M G T E G T I T A A
R


AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Exon 4
V A L A V L T H T P T L R A R V G S P
A
I H L H C A F A A P P S S F V L E W R H Q N R G A G R
P
V L L A Y D S S T A R A P R A T P G A E L L L G T R
D G D G V T A V T L R L A R P S P G D E G T Y I C S V
Q
F L P H G H T Q T V L Q L H V F


AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Exon 5
E P P K V T L S P K N L V V A P G T S A E
M
L R C H V S G F Y P L D V T V T W Q R R A G G S G T
T
S
S Q S P R D T V M D S W T S G H R Q A A D G T Y S R T
R


AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Exon 6
Exon 7

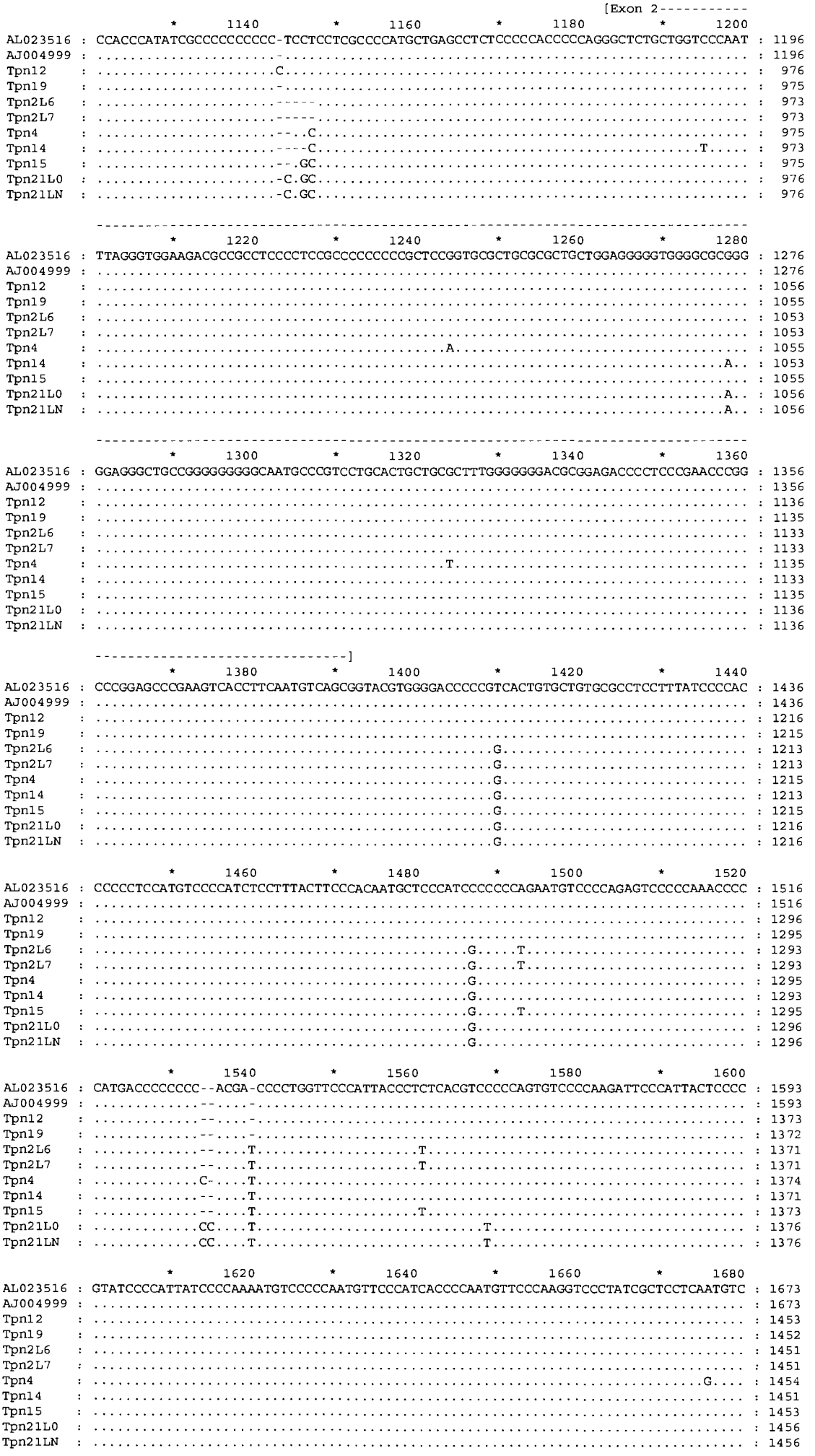

AL023516
Tpn2L6
Tpn2L7
Tpn4
Tpn12
Tpn14
Tpn15
Tpn19
Tpn21L0
Tpn21LN
Exon 8


AL023516 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4772
Tpn2L6 : TCCTATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCAGTGAAGACATCTGGGTCCCACTGAGG : 4763
Tpn2L7 : TCCTATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCAGTGAAGACATCTGGGTCCCACTGAGG : 4763
Tpn4 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGTCCCACTGAAGACATCTGGGTCCCACTGAGG : 4762
Tpn12 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4772
Tpn14 : TCCTATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4759
Tpn15 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4751
Tpn19 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4772
Tpn21L0 : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4767
Tpn21LN : TCCCATGCCATATGGACACCTGGATGACATTGGGAATGGCCAGGTCCCGCCCCACTGAAGACATCTGGGTCCCACTGAGG : 4765
AL023516 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAATGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4852
Tpn2L6 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAACGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4843
Tpn2L7 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAACGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4843
Tpn4 : AAATCTGGGCCCCTTCAGGGACACCAGGGACATGATTGGGTAACGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4842
Tpn12 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAATGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4852
Tpn14 : AAATCTGGGCCCCTTCAGGGACACCAGGGACATGATTGGGTAACGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4839
Tpn15 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAACGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4831
Tpn19 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAATGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4852
Tpn21L0 : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAATGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4847
Tpn21LN : AAATCTGGGCCCCTTCAGGGACACCGGGGACATGATTGGGTAATGGTGAGGTCACTATAAAGGGAACCCTGGAGGAATTT : 4845
AL023516 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4932
Tpn2L6 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4923
Tpn2L7 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4923
Tpn4 : GTGCTGAACATGACCTGATCGCAGAGGGAAACATCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4922
Tpn12 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4932
Tpn14 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4919
Tpn15 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4911
Tpn19 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4932
Tpn21L0 : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4927
Tpn21LN : GTGCTGAACATGACCTGATCGCAGAGGGAAACGTCTGTTAACCATGAGGGGATCATGAAGGGGCAGAGGGACAGAAAGGG : 4925
AL023516 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5012
Tpn2L6 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCACTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5003
Tpn2L7 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCACTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5003
Tpn4 : GGGATATGGGATGAGAATAGGGTAACAAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGAGGTTA : 5002
Tpn12 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5012
Tpn14 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCACTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 4999
Tpn15 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCACTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 4991
Tpn19 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5012
Tpn21L0 : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5007
Tpn21LN : GGGATATGGGATGAGAATAGGGTAACGAAAAGGTAACCACGGCATCATTGGGGTGAGACCACAGGGTGAACATGGGGTTA : 5005
AL023516 : CCATGGCCATACCGCGGGATGA : 5034
Tpn2L6 : CCATGGCCATACCGCGGGATGA : 5025
Tpn2L7 : CCATGGCCATACCGCGGGATGA : 5025
Tpn4 : CCATGGCCATACCGCGGGATGA : 5024
Tpn12 : CCATGGCCATACCGCGGGATGA : 5034
Tpn14 : CCATGGCCATACCGCGGGATGA : 5021
Tpn15 : CCATGGCCATACCGCGGGATAA : 5013
Tpn19 : CCATGGCCATACCGCGGGATGA : 5034
Tpn21L0 : CCATGGCCATACCGCGGGATGA : 5029
Tpn21LN : CCATGGCCATACCGCGGGATGA : 5027

8.5 Appendix 5 Alignment of tapasin genomic sequences with AJ004999.

```
          BLB2---------]
                    *        20          *        40          *        60          *        80
AL023516 : GCGCCCGCTCCCCATGGCTGCTGCCGCCGGACGATGGGGGACCCGGCACCCCGCGCCAGCCGCGAGCCCGGGGCCGGGCG :  80
AJ004999 : ................................................................................ :  80
Tpn12    : -------------------------------------------------------------------------------- :   -
Tpn19    : -------------------------------------------------------------------------------- :   -
Tpn2L6   : -------------------------------------------------------------------------------- :   -
Tpn2L7   : -------------------------------------------------------------------------------- :   -
Tpn4     : -------------------------------------------------------------------------------- :   -
Tpn14    : -------------------------------------------------------------------------------- :   -
Tpn15    : -------------------------------------------------------------------------------- :   -
Tpn21L0  : -------------------------------------------------------------------------------- :   -
Tpn21LN  : -------------------------------------------------------------------------------- :   -

                    *       100          *       120          *       140          *       160
AL023516 : GGGAGGGCCAATGAGCGGCGGCGGCGCCGGGCGGCGTCACCCGTCTCCGGGCAGAGTTCGACGCCGAGGGGCTTCCTCTG :  160
AJ004999 : ..............................................A..........TG.....C......C.T.... :  160
Tpn12    : -------------------------------------------------------------------------------- :   -
Tpn19    : -------------------------------------------------------------------------------- :   -
Tpn2L6   : -------------------------------------------------------------------------------- :   -
Tpn2L7   : -------------------------------------------------------------------------------- :   -
Tpn4     : -------------------------------------------------------------------------------- :   -
Tpn14    : -------------------------------------------------------------------------------- :   -
Tpn15    : -------------------------------------------------------------------------------- :   -
Tpn21L0  : -------------------------------------------------------------------------------- :   -
Tpn21LN  : -------------------------------------------------------------------------------- :   -

                    *       180          *       200          *       220          *       240
AL023516 : ACACTCTGCCCCCGTCCCACCGCCCCACGGGTGTCCCCACGGCCCTCGTCCAACCCGATGGTCCTCCAATGACCTCCATG :  240
AJ004999 : ..................................................C............................. :  240
Tpn12    : ----------------------------------------------------------...................... :   19
Tpn19    : ----------------------------------------------------------...................... :   19
Tpn2L6   : ----------------------------------------------------------...................... :   19
Tpn2L7   : ----------------------------------------------------------...................... :   19
Tpn4     : ----------------------------------------------------------...................... :   19
Tpn14    : ----------------------------------------------------------...................... :   19
Tpn15    : ----------------------------------------------------------...................... :   19
Tpn21L0  : ----------------------------------------------------------...................... :   19
Tpn21LN  : ----------------------------------------------------------...................... :   19

                    *       260          *       280          *       300          *       320
AL023516 : GTCATCCAGTGCTCACCCCATGGTATGGCCATCACCCTGCGGTCACCCAATGGTCTTATCAAAGTTATGCCATGGTCACC :  320
AJ004999 : ................................................................................ :  320
Tpn12    : ................................................................................ :   99
Tpn19    : ................................................................................ :   99
Tpn2L6   : ..................................................C............................. :   99
Tpn2L7   : ..................................................C............................. :   99
Tpn4     : ..............T...GC.......................................................... :   99
Tpn14    : ................................................................................ :   99
Tpn15    : ............T.T...GC.......................................................... :   99
Tpn21L0  : ..............T...GC.................A.....................................T. :   99
Tpn21LN  : ..............T...GC.................A.....................................T. :   99

                    *       340          *       360          *       380          *       400
AL023516 : TCTTTGTCACCCCATTCTCACTCCTTACTACTTCTGTTCCCTCTCGTTGTCCCCTAATGGTCACACCATTTTCTGCCCGT :  400
AJ004999 : ................................................................................ :  400
Tpn12    : ................................................................................ :  179
Tpn19    : ................................................................................ :  179
Tpn2L6   : ........................C.GT..................................................... :  179
Tpn2L7   : ........................C.GT..................................................... :  179
Tpn4     : ........................C.GT.....................C.............................. :  179
Tpn14    : ........................C.GT.....................................T.............. :  179
Tpn15    : ........................C.GT..T.....................C.......................... :  179
Tpn21L0  : ........................C.GT..................................................... :  179
Tpn21LN  : ........................C.GT..................................................... :  179

                    *       420          *       440          *       460          *       480
AL023516 : GGTCACTCAGCGTTTCCTTCGTTGCTCC-TCGTGGTCACCCCGTGTTCCTCTCTGTGATCCCTTCATGGTCATCCCATAT :  479
AJ004999 : ............................-............................................... :  479
Tpn12    : ............................-............................................... :  258
Tpn19    : ............................-............................................... :  258
Tpn2L6   : ............................-..A.........T..T...T............T.................. :  258
Tpn2L7   : ............................-..A.........T..T...T............T.................. :  258
Tpn4     : ..............TAC....A.C.AA..A.....................C.......................... :  259
Tpn14    : ..............TAC....A.C.AA..T.....................C.......................... :  259
Tpn15    : ..............TAC....A.C.AA..A.....................C.......................... :  259
Tpn21L0  : ..............TAC....A.C.AA..A.....................C.......................... :  259
Tpn21LN  : ..............TAC....A.C.AA..A.....................C.......................... :  259

                    *       500          *       520          *       540          *       560
AL023516 : CATCTACCTGGACTTGTGTGCGGCCATTGGGACGGTCCCACGCCACATCCTTATCACTAAATTGGAGGTAAGTGTGCAAA :  559
AJ004999 : ................................................................................ :  559
Tpn12    : ................................................................................ :  338
Tpn19    : ................................................................................ :  338
Tpn2L6   : ..............................................................A..A.............. :  338
Tpn2L7   : ..............................................................A..A.............. :  338
Tpn4     : ..............................T.................................A..A.............. :  339
Tpn14    : ..............................A.....T.........................A..A.............. :  339
Tpn15    : ..............................................................A..A.............. :  339
Tpn21L0  : ..............................................................A..A.............. :  339
Tpn21LN  : ..............................................................A..A.............. :  339
```

```
                  *       580         *       600         *       620         *       640
AL023516 : GAGTGGACTTTGGATAAGGAATCAGTTGGGTGGTCGCAGCCGGAGGGCTGTGGTCAATGGCTGCATGTCCCAGTGGGGAT : 639
AJ004999 : ..................................................................................... : 639
Tpn12    : ..................................................................................... : 418
Tpn19    : ..................................................................................... : 418
Tpn2L6   : ..................................................................................... : 418
Tpn2L7   : ..................................................................................... : 418
Tpn4     : ..................................................................................... : 419
Tpn14    : ..................................................................................... : 419
Tpn15    : ..................................................................................... : 419
Tpn21L0  : ...................................A.................................................. : 419
Tpn21LN  : ...................................A.................................................. : 419

                  *       660         *       680         *       700         *       720
AL023516 : GGCGACAAATGGGGTCCCCGGGGGGTCCGTCTGGGAGCGGTGCTCCTCTGCATCTTTGTCAAAGACCAACGAAGAGCACT : 719
AJ004999 : ..................................................................................... : 719
Tpn12    : ..................................................................................... : 498
Tpn19    : ..................................................................................... : 498
Tpn2L6   : ..................................T..............................T..................... : 498
Tpn2L7   : ..................................T..............................T..................... : 498
Tpn4     : ....................T.............................................................. : 499
Tpn14    : ..................................T..............................T..................... : 499
Tpn15    : ..................................................................G...T..................... : 499
Tpn21L0  : ..................................................................T..................... : 499
Tpn21LN  : ..................................................................T..................... : 499

                  *       740         *       760         *       780         *       800
AL023516 : GAAGACCACCAAAGACCACTGAAAATCACTGAAGATCACAAAGATCCCCC--GAAGATCCCTCCATGCAAACGCAGAAGG : 797
AJ004999 : ..................................................--.............................. : 797
Tpn12    : ..................................................--.............................. : 576
Tpn19    : ..................................................--.............................. : 576
Tpn2L6   : ..........................................C.....CC.............................. : 578
Tpn2L7   : ..........................................C.....CC.............................. : 578
Tpn4     : ..................................................--.............................. : 577
Tpn14    : ..........................................C.....--.............................. : 577
Tpn15    : ..........................................C.....--.............................. : 577
Tpn21L0  : ..................................................--.............................. : 577
Tpn21LN  : ..................................................--.............................. : 577

                  *       820         *       840         *       860         *       880
AL023516 : GTCCAATATGTGCTGGAACCTCATCCTTACGTCACCCATGGATCCGGGGTGGGTGGCAGTGGCTGTGTTTAGGTCGGCCT : 877
AJ004999 : ..................................................................................... : 877
Tpn12    : ..................................................................................... : 656
Tpn19    : ..................................................................................... : 656
Tpn2L6   : ..................................................................................... : 658
Tpn2L7   : ..................................................................................... : 658
Tpn4     : ..................................................................................... : 657
Tpn14    : .....................................................................T..... : 657
Tpn15    : ..................................................................................... : 657
Tpn21L0  : ..................................................................................... : 657
Tpn21LN  : ..................................................................................... : 657

                  *       900         *       920         *       940         *       960
AL023516 : GTGGGGAAAGCCGGGTTGTCCCACCCATGTCCCCTCTTCCAACACTGTTCCTGAATGAGTTTTCCCTCTCCGACCCTTTT : 957
AJ004999 : ..................................................................................... : 957
Tpn12    : ..................................................................................... : 736
Tpn19    : ..................................................................................... : 736
Tpn2L6   : ..............C.................................T..................................... : 738
Tpn2L7   : ..............C.................................T..................................... : 738
Tpn4     : ..............C.................................T..................................... : 737
Tpn14    : ..............C..................................................................... : 737
Tpn15    : ..............C.................................T..................................... : 737
Tpn21L0  : ..............C..................................................................... : 737
Tpn21LN  : ..............C..................................................................... : 737

                                    [TATA-box-]
                  *       980         *      1000         *      1020         *      1040
AL023516 : TTTAATGGGTTTCAGGGATTTAAAATTAATATTGACGAAGTGACGGAGGGGGTGGGGCCACAGCGGAGCCGAAAGCGAAA : 1037
AJ004999 : ..................................................................................... : 1037
Tpn12    : ..................................................................................... : 816
Tpn19    : ..................................................................................... : 816
Tpn2L6   : ....................T.................................................................. : 818
Tpn2L7   : ....................T.................................................................. : 818
Tpn4     : ....................T........A........................................................... : 817
Tpn14    : ....................T.................................................................. : 817
Tpn15    : ....................T...C.............................................................. : 817
Tpn21L0  : ..................................................................................... : 817
Tpn21LN  : ..................................................................................... : 817

                        [Exon 1----------------------]
                  *      1060         *      1080         *      1100         *      1120
AL023516 : GCAGCGGAGAGCAATGGCTGCGGGGCTGCGGCTGCTGCTGGCGGGTGAGACCCGACCCCCCCCGGCCCCCTCATGTCCCA : 1117
AJ004999 : ..................................................................................... : 1117
Tpn12    : ..................................................................................... : 896
Tpn19    : ..................................................................................... : 896
Tpn2L6   : ...........G.......................................................................... : 898
Tpn2L7   : ...........G.......................................................................... : 898
Tpn4     : ...........G.......................................................................... : 897
Tpn14    : ...........G.......................................................................... : 897
Tpn15    : ...........G.......................................................................... : 897
Tpn21L0  : ...........G.......................................................................... : 897
Tpn21LN  : ...........G.......................................................................... : 897
```

