

Structural, biophysical and evolutionary aspects of chicken MHC class I molecules

Inaugural – Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

Submitted to
the Department of Biology, Chemistry and Pharmacy of
Freie Universität Berlin

Hee Chee Seng
from Melaka, Malaysia

2011



Learning without thought is labour lost; thought without learning is perilous.

Quote: Confucius (551 - 479 B.C.)

Translation: James Legge (1867)

Calligraphy: 陈新颖

This thesis is based on research conducted from 2007 to 2011 at the Institut für Immungenetik, Charité - Universitätsmedizin Berlin, in Berlin, Germany, under the supervision of Prof. Dr. rer. nat. Andreas Ziegler and Dr. rer. nat. Barbara Uchanska-Ziegler.

Reviewers

First Reviewer: Prof. Dr. Andreas Ziegler
Institut für Immungenetik
Charité - Universitätsmedizin Berlin
Freie Universität Berlin

Second Reviewer: Prof. Dr. Markus Wahl
Institut für Chemie und Biochemie
Freie Universität Berlin

Date of Defence: 12 October 2011

Declaration

I hereby declare that the work presented in this thesis has been conducted independently and without any inappropriate support, and that all sources of content, experimental or intellectual, are suitably referenced and acknowledged.

I further declare that this thesis has not been submitted before, either in the same or a different form, to this or any other university for a degree.

Hee Chee Seng Berlin, 1 July 2011

Acknowledgement

My deepest gratitude goes to Professor Andreas Ziegler, who granted me the opportunity to work in his institute and guided me throughout the entire Ph.D. journey. So much I have learned from him in this period, not only in the field of science but also in many aspects of life. I most appreciate the writing skill I learned from him, which is timely and imperative. I am also thankful to Professor Marcus Wahl for chairing and organizing the thesis evaluation committee. I am very grateful to Dr. Barbara Uchanska-Ziegler for her guidance and often very creative and useful advice. The countless highly stimulating and mind-boggling sessions have indeed taught me some very important lessons in life. I'd also like to thank our collaborators from Max-Delbrück Centrum-Berlin, Professor Oliver Daumke for introducing me into the fascinating world of crystallography and his team especially Gao Song, for the help in research and friendship; from City of Hope, Duarte, USA, Professor Marcia Miller for many inspiring discussions; Bernhard Loll from Institut für Chemie und Biochemie, FU Berlin, for your professionalism and friendship; Heinz Fabian from Robert Koch-Institut, Berlin and Peter Schmieder from Leibniz-Institut für Molekulare Pharmakologie, Berlin for their spectroscopic contributions and discussions. Ich möchte mich auch herzlich bei meinen Institutskollegen (und Freunden!) Rolf Misselwitz, Hans Huser, Alexander Ziegler, Armin Volz, Christina Schnick, Angelika Zank und Carolin Backhaus für ihre Unterstützung bei meinen Experimenten und die sehr angenehme Atmosphäre im Labor bedanken. No word can express how gratified I am to my comrades-cum-buddies, Thomas Kellermann (also for the German translation of the Summary in this thesis), Pravin Kumar and Pablo Santos, it was really fun and stimulating working with you guys and definitely one of the most memorable moments (Mensa and Kaffee ToGo!) to me in Berlin. I'm also grateful to Berliner Krebsgesellschaft, Deutsche Forschungsgemeinschaft, the Senate of Berlin and Volkswagen Stiftung for the stipends that kept me alive in Berlin. I'm indebted to my friends in Berlin, Lijia, Lina, Lothar, Monica, Paolo, Pricilla, Sandra and members of Prenzelberg Piranhas, for the friendships and many wonderful get-together sessions. I'm particularly thankful to my friends and buddies in Malaysia for their "long-distance" support and friendships, (alphabetically ordered!) Asha, Chee Sim, Eric Chong, Dilbag, Jeff Mohd., Helena Chin, Honkie, Hooi Ling, Keat23, Kit, Kumar, Paula, Ron, Ruzana, Suresh, Thokie and Vincent. A special thank to Jen, for your love and support in the initial days. Last but not least, mom and uncle Soon, without your support and blessings, nothing would have been possible. Also in the family, Way and YenC, Sis and Ching-Yi, my god-parents and god-sisters, Owen and Jing Ling, I thank you all for your support and love.

Summary

One of the most interesting features that separates chicken from most mammalian immune systems investigated so far is that its major histocompatibility complex (MHC) maps to two genetically unlinked regions on the same chromosome. The core MHC has been termed *MHC-B*, while the other region is called *MHC-Y*. Each of the regions harbours two MHC class I loci of which only one is dominantly expressed, i.e. *BF2* and *YF1*, respectively. *BF2* has been studied extensively but much less is known about the *YF1* molecule. The major focus of this work was to understand the possible function of *YF1* molecule in the chicken immune system.

The first part of this dissertation describes the structural and biochemical analyses of the chicken *YF1*7.1* molecule. The high resolution *YF1*7.1* structures possess the typical architecture of a classical MHC class I molecule but contain a hydrophobic groove binding non-peptidic ligands. Only polyethyleneglycol (PEG) could be modelled into the electron density within the binding groove although self-lipids were added during complex formation. However, lipid-binding assay using native isoelectric focusing revealed that *YF1*7.1* is also able to bind bacterial lipids, a feature that is unprecedented among classical MHC class I molecules. This lipid-presenting molecule might aid the chicken immune system to recognize a diverse ligand repertoire with a minimal number of MHC class I molecules. Furthermore, comparison of *YF1*7.1* to other MHC class I structures revealed a structural feature in chickens that appears to be shared by non-mammalian but not by mammalian vertebrates.

The second part of this work focuses on β_2 -microglobulin (β_2m) and its evolutionary role in stabilizing a wide range of MHC class I and class I-like heavy chains (HC). The biophysical characterization and comparison of free chicken β_2m to monomeric β_2m of other vertebrates showed that chicken β_2m is structurally nearly identical but thermodynamically different to β_2m of other species so far investigated. This is likely due to fewer intra-molecular contacts within this molecule. A comparison of the free and HLA-B*27:05 HC-bound human β_2m in solution reveals a number of β_2m residues involved in interface interactions exhibiting considerable dynamics in the unbound form, but gaining stability upon binding to the HC. The analyses of molecular interactions in selected MHC class I and class I-like molecules associated with β_2m reveals a number of conserved β_2m -HC contacts established by residues that are retained from mammals to chicken. The presence of these evolutionarily conserved

interface interactions demonstrate that they are crucial for the stabilization of all MHC class I complexes. Additionally, these analyses identified $\beta_2\text{m}$ -HC interactions that are characteristic for a particular vertebrate taxon or a specific group of molecules.

The last part of this dissertation describes an optimized procedure and examples of embedding 3D models of macromolecular structures into electronic Portable Document Format (PDF) files for scientific publication. Interactive PDF-embedded 3D models are especially useful to present complex and diminutive structural details. This feature facilitates readers without prior knowledge of 3D manipulation programs to visualize images discussed within an article in 3D, that are otherwise difficult to be shown in a conventional 2D publication.

Zusammenfassung

Eines der interessantesten Merkmale, welches das Immunsystem von Hühnern von dem der meisten Säugetiere unterscheidet, ist die Tatsache, dass der Haupthistokompatibilitätskomplex (*major histocompatibility complex* oder MHC) in zwei genetisch unverbundenen Regionen desselben Chromosoms liegt. Der Kern-MHC-Bereich wird MHC-B, der andere MHC-Y genannt. Jeder der beiden Bereiche beinhaltet zwei MHC-Klasse-I-Loci, von denen jeweils nur einer dominant exprimiert wird (z.B. BF2 und YF1). BF2 wurde bereits gut untersucht, während über das YF1-Molekül fast nichts bekannt war. Der Hauptfokus dieser Arbeit bestand darin, die potentielle Rolle von YF1 im Immunsystem von Hühnern zu verstehen.

Der erste Teil dieser Dissertation beschreibt die strukturelle und biochemische Analyse des Hühner-YF1*7.1-Moleküls. Die hochaufgelöste Röntgenkristallstruktur von YF1*7.1 zeigt die typische Architektur eines klassischen MHC-Klasse-I-Komplexes. Das Molekül weist aber eine hydrophobe Bindungsgrube auf, welche nicht-peptidische Liganden bindet. Nur Polyethylenglykol (PEG) konnte in die Elektronendichte in der Bindungsgrube modelliert werden, obwohl Eigen-Lipide während der Komplexbildung hinzugegeben wurden. Mittels eines Lipid-Bindungsassays, der auf nativer isoelektrischer Fokussierung basiert, konnte jedoch gezeigt werden, dass auch YF1*7.1 in der Lage ist, bakterielle Lipide zu binden. Diese Eigenschaft wurde noch nie für klassische MHC-Klasse-I-Moleküle beschrieben. Dieses Lipid-bindende Molekül könnte dem Hühner-Immunsystem helfen, eine maximale Diversität von Liganden mit einer minimalen Anzahl an MHC-Klasse-I-Molekülen präsentieren zu können. Darüberhinaus führte der Vergleich von YF1*7.1 mit anderen MHC-Klasse-I-Strukturen zur Identifizierung eines Strukturmerkmals, welches alle klassischen MHC-Klasse-I Moleküle von Nicht-Säugetieren gemeinsam zu haben scheinen.

Der zweite Teil dieser Arbeit konzentriert sich auf β_2 -Mikroglobulin (β_2m) und seine evolutionäre Rolle bei der Stabilisierung einer großen Vielfalt von klassischen und nicht-klassischen MHC-Klasse-I-Molekülen. Die biophysikalische Charakterisierung und der Vergleich von ungebundenem Hühner- β_2m mit anderen Vertebraten- β_2m -Molekülen zeigte eine strukturell fast komplette Übereinstimmung, ergab aber thermodynamische Unterschiede. Dies könnte an der geringeren Anzahl intramolekularer Kontakte im Hühner- β_2m liegen. Der Vergleich des freien und an die schwere Kette (heavy chain, HC) gebundenen menschlichen

β_2m in Lösung zeigt, daß einige β_2m -Aminosäurereste, welche an *Interface*-Wechselwirkungen beteiligt sind, in der freien Form eine deutliche Dynamik aufweisen, jedoch nach HC-Bindung stabilisiert werden. Die Untersuchung der molekularen Interaktionen bei ausgewählten klassischen und nicht-klassischen MHC-Klasse-I Molekülen, die β_2m binden, enthüllte mehrere konservierte Kontakte zwischen β_2m und HC, die durch Aminosäuren gebildet werden, welche von Säugern bis zu Hühnern erhalten geblieben sind. Die Präsenz solcher evolutionär konservierten *Interface*-Wechselwirkungen zeigt, dass sie sehr wichtig für die Stabilisierung aller MHC-Klasse-I-Komplexe sein müssen. Zusätzlich konnten diese Untersuchungen β_2m -HC-Wechselwirkungen identifizieren, die spezifisch für eine bestimmte Tierklasse, Spezies oder Molekülgruppe sind, eine Eigenschaft, welche sich auch mit den evolutionären Entwicklungen oder den biologischen Funktionen der jeweiligen Moleküle in Übereinstimmung befindet.

Der letzte Teil dieser Dissertation beschreibt ein optimiertes Vorgehen, um interaktiv manipulierbare dreidimensionale (3D) Modelle von makromolekularen Strukturen in elektronische *Portable Document Format* (PDF) Dateien für wissenschaftliche Publikationen einzubetten. Derartige Modelle sind besonders nützlich, um komplexe strukturelle Details zu veranschaulichen, insbesondere auch vergleichend. Durch diese Herangehensweise können Leser, welche 3D-Bearbeitungsprogramme für Kristallographen nicht beherrschen, die in einem Artikel diskutierten Bilder direkt in 3D ansehen, was in einer konventionellen 2D-Publikation nicht möglich ist.

List of publications

Hee, C.S.*, Gao, S.*, Miller, M.M., Goto, R.M., Ziegler, A., Daumke, O., Uchanska-Ziegler, B. (2009) Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 422-425.

Kumar, P., Ziegler, A., Grahn, A., **Hee, C.S.**, Ziegler, A. (2010) Leaving the structural ivory tower, assisted by interactive 3D PDF. *Trends Biochem. Sci.* 35, 419-422.

Loll, B., Rückert, C., **Hee, C.S.**, Saenger, W., Uchanska-Ziegler, B., Ziegler, A. (2010) Loss of recognition by cross-reactive T cells and its relation to a C-terminus-induced conformational reorientation of an HLA-B*2705-bound peptide. *Prot. Sci.* 20, 278-290.

Hee, C.S.*, Gao, S.*, Loll, B., Miller, M.M., Uchanska-Ziegler, B., Daumke, O., Ziegler, A. (2010) Structure of a classical MHC class I molecule that binds "non-classical" ligands. *PLoS Biol.* 8, e1000557.

Uchanska-Ziegler, B., Loll, B., Fabian, H., **Hee, C.S.**, Saenger, W., Ziegler, A. (2011) HLA class I-associated diseases with a suspected autoimmune etiology: HLA-B27 subtypes as a model system. *Eur. J. Cell Biol.* (In press)

Hee, C.S., Fabian, H., Uchanska-Ziegler, B., Ziegler, A., Loll, B. Comparative biophysical characterization of chicken β_2 -microglobulin. (Submitted)

Hee, C.S., Beerbaum, M., Diehl, A., Loll, B., Schmieder, P., Uchanska-Ziegler, B., Ziegler, A. (2011) Dynamics of free versus complexed β_2 -microglobulin and the evolution of interfaces in MHC class I molecules. (Submitted)

This is a cumulative dissertation primarily based on the above publications.