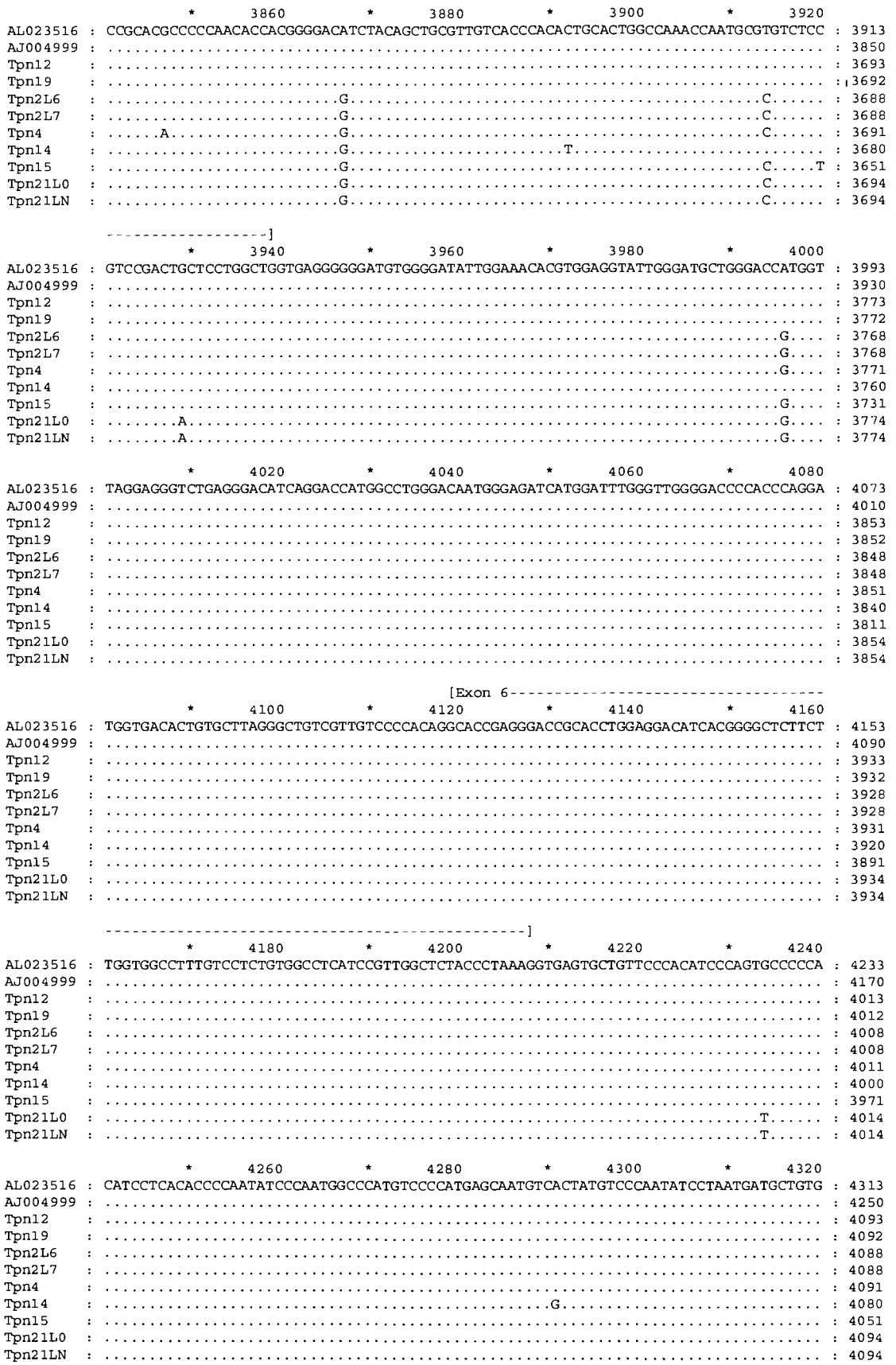

[Exon 2----------
* 1140 * 1160 * 1180 * 1200
AL023516 : CCACCCATATCGCCCCCCCCCC-TCCTCCTCGCCCCATGCTGAGCCTCTCCCCCACCCCCAGGGCTCTGCTGGTCCCAAT : 1196
AJ004999 : ..... : 1196
Tpn12 : .....C..... : 976
Tpn19 : ..... : 975
Tpn2L6 : ..... : 973
Tpn2L7 : ..... : 973
Tpn4 : .....--..C..... : 975
Tpn14 : .....----C.....T..... : 973
Tpn15 : .....--.GC..... : 975
Tpn21L0 : .....-C.GC..... : 976
Tpn21LN : .....-C.GC..... : 976
* 1220 * 1240 * 1260 * 1280
AL023516 : TTAGGGTGGAAGACGCCGCCTCCCCTCCGCCCCCCCCCGCTCCGGTGCGCTGCGCGCTGCTGGAGGGGGTGGGGCGCGGG : 1276
AJ004999 : ..... : 1276
Tpn12 : ..... : 1056
Tpn19 : ..... : 1055
Tpn2L6 : ..... : 1053
Tpn2L7 : ..... : 1053
Tpn4 : .....A..... : 1055
Tpn14 : .....A.. : 1053
Tpn15 : ..... : 1055
Tpn21L0 : .....A.. : 1056
Tpn21LN : .....A.. : 1056
* 1300 * 1320 * 1340 * 1360
AL023516 : GGAGGGCTGCCGGGGGGGGGCAATGCCCGTCCTGCACTGCTGCGCTTTGGGGGGGACGCGGAGACCCCTCCCGAACCCGG : 1356
AJ004999 : ..... : 1356
Tpn12 : ..... : 1136
Tpn19 : ..... : 1135
Tpn2L6 : ..... : 1133
Tpn2L7 : ..... : 1133
Tpn4 : .....T..... : 1135
Tpn14 : ..... : 1133
Tpn15 : ..... : 1135
Tpn21L0 : ..... : 1136
Tpn21LN : ..... : 1136
------------------------------]
* 1380 * 1400 * 1420 * 1440
AL023516 : CCCGGAGCCCGAAGTCACCTTCAATGTCAGCGGTACGTGGGGACCCCCGTCACTGTGCTGTGCGCCTCCTTTATCCCCAC : 1436
AJ004999 : ..... : 1436
Tpn12 : ..... : 1216
Tpn19 : ..... : 1215
Tpn2L6 : .....G..... : 1213
Tpn2L7 : .....G..... : 1213
Tpn4 : .....G..... : 1215
Tpn14 : .....G..... : 1213
Tpn15 : .....G..... : 1215
Tpn21L0 : .....G..... : 1216
Tpn21LN : .....G..... : 1216
* 1460 * 1480 * 1500 * 1520
AL023516 : CCCCCTCCATGTCCCCATCTCCTTTACTTCCCACAATGCTCCCATCCCCCCCAGAATGTCCCCAGAGTCCCCCAAACCCC : 1516
AJ004999 : ..... : 1516
Tpn12 : ..... : 1296
Tpn19 : ..... : 1295
Tpn2L6 : .....G.....T..... : 1293
Tpn2L7 : .....G.....T..... : 1293
Tpn4 : .....G..... : 1295
Tpn14 : .....G..... : 1293
Tpn15 : .....G.....T..... : 1295
Tpn21L0 : .....G..... : 1296
Tpn21LN : .....G..... : 1296
* 1540 * 1560 * 1580 * 1600
AL023516 : CATGACCCCCCCC--ACGA-CCCCTGGTTCCCATTACCCTCTCACGTCCCCCAGTGTCCCCAAGATTCCCATTACTCCCC : 1593
AJ004999 : .....--.....- : 1593
Tpn12 : .....--.....- : 1373
Tpn19 : .....--.....- : 1372
Tpn2L6 : .....--....T.....T..... : 1371
Tpn2L7 : .....--....T.....T..... : 1371
Tpn4 : .....C-....T..... : 1374
Tpn14 : .....--....T..... : 1371
Tpn15 : .....--....T.....T..... : 1373
Tpn21L0 : .....CC....T.....T..... : 1376
Tpn21LN : .....CC....T.....T..... : 1376
* 1620 * 1640 * 1660 * 1680
AL023516 : GTATCCCCATTATCCCCAAAATGTCCCCCAATGTTCCCATCACCCCAATGTTCCCAAGGTCCCTATCGCTCCTCAATGTC : 1673
AJ004999 : ..... : 1673
Tpn12 : ..... : 1453
Tpn19 : ..... : 1452
Tpn2L6 : ..... : 1451
Tpn2L7 : ..... : 1451
Tpn4 : .....G.... : 1454
Tpn14 : ..... : 1451
Tpn15 : ..... : 1453
Tpn21L0 : ..... : 1456
Tpn21LN : ..... : 1456

```
                  *      1700         *      1720         *      1740         *      1760
AL023516 : GCTATGATCCCTATTCCCAAAATGTCACCAATGTCCCCAAAATCCCCATTATCTCCCACCTCTCCAAAGTCCCCAAGATC : 1753
AJ004999 : ..................................................................................... : 1753
Tpn12    : ..................................................................................... : 1533
Tpn19    : ..................................................................................... : 1532
Tpn2L6   : ..................................................................................... : 1531
Tpn2L7   : ..................................................................................... : 1531
Tpn4     : ..................................................................................... : 1534
Tpn14    : ..................................................................................... : 1531
Tpn15    : ..................................................................................... : 1533
Tpn21L0  : ..................................................................................... : 1536
Tpn21LN  : ..................................................................................... : 1536

                  *      1780         *      1800         *      1820         *      1840
AL023516 : CCCATTACCCCCAATATCCTCATTACACCCCAAATGTCCCCAATGTCCCCTCCATGTCCCCCAGAGACCCCATTAGCCCC : 1833
AJ004999 : ..................................................................................... : 1833
Tpn12    : ..................................................................................... : 1613
Tpn19    : ..................................................................................... : 1612
Tpn2L6   : ....................................T................................................ : 1611
Tpn2L7   : ....................................T................................................ : 1611
Tpn4     : ........................................................T........................ : 1614
Tpn14    : ..................................................................................... : 1611
Tpn15    : ....................................T................................................ : 1613
Tpn21L0  : ..................................................................................... : 1616
Tpn21LN  : ..................................................................................... : 1616

                  *      1860         *      1880         *      1900         *      1920
AL023516 : AATAGCTCCCAAACTGTCCCCAGTGTCCCCATTAACCCCAAAATGACCCCATTACGCCCCACACCCCTCCCAACCCCATG : 1913
AJ004999 : ..................................................................................... : 1913
Tpn12    : ..................................................................................... : 1693
Tpn19    : ..................................................................................... : 1692
Tpn2L6   : ..................................................................................... : 1691
Tpn2L7   : ..................................................................................... : 1691
Tpn4     : ..................................................................................... : 1694
Tpn14    : ..................................................................................... : 1691
Tpn15    : ..................................................................................... : 1693
Tpn21L0  : ..................................................................................... : 1696
Tpn21LN  : ..................................................................................... : 1696

                                                                    [Exon 3----------------------------
                  *      1940         *      1960         *      1980         *      2000
AL023516 : CCCTCAGACCCCTTCATCCCTCTCACTCCTCTCTCCCTCGCAGACCCCTGGGGGACTCTGACCCCACTCGGGGTCCCCCC : 1993
AJ004999 : ..................................................................................... : 1993
Tpn12    : ..................................................................................... : 1773
Tpn19    : ..................................................................................... : 1772
Tpn2L6   : ....................................C................................................ : 1771
Tpn2L7   : ....................................C................................................ : 1771
Tpn4     : ....................................C................................................ : 1774
Tpn14    : ..................................................................................... : 1771
Tpn15    : ....................................C................................................ : 1773
Tpn21L0  : ..................................................................................... : 1776
Tpn21LN  : ..................................................................................... : 1776

         ----------------------------------------------------------------------------------------
                  *      2020         *      2040         *      2060         *      2080
AL023516 : CCGGACTCCCCCCAGCTGCGAACTGAACCCCACGAACCCCCAGACCGGCTCTGACCCATGGAGCCGCCCTCTGCACCCCG : 2073
AJ004999 : ..................................................................................... : 2073
Tpn12    : ..................................................................................... : 1853
Tpn19    : ..................................................................................... : 1852
Tpn2L6   : .........T............................................................................ : 1851
Tpn2L7   : .........T............................................................................ : 1851
Tpn4     : ..................................................................................... : 1854
Tpn14    : ..................................................................................... : 1851
Tpn15    : .........T............................................................................ : 1853
Tpn21L0  : ..................................................................................... : 1856
Tpn21LN  : ..................................................................................... : 1856

         ----------------------------------------------------------------------------------------
                  *      2100         *      2120         *      2140         *      2160
AL023516 : ACGCCCGCAGCCCCCCAACCGCGGGGGGGCAGTGGTGGGTGGCGGCGGTGGGGACCCCGCAGTACGGTGTCACTGCGCTG : 2153
AJ004999 : ..................................................................................... : 2153
Tpn12    : ..................................................................................... : 1933
Tpn19    : ..................................................................................... : 1932
Tpn2L6   : ..................................................................................... : 1931
Tpn2L7   : ..................................................................................... : 1931
Tpn4     : ........................................................C............................ : 1934
Tpn14    : ..................................................................................... : 1931
Tpn15    : ..................................................................................... : 1933
Tpn21L0  : ..................................................................................... : 1936
Tpn21LN  : ..................................................................................... : 1936

         ----------------------------------------]
                  *      2180         *      2200         *      2220         *      2240
AL023516 : CTGCAGGGGGGGATGGGCACAGAAGGAACCATCACTGCCGCCGGTAAGGGGGAACTTGGGGTGTCCCTCCCTGGGTGTCC : 2233
AJ004999 : ..................................................................................... : 2233
Tpn12    : ..................................................................................... : 2013
Tpn19    : ..................................................................................... : 2012
Tpn2L6   : ..........................................................C.....................C.... : 2011
Tpn2L7   : ..........................................................C.....................C.... : 2011
Tpn4     : ........................A.....................................................C.... : 2014
Tpn14    : ..................................................................................... : 2011
Tpn15    : ..................................................................................... : 2013
Tpn21L0  : ..................................................................................... : 2016
Tpn21LN  : ..................................................................................... : 2016
```

```
                    *      2260          *      2280          *      2300          *      2320
AL023516 : CCATGTCCCTATCTGTCCCCAGTGTGTCCCCATTTGTCCCCTCCTCTGCATGTGTCCCAATGTCTCCATACATCCCTAT : 2313
AJ004999 : ................................................................................ : 2313
Tpn12    : ................................................................................ : 2093
Tpn19    : ................................................................................ : 2092
Tpn2L6   : ...........................C.....................................C.............. : 2091
Tpn2L7   : ...........................C.....................................C.............. : 2091
Tpn4     : .................................................................C.............. : 2094
Tpn14    : ................................................................................ : 2091
Tpn15    : ...........................C.....................................C.............. : 2093
Tpn21L0  : .....................C.......................................................... : 2096
Tpn21LN  : .....................C.......................................................... : 2096

                    *      2340          *      2360          *      2380          *      2400
AL023516 : AATAACCATATGTCCCCACTCATCCCCATATTCCCCATGTGTCCCCATATCCCCACACATCCCAGTGTGCCCCAACACAT : 2393
AJ004999 : ................................................................................ : 2393
Tpn12    : ................................................................................ : 2173
Tpn19    : ................................................................................ : 2172
Tpn2L6   : .................................................................C............G. : 2171
Tpn2L7   : .................................................................C............G. : 2171
Tpn4     : ..................................T..............................C.............. : 2174
Tpn14    : .................................................................C............G. : 2171
Tpn15    : .................................................................C............G. : 2173
Tpn21L0  : ................................................................................ : 2176
Tpn21LN  : ................................................................................ : 2176

                    *      2420          *      2440          *      2460          *      2480
AL023516 : CCCCATGTGCCCCCCCCCATGCATCACTACCATCCCCCTATCCCCCAAGTGTCCCTGTGTCCCTGCAGTTTCTCCCTGTC : 2473
AJ004999 : ................................................................................ : 2473
Tpn12    : ................................................................................ : 2253
Tpn19    : ................................................................................ : 2252
Tpn2L6   : ..............---............................................................... : 2248
Tpn2L7   : ..............---............................................................... : 2248
Tpn4     : ..............---.........................................................T..... : 2251
Tpn14    : .........T....---............................................................... : 2248
Tpn15    : ..............---............................................................... : 2250
Tpn21L0  : ................................................................................ : 2256
Tpn21LN  : ................................................................................ : 2256

                    *      2500          *      2520          *      2540          *      2560
AL023516 : CTCATGTGTTCCCATGTCTCCATGTCACTGTGTCCCCGTGTCCCCACACATCACCATGCCCCCCATGCATCCCTGCGTCC : 2553
AJ004999 : .................................................................--------------- : 2538
Tpn12    : ................................................................................ : 2333
Tpn19    : ................................................................................ : 2332
Tpn2L6   : ............................................................T..T................ : 2328
Tpn2L7   : ............................................................T..T................ : 2328
Tpn4     : ...............................................................T................ : 2331
Tpn14    : .......................--------.....T..........................T................ : 2320
Tpn15    : ..................................................................AA............ : 2330
Tpn21L0  : ................................................................................ : 2336
Tpn21LN  : ................................................................................ : 2336

                    *      2580          *      2600          *      2620          *      2640
AL023516 : CCCACACATCCCCATAATCCCCATATTTCCTCATCTCTCCATGTCCCTGCAGCGCCCCCATGTCCCTTCACCTCTCCATG : 2633
AJ004999 : ----------------------------------------------.................................. : 2572
Tpn12    : ................................................................................ : 2413
Tpn19    : ................................................................................ : 2412
Tpn2L6   : ................................................................................ : 2408
Tpn2L7   : ................................................................................ : 2408
Tpn4     : .......................................................................T........ : 2411
Tpn14    : .G.............................................................................. : 2400
Tpn15    : ...................................................T............................ : 2410
Tpn21L0  : ...............................T.............T.................................. : 2416
Tpn21LN  : ...............................T.............T.................................. : 2416

                    *      2660          *      2680          *      2700          *      2720
AL023516 : TCCCCCAGTGTCCCCATATCCCCTCATTGTCCCCATGCCCCCTCACCTCCCCGTGTCCCCCGTGTCCCTATGTCCCCCTG : 2713
AJ004999 : ..............-..........................................................T...... : 2651
Tpn12    : ................................................................................ : 2493
Tpn19    : ................................................................................ : 2492
Tpn2L6   : ..................................................T............................. : 2488
Tpn2L7   : ..................................................T............................. : 2488
Tpn4     : ......................................T..............T.......................... : 2491
Tpn14    : ................................................................................ : 2480
Tpn15    : ..................................................T.........A................... : 2490
Tpn21L0  : ................................................................................ : 2496
Tpn21LN  : ................................................................................ : 2496

                    *      2740          *      2760          *      2780          *      2800
AL023516 : GTGTTTCCATGTCCCCTCATGCCCCCATGTCCCCTCATGTCCCCATATCCCCCAGTGTCCCCATGTCCCTTCACCTCCCC : 2793
AJ004999 : ................................................................................ : 2731
Tpn12    : ................................................................................ : 2573
Tpn19    : ................................................................................ : 2572
Tpn2L6   : .....C................................C......................................... : 2568
Tpn2L7   : .....C................................C......................................... : 2568
Tpn4     : ......................................C......................................... : 2571
Tpn14    : .....C................................C......................................... : 2560
Tpn15    : .....C................................C.................C....................... : 2570
Tpn21L0  : ......................................C......................................... : 2576
Tpn21LN  : ......................................C......................................... : 2576
```

```
                  *      2820         *      2840         *      2860         *      2880
AL023516 : ATGTCCCCCAATATTCCCATATCCCCTCACCTCCCCATTTCCCCCCGATGTTCCCATGTCCCCTCACCTCCCCATGTCTT : 2873
AJ004999 : ...............................G.............................G.............. : 2811
Tpn12    : ..................................................................................... : 2653
Tpn19    : ..................................................................................... : 2652
Tpn2L6   : ..................................................................................... : 2648
Tpn2L7   : ..................................................................................... : 2648
Tpn4     : .....................................................................C. : 2651
Tpn14    : ..................................................................A.....C. : 2640
Tpn15    : ..........G..........................................................C. : 2650
Tpn21L0  : .....................................................................C. : 2656
Tpn21LN  : .....................................................................C. : 2656

               [Exon 4-------------------------------------------------------------------
                  *      2900         *      2920         *      2940         *      2960
AL023516 : CACAGTGGCCCTGGCGGTGCTCACCCACACCCCGACCCTCCGGGCCCGTGTGGGGTCCCCCATCCACCTGCACTGCGCCT : 2953
AJ004999 : ..................................................................................... : 2891
Tpn12    : ..................................................................................... : 2733
Tpn19    : ..................................................................................... : 2732
Tpn2L6   : ..................................................................................... : 2728
Tpn2L7   : ..................................................................................... : 2728
Tpn4     : ..................................................................................... : 2731
Tpn14    : ..................................................................................... : 2720
Tpn15    : ..................................................................................... : 2730
Tpn21L0  : ................................G.................................................... : 2736
Tpn21LN  : ................................G.................................................... : 2736

               ---------------------------------------------------------------------------
                  *      2980         *      3000         *      3020         *      3040
AL023516 : TCGCTGCCCCCCCATCCTCCTTTGTCCTCGAGTGGCGTCACCAGAACAGGGGTGCGGGGAGGGTCCTGCTGGCCTATGAC : 3033
AJ004999 : ..................................................................................... : 2971
Tpn12    : ..................................................................................... : 2813
Tpn19    : ..................................................................................... : 2812
Tpn2L6   : ..................................................................................... : 2808
Tpn2L7   : ..................................................................................... : 2808
Tpn4     : ....C........G...C..................................................................... : 2811
Tpn14    : ....C........G...C..................................................................... : 2800
Tpn15    : ....C............C.....................................................G.............. : 2810
Tpn21L0  : ....C........G...C..................................................................... : 2816
Tpn21LN  : ....C........G...C..................................................................... : 2816

               ---------------------------------------------------------------------------
                  *      3060         *      3080         *      3100         *      3120
AL023516 : AGTTCCACCGCCCGCGCCCCCCGCGCCACCCCCGGGGCCGAACTGCTGCTGGGGACACGGGATGGGGACGGGGTGACAGC : 3113
AJ004999 : ............................CA....................................................... : 3051
Tpn12    : ..................................................................................... : 2893
Tpn19    : ..................................................................................... : 2892
Tpn2L6   : ..................................................................................... : 2888
Tpn2L7   : ..................................................................................... : 2888
Tpn4     : ..................................................................................... : 2891
Tpn14    : ..................................................................................... : 2880
Tpn15    : ..................................................................................... : 2890
Tpn21L0  : ..................................................................................... : 2896
Tpn21LN  : ..................................................................................... : 2896

               ---------------------------------------------------------------------------
                  *      3140         *      3160         *      3180         *      3200
AL023516 : GGTGACACTGCGGCTGGCGCGGCCATCACCGGGGGATGAGGGCACCTACATCTGCTCCGTGTTCCTGCCCCACGGGCACA : 3193
AJ004999 : ..................................................................................... : 3131
Tpn12    : ..................................................................................... : 2973
Tpn19    : ..................................................................................... : 2972
Tpn2L6   : ..................................................................................... : 2968
Tpn2L7   : ..................................................................................... : 2968
Tpn4     : ..................................................................................... : 2971
Tpn14    : A.................A.A.................................................................. : 2960
Tpn15    : ..................................................................................... : 2970
Tpn21L0  : A.................A.A.................................................................. : 2976
Tpn21LN  : A.................A.A.................................................................. : 2976

               ----------------------------]
                  *      3220         *      3240         *      3260         *      3280
AL023516 : CACAGACAGTGCTGCAGCTCCACGTCTTTGGTGCGTCCATGTGGGGCAGGCGGTGTTCCTATGGGGTGTGGGGTTGGGCA : 3273
AJ004999 : ..................................................................................... : 3211
Tpn12    : ..................................................................................... : 3053
Tpn19    : ..................................................................................... : 3052
Tpn2L6   : ..................................................................................... : 3048
Tpn2L7   : ..................................................................................... : 3048
Tpn4     : ..................................................................................... : 3051
Tpn14    : ..................................................................................... : 3040
Tpn15    : ..................................................................................... : 3050
Tpn21L0  : ..................................................................................... : 3056
Tpn21LN  : ..................................................................................... : 3056
```

```
                  *      3300         *      3320         *      3340         *      3360
AL023516 : GTGTTCCTACGGAGTGTGTATGACTGGGTGGTATTCCTATTGGTCAGATAGGACATATGGGAGCAGGCGGTATTCCTATG : 3353
AJ004999 : ..............................................................................  : 3291
Tpn12    : ..............................................................................  : 3133
Tpn19    : ..............................................................................  : 3132
Tpn2L6   : ..............................................................................  : 3128
Tpn2L7   : ..............................................................................  : 3128
Tpn4     : .........T....................................................................  : 3131
Tpn14    : .........T....................................................................  : 3120
Tpn15    : .........T....................................................................  : 3130
Tpn21L0  : .........T....................................................................  : 3136
Tpn21LN  : .........T....................................................................  : 3136

                  *      3380         *      3400         *      3420         *      3440
AL023516 : GGGGCTGTAGGGTGGATGGGACTGGGTGATATTCCTGTGGGGGGCTGTAGGGTGGATGGGACTGGGTGGTATTCCTATGG : 3433
AJ004999 : ...-..........................................................................  : 3370
Tpn12    : ..............................................................................  : 3213
Tpn19    : ..............................................................................  : 3212
Tpn2L6   : ..............................................................................  : 3208
Tpn2L7   : ..............................................................................  : 3208
Tpn4     : ..................................TT..........................................  : 3211
Tpn14    : ..............................................................................  : 3200
Tpn15    : ..............................................................C...............  : 3210
Tpn21L0  : ..............................................................................  : 3216
Tpn21LN  : ..............................................................................  : 3216

                  *      3460         *      3480         *      3500         *      3520
AL023516 : AGGCTATAGGGTGGATGGGACCGGGTGGTATTCCTATGAGGACTATAGGATGGGGTGGCATCATCCCATAGTTCACCTGT : 3513
AJ004999 : ..............................................................................  : 3450
Tpn12    : ..............................................................................  : 3293
Tpn19    : ..............................................................................  : 3292
Tpn2L6   : ..............................................................................  : 3288
Tpn2L7   : ..............................................................................  : 3288
Tpn4     : ..............................................................................  : 3291
Tpn14    : ..............................................................................  : 3280
Tpn15    : .........----------------------------------------....A.....A.....................  : 3251
Tpn21L0  : ..............................................A...............................  : 3296
Tpn21LN  : ..............................................A...............................  : 3296

                  *      3540         *      3560         *      3580         *      3600
AL023516 : AGGTTTATAGGGGGGGATGAGCCCTATAGGAGGGGCCCTATGGGCTATATGGGACCGATGTCCCCCCACATGTCTCCCAG : 3593
AJ004999 : ............................C.................................................  : 3530
Tpn12    : ..............................................................................  : 3373
Tpn19    : ..............................................................................  : 3372
Tpn2L6   : ..............................................................................  : 3368
Tpn2L7   : ..............................................................................  : 3368
Tpn4     : ..............................................................................  : 3371
Tpn14    : ..............................................................................  : 3360
Tpn15    : ..............................................................................  : 3331
Tpn21L0  : ................................A.............................--.................  : 3374
Tpn21LN  : ................................A.............................--.................  : 3374

           [Exon 5----------------------------------------------------------------------------
                  *      3620         *      3640         *      3660         *      3680
AL023516 : AGCCCCCCAAGGTGACGCTGTCCCCGAAGAACCTGGTGGTGGCCCCGGGGACGTCAGCAGAGCTACGCTGCCACGTGTCT : 3673
AJ004999 : ..............................................................................  : 3610
Tpn12    : ..............................................................................  : 3453
Tpn19    : ..............................................................................  : 3452
Tpn2L6   : ..............................................T...............................  : 3448
Tpn2L7   : ..............................................T...............................  : 3448
Tpn4     : ..............................................T...............................  : 3451
Tpn14    : ..............................................................................  : 3440
Tpn15    : ..............................................A...............................  : 3411
Tpn21L0  : ..............................................T.......T.......................  : 3454
Tpn21LN  : ..............................................T.......T.......................  : 3454

           ----------------------------------------------------------------------------
                  *      3700         *      3720         *      3740         *      3760
AL023516 : GGCTTCTACCCCTTGGATGTGACGGTGACGTGGCAGCGCCGCGCCGGGGGCTCGGGGACATCACAGTCACCCAGGGACAC : 3753
AJ004999 : ..............................................................................  : 3690
Tpn12    : ..............................................................................  : 3533
Tpn19    : ..............................................................................  : 3532
Tpn2L6   : ..............................................................G...............  : 3528
Tpn2L7   : ..............................................................G...............  : 3528
Tpn4     : ..............................................A...........G.....................  : 3531
Tpn14    : ..............................................A.....A.....G.....................  : 3520
Tpn15    : ..............................................A.....A.....G.....................  : 3491
Tpn21L0  : ..............................................A.....A.....................  : 3534
Tpn21LN  : ..............................................A.....A.....................  : 3534

           ----------------------------------------------------------------------------
                  *      3780         *      3800         *      3820         *      3840
AL023516 : AGTGATGGACAGCTGGACTTCAGGTCACCGCCAGGCAGCCGATGGAACCTACAGCCGGACGGCGGCAGCACGGCTGATCC : 3833
AJ004999 : ..............................................................................  : 3770
Tpn12    : ..............................................................................  : 3613
Tpn19    : ..............................................................................  : 3612
Tpn2L6   : ..............................................................................  : 3608
Tpn2L7   : ..............................................................................  : 3608
Tpn4     : ..............................................................A...............  : 3611
Tpn14    : ..............................................................A...............  : 3600
Tpn15    : ..............................................................................  : 3571
Tpn21L0  : ..............................................................................  : 3614
Tpn21LN  : ..............................................................................  : 3614
```