*Equal contribution

Abbreviation

2D	Two-dimensional
3D	Three-dimensional
Å	Angstrom (0.1 nm)
β ₂ m	β ₂ -microglobulin
BF	Class I gene in the chicken or avian <i>MHC-B</i> region
BL	Class II gene in the chicken or avian <i>MHC-B</i> region
bp	Base pair
C1	Constant-1 domain
CD	Circular dichroism spectroscopy
CD1, CD4, CD8, CD94	Cluster of differentiation 1, 4, 8 or 94 (cell surface antigens)
CMV	Cytomegalovirus
CTL	Cytotoxic T lymphocyte
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
EPCR	Endothelial protein C receptor
FcRn	neonatal Fc receptor
FISH	Fluorescence <i>in situ</i> hybridization
FPLC	Fast Performance Liquid Chromatography
HC	Heavy chain
HFE	Hereditary hemochromatosis
HLA	Human leukocyte antigen
HSP	Heat shock protein
IEF	Isoelectric focusing
IgG	Immunoglobulin G
IgSF	Immunoglobulin superfamily
kb	Kilobase (one thousand base pairs)
KIR	Killer cell immunoglobulin-like receptor
KLRK1	Killer cell lectin-like receptor subfamily K, member 1
LRC	Leukocyte receptor complex
MAIT	Mucosal-associated invariant T cells
MDV	Marek's disease virus

MIC	MHC class I chain-related gene
MILL	MHC class I-like located near the leukocyte receptor complex
MHC	Major histocompatibility complex
MR1	MHC class I-related gene 1
MYA	Million years ago
NK	Natural killer
NKG2D	Also known as KLRK1 (a cell surface receptor)
NOR	Nucleolus organizer region
PEG	Polyethyleneglycol
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDF	Portable Document Format
PIP	Prolactin-inducible protein
RAE	Retinoic acid early transcript
RFLP	Restriction fragment length polymorphism
RSV	Rous sarcoma virus
Rfp-Y	Restriction fragment pattern Y
TAP	Transporter associated with antigen processing
TCR	T cell receptor
ULBP	UL-16 binding protein
YF	Class I gene in the chicken or avian <i>MHC-Y</i> region
YL	Class II gene in the chicken or avian <i>MHC-Y</i> region
ZAG	Zinc- α 2-glycoprotein

Table of Contents

Reviewers	2
Declaration	3
Acknowledgement	4
Summary	5
Zusammenfassung (German)	7
List of Publications	9
Abbreviations	10
Table of Contents	12
<hr/>	
1. Introduction	14
1.1 The Major Histocompatibility Complex	14
1.2 The Chicken MHC	15
1.3 MHC Class I Antigens	18
1.3.1 Classical MHC Class I Antigens	18
1.3.2 Non-classical MHC Class I Antigens	19
1.4 Architecture of MHC Class I Molecules	24
1.5 Chicken MHC Class I Antigens	27
1.5.1 Chicken Class I Antigens of the <i>MHC-B</i> region	27
1.5.2 Chicken Class I Antigens of the <i>MHC-Y</i> region	28
1.6 β_2 -microglobulin	29
1.7 Scope and Objectives of this Study	30
<hr/>	
2. The Structure of the Chicken YF1*7.1 Molecule and its Ligands	32
2.1 Summary	32
2.2 Publications	32
2.2.1 Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1	33
2.2.2 Structure of a classical MHC class I molecule that binds "non-classical" ligands	38
<hr/>	
3. Comparative Biophysical Characterizations of Human and Chicken β_2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules	56
3.1 Summary	56
3.2 Manuscripts	56
3.2.1 Comparative biophysical characterization of chicken β_2 -microglobulin	57
3.2.2 Dynamics of free versus complexed β_2 -microglobulin and the evolution of interfaces in MHC class I molecules	86
<hr/>	
4. PDF-embedded Interactive 3D Images	117
4.1 Summary	117
4.2 Publications	117

4.2.1 Leaving the structural ivory tower, assisted by interactive 3D PDF	118
4.2.2 Loss of recognition by cross-reactive T cells and its relation to a C-terminus-induced conformational reorientation of an HLA-B*2705-bound peptide	136
4.2.3 HLA class I-associated diseases with a suspected autoimmune etiology: HLA-B27 subtypes as a model system	149
5. Discussion	162
5.1 <i>MHC-Y</i> and Class I <i>YF</i> of Other Avian Species	162
5.2 Evolution of MHC Class I and β_2m Genes	168
5.3 Concluding remarks	173
6. References	174

Figures S2 (section 2.2.2, p. 51), 6 (section 3.2.1, p. 85), 1 and 2 (section 4.2.1, p. 119-120), Supplement S3 (section 4.2.1, p. 135), 3 (section 4.2.2, p. 142) and 5 (section 4.2.3, p. 157) contain embedded 3D imagery which can be viewed in the PDF version of this dissertation available on the attached CD using the freely available Adobe Reader X version 10.1.0.

1. Introduction

1.1 The Major Histocompatibility Complex

The major histocompatibility complex (MHC) is a cluster of genes found in all jawed vertebrates (Fig. 1.1) (Flajnik and Kasahara, 2001; Kelley et al., 2005). Historically, the MHC derives its name from genetic studies of transplant incompatibility, initially called “major transplantation antigens” (Klein, 1986), that first revealed an MHC in mice, the H-2 system (Gorer, 1938; Snell and Higgins, 1951), and later the human leukocyte antigens (HLA) (Dausset, 1958). The MHC is called a “complex” because its genes in humans and rodents are found closely clustered within a single chromosomal segment. The gene-dense MHC region harbors a collection of immune related genes that involved in innate and adaptive immune responses (Beck et al., 1999; Horton et al., 2004). MHC molecules control the immune response through self/non-self recognition, participate in mate selection and reproduction (Ziegler, 1997; Ziegler et al., 2010), as well as neuronal communication (Huh et al., 2000; Ribic et al., 2010). The MHC has also been associated with autoimmune conditions (Singer et al., 1997; Fernando et al., 2008) and resistance to infection (Carrington and Bontrop, 2002; Zekarias et al., 2002).

An MHC is found in all jawed vertebrates studied to date, from mammals to cartilaginous fishes (Flajnik and Kasahara, 2001). In the mammalian model, the MHC is divided into regions with similar function, including class I, class II, class III (Klein, 1976), extended class I and extended class II regions (Fig. 1.1) (Totaro et al., 1996; Herberg et al., 1998; Stephens et al., 1999; Horton et al., 2004). The number and the organization of genes in each region vary between species. The class I region contains classical class I gene(s) with their products presenting peptides. These peptides are mostly derived from endogenous proteins and they are presented to the receptors in CD8⁺ T lymphocytes (cytotoxic T-cells, CTL) through T-cell receptor (TCR) (Rudolph et al., 2006) or natural killer (NK) cells via killer cell immunoglobulin-like receptor (KIR) (Boyington et al., 2001; Vilches and Parham, 2002). The class II molecules encoded in the class II region can present peptide antigens that are primarily derived from exogenous proteins to CD4⁺ T cells (Klein, 1986; Murphy et al., 2008). The class III region contains a diverse group of immune and non-immune related genes. Some of the class III molecules participate in cell lysis and mediate inflammatory responses. Examples include the complement proteins, heat shock proteins and inflammatory cytokines (Beck and Trowsdale, 2000).

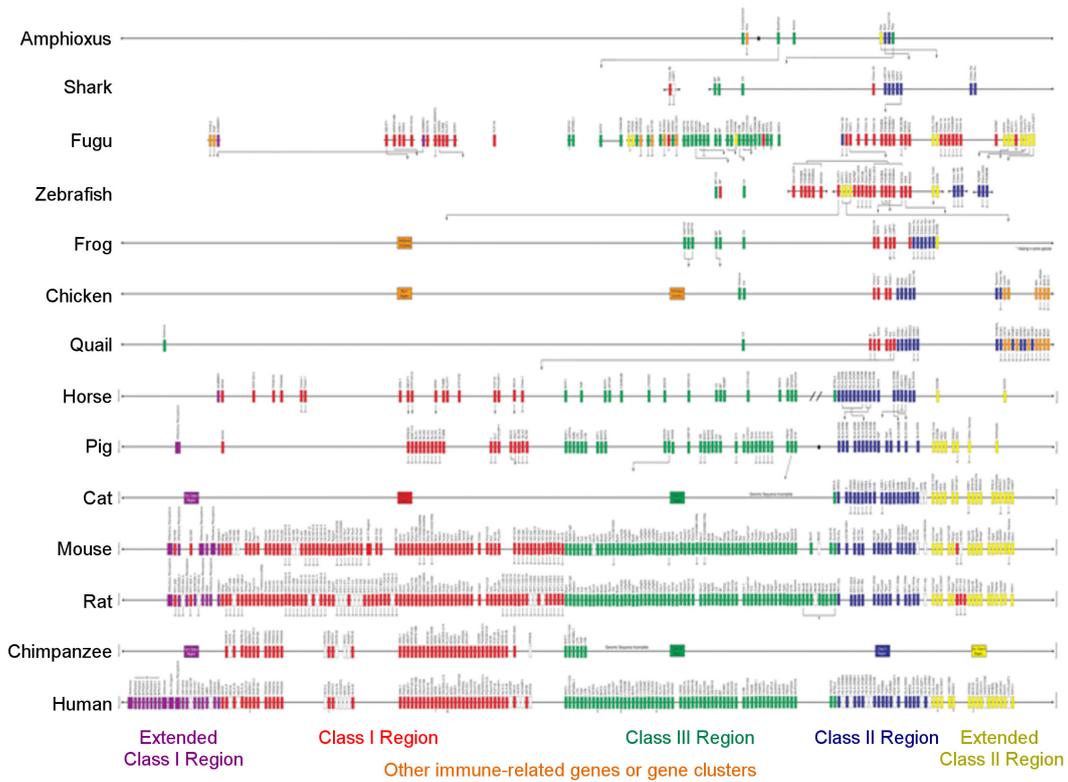


Figure 1.1. Comparative genomics of major histocompatibility complexes. The genes are coloured according to the regions they belonged to as indicated at the bottom of the figure. Larger boxes represent clusters containing multiple genes. (Adapted and modified from Kelley et al., 2005)

1.2 The Chicken MHC

As shown in Fig. 1.1, MHC regions vary in size and gene organization between mammalian and non-mammalian vertebrates. The “gold standard” for comparative genomic analyses is the human MHC genomic sequence of about 3.6 Mb in length with approximately 224 coding and non-coding sequences (Horton et al., 2004). The first completed MHC genomic sequence of a non-mammalian vertebrate was the chicken MHC (historically termed *B* locus and now *MHC-B*) which has been described as “minimal essential” because it harbours only 19 genes with a total length of about 92 kb (Kaufman et al., 1999b). However, a 242 kb region (expanded beyond the initial 92 kb) of *MHC-B* was sequenced revealing the presence of additional 27 genes within the 242 kb region (Fig. 1.2) (Shiina et al., 2007). Today, the size of the *MHC-B* region has been estimated to be about 0.5 Mb (Delany et al., 2009).

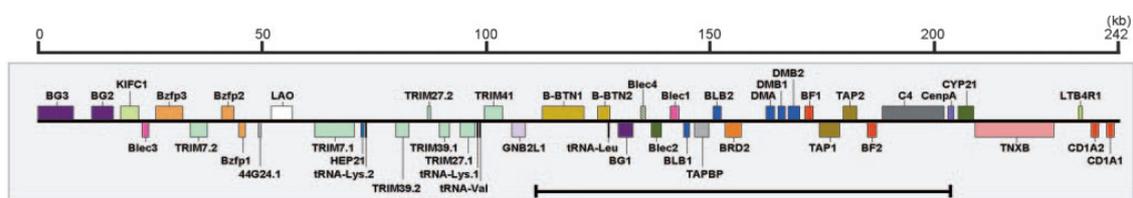


Figure 1.2. The 242 kb chicken *MHC-B* genomic organization. Genes belonging to the same family are indicated by the same coloured boxes. Upper boxes indicate gene orientation from left to right and opposite direction for lower boxes. Bar with blocked ends denotes the 92 kb region previously sequenced (adapted and modified from Shiina et al., 2007).