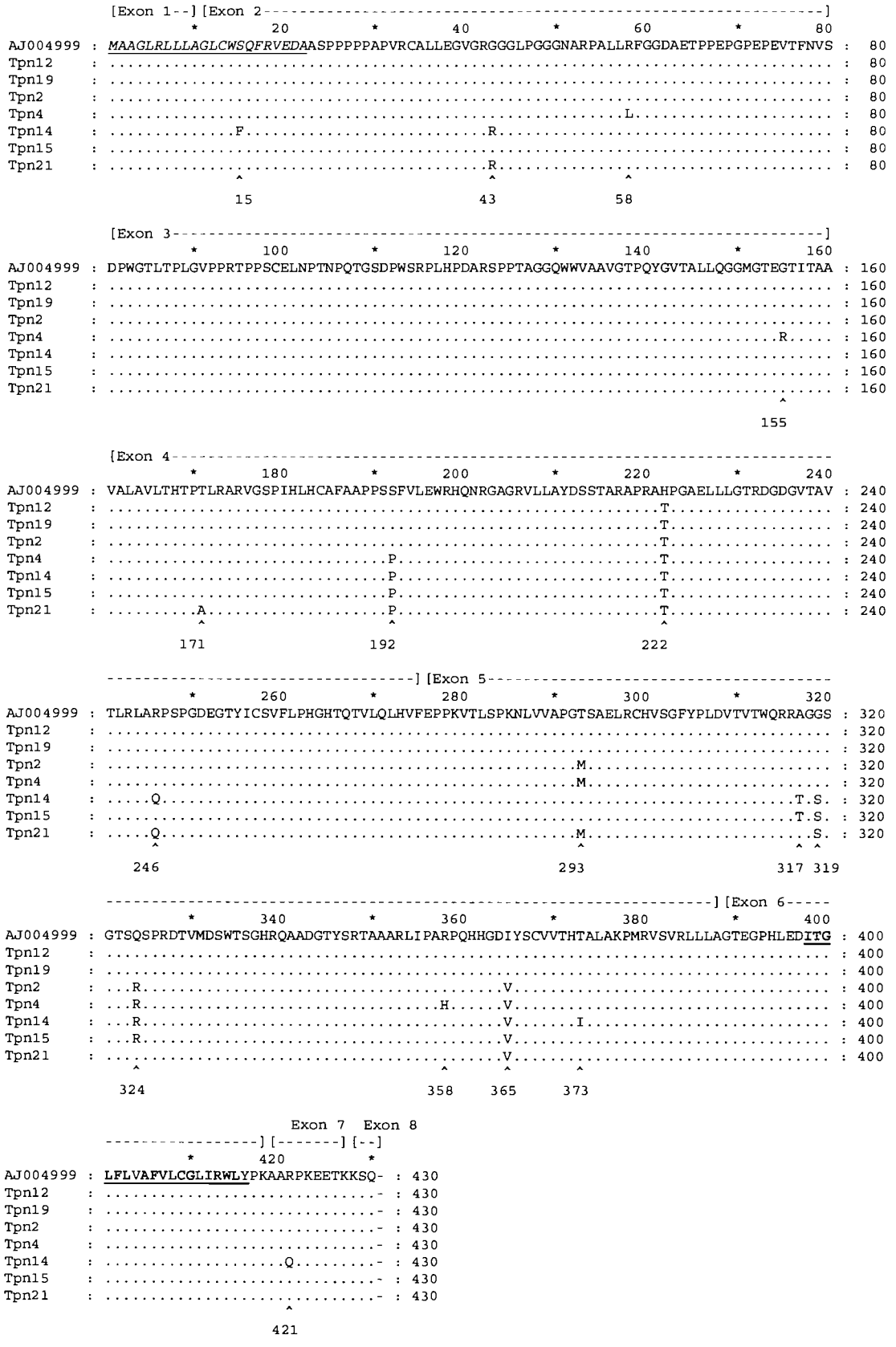

* 3860 * 3880 * 3900 * 3920
AL023516 : CCGCACGCCCCCAACACCACGGGGACATCTACAGCTGCGTTGTCACCCACACTGCACTGGCCAAACCAATGCGTGTCTCC : 3913
AJ004999 : ..... : 3850
Tpn12 : ..... : 3693
Tpn19 : ..... : 3692
Tpn2L6 : .....G.....C..... : 3688
Tpn2L7 : .....G.....C..... : 3688
Tpn4 : .....A.....G.....C..... : 3691
Tpn14 : .....G.....T..... : 3680
Tpn15 : .....G.....C.....T : 3651
Tpn21L0 : .....G.....C..... : 3694
Tpn21LN : .....G.....C..... : 3694
]
* 3940 * 3960 * 3980 * 4000
AL023516 : GTCCGACTGCTCCTGGCTGGTGAGGGGGGATGTGGGGATATTGGAAACACGTGGAGGTATTGGGATGCTGGGACCATGGT : 3993
AJ004999 : ..... : 3930
Tpn12 : ..... : 3773
Tpn19 : ..... : 3772
Tpn2L6 : .....G.... : 3768
Tpn2L7 : .....G.... : 3768
Tpn4 : .....G.... : 3771
Tpn14 : ..... : 3760
Tpn15 : .....G.... : 3731
Tpn21L0 : .....A.....G.... : 3774
Tpn21LN : .....A.....G.... : 3774
* 4020 * 4040 * 4060 * 4080
AL023516 : TAGGAGGGTCTGAGGGACATCAGGACCATGGCCTGGGACAATGGGAGATCATGGATTTGGGTTGGGGACCCCACCCAGGA : 4073
AJ004999 : ..... : 4010
Tpn12 : ..... : 3853
Tpn19 : ..... : 3852
Tpn2L6 : ..... : 3848
Tpn2L7 : ..... : 3848
Tpn4 : ..... : 3851
Tpn14 : ..... : 3840
Tpn15 : ..... : 3811
Tpn21L0 : ..... : 3854
Tpn21LN : ..... : 3854
[Exon 6
* 4100 * 4120 * 4140 * 4160
AL023516 : TGGTGACACTGTGCTTAGGGCTGTCGTTGTCCCCACAGGCACCGAGGGACCGCACCTGGAGGACATCACGGGGCTCTTCT : 4153
AJ004999 : ..... : 4090
Tpn12 : ..... : 3933
Tpn19 : ..... : 3932
Tpn2L6 : ..... : 3928
Tpn2L7 : ..... : 3928
Tpn4 : ..... : 3931
Tpn14 : ..... : 3920
Tpn15 : ..... : 3891
Tpn21L0 : ..... : 3934
Tpn21LN : ..... : 3934
]
* 4180 * 4200 * 4220 * 4240
AL023516 : TGGTGGCCTTTGTCCTCTGTGGCCTCATCCGTTGGCTCTACCCTAAAGGTGAGTGCTGTTCCCACATCCCAGTGCCCCCA : 4233
AJ004999 : ..... : 4170
Tpn12 : ..... : 4013
Tpn19 : ..... : 4012
Tpn2L6 : ..... : 4008
Tpn2L7 : ..... : 4008
Tpn4 : ..... : 4011
Tpn14 : ..... : 4000
Tpn15 : ..... : 3971
Tpn21L0 : .....T..... : 4014
Tpn21LN : .....T..... : 4014
* 4260 * 4280 * 4300 * 4320
AL023516 : CATCCTCACACCCCAATATCCCAATGGCCCATGTCCCCATGAGCAATGTCACTATGTCCCAATATCCTAATGATGCTGTG : 4313
AJ004999 : ..... : 4250
Tpn12 : ..... : 4093
Tpn19 : ..... : 4092
Tpn2L6 : ..... : 4088
Tpn2L7 : ..... : 4088
Tpn4 : ..... : 4091
Tpn14 : .....G..... : 4080
Tpn15 : ..... : 4051
Tpn21L0 : ..... : 4094
Tpn21LN : ..... : 4094

```
                                                                                    [Exon 7-----------------]
                  *       4340         *       4360         *       4380         *       4400
AL023516 : TACCCATGTGTCCCCATGTCCCTATTCCACTCACTCTTTCTCTCCCCTCAGCTGCACGACCCAAAGAGGAAACCAAGGTA : 4393
AJ004999 : ................................................................................ : 4330
Tpn12    : ................................................................................ : 4173
Tpn19    : ................................................................................ : 4172
Tpn2L6   : ................................................................................ : 4168
Tpn2L7   : ................................................................................ : 4168
Tpn4     : ................................................................................ : 4171
Tpn14    : .........................................................A...................... : 4160
Tpn15    : ...................T............................................................ : 4131
Tpn21L0  : ....................T........................................................... : 4174
Tpn21LN  : ....................T........................................................... : 4174

                  *       4420         *       4440         *       4460         *       4480
AL023516 : ACATTCCTCCCCAAAAACCCCAAATCCCCCCAAAACACCTCCAAGCACCCCAAAACTCACCATTCTCATTCCCCCCCCC : 4473
AJ004999 : ..............................-................................................. : 4409
Tpn12    : ................................................................................ : 4253
Tpn19    : ................................................................................ : 4252
Tpn2L6   : ....C..C........................................................................ : 4248
Tpn2L7   : ....C..C........................................................................ : 4248
Tpn4     : ...........................................................................A.... : 4251
Tpn14    : ....C..C........................................................................ : 4240
Tpn15    : ....C..C...A.....A...................................................C.......... : 4211
Tpn21L0  : ....C..C................C....................................................... : 4254
Tpn21LN  : ....C..C................C....................................................... : 4254

                                         [Exon 8----]
                  *       4500         *       4520         *       4540         *       4560
AL023516 : CCCCCCCCCCCC--ATGCCTTGCAGAAATCGCAGTGACCTCCACTCCAGCTCTCAGCACCTC------------------ : 4533
AJ004999 : ............CC................................................------------------ : 4471
Tpn12    : ...........---................................................------------------ : 4312
Tpn19    : ............--................................................------------------ : 4312
Tpn2L6   : .......-------................................................------------------ : 4303
Tpn2L7   : .......-------................................................------------------ : 4303
Tpn4     : ...-----------................A...............................------------------ : 4302
Tpn14    : ...........---.......................T........................------------------ : 4299
Tpn15    : ....G......---........T.......................................CACTGCAGCTCTCAGCAC : 4288
Tpn21L0  : .....---------................................................------------------ : 4307
Tpn21LN  : ...-----------................................................------------------ : 4305

                [polA]        polyA_Addition      -
                  *       4580         *       4600         *       4620         *       4640
AL023516 : ---AGCTCCAGATAAAGAGTTTTTCACCCCAAAGTTCTCTCTGTGTGGTGGTGTCCCCACAGATCTGGGTGCAGAGGGGG : 4610
AJ004999 : ---.................................A.A.A....................................... : 4548
Tpn12    : ---............................................................................. : 4389
Tpn19    : ---............................................................................. : 4389
Tpn2L6   : ---............................................................................. : 4380
Tpn2L7   : ---............................................................................. : 4380
Tpn4     : ---............................................................................. : 4379
Tpn14    : ---............................................................................. : 4376
Tpn15    : CTC......................G...................................................... : 4368
Tpn21L0  : ---...........................................................G................. : 4384
Tpn21LN  : ---...........................................................G................. : 4382

                  *       4660         *       4680         *       4700         *       4720
AL023516 : GAGAAATGGGGGCAAACTGGGAGCAGTGGGAGCAGTGGGAGGAAGTCCTGGGTTGGTGAGGCAGATGAGTGGCACCTGGG : 4690
AJ004999 : ................................................................................ : 4628
Tpn12    : ................................................................................ : 4469
Tpn19    : .....C.......................................................................... : 4469
Tpn2L6   : .....C..........................................................C............... : 4460
Tpn2L7   : .....C..........................................................C............... : 4460
Tpn4     : .....C..........................................................C............... : 4459
Tpn14    : .....C..........................................................C......A........ : 4456
Tpn15    : .....C..........................................................C............... : 4448
Tpn21L0  : .....C..........................................................C......A........ : 4464
Tpn21LN  : .....C..........................................................C......A........ : 4462

                  *       4740         *       4760         *       4780         *       4800
AL023516 : GACATCTGGGTGCCATCCCTTGTGGACATCTGGGTGACACTGCGATTGCATTGGGTGACATTGGGATCC-CATCCCAAGC : 4769
AJ004999 : ...........................................-.....C...................T..GGT..CTG : 4707
Tpn12    : .....................................................................-.......... : 4548
Tpn19    : .....................................................................-.......... : 4548
Tpn2L6   : .....................................................................-.......... : 4539
Tpn2L7   : .....................................................................-.......... : 4539
Tpn4     : ...........................................A.........................-.......... : 4538
Tpn14    : .....................................................................-.......... : 4535
Tpn15    : .....................................T...............................-.......... : 4527
Tpn21L0  : .....................................................................-.......... : 4543
Tpn21LN  : .....................................................................-.......... : 4541

                  *       4820         *       4840         *       4860         *       4880
AL023516 : TGGACATCGGGATGCCATGAAGGATGGCCACGTCCCATCACATGTGGGCATCTGGGTGACATCGAGGATGGTCATTTCCC : 4849
AJ004999 : CA.----------------------------------------------------------------------------- : 4710
Tpn12    : ................................................................................ : 4628
Tpn19    : ................................................................................ : 4628
Tpn2L6   : ................................................................................ : 4619
Tpn2L7   : ................................................................................ : 4619
Tpn4     : ..............................................................AG................ : 4618
Tpn14    : ....................................G........................................... : 4615
Tpn15    : ................................................................G............... : 4607
Tpn21L0  : ................................................................G............... : 4623
Tpn21LN  : ................................................................G............... : 4621
```

Appendix 5 Alignment of tapasin genomic sequences with AJ004999.

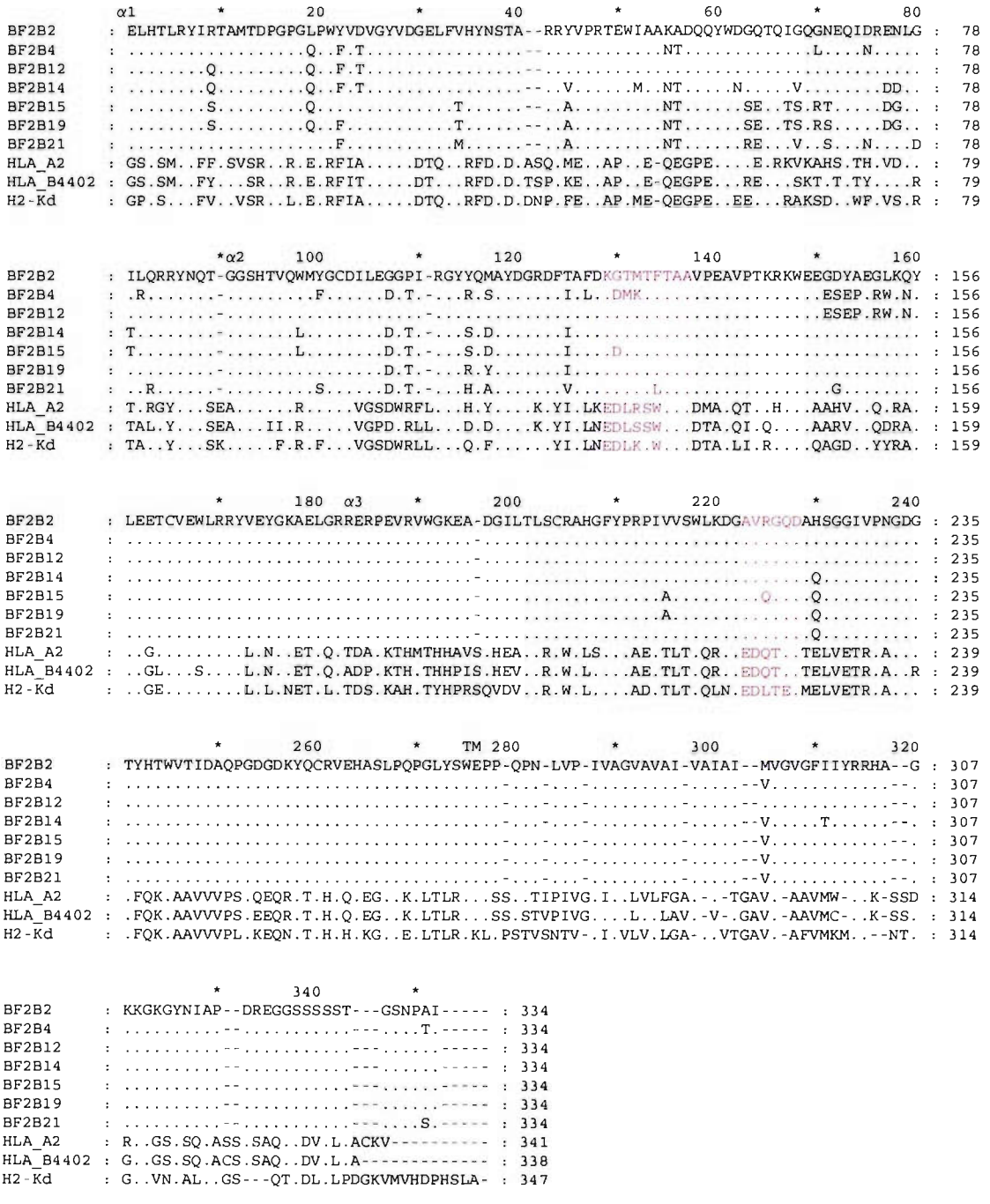
Consensus genomic sequences were produced for each haplotype and aligned with AL023516 and AJ004999 sequences using ClustalX and Genedoc analysis programs. Nucleotide identity with the AL023516 sequence is represented by dots (.), gaps introduced to maximize alignment are shown by dashes (-). Exons, BlB2 gene, putative TATA box, polyadenylation site and addition site (predicted by Frangoulis *et al*) are indicated above the alignment.

8.6 Appendix 6 Alignment of amino acid sequences of tapasin with AJ004999



Polypeptide sequences were deduced from the cDNA sequences from the seven haplotypes sequenced at Compton, and aligned with the AJ004999 peptide sequence using GCG and GeneDoc analysis programs. Amino acid identity with the AJ004999 peptide sequence is represented by dots (.). Exon boundaries are indicated above the alignment. The predicted signal sequences is italicised and underlined, and predicted transmembrane domain is underlined and in bold. The position of residues which vary between haplotypes is shown beneath the alignment.

8.7 Appendix 7 Alignment of the amino acid sequences of chicken BF2 and mammalian class I molecules



Appendix 7 Alignment of the amino acid sequences of chicken BF2 and mammalian class I molecules

Chicken BF2 protein sequences were aligned with those of HLA-A2 (K02883), HLA-B4402 (M24038) and H2-K^d (X01815), all with signal sequences removed (Kaufman et al., 1992). Dots (.) indicate identities, dashes (-) indicate gaps introduced to maximize alignments. Numbers to the right show the amino acid number of the right hand amino acid. Numbers above the alignment indicate alignment spacing, and strictly not amino acid number. Domains are indicated above the alignment by $\alpha 1$, $\alpha 2$, $\alpha 3$ representing exons 2, 3 and 4, and TM representing the transmembrane and cytoplasmic domains encoded by exons 5-8 (boundaries are aligned with the second character). Residues that align with those predicted to interact with mammalian tapasin are shown in red text (mammalian residues 128-136 align with chicken residues 125-133, and mammalian residues 222-227 align with chicken residues 218-223).

8.8 Appendix 8 Amino acid alignments of mature chicken BF1 proteins

```
                   *        20         *        40         *        60         *        80
BF1B2  : ELHTLRYISTAMTDPGPGQPWYVDVGYVDGELFTHYNSTARRAVPRTEWIAANTDQQYWDRETQIAQGNEQIDRENLDIR :  80
BF1B4  : ........H........................V........Y......M..KA......GQ...G.R..RSVKVS..TL :  80
BF1B12 : ...S...VH.........L..F...........V........Y......M..........GQ...G....RSVEVS.NTL :  80
BF1B19 : ...S...VH.........L..F...........V........Y......M..........GQ...G....RSVKVS.NTL :  80
BF1B21 : ........H........................V........Y......M..KA......GQ...G.R..RSVKVS..TL :  80

                   *       100         *       120         *       140         *       160
BF1B2  : QQRHNQTGGSHTAQWMYGCDILEDGTIRGYHQMACDGRDFIALAEDMKTFTAAVPEAVPTKRKWEEGGYAERKKQYLEET : 160
BF1B4  : .E.Y........V...F.............R.V.Y..K....FDK.......................V..GW.S..... : 160
BF1B12 : .E.Y........V.L.................T.Y.......FDKGTM....................V...W.S..... : 160
BF1B19 : .E.Y........V.L.................T.Y.......FDKGTM....................V...W.S..... : 160
BF1B21 : .E.Y........V...F.............R.V.Y..K....FDK.......................V..GW.S..... : 160

                   *       180         *       200         *       220         *       240
BF1B2  : CVEGLRRYVEYGKAELGRRERPEVRVWGKEADGILTLSCRAHGFYPRPIVVSWLKDGAVRGQDAQSGGIVPNGDGTYHTW : 240
BF1B4  : ...W.............................................A...................M.......... : 240
BF1B12 : .................................................A.............................. : 240
BF1B19 : .................................................A.............................. : 240
BF1B21 : ...W.................................................................M.......... : 240

                   *       260         *       280         *       300         *       320
BF1B2  : VTIDAQPGDGDKYQCRVEHASLPQPGLYSWEPPQPNLVPIVAGVAVAIVAIAIVVGVGFILYRCHAGKKGKGYNIAPDRE : 320
BF1B4  : ....V.......................................................I..R................ : 320
BF1B12 : ............................................................I..R................ : 320
BF1B19 : ............................................................I..R................ : 320
BF1B21 : ............................................................I..R................ : 320

                   *
BF1B2  : DGSSSSSTGSNPSI- : 334
BF1B4  : G.............- : 334
BF1B12 : G...........A.- : 334
BF1B19 : G...........A.- : 334
BF1B21 : G.............- : 334
```

8.9 Appendix 9 Amino acid alignments of chicken TAP1 proteins

```
                   *        20         *        40         *        60         *        80
B2_TAP1  : MGTMGAGRRLLLSLSPERRRCAAVMGLMAASALGEMAVPYYMGRASDWVAREDELAAILPMVLLGLSSAVTELVCDVTFV :  80
B4_TAP1  : .....................................................K.......................... :  80
B12_TAP1 : ..K............................................................................. :  80
B14_TAP1 : ................................................................................ :  80
B15_TAP1 : ................................................................................ :  80
B19_TAP1 : ..K............................................................................. :  80
B21_TAP1 : ................................................................................ :  80

                   *       100         *       120         *       140         *       160
B2_TAP1  : GTLSRTQSRLQRRVFAAVLRQSITELRADGAGDVAMRVTRDAEDVREALGEALSLLLWYLARGLCLFATMAWLSPRMALL : 160
B4_TAP1  : ..................................................K............................. : 160
B12_TAP1 : ................................................................................ : 160
B14_TAP1 : ...........H.................................................................... : 160
B15_TAP1 : ..................................T............................................. : 160
B19_TAP1 : ................................................................................ : 160
B21_TAP1 : ................................................................................ : 160

                   *       180         *       200         *       220         *       240
B2_TAP1  : TVLALPLLLALPRAVGHFRQALAPQMQKAQARASEVAVETFQAMATVRSFANEDGAAAHYRQRLQQSHRLEKKDVALYTA : 240
B4_TAP1  : .A.............................................................................. : 240
B12_TAP1 : .A.............................................................................. : 240
B14_TAP1 : .A.............................................................................. : 240
B15_TAP1 : .A...............Q.............................................................. : 240
B19_TAP1 : .A.............................................................................. : 240
B21_TAP1 : .A.............................................................................. : 240

                   *       260         *       280         *       300         *       320
B2_TAP1  : SLWTSGFSALALKMGILYYGGQLVAAGTVSTGDLVTFLLYQIQFTDVLRVLLDYFPTLMKAVGSSEKIFEFLDREPQVAP : 320
B4_TAP1  : ................................................................................ : 320
B12_TAP1 : .................................................E.............................. : 320
B14_TAP1 : .................................................E.............................. : 320
B15_TAP1 : .................................................E.............................. : 320
B19_TAP1 : .................................................E.............................. : 320
B21_TAP1 : ................................................................................ : 320

                   *       340         *       360         *       380         *       400
B2_TAP1  : SGTMAPADLQGHLQLEDVWFSYPGRQEPVLKGVSLELRPGEVLALLGPPGAGKSTLVALVSRLHQPTAGRLLLDGHPLPA : 400
B4_TAP1  : ................................................................................ : 400
B12_TAP1 : ................................................................................ : 400
B14_TAP1 : ................................................................................ : 400
B15_TAP1 : ................................................................................ : 400
B19_TAP1 : ................................................................................ : 400
B21_TAP1 : ......T......................................................................... : 400
```

```
                  *       420         *       440         *       460         *       480
B2_TAP1  : YQHSYLCRQVAVVPQEPLLFARSLHANISYGLGGCSRAQVTAAARRVGAHDFITRLPQGYDTEVGELGGQLSGGQRQAVA : 480
B4_TAP1  : ................................................................................ : 480
B12_TAP1 : ..........................................Q..................................... : 480
B14_TAP1 : ..........................................L..................................... : 480
B15_TAP1 : ................................................................................ : 480
B19_TAP1 : ..........................................Q..................................... : 480
B21_TAP1 : ....................................W........................................... : 480

                  *       500         *       520         *       540         *       560
B2_TAP1  : IARALLRDPRILILDEHTSALDTESQQQVEQEILAAKGSGRAVLMVTGRAALAARAQRVVVLEGGEVRQEGPPHEVLRPG : 560
B4_TAP1  : ................................................................................ : 560
B12_TAP1 : ..........................................E..................................... : 560
B14_TAP1 : ......Q......................................................................... : 560
B15_TAP1 : ................................................................................ : 560
B19_TAP1 : ..........................................E..................................... : 560
B21_TAP1 : ................................................................................ : 560

                  *       580         *
B2_TAP1  : SLLRDWGQQGAPGEGDRG-SG---GEG~~~ : 583
B4_TAP1  : ..................-.----...--- : 583
B12_TAP1 : ..................IA.VMD...RGW : 590
B14_TAP1 : ..................-.----...--- : 583
B15_TAP1 : ..................-.----...--- : 583
B19_TAP1 : ..................IA.VMD...RGW : 590
B21_TAP1 : ..................-.----...--- : 583
```

8.10 Appendix 10 Amino acid alignments of chicken TAP2 proteins

```
                  *       20          *       40          *       60          *       80
B2_TAP2  : MAMPPYILRLSCTLLLADLALMLALAHFFPALAHLGWVGSWLEAGLRLLVLGGAGQLLAPRGPHGAAVLLSLGPAIFLTL :  80
B4_TAP2  : ..............................................................R................. :  80
B12_TAP2 : ..............................................................R................. :  80
B14_TAP2 : ........H...................................................................... :  80
B15_TAP2 : ........H.....................S................................................. :  80
B19_TAP2 : ..............................................................R................. :  80
B21_TAP2 : ................................................................................ :  80

                  *       100         *       120         *       140         *       160
B2_TAP2  : RGYVGLPGAAPVLLAMATPSWLVLTHGTAVVALLTWSLLVPTVATGAKEAEAWVPLRRLLALAWPEWPFLGCAFLFLALA : 160
B4_TAP2  : ................................................................................ : 160
B12_TAP2 : ................................................................................ : 160
B14_TAP2 : ................................................................................ : 160
B15_TAP2 : ................................................................................ : 160
B19_TAP2 : ................................................................................ : 160
B21_TAP2 : ................................................................................ : 160

                  *       180         *       200         *       220         *       240
B2_TAP2  : ALGETSVPYCTGRALDVLRQGDGLAAFTAAVGLMCLASASSSLFAGCRGGLFTFIRFRFILRTRDQLFSSLVYRDLAFFQ : 240
B4_TAP2  : ..............................................................V................. : 240
B12_TAP2 : ................................................................................ : 240
B14_TAP2 : ..........................................................G...V................. : 240
B15_TAP2 : ..........................................................G...V................. : 240
B19_TAP2 : ................................................................................ : 240
B21_TAP2 : ..............................................................V................. : 240

                  *       260         *       280         *       300         *       320
B2_TAP2  : KTTAAELASRLTTDVTLASNVLALNINVMLRNLGQVLGLCAFMLGLSPRLTMLALLEVPLAVTARKVYDTRHQMLQRAVL : 320
B4_TAP2  : ................................................................................ : 320
B12_TAP2 : ................................................................................ : 320
B14_TAP2 : ......................T......................................................... : 320
B15_TAP2 : .......S.............T......................................................... : 320
B19_TAP2 : ................................................................................ : 320
B21_TAP2 : ................................................................................ : 320

                  *       340         *       360         *       380         *       400
B2_TAP2  : DAAADTGAAVQESISSIEMVRVFNGEEEEEHRYSQVLDRTLRLRDQRDTERAIFLLIRRVLQLAVQALVLYCGHQQLREG : 400
B4_TAP2  : ..............................................................Q................. : 400
B12_TAP2 : ..............................................................Q.................H.. : 400
B14_TAP2 : ..............................................................Q................. : 400
B15_TAP2 : ..............................................................Q................. : 400
B19_TAP2 : ..............................................................Q.................H.. : 400
B21_TAP2 : ....N.......................Y..................................Q................. : 400

                  *       420         *       440         *       460         *       480
B2_TAP2  : TLTAGSLVAFILYQTKAGSCVQALAYSYGDLLSNAAAACKVFDYLNWERAVGAGGTYVPTRLRGHVTFHRVSFAYPTRPE : 480
B4_TAP2  : ..................................V.........D...P..............I.............. : 480
B12_TAP2 : .....G.......................................................................... : 480
B14_TAP2 : ..............N.............................D...........M.......I.............. : 480
B15_TAP2 : ..............N.............................D...........M.......I.............. : 480
B19_TAP2 : .....G.......................................................................C.. : 480
B21_TAP2 : ..................................V.........D................................... : 480
```

```
                    *       500         *       520         *       540         *       560
B2_TAP2   : RLVLQDVTFELRPGEVTALAGLNGSGKSTCVALLERFYEPGAGEVLLDGVPLRDYEHRYLHRQVALVGQEPVLFSGSIRD : 560
B4_TAP2   : ............S.................................................................... : 560
B12_TAP2  : ............................................................................... : 560
B14_TAP2  : ............................................................................... : 560
B15_TAP2  : ............................................................................... : 560
B19_TAP2  : ............................................................................... : 560
B21_TAP2  : ............................................................................... : 560

                    *       580         *       600         *       620         *       640
B2_TAP2   : NIAYGMEDCEEEEIIAAARAAGALGFISALEQGFGTDVGERGGQLSAGQKQRIAIARALVRHPTVLILDEATSALDGDSD : 640
B4_TAP2   : ...........................................................R.................... : 640
B12_TAP2  : ............................................................................... : 640
B14_TAP2  : .........K...................................................................... : 640
B15_TAP2  : ...........................................................R.................... : 640
B19_TAP2  : .........K..................................................R.................... : 640
B21_TAP2  : ..................K........................................R.................... : 640

                    *       660         *       680         *       700
B2_TAP2   : AMLQQWVRNGGDRTVLLITHQPRMLEKADRIVVLEHGTVAEMGTPAELRTRGGPYSRLLQH : 701
B4_TAP2   : ............................V.............................. : 701
B12_TAP2  : ............................................................ : 701
B14_TAP2  : ............................................................ : 701
B15_TAP2  : ............................................................ : 701
B19_TAP2  : ............................V.............................. : 701
B21_TAP2  : ............................V.............................. : 701
```

8.11 Appendix 11 Amino acid alignments of chicken BLB proteins

BLB1

```
                      *         20          *         40          *         60          *         80
BlB1 2b1b2  : FFQWTFKAECHYLNGTERARFLERHIYNRQQFMHFDSDVGKYVADTPLGERQAEIWNSNAEILEDEMNAVDTFCRHNYGV : 80
BlB1 4b1b2  : ...................V.Y.V.YV....EYA........H....A...P...Y.........NR..E...Y....... : 80
BlB1 12b1b2 : ..FCGAIF..........V...D.E......YA.................P...Y......FM.NR..E........... : 80
BlB1 14b1b2 : ..FYGMI...........V.L...Q.......A..........G..A.......Y...D..Y..NR..E..R........ : 80
BlB1 15b1b2 : ..FYGVIF..QF......V.L...Q......TS...............D.Y.V.....D.....NR..E...V....... : 80
BlB1 19b1b2 : ..FCGAIF..........V...D.E......YA.................P...Y......FM.NR..E........... : 80
BlB1 21b1b2 : ...................V.Y.V.YV....EYA........H........P...Y.........NR..E...Y....... : 80


BlB1 2b1b2  : GESFTVQRS : 89
BlB1 4b1b2  : V........ : 89
BlB1 12b1b2 : ......... : 89
BlB1 14b1b2 : ......... : 89
BlB1 15b1b2 : D........ : 89
BlB1 19b1b2 : ......... : 89
BlB1 21b1b2 : V........ : 89
```

BLB2

```
                      *         20          *         40          *         60          *         80
BlB2 2b1b2  : FFFCGAIFECHYLNGTERVRYLQRYIYNRQQLVHFDSDVGKFVADTPLGEPQAEYWNSNAELLENIMNIADGSCRHNYGI : 80
BlB2 4b1b2  : ...Y.R.A............F.D.Q.......M.........................D..FM.IKR.EV.TF......V : 80
BlB2 12b1b2 : .......S.........................FT............S...................R..EV.RF......G : 80
BlB2 14b1b2 : ...Y.M.L.............D.Q......FM........Y...................FM..R..EV.TF......V : 80
BlB2 15b1b2 :  ..H.V.A..........A.F.A.H......YA........Y.....................L..TV..P........ : 80
BlB2 19b1b2 : .......S.....................YA............S...................R..EV.RF......G : 80
BlB2 21b1b2 : ...Y.K.G............F.D.Q......FA.................R..............L..EV.RV....... : 80


BlB2 2b1b2  : LESFTVQRS : 89
BlB2 4b1b2  : N.P...... : 89
BlB2 12b1b2 : V........ : 89
BlB2 14b1b2 : G........ : 89
BlB2 15b1b2 : ......... : 89
BlB2 19b1b2 : V........ : 89
BlB2 21b1b2 : ......... : 89
```

References:

(1999). Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature *401*, 921-923.

Afanassieff, M., Goto, R. M., Ha, J., Sherman, M. A., Zhong, L., Auffray, C., Coudert, F., Zoorob, R., and Miller, M. M. (2001). At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. J Immunol *166*, 3324-3333.

Alexander, J., Payne, J. A., Murray, R., Frelinger, J. A., and Cresswell, P. (1989). Differential transport requirements of HLA and H-2 class I glycoproteins. Immunogenetics *29*, 380-388.

Bacon, L. D., Hunter, D. B., Zhang, H. M., Brand, K., and Etches, R. (2004). Retrospective evidence that the MHC (B haplotype) of chickens influences genetic resistance to attenuated infectious bronchitis vaccine strains in chickens. Avian Pathol *33*, 605-609.

Bacon, L. D., Witter, R. L., Crittenden, L. B., Fadly, A., and Motta, J. (1981). B-haplotype influence on Marek's disease, Rous sarcoma, and lymphoid leukosis virus-induced tumors in chickens. Poult Sci *60*, 1132-1139.

Bangia, N., Lehner, P. J., Hughes, E. A., Surman, M., and Cresswell, P. (1999). The N-terminal region of tapasin is required to stabilize the MHC class I loading complex. Eur J Immunol *29*, 1858-1870.

Barber, L. D., Howarth, M., Bowness, P., and Elliott, T. (2001). The quantity of naturally processed peptides stably bound by HLA-A*0201 is significantly reduced in the absence of tapasin. Tissue Antigens *58*, 363-368.

Barnden, M. J., Purcell, A. W., Gorman, J. J., and McCluskey, J. (2000). Tapasin-mediated retention and optimization of peptide ligands during the assembly of class I molecules. J Immunol *165*, 322-330.

Bennett, E. M., Bennink, J. R., Yewdell, J. W., and Brodsky, F. M. (1999). Cutting edge: adenovirus E19 has two mechanisms for affecting class I MHC expression. J Immunol *162*, 5049-5052.

Bernot, A., Zoorob, R., and Auffray, C. (1994). Linkage of a new member of the lectin supergene family to chicken Mhc genes. Immunogenetics *39*, 221-229.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. Nature *329*, 506-512.

Bloom, S. E., and Bacon, L. D. (1985). Linkage of the major histocompatibility (B) complex and the nucleolar organizer in the chicken. Assignment to a microchromosome. J Hered *76*, 146-154.