Besides the small size of the MHC, the gene arrangement in the chicken *MHC-B* region is also different from that of mammalian MHC as first predicted based on genetic and biochemical evidence (Ziegler and Pink, 1976). In chicken, the class III region is located outside the class I region as opposed to the situation in mammals where the class III loci are located between the class I and II regions (Fig. 1.1) (Kaufman et al., 1999b). Several genes are located differently in chicken and human MHC. For example, the tapasin gene (*TAPBP*) maps within the chicken class II but not the extended class II region (Fig. 1.2) (Frangoulis et al., 1999). Furthermore, the *TAP* genes (*TAP1* and *TAP2*) in the chicken are found in the class I as opposed to the class II region in humans, in agreement with the functional co-evolution of both types of loci (Kaufman et al., 1999a). The *MHC-B* region is also different from typical mammalian MHC by the absence of proteasome genes (*PSMB*). However, it contains putative NK cell receptor-related genes such as lectin-related sequences (*Blec1* and *Blec2*) (Kaufman et al., 1999b). On the other hand, many genes present in the MHC of mammals are absent in chicken, including the class II alpha gene which is located 5.6 cM away from the *MHC-B*, but resides on the same chromosome (Juul-Madsen et al., 2000; Salomonsen et al., 2003). Some immune-related genes are located at the vicinity of the *MHC-B* region including the two *CDI* genes (*CDIA1* and *CDIA2* in Fig. 1.2, now known as *CDI-1* or *CDI-2*, respectively) (Maruoka et al., 2005; Salomonsen et al., 2005; Shiina et al., 2007) and a group of highly polymorphic immunoglobulin superfamily (IgSF) genes encoding the BG antigens. These are well-expressed on erythrocytes and other tissues (Pink et al., 1977; Miller et al., 1990; Miller et al., 1991; Kaufman and Salomonsen, 1993). The BG gene (*BG1*) was originally identified in the *MHC-B* (Pink et al., 1977; Kaufman et al., 1999b) and two additional BG genes (*BG2* and *BG3*) were discovered recently in the extended *MHC-B* upstream of a cluster of *TRIM* genes (Fig. 1.2) (Shiina et al., 2007). The function of BG antigens has not yet been fully elucidated, but the polymorphic nature of BG genes is of particular interest because recombinant chicken lines showed differential resistance to

infection by Marek's Disease Virus (MDV) and Rous Sarcoma Virus (RSV) (Goto et al., 2009). *BG* genes have also been found in other avian species such as turkey (Chaves et al., 2009; Bauer and Reed, 2011), quail (Shiina et al., 2004) and ring-neck pheasant (Jarvi et al., 1996).

One of the major differences between the human and the chicken MHC is that part of the latter has also been found on a separate region of the same chromosome. This region was initially coined restriction fragment pattern Y (*Rfp-Y*) (Briles et al., 1993) and is now called *MHC-Y*. The *MHC-B* and *MHC-Y* regions are genetically unlinked (Briles et al., 1993; Miller et al., 1994) and it was previously reported that the lack of linkage was due to their location on opposite sides of the nucleolus organizer region (*NOR*) coupled with enhanced recombination involving the rRNA gene repeats of the *NOR* (Miller et al., 1996). Recently, however, Delany and co-workers (2009) established the chromosomal location of these regions using multicolor fluorescence *in situ* hybridization (FISH). This study revealed that the *MHC-B* and *MHC-Y* regions are located on the same side of the *NOR*, rather than opposite ends as previously proposed (Fig. 1.3). These workers suggested also that a large GC-rich region separates *MHC-B* and *MHC-Y*, forming the basis for the lack of genetic linkage between the two regions. The *MHC-Y* genomic sequence has not been fully assembled, but it has been predicted to encompass about 0.4 Mb (Delany et al., 2009).

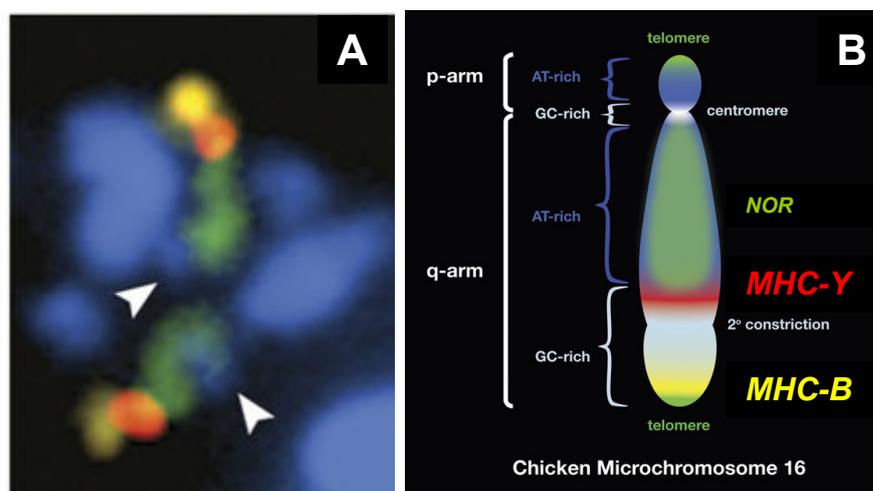


Figure 1.3. Architectural features and gene complex order of chicken microchromosome 16. (A) Probes (*NOR*, green; *MHC-Y*, red; *MHC-B*, yellow) were hybridized to mitotic prometaphase. Chromosomes were stained by DAPI (blue). The white arrows indicate the p-arms of the chromosomes. The result permits an unambiguous assignment of the q-arm order: centromere – *NOR* – *MHC-Y* – *MHC-B* – telomere. (B) Cytogenomic model of chicken microchromosome 16 based on the FISH results in (A). (Figures adapted and modified from Delany et al., 2009).

1.3 MHC Class I Antigens

MHC class I antigens can be divided into two broad groups, the classical class I (also known as class Ia) and the non-classical class I (or class Ib) molecules. Several criteria are used to categorize an MHC class I molecule, including the polymorphism of the binding groove, its expression level and tissue distribution, its ligand and whether it influences the fate of graft in transplantation (Klein, 1986). Classical class I molecules are ubiquitously expressed and possess a polymorphic binding groove that specifically presents peptides. On the other hand, non-classical class I molecules are less polymorphic, exhibit weak or tissue-restricted expression and have minimal influence on graft rejection. However, this dichotomous categorization of MHC class I molecules as classical and non-classical is conceptually misleading despite its historical basis and practical usefulness (Shawar et al., 1994). The classical class I loci represent only the tip of a genetic iceberg of class I genes, as many additional loci were not detected using classical techniques because of lacking polymorphism and low expression levels in most tissues. As a result, serological and cellular responses to these class I gene products are often difficult to identify, leading to their categorization as pseudogenes (Shawar et al., 1994). The definition of class I molecules gets more complicated when “non-classical” class I molecules are found to exhibit also “classical” features including polymorphisms (e.g. MICA, mouse H-2Qa1 and chicken YF1), show ubiquitous expression (e.g. mouse H-2Qa2 and chicken YF1), and present peptides just like classical class I molecules (H-2Qa2, HLA-E and HLA-G). The chicken YF1*7.1 molecule might be the most peculiar of all class I molecules so far discovered, as its recently determined crystal structure revealed, for the first time, a “classical” class I characteristics although it binds “non-classical” ligands such as lipids (Hee et al., 2010). Therefore, it is imperative to consider the variation of class I molecules as continuous rather than dichotomous.

1.3.1 Classical MHC Class I Antigens

Human and mice possess three classical MHC class I genes each, namely *HLA-A*, *HLA-B* and *HLA-C* in human and *H-2K*, *H-2D* and *H-2L* in mice that are highly polymorphic and widely expressed in most tissues. They present antigenic peptides to $\alpha\beta$ TCR expressed by T cells (Davis and Bjorkman, 1988). Human class I genes evolved through duplication in modules followed by diversification, co-evolution, and sequence exchange (Shiina et al., 1999c), which is in line with findings in primates but not in rodents. This indicates that the class I genes of human and mice are not orthologous but display higher homology to the non-classical genes of their own species.

In contrast, MHC class I genes in lower vertebrates are poorly understood and the number of gene loci varies between species. For example, in avian species, both chicken and turkey possess two class I loci in each of their *MHC-B* and *MHC-Y* regions, respectively (Kaufman et al., 1999b; Chaves et al., 2007), but duck and quail possess five and seven class I genes, respectively (Shiina et al., 2004; Moon et al., 2005). To date, only one class I gene from the *MHC-B* (*BF2*) and the *MHC-Y* (*YF1*) regions can be assigned with confidence as classical class I antigens in chicken, based on polymorphism, expression levels and 3-dimensional structure (see Chapter 1.5 for details). In contrast to most other species, there is only one class I gene in an amphibian, the frog *Xenopus* (Shum et al., 1993), which is in linkage disequilibrium with *TAP1*, *TAP2* and *LMP7* (Flajnik et al., 1999). However, several class I genes have been defined as non-classical in frog, including a group on the same chromosome as the MHC (Flajnik et al., 1993), apparently similar to the class I genes in the *MHC-Y* region of chicken. Axolotl, another amphibian species, possesses multiple class I sequences, suggesting duplication and rapid evolution in their vertebrate taxon (Sammut et al., 1999). The MHC class I genes in teleost (bony fish) species, like the birds, are highly diverse. Some species have limited diversity and less duplication, e.g. rainbow trout (Hansen and Kaattari, 1996), salmon (Miller and Withler, 1997) and brown trout (Shum et al., 2001), while others have a higher number of class I genes, e.g. swordtail fish (Figueroa et al., 2001), cichlid (Murray et al., 2000) and cod (Persson et al., 1999). Cartilaginous fish are the earliest jawed vertebrates with an ancestral MHC organization (Flajnik et al., 1999). To date, four class I genes have been isolated from the nurse shark, one classical and three non-classical genes as assigned by the authors (Bartl et al., 1997).

1.3.2 Non-classical MHC Class I Antigens

An MHC region may also encode several structurally related non-classical MHC class I molecules that are less polymorphic, exhibit a restricted tissue distribution and usually lack an influence on graft transplantation. Non-classical class I genes have been meticulously separated into “young”, “middle-aged” and “old” groups based on their evolutionary relationships (Fig. 1.4 and Table 1.1) (Rodgers and Cook, 2005). Young non-classical class I genes diverged 5-20 million years ago (MYA) from classical class I genes of their respective lineages, while middle-aged genes arose more than 65 MYA, early in mammalian radiation and old class I genes evolved early in vertebrate evolution and are often found outside of the MHC.

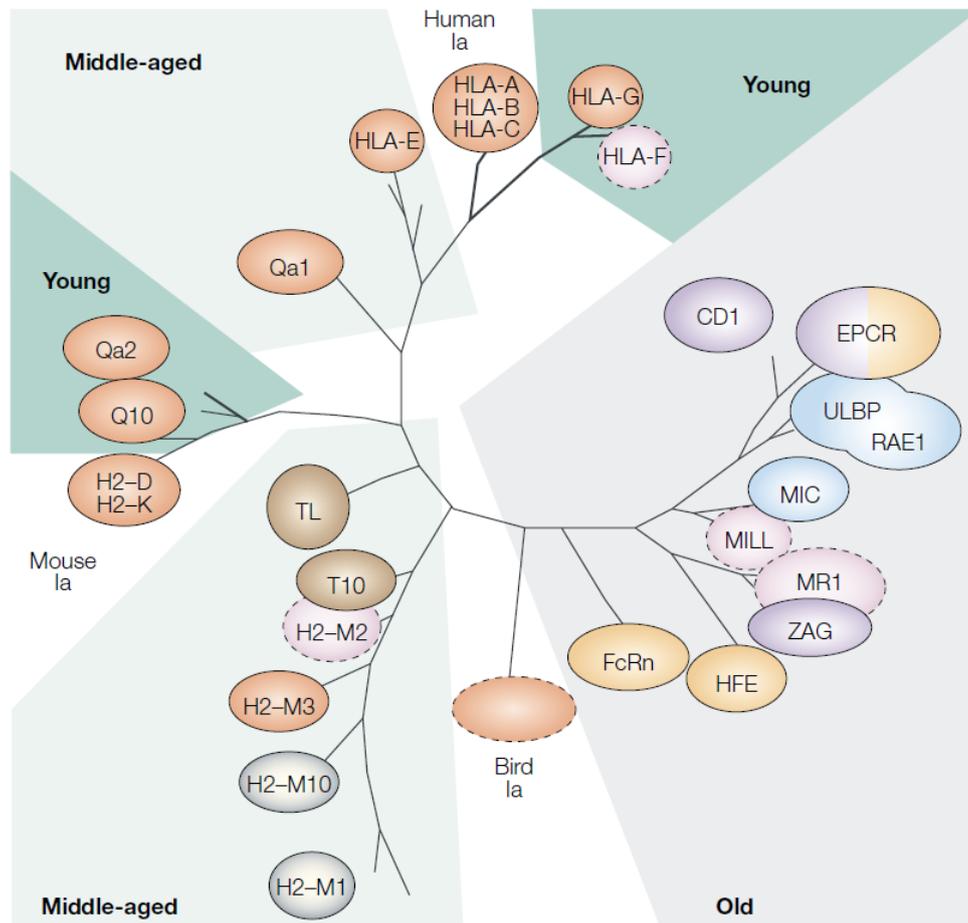


Figure 1.4. Phylogenetic relationships of MHC class I genes. Non-classical class I molecules are grouped as “young”, “middle-aged” or “old”. (Adapted from Rodgers and Cook, 2005)

The young non-classical class I genes that are closely related to classical class I genes include the human *HLA-F* and *HLA-G* as well as the mouse *H-2Qa2*, *H-2Q10* and *H-2B1* genes (Fig. 1.4 and Table 1.1). The binding motifs and the diversity of peptides presented by young non-classical MHC class I molecules are similar to those of classical MHC class I molecules (Munz et al., 1999; Zappacosta et al., 2000; He et al., 2001). *HLA-G*, *H-2Qa2* and *H-2Q10* bind a diverse set of peptides using two or three anchors motifs, with both termini buried in the peptide-binding groove, akin to classical class I molecules (Walpole et al., 2010). However, the expression of these young non-classical class I molecules is tissue-specific, unlike their classical class I counterparts, which are expressed ubiquitously (Zappacosta et al., 2000; Comiskey et al., 2003; Ishitani et al., 2003).

Table 1.1. Classification of classical and non-classical MHC class I molecules. (Adapted and modified from Rodgers and Cook, 2005).

	MHC-linked?	Ligand	Binds TCR?	Binds NK-cell receptor?	Binds β_2m?	Contains α_3-domain?
Classical MHC class I						
Human HLA-A, HLA-B, HLA-C	Yes	Peptide	Yes	LIRs, KIRs	Yes	Yes
Mouse H-2K, H-2D, H-2L	Yes	Peptide	Yes	Ly49	Yes	Yes
Chicken BF2	Yes	Peptide	ND	ND	Yes	Yes
Chicken YF1	No	Lipids	ND	ND	Yes	Yes
Non-classical MHC class I (young)						
Human HLA-F	Yes	Peptide?	ND	LIR1	Yes	Yes
Human HLA-G	Yes	Peptide	Yes	LIRs	Yes	Yes
Mouse H-2Q10	Yes	Peptide	Yes	LIRs	Yes	Yes
Mouse H-2Qa2	Yes	Peptide	Yes	Yes but undefined	Yes	Yes
Non-classical MHC class I (middle-aged)						
Human HLA-E	Yes	Peptide	Yes	CD94-NKG2A/C	Yes	Yes
Mouse H-2Qa1	Yes	Peptide	Yes	CD94-NKG2A/C	Yes	Yes
Mouse H-2M3	Yes	<i>N</i> -formyl peptide	Yes	ND	Yes	Yes
Mouse T10 and H-2TL	Yes	None	Yes	ND	Yes	Yes
Mouse H-2M1 and H-2M10	Yes	Odorant?	ND	ND	Yes	Yes
Non-classical MHC class I (old)						
Human MIC	Yes	None	ND	NKG2D	No	Yes
Mouse MILL	No	None	ND	ND	No	Yes
Human HFE	Yes	Transferrin	No	ND	Yes	Yes
Mouse HFE	No	Transferrin	No	ND	Yes	Yes
Human and mouse CD1	No	Lipid & lipopeptide	Yes	Yes but undefined	Yes	Yes
Human and mouse EPCR		Lipid and protein C	ND	ND	No	No
Human and mouse MR1	No	Lipid?	Yes	ND	Yes	Yes
Human and mouse ULBP (or RAE1)	No	None	ND	NKG2D	No	No
Human and mouse FcRn	No	IgG	ND	ND	Yes	Yes
Human and mouse ZAG	No	Lipid?	ND	ND	No	Yes

Only one non-classical class I gene is classified as middle-aged in human, i.e. *HLA-E*, in contrast to at least seven of such genes with diverse functions in mouse (Table 1.1 and Fig. 1.4). For example, mouse H-2Qa1 (also known as H-2T23) and H-2T10 interact with TCR without binding a peptide or ligand (Wingren et al., 2000; Liu et al., 2003), whereas H-2M1 and H-2M10 appear to interact with distinct vomeronasal receptor type II molecule (Loconto et al., 2003). Human HLA-E and murine H-2Qa1 are specialized to present ligands derived from certain MHC class I signal peptides, from TCR variable regions or from a conserved region of heat shock protein 60 (HSP60) and its bacterial ortholog GroEL. HLA-E and H-2Qa1 regulate NK cells by acting as a ligand for CD94/NKG2 receptors that inhibit NK cell cytotoxicity (Borrego et al., 1998). Murine H-2M3 is specialized in binding *N*-formylated peptides (Wang et al., 1995), as shown in mice infected with bacterium *Listeria monocytogenes* and *Mycobacterium tuberculosis* (D'Orazio et al., 2003; Urdahl et al., 2003).

Old non-classical class I genes (also known as MHC class I-like loci) might have diverged to evolve different functions early in vertebrate evolution. Possibly as a consequence, they are usually no longer linked to the MHC. The best studied molecules of this group are the CD1 molecules. They have a narrow, sometimes maze-like and hydrophobic binding groove that is optimized to bind and present lipids to TCR mounted on invariant CD1-restricted natural killer T cells (Brigl and Brenner, 2004). Noteworthy, it has also been claimed that the murine CD1d molecule can bind peptides with a hydrophobic motif and that the complex can be recognized by CD8⁺ T cells (Castano et al., 1995; Brossay et al., 1998). However, no crystal structure of peptide-bound CD1 molecule has been reported to date.

Other old non-classical class I genes include *HFE* (also known as hereditary hemochromatosis), that is encoded in the extended *HLA* class I region (Feder et al., 1996), but maps outside of the *H-2* complex in mice. HFE interacts with the transferrin receptor and plays a role in iron metabolism (Feder et al., 1998). The crystal structure of HFE shows the inability of its narrowed binding groove to bind a peptide (Lebron et al., 1998), but, paradoxically, it has been shown to be recognized by CD8⁺ T-cells in mice lacking classical class I molecules (Rohrlich et al., 2005). Like *HFE*, another group of highly polymorphic non-classical class I genes are located within the *HLA* complex called MHC class I chain-related genes (*MIC*) (Bahram et al., 1994). MICA and the closely related MICB can be recognized by intestinal epithelial T cells that might have a specialized role in reporting stress (Groh et al., 1998). MICA has also been shown to engage with NKG2D (also known as

KLRK1) on NK cells and T-cells (Bauer et al., 1999). This was later supported by a crystal structure of the MICA-NKG2D complex (Li et al., 2001). MICA may promote anti-tumour NK- and T-cell responses (Groh et al., 2002). Similarly, a protein encoded by human cytomegalovirus (CMV) named UL16 which binds MICB, mimics the effect of HLA class I molecules and prevents the interaction with the NKG2D receptor, hence protecting virus-infected cells from immune surveillance (Cosman et al., 2001; Kubin et al., 2001; Sutherland et al., 2001). Rodents do not have *MIC* genes, but a closely related gene family to *MIC* has been identified in mice and rats. These MHC class I-like loci (*MILL*) are located near the leukocyte receptor complex (LRC) that emerged before the radiation of mammals (Watanabe et al., 2004). The functional role of *MILL* genes has yet to be established, but it has been reported that the cell surface expression of *MILL* is β_2 -microglobulin (β_2m)-associated although exogenously added peptide and TAP do not appear to be required (Kajikawa et al., 2006).

Another two MHC class I-related genes located at the vicinity of *HLA* of chromosome 6 are CMV UL-16 binding proteins (*ULBP*), also known as retinoic acid early transcript (*RAE*) (Cosman et al., 2001; Radosavljevic et al., 2002). *ULBP* are frequently expressed by tumour cells that serve as the ligands for NKG2D and mediate cytotoxicity of NK cells and $\alpha\beta$ -T-cells (Cao et al., 2007). Another group of MHC class I-related genes are simply called MHC class I-related (*MRI*). In contrast to *MIC*, *MRI* is monomorphic and has been mapped to chromosome 1 (Hashimoto et al., 1995). *MR1* is involved in the development and expansion of a small population of T cells called mucosal-associated invariant T cells (MAIT) that are preferentially located in the gut lamina propria and therefore may be involved in monitoring commensal flora or serve as a distress signal (Treiner et al., 2003). Other old non-classical class I molecules are structurally related to classical class I molecules but possess diverse functions. The endothelial protein C receptor (EPCR), zinc- α_2 -glycoprotein (ZAG) and neonatal Fc receptor (FcRn) have similar folds to that of classical class I molecules but do not bind peptide antigen (Burmeister et al., 1994; Oganessian et al., 2002). EPCR is a receptor for activated protein C, is a serine protease involved in blood coagulation (Ye et al., 1999; Riewald et al., 2002). ZAG may bind lipids and stimulates lipid degradation in adipocytes, causing the extensive fat losses associated with cancers (Hassan et al., 2008b). The intestinal epithelium of neonatal mice and rats expresses FcRn that binds to the Fc region of immunoglobulin G (IgG) in mother's milk and transfers antibodies to the foetus, helping the newborn animals to acquire passive immunity (Story et al., 1994). FcRn also protects IgG

from degradation, making IgG the longest lived of all plasma proteins (Waldmann and Strober, 1969; Junghans and Anderson, 1996).

1.4 Architecture of MHC Class I Molecules

Classical class I molecules are trimeric proteins consisting of a heavy chain (HC) which is non-covalently associated with the light chain, β_2m , as well as a peptide (Fig. 1.5A) (Bjorkman et al., 1987). The HC traverses the cell membrane and is divided into three extracellular domains (α_1 , α_2 and α_3 with about 90 amino acids per domain), a transmembrane region (~ 40 amino acids) and a short cytoplasmic tail (~ 30 amino acids). Both HC and β_2m are members of the Ig superfamily that share a disulfide-bonded domain structure with antibodies (Hood et al., 1985). The α_1 and α_2 domains form a long groove between the α_1 - and α_2 - helices which serve as the binding site for peptides (Fig. 1.5) (Saper et al., 1991).

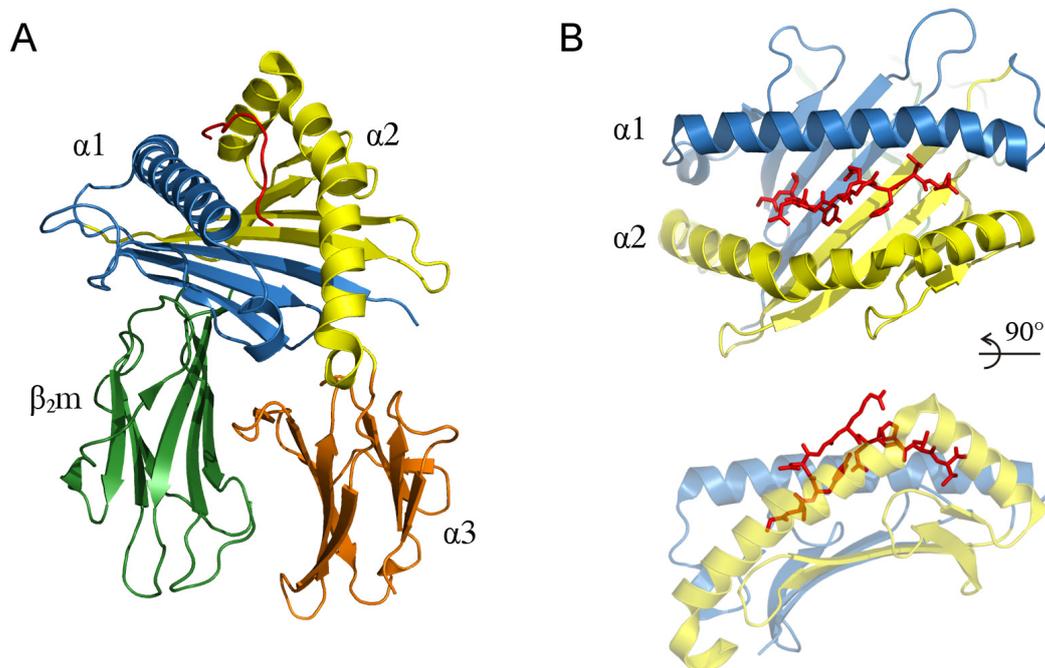


Figure 1.5. Structure of a classical MHC class I complex. Here, HLA-A2 (PDB entry 1IEF) is used to render the images. **(A)** Overall scaffold of an MHC class I complex shown in ribbon representation with each domain labelled and coloured (blue for α_1 domain, yellow for α_2 domain, orange for α_3 domain and green for β_2m). The peptide is shown as red lines in the α_1 - α_2 binding groove. **(B)** Two views of the peptide-binding groove: top view looking down into the binding groove (top) and side view looking through the α_2 -helix (semi-transparent) towards the α_1 -helix. A peptide is shown in stick representation (red).

The peptide binding groove can be subdivided into six pockets, termed pocket A to F (Deres et al., 1993; Matsui et al., 1993). HC residues located in the binding groove are highly polymorphic; hence, different alleles produce binding grooves with different binding capabilities. The groove is closed at both ends, limiting the size of suitable peptides to 8-12 amino acids (Rammensee et al., 1995), although longer peptides can also be bound (Urban et al., 1994; Bell et al., 2009). The side chains of peptides are held at the N- and C-termini by anchoring into the pockets of the binding groove and the non-anchoring amino acids (usually in the middle of the peptide) bulge out to interact with the TCR (Fig. 1.5B). Anchor residues of a peptide are conserved and specific for a particular class I allele while the non-anchoring residues vary considerably, therefore allowing numerous peptides to be presented by a given MHC class I allele. The relatively non-polymorphic $\alpha 3$ -domain is close to the membrane and non-covalently associated with β_2m . Both are located beneath the binding groove (Fig. 1.5A).