Briles, W. E., Bumstead, N., Ewert, D. L., Gilmour, D. G., Gogusev, J., Hala, K., Koch, C., Longenecker, B. M., Nordskog, A. W., Pink, J. R., *et al.* (1982). Nomenclature for chicken major histocompatibility (B) complex. Immunogenetics *15*, 441-447.

Briles, W. E., Goto, R. M., Auffray, C., and Miller, M. M. (1993). A polymorphic system related to but genetically independent of the chicken major histocompatibility complex. Immunogenetics *37*, 408-414.

Briles, W. E., Mc, G. W., and Irwin, M. R. (1950). On multiple alleles effecting cellular antigens in the chicken. Genetics *35*, 633-652.

Brodsky, F. M., Bodmer, W. F., and Parham, P. (1979). Characterization of a monoclonal anti-beta 2-microglobulin antibody and its use in the genetic and biochemical analysis of major histocompatibility antigens. Eur J Immunol *9*, 536-545.

Bulik, S., Peters, B., and Holzhutter, H. G. (2005). Quantifying the contribution of defective ribosomal products to antigen production: a model-based computational analysis. J Immunol *175*, 7957-7964.

Cardozo, C., and Kohanski, R. A. (1998). Altered properties of the branched chain amino acid-preferring activity contribute to increased cleavages after branched chain residues by the "immunoproteasome". J Biol Chem *273*, 16764-16770.

Cascio, P., Hilton, C., Kisselev, A. F., Rock, K. L., and Goldberg, A. L. (2001). 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. Embo J *20*, 2357-2366.

Cattley, S. K., Longman, N., Dawkins, R. L., Gaudieri, S., Kulski, J. K., and Leelayuwat, C. (1999). Phylogenetic analysis of primate MIC (PERB11) sequences suggests that the representation of the gene family differs in different primates: comparison of MIC (PERB11) and C4. Eur J Immunogenet *26*, 233-238.

Chen, M., Stafford, W. F., Diedrich, G., Khan, A., and Bouvier, M. (2002). A characterization of the lumenal region of human tapasin reveals the presence of two structural domains. Biochemistry *41*, 14539-14545.

Chun, T., Grandea, A. G., 3rd, Lybarger, L., Forman, J., Van Kaer, L., and Wang, C. R. (2001). Functional roles of TAP and tapasin in the assembly of M3-N-formylated peptide complexes. J Immunol *167*, 1507-1514.

Clare, R. A., Strout, R. G., Taylor, R. L., Jr., Collins, W. M., and Briles, W. E. (1985). Major histocompatibility (B) complex effects on acquired immunity to cecal coccidiosis. Immunogenetics *22*, 593-599.

Copeman, J., Bangia, N., Cross, J. C., and Cresswell, P. (1998). Elucidation of the genetic basis of the antigen presentation defects in the mutant cell line .220 reveals polymorphism and alternative splicing of the tapasin gene. Eur J Immunol *28*, 3783-3791.

Crone, M., Simonsen, M., Skjodt, K., Linnet, K., and Olsson, L. (1985). Mouse monoclonal antibodies to class I and class II antigens of the chicken MHC. Evidence for at least two class I products of the B complex. Immunogenetics *21*, 181-187.

Daniel, S., Caillat-Zucman, S., Hammer, J., Bach, J. F., and van Endert, P. M. (1997). Absence of functional relevance of human transporter associated with antigen processing polymorphism for peptide selection. J Immunol *159*, 2350-2357.

Day, P. M., Esquivel, F., Lukszo, J., Bennink, J. R., and Yewdell, J. W. (1995). Effect of TAP on the generation and intracellular trafficking of peptide-receptive major histocompatibility complex class I molecules. Immunity *2*, 137-147.

Day, P. M., Yewdell, J. W., Porgador, A., Germain, R. N., and Bennink, J. R. (1997). Direct delivery of exogenous MHC class I molecule-binding oligopeptides to the endoplasmic reticulum of viable cells. Proc Natl Acad Sci U S A *94*, 8064-8069.

De Silva, A. D., Boesteanu, A., Song, R., Nagy, N., Harhaj, E., Harding, C. V., and Joyce, S. (1999). Thermolabile H-2Kb molecules expressed by transporter associated with antigen processing-deficient RMA-S cells are occupied by low-affinity peptides. J Immunol *163*, 4413-4420.

Dick, T. P., Bangia, N., Peaper, D. R., and Cresswell, P. (2002). Disulfide bond isomerization and the assembly of MHC class I-peptide complexes. Immunity *16*, 87-98.

Dick, T. P., Ruppert, T., Groettrup, M., Kloetzel, P. M., Kuehn, L., Koszinowski, U. H., Stevanovic, S., Schild, H., and Rammensee, H. G. (1996). Coordinated dual cleavages induced by the proteasome regulator PA28 lead to dominant MHC ligands. Cell *86*, 253-262.

Diedrich, G., Bangia, N., Pan, M., and Cresswell, P. (2001). A role for calnexin in the assembly of the MHC class I loading complex in the endoplasmic reticulum. J Immunol *166*, 1703-1709.

Duch, M., Tolstrup, A., Dalum, I., Jespersen, T., Mouritsen, S., and Pedersen, F. S. (1999). Functional testing of a bicistronic retroviral vector for intracellular peptide production. Biotechniques *26*, 1032-1034, 1036.

Edwards, J. C., Bowness, P., and Archer, J. R. (2000). Jekyll and Hyde: the transformation of HLA-B27. Immunol Today *21*, 256-260.

Elliott, T. (1997). How does TAP associate with MHC class I molecules? Immunol Today *18*, 375-379.

Elliott, T. (2006). The 'chop-and-change' of MHC class I assembly. Nat Immunol *7*, 7-9.

Elliott, T., Cerundolo, V., Elvin, J., and Townsend, A. (1991). Peptide-induced conformational change of the class I heavy chain. Nature *351*, 402-406.

Elliott, T., and Williams, A. (2005). The optimization of peptide cargo bound to MHC class I molecules by the peptide-loading complex. Immunol Rev *207*, 89-99.

Farmery, M. R., Allen, S., Allen, A. J., and Bulleid, N. J. (2000). The role of ERp57 in disulfide bond formation during the assembly of major histocompatibility complex class I in a synchronized semipermeabilized cell translation system. J Biol Chem *275*, 14933-14938.

Ford, S., Antoniou, A., Butcher, G. W., and Powis, S. J. (2004). Competition for access to the rat major histocompatibility complex class I peptide-loading complex reveals optimization of peptide cargo in the absence of transporter associated with antigen processing (TAP) association. J Biol Chem *279*, 16077-16082.

Flajnik, M. F., and Kasahara, M. (2001). Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. Immunity *15*, 351-362.

Frangoulis, B., Park, I., Guillemot, F., Severac, V., Auffray, C., and Zoorob, R. (1999). Identification of the Tapasin gene in the chicken major histocompatibility complex. Immunogenetics *49*, 328-337.

Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A., and Wilson, I. A. (1992). Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. Science *257*, 919-927.

Fulton, J. E., Thacker, E. L., Bacon, L. D., and Hunt, H. D. (1995). Functional analysis of avian class I (BFIV) glycoproteins by epitope tagging and mutagenesis in vitro. Eur J Immunol *25*, 2069-2076.

Furukawa, H., Kashiwase, K., Yabe, T., Ishikawa, Y., Akaza, T., Tadokoro, K., Tohma, S., Inoue, T., Tokunaga, K., Yamamoto, K., and Juji, T. (1998). Polymorphism of TAPASIN and its linkage disequilibria with HLA class II genes in the Japanese population. Tissue Antigens *52*, 279-281.

Gao, B., Adhikari, R., Howarth, M., Nakamura, K., Gold, M. C., Hill, A. B., Knee, R., Michalak, M., and Elliott, T. (2002). Assembly and antigen-presenting function of MHC class I molecules in cells lacking the ER chaperone calreticulin. Immunity *16*, 99-109.

Garbi, N., Tan, P., Diehl, A. D., Chambers, B. J., Ljunggren, H. G., Momburg, F., and Hammerling, G. J. (2000). Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. Nat Immunol *1*, 234-238.

Garbi, N., Tanaka, S., Momburg, F., and Hammerling, G. J. (2006). Impaired assembly of the major histocompatibility complex class I peptide-loading complex in mice deficient in the oxidoreductase ERp57. Nat Immunol *7*, 93-102.

Garbi, N., Tiwari, N., Momburg, F., and Hammerling, G. J. (2003). A major role for tapasin as a stabilizer of the TAP peptide transporter and consequences for MHC class I expression. Eur J Immunol *33*, 264-273.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E., and Wiley, D. C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134-141.

Germain, R. N., Bentley, D. M., and Quill, H. (1985). Influence of allelic polymorphism on the assembly and surface expression of class II MHC (Ia) molecules. Cell *43*, 233-242.

Gileadi, U., Moins-Teisserenc, H. T., Correa, I., Booth, B. L., Jr., Dunbar, P. R., Sewell, A. K., Trowsdale, J., Phillips, R. E., and Cerundolo, V. (1999). Generation of an immunodominant CTL epitope is affected by proteasome subunit composition and stability of the antigenic protein. J Immunol *163*, 6045-6052.

Gobin, S. J., Keijsers, V., van Zutphen, M., and van den Elsen, P. J. (1998). The role of enhancer A in the locus-specific transactivation of classical and nonclassical HLA class I genes by nuclear factor kappa B. J Immunol *161*, 2276-2283.

Grandea, A. G., 3rd, Androlewicz, M. J., Athwal, R. S., Geraghty, D. E., and Spies, T. (1995). Dependence of peptide binding by MHC class I molecules on their interaction with TAP. Science *270*, 105-108.

Grandea, A. G., 3rd, Golovina, T. N., Hamilton, S. E., Sriram, V., Spies, T., Brutkiewicz, R. R., Harty, J. T., Eisenlohr, L. C., and Van Kaer, L. (2000). Impaired assembly yet normal trafficking of MHC class I molecules in Tapasin mutant mice. Immunity *13*, 213-222.

Grandea, A. G., 3rd, Lehner, P. J., Cresswell, P., and Spies, T. (1997). Regulation of MHC class I heterodimer stability and interaction with TAP by tapasin. Immunogenetics *46*, 477-483.

Greenwood, R., Shimizu, Y., Sekhon, G. S., and DeMars, R. (1994). Novel allele-specific, post-translational reduction in HLA class I surface expression in a mutant human B cell line. J Immunol *153*, 5525-5536.

Grzybowska, E. A., Wilczynska, A., and Siedlecki, J. A. (2001). Regulatory functions of 3'UTRs. Biochem Biophys Res Commun *288*, 291-295.

Guex, N., and Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis *18*, 2714-2723.

Guild, B. C., Erikson, R. L., and Strominger, J. L. (1983). HLA-A2 and HLA-B7 antigens are phosphorylated in vitro by rous sarcoma virus kinase (pp60v-src) at a tyrosine residue encoded in a highly conserved exon of the intracellular domain. Proc Natl Acad Sci U S A *80*, 2894-2898.

Guild, B. C., and Strominger, J. L. (1984). Human and murine class I MHC antigens share conserved serine 335, the site of HLA phosphorylation in vivo. J Biol Chem *259*, 9235-9240.

Guillemot, F., Billault, A., Pourquie, O., Behar, G., Chausse, A. M., Zoorob, R., Kreibich, G., and Auffray, C. (1988). A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer. Embo J *7*, 2775-2785.

Hammer, G. E., Gonzalez, F., Champsaur, M., Cado, D., and Shastri, N. (2006). The aminopeptidase ERAAP shapes the peptide repertoire displayed by major histocompatibility complex class I molecules. Nat Immunol *7*, 103-112.

Hansen, T. H., Lybarger, L., Yu, L., Mitaksov, V., and Fremont, D. H. (2005). Recognition of open conformers of classical MHC by chaperones and monoclonal antibodies. Immunol Rev *207*, 100-111.

Harris, M. R., Lybarger, L., Yu, Y. Y., Myers, N. B., and Hansen, T. H. (2001). Association of ERp57 with mouse MHC class I molecules is tapasin dependent and mimics that of calreticulin and not calnexin. J Immunol *166*, 6686-6692.

Harris, M. R., Yu, Y. Y., Kindle, C. S., Hansen, T. H., and Solheim, J. C. (1998). Calreticulin and calnexin interact with different protein and glycan determinants during the assembly of MHC class I. J Immunol *160*, 5404-5409.

Hedley, M. L., Ozato, K., Maryanski, J., Tucker, P. W., and Forman, J. (1989). Expression of mutant H-2d proteins encoded by class I genes which alternatively process the 5' end of their transcripts. J Immunol *143*, 1026-1031.

Herberg, J. A., Sgouros, J., Jones, T., Copeman, J., Humphray, S. J., Sheer, D., Cresswell, P., Beck, S., and Trowsdale, J. (1998). Genomic analysis of the Tapasin gene, located close to the TAP loci in the MHC. Eur J Immunol *28*, 459-467.

Hill, A. V., Allsopp, C. E., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., Bennett, S., Brewster, D., McMichael, A. J., and Greenwood, B. M. (1991). Common west African HLA antigens are associated with protection from severe malaria. Nature *352*, 595-600.

Hillier, L. W., Miller, W., Birney, E., Warren, W., Hardison, R. C., Ponting, C. P., Bork, P., Burt, D. W., Groenen, M. A., Delany, M. E., *et al.* (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature *432*, 695-716.

Howarth, M., Williams, A., Tolstrup, A. B., and Elliott, T. (2004). Tapasin enhances MHC class I peptide presentation according to peptide half-life. Proc Natl Acad Sci U S A *101*, 11737-11742.

Hughes, E. A., and Cresswell, P. (1998). The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex. Curr Biol *8*, 709-712.

Hunt, H. D., and Fulton, J. E. (1998). Analysis of polymorphisms in the major expressed class I locus (B-FIV) of the chicken. Immunogenetics *47*, 456-467.

Hutchison, C. A., 3rd, Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P., and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. J Biol Chem *253*, 6551-6560.

Jacob, J. P., Milne, S., Beck, S., and Kaufman, J. (2000). The major and a minor class II beta-chain (B-LB) gene flank the Tapasin gene in the B-F /B-L region of the chicken major histocompatibility complex. Immunogenetics *51*, 138-147.

Joly, E., Le Rolle, A. F., Gonzalez, A. L., Mehling, B., Stevens, J., Coadwell, W. J., Hunig, T., Howard, J. C., and Butcher, G. W. (1998). Co-evolution of rat TAP transporters and MHC class I RT1-A molecules. Curr Biol *8*, 169-172.

Juul-Madsen, H. R., Zoorob, R., Auffray, C., Skjodt, K., and Hedemand, J. E. (1997). New chicken Rfp-Y haplotypes on the basis of MHC class II RFLP and MLC analyses. Immunogenetics *45*, 345-352.

Karre, K., Ljunggren, H. G., Piontek, G., and Kiessling, R. (1986). Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. Nature *319*, 675-678.

Kasahara, M., Kandil, E., Salter-Cid, L., and Flajnik, M. F. (1996). Origin and evolution of the class I gene family: why are some of the mammalian class I genes encoded outside the major histocompatibility complex? Res Immunol *147*, 278-284; discussion 284-275.

Katsanis, N., Fitzgibbon, J., and Fisher, E. M. (1996). Paralogy mapping: identification of a region in the human MHC triplicated onto human chromosomes 1 and 9 allows the prediction and isolation of novel PBX and NOTCH loci. Genomics *35*, 101-108.

Kaufman, J. (1999). Co-evolving genes in MHC haplotypes: the "rule" for nonmammalian vertebrates? Immunogenetics *50*, 228-236.

Kaufman, J. (2000). The simple chicken major histocompatibility complex: life and death in the face of pathogens and vaccines. Philos Trans R Soc Lond B Biol Sci *355*, 1077-1084.

Kaufman, J., Andersen, R., Avila, D., Engberg, J., Lambris, J., Salomonsen, J., Welinder, K., and Skjodt, K. (1992). Different features of the MHC class I heterodimer have evolved at different rates. Chicken B-F and beta 2-microglobulin sequences reveal invariant surface residues. J Immunol *148*, 1532-1546.

Kaufman, J., Jacob, J., Shaw, I., Walker, B., Milne, S., Beck, S., and Salomonsen, J. (1999a). Gene organisation determines evolution of function in the chicken MHC. Immunol Rev *167*, 101-117.

Kaufman, J., Milne, S., Gobel, T. W., Walker, B. A., Jacob, J. P., Auffray, C., Zoorob, R., and Beck, S. (1999b). The chicken B locus is a minimal essential major histocompatibility complex. Nature *401*, 923-925.

Kaufman, J., and Salomonsen, J. (1992). B-G: we know what it is, but what does it do? Immunol Today *13*, 1-3.

Kaufman, J., and Salomonsen, J. (1997). The "minimal essential MHC" revisited: both peptide-binding and cell surface expression level of MHC molecules are polymorphisms selected by pathogens in chickens. Hereditas *127*, 67-73.

Kaufman, J., Skjodt, K., and Salomonsen, J. (1991). The B-G multigene family of the chicken major histocompatibility complex. Crit Rev Immunol *11*, 113-143.

Kaufman, J., Volk, H., and Wallny, H. J. (1995). A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. Immunol Rev *143*, 63-88.