To date, there are more than 200 MHC class I structures (both classical and non-classical) deposited in the Protein Data Bank (PDB). The available classical class I molecules include human HLA-A, HLA-B, HLA-C, macaque Mamu-A, bovine BoLA, mouse H-2K, H-2D, H-2L, rat RT1-A, chicken BF2 and YF1. All classical peptide-binding MHC class I molecules determined to date share an identical overall architecture as shown in Fig. 1.5 (Bjorkman et al., 1987; Wang et al., 1996; Koch et al., 2007; Hee et al., 2010). On the other hand, the structures of non-classical class I molecules vary according to their “age” as described in Section 1.3.2. The overall structures of “young” and “middle-age” non-classical class I molecules are highly similar to the classical peptide-binding molecules, with key differences within the binding grooves. For instance, HLA-E and HLA-G possess more anchor sites than their classical class I counterparts (Sullivan et al., 2006). The crystal structure of HLA-E also revealed a relatively hydrophobic binding groove which is specialized in binding hydrophobic leader peptides from certain HLA class I molecules (Petrie et al., 2008). Mouse H-2M3 selectively binds *N*-formylated peptides and possesses an interesting adaptation of the binding groove: the A pocket of other MHC class I molecules, which accommodates the positively charged amino (N) terminus of peptides, is closed in H-2M3, and the entire peptide is “frame-shifted” by one position so that the N-formyl group engages the first peptide bond site (Wang et al., 1996). H-2Qa2, another “young” non-classical molecule, has a shallow and hydrophobic binding groove that makes few specific contacts and exhibits a poor shape complementarity to the peptide. The structure implies a “conformational selection” instead of specific interaction by H-2Qa2 for peptides (He et al., 2001).

Some “old” non-classical class I molecules still maintain the overall class I scaffold and bind β_2m while some have retained only the $\alpha 1$ - $\alpha 2$ domains but are unable to bind a peptide or ligand. The best studied “old” non-classical class I molecule with most structures determined is CD1 (Silk et al., 2008). The structure of human CD1a (Zajonc et al., 2003), CD1b (Garcia-Alles et al., 2006), CD1c (Scharf et al., 2010), CD1d (Koch et al., 2005), mouse CD1d (Zeng et al., 1997), bovine CD1b3 (Girardi et al., 2010) and two chicken CD1, CD1-1 (Dvir et al., 2010) and CD1-2 (Zajonc et al., 2008) molecules have been solved. CD1 molecules exhibit the typical class I scaffold and are non-covalently associated with β_2m . However, they possess narrow, deep and tunnel-like hydrophobic binding grooves that specialize in binding the lipidic tails of glycolipids or lipopeptides, with the polar moieties presented to T cells (Borg et al., 2007). On the other hand, HFE and FcRn are associated with β_2m , but do not bind peptides or other small ligands due to the narrowing of the binding grooves (Lebron et al., 1998; Vaughn and Bjorkman, 1998). MIC and ZAG are two examples of “old” non-classical class I molecules that possess the $\alpha 1$ - $\alpha 2$ - $\alpha 3$ -domains but do not bind β_2m or peptides (Li et al., 1999; Sanchez et al., 1999). The structure of MICA in complex with the NK cell receptor NKG2D revealed that the interaction is analogous to that observed in TCR-MHC class I complexes, and interestingly, the partially disordered $\alpha 2$ -helix in the structure of MICA alone, is ordered in the complex and forms parts of the NKG2D interface (Li et al., 2001). The binding grooves of MICA or MICB are too narrow to bind peptides, but the binding groove of ZAG is large and open. In the crystal structure, it clearly contains a non-peptidic compound (Sanchez et al., 1999). Intriguingly, a recent report on the crystal structure of a complex purified from human seminal plasma between ZAG and prolactin-inducible protein (PIP) revealed that PIP associated in a different binding mode with ZAG at the location typically occupied by β_2m in classical class I molecule (Hassan et al., 2008a). Some “old” non-classical MHC class I molecules exist only as an isolated $\alpha 1$ - $\alpha 2$ domain platform such as EPCR and RAE-1. The binding groove of EPCR is strikingly similar to that of CD1d, and is also capable to bind phospholipid, which is crucial for the binding of protein C (Oganessian et al., 2002). RAE-1 is unable to bind ligand in its binding groove due to narrowing between of the $\alpha 1$ and $\alpha 2$ -helices. The RAE-1 structure also revealed non-canonical disulfide bonds and promiscuous binding of NKG2D when compared to the structure of MICA-NKG2D (Li et al., 2002).

1.5 Chicken MHC Class I Antigens

1.5.1 Chicken Class I Antigens of the *MHC-B* region

The chicken provides one of the best examples of linkage between MHC genetics and resistance to infections. The chicken *MHC-B* region is in strikingly strong association with the growth of tumours following infection with the oncogenic Marek's disease virus (MDV) (Briles et al., 1977; Bacon et al., 2001) and Rous sarcoma virus (RSV) (Plachy and Benda, 1981; White et al., 1994). For example, birds with the *B21* haplotype remained tumour-free after MDV infection while other haplotypes confer only moderate resistance (*B2* haplotype) or are highly susceptible to MDV infection (*B13*, *B15* and *B19* haplotypes). The B haplotypes have also been reported to have influence on the efficacy of the vaccines used to protect chickens from Marek's disease (Bacon and Witter, 1992; Bacon and Witter, 1995; Lee et al., 2004). The genetic basis of the MHC-linked MDV infection and vaccine efficacy remains unresolved, but it has been postulated that polymorphic loci within the *MHC-B* region might be responsible for the observed differences (Kaufman et al., 1995; Kaufman et al., 1999a). There are a number of polymorphic loci within *MHC-B* that could contribute to differential reactions upon infection or vaccination, including class I genes.

Two class I genes are part of *MHC-B*, namely *BF1* and *BF2* (Kaufman et al., 1999b). The *BF2* locus is predominantly expressed, whereas the *BF1* locus is the less well-expressed. The mRNA expression levels of *BF2* and *BF1* differ by as much as 10-fold (Kaufman et al., 1995). The reason for the differential expression lies in the sequence of the distal region of *BF1* which includes an enhancer A element that is highly diverged in some haplotypes (*B2*, *B4*, *B6* and *B21*), disrupted (*B14* and *B15*) or deleted entirely (*B12* and *B19*) (Kroemer et al., 1990). The single predominantly expressed *BF2* gene together with the "minimal essential" organization of the *MHC-B* region make the chicken one of the best candidates for MHC-linked disease association studies.

Some studies explained the striking disease association by studying the peptide anchor motif that is required by the predominantly expressed *BF2* molecule in different haplotypes. The class I molecules of RSV-susceptible haplotypes such as *B4* and *B15* have been predicted to have a lower number of peptides that can be presented due to the "stringent" selection of peptide binding motifs, while the class I molecule of the RSV-resistant haplotype *B12* could bind to a larger number of viral peptides because of the less restricted peptide binding motif (Wallny et al., 2006). Another study examined the peptide motif of two class I molecules

with their haplotypes differentially associated with MDV infections: the susceptible *B13* (same as *B4*) and resistant haplotype *B21*. The results are in agreement with the previous study, as the class I molecule of the *B13* haplotype (MDV-susceptible) has a more restricted peptide binding motif able to present only a small number of viral peptides compared to the class I molecule of *B21* (MDV-resistance) which has a loosely defined motif that could bind to a vast variety of peptide sequences (Sherman et al., 2008).

The two crystal structures of BF2*2101 bound to peptides with different length and chemical nature illustrate the promiscuous peptide binding mode of this molecule (Koch et al., 2007). The binding groove of BF2*2101 has a large central cavity which allows a conformationally flexible arginine side chain at the floor of the binding groove to interact with different peptide sequences. These features of the BF2*2101 molecule enlarge the pool of interacting peptides since peptides with different lengths and sequences can be accommodated by its binding groove.

1.5.2 Chicken Class I Antigens of the *MHC-Y* region

The *MHC-Y* region has been associated with allograft rejection (Pharr et al., 1997; Thoraval et al., 2003), resistance to MDV infection (Wakenell et al., 1996; Pharr et al., 1997) and susceptibility to RSV infection (LePage et al., 2000; Praharaj et al., 2004). However, there is no definitive proof for these associations as yet. Recognition of YF molecules by T cell has not been demonstrated, but preliminary evidence suggests that YF molecules may be recognized by NK cells (Foster et al., 2005).

As in the *MHC-B* region, two class I genes (*YF1* and *YF2*) have been mapped to the *MHC-Y* region (Miller et al., 1994). More than 17 *MHC-Y* haplotypes have been identified so far (Miller et al., 2004) and some *MHC-Y* haplotypes are known to express two and potentially more class I genes (Thoraval et al., 2003). It has also been reported that at least one MHC class I locus, *YF1* (previously known as *YFV*), is ubiquitously transcribed in virtually all organs of both adult and embryonic chickens and is associated with β_2m as a mature protein (Afanassieff et al., 2001). In a separate study, Hunt and colleagues concluded that the YF1*7.1 is dynamically expressed on the surface of blood cells within the spleen, both pre- and post-hatching (Hunt et al., 2006).

Similar to BF2 molecules, YF1 is highly polymorphic in the region that might bind small ligands. YF1 alleles differ as much as 21% in their $\alpha 1$ domain which is comparable to the value of 27% for BF2 alleles, while the $\alpha 2$ -domains of both molecules are equally less variable by the value of $\sim 10\%$. In the antigen binding region ($\alpha 1$ - and $\alpha 2$ -domains), YF1 shares about 57% sequence identity with BF2 molecules and about 44% with the human HLA-A2 (Hunt and Fulton, 1998; Afanassieff et al., 2001). The polymorphism and the relatively high ratios of non-synonymous to synonymous substitutions in the YF1 binding grooves indicate that this region may be under selection for diversity (Afanassieff et al., 2001). However, little is known about the ligand and immunological function of YF1, despite the fact that YF1 shares many features with classical MHC class I molecules such as the BF2 molecules.

1.6 β_2 -microglobulin

β_2m was originally discovered in the urine of patients with chronic kidney dysfunction caused by cadmium poisoning (Berggard and Bearn, 1968). Free β_2m is found in most of the body fluids (Peterson et al., 1977), including serum, urine, milk and colostrums (Cejka et al., 1976). More than 99% of the circulating β_2m is eliminated by the kidney and in the case of renal failure the concentration of serum β_2m increases up to 60-fold (Floege and Ehlerding, 1996). Due to the abundance of free β_2m , amyloid fibrils are formed that may be deposited in the synovia of carpal tunnels (Gejyo et al., 1985) leading to the development of dialysis-related amyloidosis. The formation of amyloid fibrils from soluble β_2m has not been fully understood but several factors promoting amyloid formation have been identified. These include protein truncation, acidic conditions, the presence of amyloid seeds or the influence of copper ions (Heegaard, 2009). Two recent publications provide atomic insights into amyloid formation by employing structural approaches. Radford and co-workers determined the solution structure of a truncated β_2m variant and discovered a rare conformer with increased amyloidogenic potential (Eichner et al., 2011). On the other hand, Eisenberg and colleagues solved the crystal structures of a covalently bound, domain-swapped dimer and the fibrillizing segment located in the hinge loop that fit well into a fibril model are consistent with biochemical and biophysical findings (Liu et al., 2011). Little is known about the amyloidogenic behaviour of β_2m from other species. However, it has been shown that by swapping seven residues of human β_2m into mouse β_2m , the hybrid is transformed into an amyloid-prone state (Ivanova et al., 2004).

Since it is non-covalently bound to MHC class I or class I-like heavy chains, β_2m has been coined the light chains of these complexes (Fig. 1.5). It contains 99 residues that fold into a seven-stranded β -sandwich stabilized by one intra-molecular disulfide bond (Becker and Reeke, Jr., 1985). β_2m serves as a chaperone that stabilizes MHC class I heavy chains (Kraugel et al., 1979; Lancet et al., 1979). It is also involved in the function of class I molecules in the context of recognition by antibodies, T cells and NK cells (Michaelsson et al., 2001; Wang et al., 2002). In contrast to MHC class I genes, the polymorphism of β_2m is very limited: seven alleles are known in mouse (Gasser et al., 1985; Hermel et al., 1993) but only one has been identified in human and chicken (Riegert et al., 1996). The three-dimensional (3D) structures of free monomeric β_2m from different species are highly similar although they share less than 50% of their amino acid sequences (Becker and Reeke, Jr., 1985; Trinh et al., 2002; Chen et al., 2010). Owing to the high structural similarity, cross-species association of β_2m with MHC class I heavy chain is possible (Kubota, 1984; Bernabeu et al., 1985). Recent structural analyses provide first insights into cross-species interactions and explain why in some cases β_2m from another species has a higher affinity for a given heavy chain than the native β_2m . For example, human β_2m has been shown to enhance the stability of two very distinct heavy chains, the mouse H-2D^b and chicken CD1-2 (Achour et al., 2006; Zajonc et al., 2008).

1.7 Scope and Objectives of this Study

This doctoral work aimed at elucidating (a) the possible function of the chicken MHC class I molecule YF1*7.1, (b) the biophysical characteristics of chicken β_2m , dynamics of free and HC-bound human β_2m , and the evolutionary role of β_2m as a subunit of MHC class I complexes. The third aim was (c) to simplify and optimize the method of communicating 3D structural data in a scientific publication.

(a) The major focus of this study aimed at understanding the immunological function of the dominantly expressed MHC class I gene *YF1* by using biophysical approaches. *YF1* is interesting because it shares many features of classical MHC class I genes, such as polymorphism and non-restricted tissue distribution. In addition, *MHC-Y* haplotypes have been associated with allograft rejection and viral infection although the basis of these associations remains to be determined. The YF1-directed were expected to shed light on natural ligands of this molecule and to provide hints on the role of YF1 in the immune system of the chicken.

(b) The second part of this doctoral study had two foci. First, the characterization of free chicken β_2m by structural and biophysical approaches was attempted, followed by comparison with β_2m from other species. The results were expected to identify characteristics of chicken β_2m that might distinguish it from β_2m of other vertebrates, possibly providing hints for the suitability of chicken β_2m as a model for amyloidosis. Second, comparisons of the molecular dynamics of free and HC-bound human β_2m in solution were carried out, as well as an extensive analysis of the interactions between the interfaces of β_2m and HC of selected MHC class I and class I-like molecules across species. The results were expected to improve our understanding on how monomorphic β_2m stabilizes a vast variety of MHC class I complexes, ranging from classical to non-classical.

(c) The third aim of this doctoral study was to develop and optimize a simplified method to embed 3D information within a PDF, as the conventional publishing in 2D is inadequate to present complicated 3D structural information. The recent advancement in PDF-related software and programs now permits an incorporation of 3D models into a PDF, thereby enhancing the communication of structural information, especially to the readers unfamiliar with molecular visualization programs. However, the method to embed 3D imagery was not straightforward as it required multiple processes involving several programs. I aimed to optimize a “user-friendly” method to embed structural 3D information within a PDF for the benefit of the structural biology community as well as scientists from other disciplines.

2. The Structure of the Chicken YF1*7.1 Molecule and its Ligands

2.1 Summary

The function of *YF1* has been puzzling since the discovery of this MHC class I gene outside of the core chicken *MHC-B* region. The objective of the studies described in the following articles was to understand the role of YF1 in the chicken immune system. A structural approach was chosen as the first attempt to gain insights into the molecular properties of YF1. The first part of the work involved extensive experiments to identify the right expression vectors for YF1*7.1 and chicken β_2m , since the vectors employed for human HC and β_2m turned out to be unsuitable for the corresponding chicken proteins. This was then followed by optimizing the reconstitution, purification and crystallization conditions of the protein complexes. The second part of this study were devoted to the X-ray crystallographic analyses of various YF1*7.1:ligand complexes and a computational analysis of the structures, particularly in the antigen binding grooves that contained various ligands. This was followed by lipid binding assays to assess the ability of YF1*7.1 to bind self/nonself-lipids, and thermostability analyses to determine its thermodynamic behaviour. The results reveal that YF1*7.1 exhibit the classical MHC class I architecture but possesses a hydrophobic binding groove containing non-peptidic ligands. Lipid binding assays showed that YF1*7.1 is capable of binding bacterial lipids. This unprecedented characteristic of YF1*7.1 suggests that it might aid the chicken to enlarge its otherwise very small repertoire of antigen-presenting MHC class I molecules. Comparative analyses of YF1*7.1 and other MHC class I molecules also revealed a distinctive structural feature of the complex in chickens that appears to be shared by non-mammalian but not by mammalian vertebrates.

2.2 Publications

2.2.1 Hee, C.S., Gao, S., Miller, M.M., Goto, R.M., Ziegler, A., Daumke, O., Uchanska-Ziegler, B. (2009). Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 65, 422-425.

2.2.2 Hee, C.S., Gao, S., Loll, B., Miller, M.M., Uchanska-Ziegler, B., Daumke, O., Ziegler, A. (2010). Structure of a classical MHC class I molecule that binds "non-classical" ligands. *PLoS Biol.* 8, e1000557.

Acta Crystallographica Section F

**Structural Biology
and Crystallization
Communications**

ISSN 1744-3091

Editors: **H. M. Einspahr and M. S. Weiss**

Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1

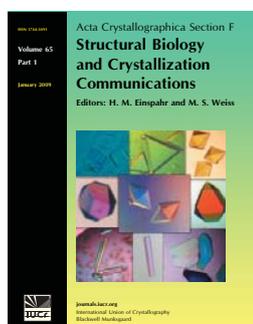
Chee Seng Hee, Song Gao, Marcia M. Miller, Ronald M. Goto, Andreas Ziegler, Oliver Daumke and Barbara Uchanska-Ziegler

Acta Cryst. (2009). **F65**, 422–425

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section F: Structural Biology and Crystallization Communications is a rapid all-electronic journal, which provides a home for short communications on the crystallization and structure of biological macromolecules. It includes four categories of publication: protein structure communications; nucleic acid structure communications; structural genomics communications; and crystallization communications. Structures determined through structural genomics initiatives or from iterative studies such as those used in the pharmaceutical industry are particularly welcomed. *Section F* is essential for all those interested in structural biology including molecular biologists, biochemists, crystallization specialists, structural biologists, biophysicists, pharmacologists and other life scientists.

Crystallography Journals **Online** is available from journals.iucr.org

crystallization communications

Acta Crystallographica Section F

Structural Biology
and Crystallization
Communications

ISSN 1744-3091

Chee Seng Hee,^{a,†} Song Gao,^{b,†}
Marcia M. Miller,^c Ronald M.
Goto,^c Andreas Ziegler,^{a*}
Oliver Daumke^b and Barbara
Uchanska-Ziegler^a^aInstitut für Immunogenetik, Charité-
Universitätsmedizin Berlin, Freie Universität
Berlin, Thielallee 73, 14195 Berlin, Germany,^bMax-Delbrück-Centrum für Molekulare
Medizin, Robert-Rössle-Strasse 10,
13125 Berlin, Germany, and ^cDivision of
Molecular Biology, City of Hope, Beckman
Research Institute, Duarte, CA 91010, USA† These authors contributed equally to this
study.Correspondence e-mail:
andreas.ziegler@charite.deReceived 15 December 2008
Accepted 11 March 2009© 2009 International Union of Crystallography
All rights reservedExpression, purification and preliminary X-ray
crystallographic analysis of the chicken MHC class I
molecule YF1*7.1

*YF1*7.1* is an allele of a polymorphic major histocompatibility complex (MHC) class I-like locus within the chicken *Y* gene complex. With the aim of understanding the possible role of the *YF1*7.1* molecule in antigen presentation, the complex of *YF1*7.1* heavy chain and β_2 -microglobulin was reconstituted and purified without a peptide. Crystals diffracted synchrotron radiation to 1.32 Å resolution and belonged to the monoclinic space group $P2_1$. The phase problem was solved by molecular replacement. A detailed examination of the structure may provide insight into the type of ligand that could be bound by the *YF1*7.1* molecule.

1. Introduction

The chicken *Y* system (previously known as *Rfp-Y*) was initially detected as an independently segregating polymorphic major histocompatibility complex (MHC)-like locus (Briles *et al.*, 1993). The *Y* and the *B* (the chicken MHC) systems are two genetically unlinked gene clusters that map to the same microchromosome (Briles *et al.*, 1993; Miller *et al.*, 1994, 1996; Fillon *et al.*, 1996). Comparable to the situation in the *B* cluster, several MHC genes map to the *Y* cluster, including at least two class I heavy chain loci (*YF1* and *YF2*) and three class II β genes (*YLB1*, *YLB2* and *YLB3*) (Zoorob *et al.*, 1993). The sequence identity between the *Y* class I proteins is about 93%, while values of about 63% are obtained when comparing them with the chicken *B* system class I proteins (BF). The sequence identity to the human HLA-A antigen, which like BF molecules displays peptides (Madden, 1995; Afanassieff *et al.*, 2001; Wallny *et al.*, 2006; Koch *et al.*, 2007), is still about 49%, but there is only a very limited sequence identity (~20%) to mammalian and chicken MHC class I-like CD1 molecules (Miller *et al.*, 2005; Salomonsen *et al.*, 2005; Zajonc *et al.*, 2008).

YF transcripts are ubiquitously distributed in nearly all organs of both adult as well as embryonic chickens and are translated to yield a mature heavy chain that is associated with β_2 -microglobulin (β_2m ; Afanassieff *et al.*, 2001). *YF* class I antigens are dynamically expressed at levels comparable to but independently of BF class I molecules on erythrocytes, lymphocytes, granulocytes, monocytes and thrombocytes within the spleen of birds before and after hatching (Hunt *et al.*, 2006). These findings revealed that the expression of *YF* antigens is not restricted, as is observed for the products of many nonclassical class I genes in mammals (Shawar *et al.*, 1994). *YF* molecules could thus be more closely related to typical class I antigens than to nonclassical class I molecules, as already suggested by the close sequence identity between the two proteins.

YF1 antigens exhibit a number of substitutions of binding-groove residues that are usually conserved in the case of 'classical' class I proteins, indicating that *YF* molecules might present ligands that are different from the peptides bound by classical BF molecules. However, this remains speculative as ligands for *YF* molecules have so far not been identified. In addition, although the *Y* system genes have been associated with allograft rejection (Bacon & Witter, 1995; Pharr *et al.*, 1996; Thoraval *et al.*, 2003), resistance to Marek's disease virus (MDV; Wakenell *et al.*, 1996; Pharr *et al.*, 1997) and susceptibility to Rous sarcoma virus (RSV; LePage *et al.*, 2000; van der Laan *et al.*,

2004; Praharaj *et al.*, 2004), there is as yet no definitive proof for these associations and the identity of a cellular receptor for YF antigens is equally uncertain.

We have employed X-ray crystallography to shed light on the structural basis of the antigen-presenting properties of the YF1*7.1 antigen. A close examination of the molecule, in particular its ligand-binding site, may provide hints regarding the chemical nature of YF ligands and could thus help to find a function for the avian YF antigens.

2. Materials and methods

2.1. Protein preparation

The cDNA sequences encoding the extracellular domains of YF1*7.1 (residues 22–294 of the signal peptide-containing protein) and chicken β_2m (residues 21–120 of the signal peptide-containing protein) were cloned into the vector pMAL-p4x, expressed as a maltose-binding protein (MBP) fusion construct and purified using the pMAL purification system (New England Biolabs, Germany) with a number of modifications. A 1 l LB culture of *Escherichia coli* TB1 transformed with the YF1*7.1 or the β_2m construct was grown to an OD₆₀₀ of 0.5. Protein expression was induced by adding 0.4 mM isopropyl β -D-1-thiogalactopyranoside and cells were incubated for 4 h at 303 K in the case of YF1*7.1 and ~20 h at 298 K in the case of β_2m . They were harvested by centrifugation at 4500g for 10 min at 277 K.

For YF1*7.1, the cell pellet was resuspended in 30 ml lysis buffer [25% saccharose, 1 mM EDTA, 50 mM Tris–HCl pH 8.0 and 30 μ M phenylmethanesulfonyl fluoride (PMSF)] and frozen at 253 K. For preparation of inclusion bodies, the sample was thawed and incubated for 30 min on ice in the presence of 1.7 mg lysozyme, 1 μ g ml⁻¹ DNase I, 10 μ M MgCl₂ and 1 μ M MnCl₂ (all final concentrations) were then added and the cells were broken by sonication on ice for 3 min. The inclusion bodies released from the cells were collected by centrifugation at 10 000g for 10 min at 277 K and resuspended in 30 ml detergent buffer [200 mM NaCl, 1% deoxycholate, 1% (v/v) Nonidet-P40, 2 mM EDTA, 20 mM Tris–HCl pH 8.0 and 2 mM dithiothreitol (DTT)]. The sample was then subjected to sonication and centrifugation as above and the pellet was resuspended in 30 ml Triton buffer [100 mM NaCl, 1 mM EDTA, 0.5% (v/v) Triton X-100, 50 mM Tris–HCl pH 8.0 and 2 mM DTT], which was followed by another round of sonication and centrifugation. The pellet was washed three times with Triton buffer as before but without sonication. After the final washing step, the pellet was resuspended in 12 ml inclusion-body buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl and 10 mM DTT) and the protein was concentrated by centrifugation at 10 000g for 10 min at 277 K, resuspended in 3 ml urea buffer [50% (w/v) urea, 50 mM NaCl and 20 mM Tris–HCl pH 7.5] and shaken for 30 min at room temperature. After centrifugation at 10 000g for 20 min at room temperature, the supernatant was collected.

In the case of β_2m , the MBP-fusion protein was extracted from the periplasm according to the pMAL purification protocol. The cell pellet obtained by centrifugation was resuspended and incubated in 100 ml buffer containing 30 mM Tris–HCl pH 8.0, 20% sucrose, 1 mM EDTA and 0.3 mM PMSF on a shaker for 10 min. After incubation, the sample was centrifuged at 8000g for 10 min at 277 K and the pellet was resuspended in 100 ml chilled 5 mM MgSO₄ followed by incubation for 10 min in an ice bath. The sample was then centrifuged at 8000g for 10 min at 277 K; the supernatant was collected and filtered through a 0.22 μ m filter membrane (Millipore). The filtered

sample was applied onto a 10 ml self-packed amylose resin (New England Biolabs, Germany) column, washed with 100 ml column buffer (20 mM Tris–HCl pH 7.5, 200 mM NaCl, 1 mM EDTA) and eluted with 30 ml column buffer containing 10 mM maltose. The eluate was collected in 1 ml fractions and those containing protein (as determined by spectrophotometry at 280 nm) were pooled.