Kimura, A., Israel, A., Le Bail, O., and Kourilsky, P. (1986). Detailed analysis of the mouse H-2Kb promoter: enhancer-like sequences and their role in the regulation of class I gene expression. Cell *44*, 261-272.

Koch, J., Guntrum, R., Heintke, S., Kyritsis, C., and Tampe, R. (2004). Functional dissection of the transmembrane domains of the transporter associated with antigen processing (TAP). J Biol Chem *279*, 10142-10147.

Knittler, M. R., Gulow, K., Seelig, A., and Howard, J. C. (1998). MHC class I molecules compete in the endoplasmic reticulum for access to transporter associated with antigen processing. J Immunol *161*, 5967-5977.

Kroemer, G., Zoorob, R., and Auffray, C. (1990). Structure and expression of a chicken MHC class I gene. Immunogenetics *31*, 405-409.

Kulig, K., Nandi, D., Bacik, I., Monaco, J. J., and Vukmanovic, S. (1998). Physical and functional association of the major histocompatibility complex class I heavy chain alpha3 domain with the transporter associated with antigen processing. J Exp Med *187*, 865-874.

Kunisawa, J., and Shastri, N. (2003). The group II chaperonin TRiC protects proteolytic intermediates from degradation in the MHC class I antigen processing pathway. Mol Cell *12*, 565-576.

Lamont, S. J., Bolin, C., and Cheville, N. (1987). Genetic resistance to fowl cholera is linked to the major histocompatibility complex. Immunogenetics *25*, 284-289.

Lehner, P. J., Surman, M. J., and Cresswell, P. (1998). Soluble tapasin restores MHC class I expression and function in the tapasin-negative cell line .220. Immunity *8*, 221-231.

Leonhardt, R. M., Keusekotten, K., Bekpen, C., and Knittler, M. R. (2005). Critical Role for the Tapasin-Docking Site of TAP2 in the Functional Integrity of the MHC Class I-Peptide-Loading Complex. J Immunol *175*, 5104-5114.

Lewis, J. W., and Elliott, T. (1998). Evidence for successive peptide binding and quality control stages during MHC class I assembly. Curr Biol *8*, 717-720.

Lewis, J. W., Neisig, A., Neefjes, J., and Elliott, T. (1996). Point mutations in the alpha 2 domain of HLA-A2.1 define a functionally relevant interaction with TAP. Curr Biol *6*, 873-883.

Lewis, J. W., Sewell, A., Price, D., and Elliott, T. (1998). HLA-A*0201 presents TAP-dependent peptide epitopes to cytotoxic T lymphocytes in the absence of tapasin. Eur J Immunol *28*, 3214-3220.

Li, S., Paulsson, K. M., Sjogren, H. O., and Wang, P. (1999). Peptide-bound major histocompatibility complex class I molecules associate with tapasin before dissociation from transporter associated with antigen processing. J Biol Chem *274*, 8649-8654.

Li, S., Sjogren, H. O., Hellman, U., Pettersson, R. F., and Wang, P. (1997). Cloning and functional characterization of a subunit of the transporter associated with antigen processing. Proc Natl Acad Sci U S A *94*, 8708-8713.

Lindquist, J. A., Jensen, O. N., Mann, M., and Hammerling, G. J. (1998). ER-60, a chaperone with thiol-dependent reductase activity involved in MHC class I assembly. Embo J *17*, 2186-2195.

Livingstone, A. M., Powis, S. J., Diamond, A. G., Butcher, G. W., and Howard, J. C. (1989). A trans-acting major histocompatibility complex-linked gene whose alleles determine gain and loss changes in the antigenic structure of a classical class I molecule. J Exp Med *170*, 777-795.

Livingstone, A. M., Powis, S. J., Gunther, E., Cramer, D. V., Howard, J. C., and Butcher, G. W. (1991). Cim: an MHC class II-linked allelism affecting the antigenicity of a classical class I molecule for T lymphocytes. Immunogenetics *34*, 157-163.

Lizee, G., Basha, G., Tiong, J., Julien, J. P., Tian, M., Biron, K. E., and Jefferies, W. A. (2003). Control of dendritic cell cross-presentation by the major histocompatibility complex class I cytoplasmic domain. Nat Immunol *4*, 1065-1073.

Ljunggren, H. G., Stam, N. J., Ohlen, C., Neefjes, J. J., Hoglund, P., Heemels, M. T., Bastin, J., Schumacher, T. N., Townsend, A., Karre, K., and et al. (1990). Empty MHC class I molecules come out in the cold. Nature *346*, 476-480.

Lybarger, L., Wang, X., Harris, M. R., Virgin, H. W. t., and Hansen, T. H. (2003). Virus subversion of the MHC class I peptide-loading complex. Immunity *18*, 121-130.

Lybarger, L., Yu, Y. Y., Chun, T., Wang, C. R., Grandea, A. G., 3rd, Van Kaer, L., and Hansen, T. H. (2001). Tapasin enhances peptide-induced expression of H2-M3 molecules, but is not required for the retention of open conformers. J Immunol *167*, 2097-2105.

Madden, D. R., Gorga, J. C., Strominger, J. L., and Wiley, D. C. (1991). The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. Nature *353*, 321-325.

Marguet, D., Spiliotis, E. T., Pentcheva, T., Lebowitz, M., Schneck, J., and Edidin, M. (1999). Lateral diffusion of GFP-tagged H2Ld molecules and of GFP-TAP1 reports on the assembly and retention of these molecules in the endoplasmic reticulum. Immunity *11*, 231-240.

Maruoka, T., Tanabe, H., Chiba, M., and Kasahara, M. (2005). Chicken CD1 genes are located in the MHC: CD1 and endothelial protein C receptor genes constitute a distinct subfamily of class-I-like genes that predates the emergence of mammals. Immunogenetics *57*, 590-600.

Mayer, W. E., and Klein, J. (2001). Is tapasin a modified Mhc class I molecule? Immunogenetics *53*, 719-723.

Mesa, C. M., Thulien, K. J., Moon, D. A., Veniamin, S. M., and Magor, K. E. (2004). The dominant MHC class I gene is adjacent to the polymorphic TAP2 gene in the duck, Anas platyrhynchos. Immunogenetics *56*, 192-203.

Miggiano, V. C., Birgen, I., and Pink, J. R. (1974). The mixed leukocyte reaction in chickens. Evidence for control by the major histocompatibility complex. Eur J Immunol *4*, 397-401.

Miller, M. M., Bacon, L. D., Hala, K., Hunt, H. D., Ewald, S. J., Kaufman, J., Zoorob, R., and Briles, W. E. (2004). 2004 Nomenclature for the chicken major histocompatibility (B and Y) complex. Immunogenetics *56*, 261-279.

Miller, M. M., Goto, R., Bernot, A., Zoorob, R., Auffray, C., Bumstead, N., and Briles, W. E. (1994). Two Mhc class I and two Mhc class II genes map to the chicken Rfp-Y system outside the B complex. Proc Natl Acad Sci U S A *91*, 4397-4401.

Miller, M. M., Goto, R. M., Taylor, R. L., Jr., Zoorob, R., Auffray, C., Briles, R. W., Briles, W. E., and Bloom, S. E. (1996). Assignment of Rfp-Y to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. Proc Natl Acad Sci U S A *93*, 3958-3962.

Miller, M. M., Wang, C., Parisini, E., Coletta, R. D., Goto, R. M., Lee, S. Y., Barral, D. C., Townes, M., Roura-Mir, C., Ford, H. L., *et al.* (2005). Characterization of two avian MHC-like genes reveals an ancient origin of the CD1 family. Proc Natl Acad Sci U S A *102*, 8674-8679.

Mo, X. Y., Cascio, P., Lemerise, K., Goldberg, A. L., and Rock, K. (1999). Distinct proteolytic processes generate the C and N termini of MHC class I-binding peptides. J Immunol *163*, 5851-5859.

Moller, L. B., Kaufman, J., Verland, S., Salomonsen, J., Avila, D., Lambris, J. D., and Skjodt, K. (1991). Variations in the cytoplasmic region account for the heterogeneity of the chicken MHC class I (B-F) molecules. Immunogenetics *34*, 110-120.

Momburg, F., Roelse, J., Hammerling, G. J., and Neefjes, J. J. (1994a). Peptide size selection by the major histocompatibility complex-encoded peptide transporter. J Exp Med *179*, 1613-1623.

Momburg, F., Roelse, J., Howard, J. C., Butcher, G. W., Hammerling, G. J., and Neefjes, J. J. (1994b). Selectivity of MHC-encoded peptide transporters from human, mouse and rat. Nature *367*, 648-651.

Monaco, J. J., and McDevitt, H. O. (1982). Identification of a fourth class of proteins linked to the murine major histocompatibility complex. Proc Natl Acad Sci U S A *79*, 3001-3005.

Moon, D. A., Veniamin, S. M., Parks-Dely, J. A., and Magor, K. E. (2005). The MHC of the duck (Anas platyrhynchos) contains five differentially expressed class I genes. J Immunol *175*, 6702-6712.

Morrice, N. A., and Powis, S. J. (1998). A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules. Curr Biol *8*, 713-716.

Muller, C. A., Engler-Blum, G., Gekeler, V., Steiert, I., Weiss, E., and Schmidt, H. (1989). Genetic and serological heterogeneity of the supertypic HLA-B locus specificities Bw4 and Bw6. Immunogenetics *30*, 200-207.

Myers, N. B., Harris, M. R., Connolly, J. M., Lybarger, L., Yu, Y. Y., and Hansen, T. H. (2000). Kb, Kd, and Ld molecules share common tapasin dependencies as determined using a novel epitope tag. J Immunol *165*, 5656-5663.

Neefjes, J., Gottfried, E., Roelse, J., Gromme, M., Obst, R., Hammerling, G. J., and Momburg, F. (1995). Analysis of the fine specificity of rat, mouse and human TAP peptide transporters. Eur J Immunol *25*, 1133-1136.

Neisig, A., Roelse, J., Sijts, A. J., Ossendorp, F., Feltkamp, M. C., Kast, W. M., Melief, C. J., and Neefjes, J. J. (1995). Major differences in transporter associated with antigen presentation (TAP)-dependent translocation of MHC class I-presentable peptides and the effect of flanking sequences. J Immunol *154*, 1273-1279.

Neisig, A., Wubbolts, R., Zang, X., Melief, C., and Neefjes, J. (1996). Allele-specific differences in the interaction of MHC class I molecules with transporters associated with antigen processing. J Immunol *156*, 3196-3206.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng *10*, 1-6.

Nijenhuis, M., and Hammerling, G. J. (1996). Multiple regions of the transporter associated with antigen processing (TAP) contribute to its peptide binding site. J Immunol *157*, 5467-5477.

Nijenhuis, M., Schmitt, S., Armandola, E. A., Obst, R., Brunner, J., and Hammerling, G. J. (1996). Identification of a contact region for peptide on the TAP1 chain of the transporter associated with antigen processing. J Immunol *156*, 2186-2195.

Nonaka, M., Namikawa, C., Kato, Y., Sasaki, M., Salter-Cid, L., and Flajnik, M. F. (1997). Major histocompatibility complex gene mapping in the amphibian Xenopus implies a primordial organization. Proc Natl Acad Sci U S A *94*, 5789-5791.

Obst, R., Armandola, E. A., Nijenhuis, M., Momburg, F., and Hammerling, G. J. (1995). TAP polymorphism does not influence transport of peptide variants in mice and humans. Eur J Immunol *25*, 2170-2176.

Ortmann, B., Androlewicz, M. J., and Cresswell, P. (1994). MHC class I/beta 2-microglobulin complexes associate with TAP transporters before peptide binding. Nature *368*, 864-867.

Ortmann, B., Copeman, J., Lehner, P. J., Sadasivan, B., Herberg, J. A., Grandea, A. G., Riddell, S. R., Tampe, R., Spies, T., Trowsdale, J., and Cresswell, P. (1997). A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. Science *277*, 1306-1309.

Paquet, M. E., Cohen-Doyle, M., Shore, G. C., and Williams, D. B. (2004). Bap29/31 influences the intracellular traffic of MHC class I molecules. J Immunol *172*, 7548-7555.

Parham, P., Barnstable, C. J., and Bodmer, W. F. (1979). Use of a monoclonal antibody (W6/32) in structural studies of HLA-A,B,C, antigens. J Immunol *123*, 342-349.

Parham, P., Lawlor, D. A., Lomen, C. E., and Ennis, P. D. (1989). Diversity and diversification of HLA-A,B,C alleles. J Immunol *142*, 3937-3950.

Park, B., Kim, Y., Shin, J., Lee, S., Cho, K., Fruh, K., Lee, S., and Ahn, K. (2004). Human cytomegalovirus inhibits tapasin-dependent peptide loading and optimization of the MHC class I peptide cargo for immune evasion. Immunity *20*, 71-85.

Park, B., Lee, S., Kim, E., and Ahn, K. (2003). A single polymorphic residue within the peptide-binding cleft of MHC class I molecules determines spectrum of tapasin dependence. J Immunol *170*, 961-968.

Paulsson, K. M., Kleijmeer, M. J., Griffith, J., Jevon, M., Chen, S., Anderson, P. O., Sjogren, H. O., Li, S., and Wang, P. (2002). Association of tapasin and COPI provides a mechanism for the retrograde transport of major histocompatibility complex (MHC) class I molecules from the Golgi complex to the endoplasmic reticulum. J Biol Chem *277*, 18266-18271.

Peace-Brewer, A. L., Tussey, L. G., Matsui, M., Li, G., Quinn, D. G., and Frelinger, J. A. (1996). A point mutation in HLA-A*0201 results in failure to bind the TAP complex and to present virus-derived peptides to CTL. Immunity *4*, 505-514.

Peaper, D. R., Wearsch, P. A., and Cresswell, P. (2005). Tapasin and ERp57 form a stable disulfide-linked dimer within the MHC class I peptide-loading complex. Embo J.

Peitsch, M. C. (1996). ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem Soc Trans *24*, 274-279.

Peh, C. A., Burrows, S. R., Barnden, M., Khanna, R., Cresswell, P., Moss, D. J., and McCluskey, J. (1998). HLA-B27-restricted antigen presentation in the absence of tapasin reveals polymorphism in mechanisms of HLA class I peptide loading. Immunity *8*, 531-542.

Peh, C. A., Laham, N., Burrows, S. R., Zhu, Y., and McCluskey, J. (2000). Distinct functions of tapasin revealed by polymorphism in MHC class I peptide loading. J Immunol *164*, 292-299.

Petersen, J. L., Hickman-Miller, H. D., McIlhaney, M. M., Vargas, S. E., Purcell, A. W., Hildebrand, W. H., and Solheim, J. C. (2005). A charged amino acid residue in the transmembrane/cytoplasmic region of tapasin influences MHC class I assembly and maturation. J Immunol *174*, 962-969.

Pharr, G. T., Dodgson, J. B., Hunt, H. D., and Bacon, L. D. (1998). Class II MHC cDNAs in 15I5 B-congenic chickens. Immunogenetics *47*, 350-354.

Pharr, G. T., Gwynn, A. V., and Bacon, L. D. (1996). Histocompatibility antigen(s) linked to Rfp-Y (Mhc-like) genes in the chicken. Immunogenetics *45*, 52-58.

Pharr, G. T., Vallejo, R. L., and Bacon, L. D. (1997). Identification of Rfp-Y (Mhc-like) haplotypes in chickens of Cornell lines N and P. J Hered *88*, 504-512.

Plachy, J., Pink, J. R., and Hala, K. (1992). Biology of the chicken MHC (B complex). Crit Rev Immunol *12*, 47-79.

Pontarotti, P. A., Mashimo, H., Zeff, R. A., Fisher, D. A., Hood, L., Mellor, A., Flavell, R. A., and Nathenson, S. G. (1986). Conservation and diversity in the class I genes of the major histocompatibility complex: sequence analysis of a Tlab gene and comparison with a Tlac gene. Proc Natl Acad Sci U S A *83*, 1782-1786.

Powis, S. J., Howard, J. C., and Butcher, G. W. (1991). The major histocompatibility complex class II-linked cim locus controls the kinetics of intracellular transport of a classical class I molecule. J Exp Med *173*, 913-921.

Princiotta, M. F., Finzi, D., Qian, S. B., Gibbs, J., Schuchmann, S., Buttgereit, F., Bennink, J. R., and Yewdell, J. W. (2003). Quantitating protein synthesis, degradation, and endogenous antigen processing. Immunity *18*, 343-354.

Procko, E., Raghuraman, G., Wiley, D. C., Raghavan, M., and Gaudet, R. (2005). Identification of domain boundaries within the N-termini of TAP1 and TAP2 and their importance in tapasin binding and tapasin-mediated increase in peptide loading of MHC class I. Immunol Cell Biol *83*, 475-482.

Purcell, A. W., Gorman, J. J., Garcia-Peydro, M., Paradela, A., Burrows, S. R., Talbo, G. H., Laham, N., Peh, C. A., Reynolds, E. C., Lopez De Castro, J. A., and McCluskey, J. (2001). Quantitative and qualitative influences of tapasin on the class I peptide repertoire. J Immunol *166*, 1016-1027.

Reits, E., Griekspoor, A., Neijssen, J., Groothuis, T., Jalink, K., van Veelen, P., Janssen, H., Calafat, J., Drijfhout, J. W., and Neefjes, J. (2003). Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. Immunity *18*, 97-108.