The YF1*7.1-MBP and β_2m -MBP fusion proteins were mixed at a molar ratio of 1:1 (total volume of ~10 ml) and reconstituted by adding the mixture to 500 ml of a buffer containing 400 mM arginine–HCl, 2 mM EDTA, 5 mM reduced glutathione, 0.5 mM oxidized glutathione, 100 mM Tris–HCl pH 7.5 as described by Garboczi *et al.* (1992). Notably, no peptide ligand was added to the mixture. After 2 d incubation at 277 K, the YF1*7.1-MBP– β_2m -MBP complex was collected and concentrated using an Amicon Ultra-15 protein centrifugal filter with 10 kDa cutoff (Millipore) before purification by gel-filtration chromatography using a Superdex 200 16/60 column on an ÄKTA FPLC instrument (GE Healthcare, Germany).

The purified YF1*7.1-MBP– β_2m -MBP complex was then cleaved with protease factor Xa (New England Biolabs, Germany) by incubation at 296 K for 48 h to remove the fused MBP at a complex: protease ratio of 100:1 (w:w). The YF1*7.1– β_2m complex was separated from MBP using a 10 ml self-packed amylose resin column followed by purification on a Superdex 75 column using an ÄKTA FPLC instrument. The purities of the protein complexes were assessed by SDS–PAGE. The highly pure YF1*7.1– β_2m complex was used for crystallization trials at a concentration of 13 mg ml⁻¹ in a buffer containing 50 mM NaCl and 20 mM HEPES pH 7.5. The concentration of the protein complex was determined by measuring the absorption at a wavelength of 280 nm, assuming an absorption coefficient of 1.92.

2.2. Crystallization and data collection

Initial crystallization trials using the sitting-drop vapour-diffusion technique were set up at 293 K (300 nl protein solution plus 300 nl reservoir solution equilibrated against 85 μ l reservoir solution) using a Robbins Hydra II Plus One crystallization robot. The pHClear, Classics, JCSG+, Protein Complex, PEGs and PEGs II crystal screens from Qiagen were used for initial crystallization trials. Conditions producing protein crystals were optimized in a larger volume (1 μ l protein solution plus 1 μ l reservoir solution equilibrated against 500 μ l reservoir solution) prepared manually using the hanging-drop technique. Typically, crystals appeared in 1 d and reached their maximal size after 3 d. The final crystallization conditions were 0.2 M ammonium acetate, 0.1 M sodium acetate pH 5.0, 20% (w/v) polyethylene glycol (PEG) 4000. Before flash-cooling in liquid nitrogen, crystals were briefly soaked in a cryosolution containing 0.2 M ammonium acetate, 0.1 M sodium acetate pH 5.0, 20% (w/v) PEG 4000 and 25% (v/v) glycerol.

Two X-ray diffraction data sets were collected from the same crystal at 100 K on beamline 14.1 at the BESSY II synchrotron facility in Berlin, Germany using a wavelength of 0.91841 Å. The beamline was equipped with an MX225 CCD mosaic detector (Rayonics LLC). The crystal of YF1*7.1– β_2m belonged to space group *P*12₁ and diffracted to a maximal resolution of 1.32 Å. A high-resolution data set consisting of 170 images was collected using an exposure time of 10 s per image, an oscillation range of 1° and a crystal-to-detector distance of 131.4 mm. To complete the overloaded reflections from the first data set, a second data set containing the same number of images was collected using an exposure time of 1 s, an oscillation range of 1° and a crystal-to-detector distance of 288.7 mm. The two

crystallization communications

Table 1Data-collection statistics for the YF1*7.1- β_2 m complex.

Values in parentheses are for the highest resolution shell.

	Data set 1†	Data set 2‡	Merged data set
Space group	$P12_1$		
Unit-cell parameters (\AA , °)	$a = 52.80, b = 55.47, c = 63.84,$ $\alpha = 90.00, \beta = 96.85, \gamma = 90.00$		
Solvent content (%)	40.4		
Matthews coefficient‡ ($\text{\AA}^3 \text{Da}^{-1}$)	2.1		
Wavelength (\AA)	0.91841	0.91841	0.91841
Resolution (\AA)	19.11–1.32	19.75–2.50	19.75–1.32
	(1.32–1.35)	(2.50–2.56)	(1.32–1.35)
R_{merge}^{\S}	0.035 (0.462)	0.026 (0.051)	0.044 (0.461)
$R_{\text{r.i.m.}}^{\parallel}$	0.041 (0.542)	0.031 (0.060)	0.049 (0.541)
$R_{\text{p.i.m.}}^{\parallel}$	0.022 (0.286)	0.016 (0.032)	0.022 (0.286)
Unique reflections	82130 (5825)	12580 (929)	82457 (5836)
$I/\sigma(I)$	20.60 (2.95)	38.35 (23.88)	20.22 (2.99)
Completeness (%)	95.4 (92.5)	97.5 (97.7)	95.9 (92.8)
Redundancy	3.62 (3.60)	3.60 (3.62)	4.15 (3.60)

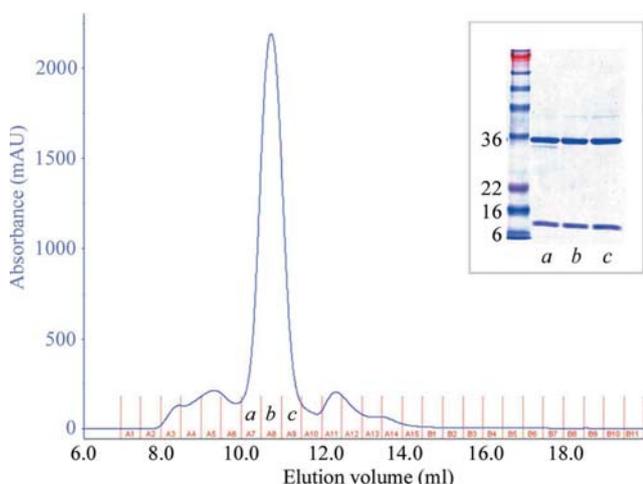
† Data sets 1 and 2 correspond to the high- and low-resolution data set, respectively. ‡ According to Matthews (1968). § $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. ¶ According to Weiss (2001).

data sets were integrated, scaled and merged using the *XDS* software package (Kabsch, 1993).

The molecular-replacement solution was obtained using the *CCP4* suite (Collaborative Computational Project, Number 4, 1994) program *MOLREP* (Vagin & Teplyakov, 1997). Refinement was performed using *REFMAC5* (Murshudov *et al.*, 1997).

3. Results and discussion

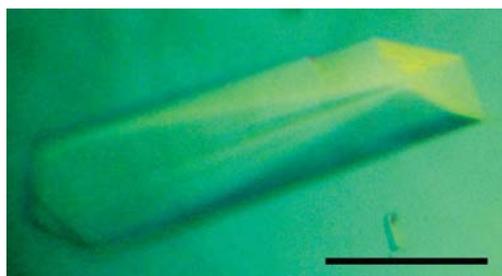
Highly pure YF1*7.1- β_2 m heterodimer was obtained following the expression of both chains in *E. coli* and a multistep purification and reconstitution procedure as described in §2 (see also Fig. 1). The YF1*7.1- β_2 m complex could be reconstituted without a peptide ligand, thus distinguishing YF from other classical MHC class I molecules, in which a peptide is essential to stabilize the cell surface-

**Figure 1**

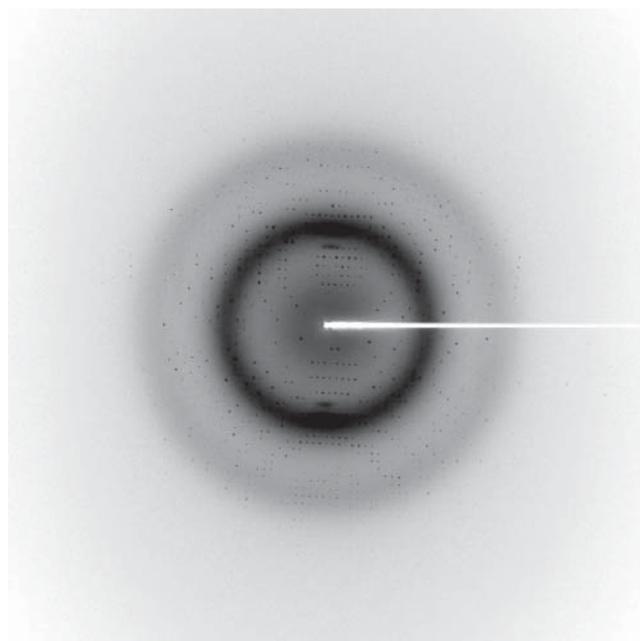
Assessment of the quality of the purified YF1*7.1- β_2 m complex using size-exclusion chromatography and SDS-PAGE under reducing conditions (inset). The blue curve represents the sample absorbance at 280 nm; numbered red fields indicate the eluted fractions. The single symmetrical peak indicates the purity and homogeneity of the sample. Inset: SDS-PAGE under reducing conditions with fractions collected from FPLC. The two bands observed in each of lanes a, b and c represent YF1*7.1 heavy chain (31.4 kDa) and β_2 m (11.6 kDa), respectively. Molecular-weight markers (kDa) are shown in the left lane.

expressed heterotrimer. Crystallization attempts at 293 K yielded several initial crystallization conditions which all contained PEG as the precipitant. After optimizing the conditions, well ordered crystals with typical dimensions of $250 \times 50 \times 30 \mu\text{m}$ (Fig. 2) could be obtained using 0.2 M ammonium acetate, 20% (w/v) PEG 4000 and 0.1 M sodium acetate pH 5.0. The crystals diffracted X-rays to a maximal resolution of 1.3 \AA using synchrotron radiation (Fig. 3) and belonged to the monoclinic space group $P12_1$ (crystallographic data and X-ray data-collection statistics are summarized in Table 1). Assuming the presence of one YF1*7.1- β_2 m complex in the asymmetric unit, the corresponding Matthews coefficient and solvent content (Matthews, 1968) were calculated to be $2.1 \text{\AA}^3 \text{Da}^{-1}$ and 40%, respectively, based on a molecular mass M_r of 45 000 for the YF complex.

Initial phases could be obtained by the molecular-replacement approach using a peptide-stripped and water-stripped model of BF2*2101 as the search model (PDB code 3bev; Koch *et al.*, 2007). The electron-density map calculated from the initial phases showed one YF1*7.1- β_2 m heterodimer in the asymmetric unit. After 20 cycles of restrained refinement, R_{work} and R_{free} values of 0.36 and 0.39, respectively, could be obtained within the resolution range

**Figure 2**

Photograph of the YF1*7.1- β_2 m crystal used for the diffraction analyses. The scale bar indicates 100 μm .

**Figure 3**

Diffraction pattern of the YF1*7.1- β_2 m crystal.

19.75–1.32 Å with a correlation coefficient of 0.841, indicating that the molecular-replacement solution was correct.

Following the recent determination of the structures of two BF-2*2101–peptide complexes (Koch *et al.*, 2007), the YF1*7.1 structure will constitute the second example of a classical MHC molecule from a nonmammalian species. However, the obvious absence of a peptide ligand from its binding groove, confirmed after the initial refinement steps, dramatically distinguishes the YF1*7.1 molecule from the BF-2*2101 antigen. A detailed investigation of the structure of the YF1*7.1 complex is currently in progress and may shed light on the type of ligand(s) that could be bound by the molecule. This might in turn allow us to obtain insight into the cellular interaction partners of YF-bearing cells.

We thank the Berliner Krebsgesellschaft for financial support of this work (Ernst von Leyden fellowship to CSH and a grant to BU-Z) and acknowledge the receipt of NRI grants 2002-35205-11628 as well as 2004-35205-14203 from the USDA CSREES. SG was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 449). We are also grateful to the staff of BESSY beamline 14.1, especially Dr J. Schulze, for their support.