Reits, E., Neijssen, J., Herberts, C., Benckhuijsen, W., Janssen, L., Drijfhout, J. W., and Neefjes, J. (2004). A major role for TPPII in trimming proteasomal degradation products for MHC class I antigen presentation. Immunity *20*, 495-506.

Reits, E. A., Griekspoor, A. C., and Neefjes, J. (2000a). How does TAP pump peptides? insights from DNA repair and traffic ATPases. Immunol Today *21*, 598-600.

Reits, E. A., Vos, J. C., Gromme, M., and Neefjes, J. (2000b). The major substrates for TAP in vivo are derived from newly synthesized proteins. Nature *404*, 774-778.

Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994). Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell *78*, 761-771.

Rodgers, J. R., Mehta, V., and Cook, R. G. (1995). Surface expression of beta 2-microglobulin-associated thymus-leukemia antigen is independent of TAP2. Eur J Immunol *25*, 1001-1007.

Rogers, S., Shaw, I., Ross, N., Nair, V., Rothwell, L., Kaufman, J., and Kaiser, P. (2003). Analysis of part of the chicken Rfp-Y region reveals two novel lectin genes, the first complete genomic sequence of a class I alpha-chain gene, a truncated class II beta-chain gene, and a large CR1 repeat. Immunogenetics *55*, 100-108.

Rogers, S. L., Gobel, T. W., Viertlboeck, B. C., Milne, S., Beck, S., and Kaufman, J. (2005). Characterization of the chicken C-type lectin-like receptors B-NK and B-lec suggests that the NK complex and the MHC share a common ancestral region. J Immunol *174*, 3475-3483.

Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., and Cresswell, P. (1996). Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. Immunity *5*, 103-114.

Salomonsen, J., Eriksson, H., Skjodt, K., Lundgreen, T., Simonsen, M., and Kaufman, J. (1991). The "adjuvant effect" of the polymorphic B-G antigens of the chicken major histocompatibility complex analyzed using purified molecules incorporated in liposomes. Eur J Immunol *21*, 649-658.

Salomonsen, J., Marston, D., Avila, D., Bumstead, N., Johansson, B., Juul-Madsen, H., Olesen, G. D., Riegert, P., Skjodt, K., Vainio, O., *et al.* (2003). The properties of the single chicken MHC classical class II alpha chain (B-LA) gene indicate an ancient origin for the DR/E-like isotype of class II molecules. Immunogenetics *55*, 605-614.

Salomonsen, J., Sorensen, M. R., Marston, D. A., Rogers, S. L., Collen, T., van Hateren, A., Smith, A. L., Beal, R. K., Skjodt, K., and Kaufman, J. (2005). Two CD1 genes map to the chicken MHC, indicating that CD1 genes are ancient and likely to have been present in the primordial MHC. Proc Natl Acad Sci U S A *102*, 8668-8673.

Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., and Goldberg, A. L. (2002). An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. Nat Immunol *3*, 1169-1176.

Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., Lepelletier, Y., Greer, F., Schomburg, L., Fruci, D., Niedermann, G., and van Endert, P. M. (2005). Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. Nat Immunol *6*, 689-697.

Schierman, L. W., and Nordskog, A. W. (1961). Relationship of blood type to histocompatibility in chickens. Science *134*, 1008-1009.

Schierman, L. W., and Nordskog, A. W. (1963). Influence of the B bloodgroup-histocompatibility locus in chickens on a graft-versus-host reaction. Nature *197*, 511-512.

Schoenhals, G. J., Krishna, R. M., Grandea, A. G., 3rd, Spies, T., Peterson, P. A., Yang, Y., and Fruh, K. (1999). Retention of empty MHC class I molecules by tapasin is essential to reconstitute antigen presentation in invertebrate cells. Embo J *18*, 743-753.

Schubert, U., Anton, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. Nature *404*, 770-774.

Schumacher, T. N., Kantesaria, D. V., Heemels, M. T., Ashton-Rickardt, P. G., Shepherd, J. C., Fruh, K., Yang, Y., Peterson, P. A., Tonegawa, S., and Ploegh, H. L. (1994a). Peptide length and sequence specificity of the mouse TAP1/TAP2 translocator. J Exp Med *179*, 533-540.

Schumacher, T. N., Kantesaria, D. V., Serreze, D. V., Roopenian, D. C., and Ploegh, H. L. (1994b). Transporters from H-2b, H-2d, H-2s, H-2k, and H-2g7 (NOD/Lt) haplotype translocate similar sets of peptides. Proc Natl Acad Sci U S A *91*, 13004-13008.

Schwartz, D. E., Tizard, R., and Gilbert, W. (1983). Nucleotide sequence of Rous sarcoma virus. Cell *32*, 853-869.

Scott, J. E., and Dawson, J. R. (1995). MHC class I expression and transport in a calnexin-deficient cell line. J Immunol *155*, 143-148.

Sesma, L., Galocha, B., Vazquez, M., Purcell, A. W., Marcilla, M., McCluskey, J., and Lopez de Castro, J. A. (2005). Qualitative and quantitative differences in peptides bound to HLA-B27 in the presence of mouse versus human tapasin define a role for tapasin as a size-dependent peptide editor. J Immunol *174*, 7833-7844.

Serwold, T., Gonzalez, F., Kim, J., Jacob, R., and Shastri, N. (2002). ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. Nature *419*, 480-483.

Sharma, J. M., and Okazaki, W. (1981). Natural killer cell activity in chickens: target cell analysis and effect of antithymocyte serum on effector cells. Infect Immun *31*, 1078-1085.

Shields, M. J., Hodgson, W., and Ribaudo, R. K. (1999). Differential association of beta2-microglobulin mutants with MHC class I heavy chains and structural analysis demonstrate allele-specific interactions. Mol Immunol *36*, 561-573.

Shields, M. J., Moffat, L. E., and Ribaudo, R. K. (1998). Functional comparison of bovine, murine, and human beta2-microglobulin: interactions with murine MHC I molecules. Mol Immunol *35*, 919-928.

Shiina, T., Shimizu, S., Hosomichi, K., Kohara, S., Watanabe, S., Hanzawa, K., Beck, S., Kulski, J. K., and Inoko, H. (2004). Comparative genomic analysis of two avian (quail and chicken) MHC regions. J Immunol *172*, 6751-6763.

Sijts, A. J., and Pamer, E. G. (1997). Enhanced intracellular dissociation of major histocompatibility complex class I-associated peptides: a mechanism for optimizing the spectrum of cell surface-presented cytotoxic T lymphocyte epitopes. J Exp Med *185*, 1403-1411.

Simonsen, M., Crone, M., Koch, C., and Hala, K. (1982). The MHC haplotypes of the chicken. Immunogenetics *16*, 513-532.

Skjodt, K., Koch, C., Crone, M., and Simonsen, M. (1985). Analysis of chickens for recombination within the MHC (B-complex). Tissue Antigens *25*, 278-282.

Skjodt, K., Welinder, K. G., Crone, M., Verland, S., Salomonsen, J., and Simonsen, M. (1986). Isolation and characterization of chicken and turkey beta 2-microglobulin. Mol Immunol *23*, 1301-1309.

Smith, J. D., Lie, W. R., Gorka, J., Kindle, C. S., Myers, N. B., and Hansen, T. H. (1992). Disparate interaction of peptide ligand with nascent versus mature class I major histocompatibility complex molecules: comparisons of peptide binding to alternative forms of Ld in cell lysates and the cell surface. J Exp Med *175*, 191-202.

Snell, G. D., and Jackson, R. B. (1958). Histocompatibility genes of the mouse. II. Production and analysis of isogenic resistant lines. J Natl Cancer Inst *21*, 843-877.

Spiliotis, E. T., Manley, H., Osorio, M., Zuniga, M. C., and Edidin, M. (2000). Selective export of MHC class I molecules from the ER after their dissociation from TAP. Immunity *13*, 841-851.

Stam, N. J., Spits, H., and Ploegh, H. L. (1986). Monoclonal antibodies raised against denatured HLA-B locus heavy chains permit biochemical characterization of certain HLA-C locus products. J Immunol *137*, 2299-2306.

Strehl, B., Seifert, U., Kruger, E., Heink, S., Kuckelkorn, U., and Kloetzel, P. M. (2005). Interferon-gamma, the functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing. Immunol Rev *207*, 19-30.

Suh, W. K., Cohen-Doyle, M. F., Fruh, K., Wang, K., Peterson, P. A., and Williams, D. B. (1994). Interaction of MHC class I molecules with the transporter associated with antigen processing. Science *264*, 1322-1326.

Suh, W. K., Derby, M. A., Cohen-Doyle, M. F., Schoenhals, G. J., Fruh, K., Berzofsky, J. A., and Williams, D. B. (1999). Interaction of murine MHC class I molecules with tapasin and TAP enhances peptide loading and involves the heavy chain alpha3 domain. J Immunol *162*, 1530-1540.

Tahara, T., Yang, S. Y., Khan, R., Abish, S., Hammerling, G. J., and Hammerling, U. (1990). HLA antibody responses in HLA class I transgenic mice. Immunogenetics *32*, 351-360.

Tan, P., Kropshofer, H., Mandelboim, O., Bulbuc, N., Hammerling, G. J., and Momburg, F. (2002). Recruitment of MHC class I molecules by tapasin into the transporter associated with antigen processing-associated complex is essential for optimal peptide loading. J Immunol *168*, 1950-1960.

Teitell, M., Cheroutre, H., Panwala, C., Holcombe, H., Eghtesady, P., and Kronenberg, M. (1994). Structure and function of H-2 T (Tla) region class I MHC molecules. Crit Rev Immunol *14*, 1-27.

Thoraval, P., Afanassieff, M., Bouret, D., Luneau, G., Esnault, E., Goto, R. M., Chausse, A. M., Zoorob, R., Soubieux, D., Miller, M. M., and Dambrine, G. (2003). Role of nonclassical class I genes of the chicken major histocompatibility complex Rfp-Y locus in transplantation immunity. Immunogenetics *55*, 647-651.

Toes, R. E., Nussbaum, A. K., Degermann, S., Schirle, M., Emmerich, N. P., Kraft, M., Laplace, C., Zwinderman, A., Dick, T. P., Muller, J., *et al.* (2001). Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. J Exp Med *194*, 1-12.

Tolstrup, A. B., Duch, M., Dalum, I., Pedersen, F. S., and Mouritsen, S. (2001). Functional screening of a retroviral peptide library for MHC class I presentation. Gene *263*, 77-84.

Tourdot, S., and Gould, K. G. (2002). Competition between MHC class I alleles for cell surface expression alters CTL responses to influenza A virus. J Immunol *169*, 5615-5621.

Tourdot, S., Nejmeddine, M., Powis, S. J., and Gould, K. G. (2005). Different MHC class I heavy chains compete with each other for folding independently of beta 2-microglobulin and peptide. J Immunol *174*, 925-933.

Towne, C. F., York, I. A., Neijssen, J., Karow, M. L., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Neefjes, J. J., and Rock, K. L. (2005). Leucine aminopeptidase is not essential for trimming peptides in the cytosol or generating epitopes for MHC class I antigen presentation. J Immunol *175*, 6605-6614.

Tsujimura, K., Obata, Y., Iwase, S., Matsudaira, Y., Ozeki, S., and Takahashi, T. (2000). The epitope detected by cytotoxic T lymphocytes against thymus leukemia (TL) antigen is TAP independent. Int Immunol *12*, 1217-1225.

Turnquist, H. R., Petersen, J. L., Vargas, S. E., McIlhaney, M. M., Bedows, E., Mayer, W. E., Grandea, A. G., 3rd, Van Kaer, L., and Solheim, J. C. (2004). The Ig-like domain of tapasin influences intermolecular interactions. J Immunol *172*, 2976-2984.

Turnquist, H. R., Schenk, E. L., McIlhaney, M. M., Hickman, H. D., Hildebrand, W. H., and Solheim, J. C. (2002). Disparate binding of chaperone proteins by HLA-A subtypes. Immunogenetics *53*, 830-834.

Turnquist, H. R., Vargas, S. E., Reber, A. J., McIlhaney, M. M., Li, S., Wang, P., Sanderson, S. D., Gubler, B., van Endert, P., and Solheim, J. C. (2001). A region of tapasin that affects L(d) binding and assembly. J Immunol *167*, 4443-4449.

Vainio, O., and Ratcliffe, M. J. (1984). Proliferation of chicken peripheral blood leukocytes in response to pokeweed mitogen is macrophage dependent. Cell Immunol *85*, 235-243.

Vainio, O., Veromaa, T., Eerola, E., Toivanen, P., and Ratcliffe, M. J. (1988). Antigen-presenting cell-T cell interaction in the chicken is MHC class II antigen restricted. J Immunol *140*, 2864-2868.

van Endert, P. M., Riganelli, D., Greco, G., Fleischhauer, K., Sidney, J., Sette, A., and Bach, J. F. (1995). The peptide-binding motif for the human transporter associated with antigen processing. J Exp Med *182*, 1883-1895.

van Endert, P. M., Tampe, R., Meyer, T. H., Tisch, R., Bach, J. F., and McDevitt, H. O. (1994). A sequential model for peptide binding and transport by the transporters associated with antigen processing. Immunity *1*, 491-500.

van Tuinen, M., and Hedges, S. B. (2001). Calibration of avian molecular clocks. Mol Biol Evol *18*, 206-213.

Vega, M. A., and Strominger, J. L. (1989). Constitutive endocytosis of HLA class I antigens requires a specific portion of the intracytoplasmic tail that shares structural features with other endocytosed molecules. Proc Natl Acad Sci U S A *86*, 2688-2692.

Walker, B. A., van Hateren, A., Milne, S., Beck, S., and Kaufman, J. (2005). Chicken TAP genes differ from their human orthologues in locus organisation, size, sequence features and polymorphism. Immunogenetics *57*, 232-247.

Wallis, J. W., Aerts, J., Groenen, M. A., Crooijmans, R. P., Layman, D., Graves, T. A., Scheer, D. E., Kremitzki, C., Fedele, M. J., Mudd, N. K., *et al.* (2004). A physical map of the chicken genome. Nature *432*, 761-764.

Wallny, H. J., Avila, D., Hunt, L. G., Powell, T. J., Riegert, P., Salomonsen, J., Skjodt, K., Vainio, O., Vilbois, F., Wiles, M. V., and Kaufman, J. (2006). Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. Proc Natl Acad Sci U S A *103*, 1434-1439.

Wearsch, P. A., Jakob, C. A., Vallin, A., Dwek, R. A., Rudd, P. M., and Cresswell, P. (2004). Major histocompatibility complex class I molecules expressed with monoglucosylated N-linked glycans bind calreticulin independently of their assembly status. J Biol Chem *279*, 25112-25121.

Williams, A. P., Bevan, S., Bunce, M., Houlston, R., Welsh, K. I., and Elliott, T. (2000). Identification of novel Tapasin polymorphisms and linkage disequilibrium to MHC class I alleles. Immunogenetics *52*, 9-11.

Williams, A. P., Peh, C. A., Purcell, A. W., McCluskey, J., and Elliott, T. (2002). Optimization of the MHC class I peptide cargo is dependent on tapasin. Immunity *16*, 509-520.

Wright, C. A., Kozik, P., Zacharias, M., and Springer, S. (2004). Tapasin and other chaperones: models of the MHC class I loading complex. Biol Chem *385*, 763-778.

Yewdell, J. W., Anton, L. C., and Bennink, J. R. (1996). Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? J Immunol *157*, 1823-1826.

Yewdell, J. W., Reits, E., and Neefjes, J. (2003). Making sense of mass destruction: quantitating MHC class I antigen presentation. Nat Rev Immunol *3*, 952-961.

York, I. A., Chang, S. C., Saric, T., Keys, J. A., Favreau, J. M., Goldberg, A. L., and Rock, K. L. (2002). The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. Nat Immunol *3*, 1177-1184.

Yu, Y. Y., Turnquist, H. R., Myers, N. B., Balendiran, G. K., Hansen, T. H., and Solheim, J. C. (1999). An extensive region of an MHC class I alpha 2 domain loop influences interaction with the assembly complex. J Immunol *163*, 4427-4433.

Zacharias, M., and Springer, S. (2004). Conformational flexibility of the MHC class I alpha1-alpha2 domain in peptide bound and free states: a molecular dynamics simulation study. Biophys J *87*, 2203-2214.

Zarling, A. L., Luckey, C. J., Marto, J. A., White, F. M., Brame, C. J., Evans, A. M., Lehner, P. J., Cresswell, P., Shabanowitz, J., Hunt, D. F., and Engelhard, V. H. (2003). Tapasin is a facilitator, not an editor, of class I MHC Peptide binding. J Immunol *171*, 5287-5295.

Zernich, D., Purcell, A. W., Macdonald, W. A., Kjer-Nielsen, L., Ely, L. K., Laham, N., Crockford, T., Mifsud, N. A., Bharadwaj, M., Chang, L., *et al.* (2004). Natural HLA class I polymorphism controls the pathway of antigen presentation and susceptibility to viral evasion. J Exp Med *200*, 13-24.

Ziegler, A., and Pink, R. (1976). Chemical properties of two antigens controlled by the major histocompatibility complex of the chicken. J Biol Chem *251*, 5391-5396.

Zinkernagel, R. M., and Doherty, P. C. (1974). Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature *248*, 701-702.