References

- Afanassieff, M., Goto, R. M., Ha, J., Sherman, M. A., Zhong, L., Auffray, C., Coudert, F., Zoorob, R. & Miller, M. M. (2001). *J. Immunol.* **166**, 3324–3333.
- Bacon, L. D. & Witter, R. L. (1995). *J. Hered.* **86**, 269–273.
- Briles, W. E., Goto, R. M., Auffray, C. & Miller, M. M. (1993). *Immunogenetics*, **37**, 408–414.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Fillon, V., Zoorob, R., Yerle, M., Auffray, C. & Vignal, A. (1996). *Cytogenet. Cell Genet.* **75**, 7–9.
- Garboczi, D. N., Hung, D. T. & Wiley, D. C. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 3429–3433.
- Hunt, H. D., Goto, R. M., Foster, D. N., Bacon, L. D. & Miller, M. M. (2006). *Immunogenetics*, **58**, 297–307.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Koch, M. *et al.* (2007). *Immunity*, **6**, 885–899.
- LePage, K. T., Briles, W. E., Kopti, F. & Taylor, R. L. (2000). *Poult. Sci.* **79**, 343–348.
- Madden, D. R. (1995). *Annu. Rev. Immunol.* **13**, 587–622.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Miller, M. M., Goto, R., Bernot, A., Zoorob, R., Auffray, C., Bumstead, N. & Briles, W. E. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4397–4401.
- Miller, M. M., Goto, R. M., Taylor, R. L., Zoorob, R., Auffray, C., Briles, R. W., Briles, W. E. & Bloom, S. E. (1996). *Proc. Natl Acad. Sci. USA*, **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., Brenner, M. B. & Dascher, C. C. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8674–8679.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Pharr, G. T., Gwynn, A. V. & Bacon, L. D. (1996). *Immunogenetics*, **45**, 52–58.
- Pharr, G. T., Vallejo, R. L. & Bacon, L. D. (1997). *J. Hered.* **88**, 504–512.
- Pinard-van der Laan, M. H., Soubieux, D., Mérat, L., Bouret, D., Luneau, G., Dambrine, G. & Thoraval, P. (2004). *Genet. Sel. Evol.* **36**, 65–81.
- Praharaj, N., Beaumont, C., Dambrine, G., Soubieux, D., Mérat, L., Bouret, D., Luneau, G., Alletru, J.-M., Pinard-van der Laan, M. H., Thoraval, P. & Mignon-Grasteau, S. (2004). *Poult. Sci.* **83**, 1479–1488.
- Salomonsen, J., Sørensen, M. R., Marston, D. A., Rogers, S. L., Collen, T., van Hateren, A., Smith, A. L., Beal, R. K., Skjødt, K. & Kaufman, J. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8668–8673.
- Shawar, S. M., Vyas, J. M., Rodgers, J. R. & Rich, R. R. (1994). *Annu. Rev. Immunol.* **12**, 839–880.
- Thoraval, P., Afanassieff, M., Bouret, D., Luneau, G., Esnault, E., Goto, R. M., Chaussé, A. M., Zoorob, R., Soubieux, D., Miller, M. M. & Dambrine, G. (2003). *Immunogenetics*, **55**, 647–651.
- Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
- Wakenell, P. S., Miller, M. M., Goto, R. M., Gauderman, W. J. & Briles, W. E. (1996). *Immunogenetics*, **44**, 242–245.
- Wallny, H. J., Avila, D., Hunt, L. G., Powell, T. J., Riegert, P., Salomonsen, J., Skjødt, K., Vainio, O., Vilbois, F., Wiles, M. V. & Kaufman, J. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 1434–1439.
- Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.
- Zajonc, D. M., Striegl, H., Dascher, C. C. & Wilson, I. A. (2008). *Proc. Natl Acad. Sci. USA*, **46**, 17925–17930.
- Zoorob, R., Bernot, A., Renoir, D. M., Choukri, F. & Auffray, C. (1993). *Eur. J. Immunol.* **23**, 1139–1145.

Structure of a Classical MHC Class I Molecule That Binds “Non-Classical” Ligands

Chee Seng Hee^{1,9}, Song Gao^{2,3,9}, Bernhard Loll⁴, Marcia M. Miller⁵, Barbara Uchanska-Ziegler¹, Oliver Daumke^{2,6}, Andreas Ziegler^{1*}

1 Institut für Immunogenetik, Charité-Universitätsmedizin Berlin, Campus Benjamin Franklin, Freie Universität Berlin, Berlin, Germany, **2** Max-Delbrück-Centrum für Molekulare Medizin, Berlin, Germany, **3** Institut für Chemie und Biochemie, Freie Universität Berlin, Berlin, Germany, **4** Institut für Chemie und Biochemie, Abteilung Strukturbiochemie, Freie Universität Berlin, Berlin, Germany, **5** Department of Molecular and Cellular Biology, City of Hope, Beckman Research Institute, Duarte, California, United States of America, **6** Institut für Medizinische Physik und Biophysik, Charité-Universitätsmedizin Berlin, Berlin, Germany

Abstract

Chicken *YF1* genes share a close sequence relationship with classical MHC class I loci but map outside of the core MHC region. To obtain insights into their function, we determined the structure of the YF1*7.1/β₂-microglobulin complex by X-ray crystallography at 1.3 Å resolution. It exhibits the architecture typical of classical MHC class I molecules but possesses a hydrophobic binding groove that contains a non-peptidic ligand. This finding prompted us to reconstitute YF1*7.1 also with various self-lipids. Seven additional YF1*7.1 structures were solved, but only polyethyleneglycol molecules could be modeled into the electron density within the binding groove. However, an assessment of YF1*7.1 by native isoelectric focusing indicated that the molecules were also able to bind nonself-lipids. The ability of YF1*7.1 to interact with hydrophobic ligands is unprecedented among classical MHC class I proteins and might aid the chicken immune system to recognize a diverse ligand repertoire with a minimal number of MHC class I molecules.

Citation: Hee CS, Gao S, Loll B, Miller MM, Uchanska-Ziegler B, et al. (2010) Structure of a Classical MHC Class I Molecule That Binds “Non-Classical” Ligands. *PLoS Biol* 8(12): e1000557. doi:10.1371/journal.pbio.1000557

Academic Editor: Christopher Dascher, Mount Sinai School of Medicine, United States of America

Received May 24, 2010; **Accepted** October 27, 2010; **Published** December 7, 2010

Copyright: © 2010 Hee et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Berliner Krebsgesellschaft (CSH and BU-Z), Land Berlin (Elsa-Neumann-Stipendium des Landes Berlin, CSH), VolkswagenStiftung (stipend to CSH, grant I/79 989 to AZ), Deutsche Forschungsgemeinschaft (SFB 449/B6/B18 to AZ, BU-Z, GS and OD), Human Frontier Science Program (OD), National Science Foundation and United States Department of Agriculture competitive grant programs (MMM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: 3D, three-dimensional; β₂m, β₂-microglobulin; DOPC, dioleoylphosphatidylcholine; DSC, differential scanning calorimetry; HC, heavy chain; IEF, isoelectric focusing; L1 and L2, unidentified ligands; MHC, major histocompatibility complex; OLA, oleic acid; PC, phosphatidylcholine; PEG, polyethyleneglycol; PLM, palmitic acid; POPC, palmitoyloleoylphosphatidylcholine

* E-mail: andreas.ziegler@charite.de

These authors contributed equally to this work.

Introduction

Although the immune systems of birds differ in several important aspects from those of mammals, for example in relying on the bursa of Fabricius, and not on bone marrow, for the production of a diverse B cell repertoire [1], the presence of a major histocompatibility complex (MHC) is a unifying feature [2]. The MHC encodes several immunologically relevant proteins, among them classical MHC class I molecules that are membrane-anchored proteins involved in the presentation of foreign or self-protein-derived peptide antigens [3]. Conversely, the products of the evolutionarily distantly related non-classical class I genes (e.g. CD1 loci) can either display non-peptidic ligands such as lipids [4] or bind entire proteins [5]. In the chicken (*Gallus gallus domesticus*), MHC genes are located on the same micro-chromosomal arm in two regions termed *MHC-B* and *MHC-Y* that are physically, but not genetically, linked due to a chromosomal segment that supports a high degree of recombination between the two regions [6].

The *MHC-B* region resembles the mammalian MHC, e.g. with regard to its influence on the rapid rejection of transplants [7], but has been termed a “minimal essential MHC” due to its small size

[8]. It plays a prominent role in genetic resistance, particularly to virally induced tumors [9]. Very near to this region are the only two CD1 genes of the chicken [10,11]. The *MHC-Y* region, on the other hand, is thought to be associated with a moderate degree of allograft rejection [12] and to influence the fate of tumors induced by Rous Sarcoma virus [13]. It contains at least one polymorphic class I locus, *YF1*, which encodes a class I heavy chain (HC) that associates with β₂-microglobulin (β₂m) and is ubiquitously transcribed in both adult and embryonic chickens. The YF1 HC is closely related to that of classical MHC-B and mammalian MHC class I HC but not to non-classical CD1 HC (Figure 1A) [14]. To obtain insights into the role of YF1 molecules in the chicken immune system, we chose a structural approach.

Results/Discussion

Basic Structural Features of the YF1*7.1 Molecule

The complex of YF1*7.1 HC and β₂m was reconstituted without adding a ligand, and the structure was determined by molecular replacement at 1.32 Å resolution, using the related BF2*2101-β₂m complex [15] as a search model (Table 1, left column). The YF1*7.1 complex exhibits the typical architecture of

Author Summary

Proteins encoded by the major histocompatibility complex (MHC) play crucial roles in vertebrate immune systems, presenting pathogen-derived protein fragments to receptors on effector cells. In contrast, some non-classical MHC class I proteins such as CD1 molecules possess a hydrophobic groove that allows them to display lipids. Chicken *MHC-Y* is a genetic region outside the core *MHC* that harbors several immune-related genes, among them YF1*7.1, which encodes a protein whose structure we solved in this study. YF1*7.1 is an MHC class I molecule that exhibits the architecture typical of classical MHC class I antigens but possesses a hydrophobic binding groove that binds non-peptidic ligands. By using lipid-binding assays, we show that this molecule can indeed bind lipids. Therefore, YF1*7.1 bridges, at least in structural terms, the traditional gap between classical and non-classical MHC class I molecules. Lipid-binding YF1 proteins might serve the chicken to enlarge its otherwise very small repertoire of antigen-presenting MHC class I molecules. Furthermore, comparative analyses of the two protein subunits of classical MHC molecules revealed a structural feature in chickens that appears to be shared by non-mammalian but not by mammalian vertebrates. This unique feature is indicative of a structure-dependent co-evolution of two genetically unlinked genes in non-mammalian species.

classical MHC class I molecules [3], with binding groove-forming, anti-parallel α 1- and α 2-helices atop of a β -sheet platform. β_2 m and the α 3-domain occupy the standard positions below the platform (Figure 1B). However, the YF1*7.1 binding cleft is narrower than that of peptide-presenting MHC class I molecules (Figure 2) and is lined by many hydrophobic residues (16 out of 30 residues forming the binding groove) (Figure S1, Table S1). Charged residues are only found above the floor and at the ends of the groove (Figures 2B, 3A). The volume of the YF1*7.1 binding groove is $\sim 1,030 \text{ \AA}^3$. This value is considerably smaller than that of typical MHC class I peptide binding grooves ($\sim 1,250$ – $1,900 \text{ \AA}^3$, see e.g. Protein Data Bank entries 1I4F and 1OF2), mammalian CD1 molecules ($\sim 1,800$ – $2,400 \text{ \AA}^3$, 2PO6 and 2H26), or chicken CD1-1 ($1,810 \text{ \AA}^3$, 3JVG) (Figure 2B). The YF1*7.1 groove is, however, larger than the miniaturized binding pocket of chicken CD1-2 ($\sim 720 \text{ \AA}^3$, 3DBX) (Figure 2B), which is thought to accommodate maximally a single alkyl chain [16].

These comparisons and its hydrophobic character indicated that the YF1*7.1 binding cleft is optimized for the presentation of medium-sized, non-peptidic ligands rather than peptides, despite the overall similarity to classical BF2 molecules of the chicken. This assumption is reinforced by the substitution of Arg9 (in BF2*2101) by Leu9 (in YF1*7.1) on the floor of the binding groove (Figures 2B, S1). Arg9 can assume different conformations that permit a promiscuous anchoring of sequence-unrelated peptides by this dominantly expressed MHC-B class I molecule [15], thus expanding the repertoire of bound peptides. In YF1*7.1, however, the homologous Leu9 residue cannot serve this purpose but contributes instead to the hydrophobic environment of the groove. Another remarkable feature of YF1*7.1 are the bridge-like contacts between several α 1- and α 2-helical residues that extend over the top of the groove, leaving only its central portion directly accessible to a ligand (Figure 3A). These interactions distinguish YF1*7.1 from BF2*2101 [15] as well as from most [3–5] but not all [17] mammalian class I molecules.

Table 1. Data collection and refinement statistics.

	YF1*7.1:L1	YF1*7.1:L2
Data collection		
Space group	<i>P</i> 12 ₁ 1	<i>P</i> 12 ₁ 1
Cell dimensions		
a, b, c (Å)	52.80, 55.47, 63.84	52.87, 55.04, 63.59
α , β , γ (°)	90.00, 96.85, 90.00	90.00, 97.04, 90.00
Resolution (Å)	20–1.32 (1.35–1.32)*	20–1.60 (1.64–1.60)
<i>R</i> _{merge} (%)	4.4 (46.1)	4.1 (29.4)
<i>I</i> / σ <i>I</i>	20.2 (3.0)	17.4 (3.8)
Completeness (%)	96.0 (92.8)	89.8 (93.2)
Redundancy	4.2 (3.6)	2.8 (2.6)
Unique reflections	82,457	43,084
Refinement		
<i>R</i> _{work} / <i>R</i> _{free} (%)	15.7/19.0	17.4/21.6
Number of atoms	3,693	3,498
Protein	3,283	3,062
Water	382	386
Ligand	20	38
Other	8	12
B-factors (Å ²)		
Overall	17.7	17.9
Protein	16.4	16.8
Water	28.1	25.8
Ligand	18.5	26.8
Other	28.6	24.3
R.m.s. deviation		
bond length (Å)	0.013	0.012
bond angle (°)	1.527	1.403

*Value in parentheses represents statistics for data in the highest resolution shell. R.m.s., root mean square.

doi:10.1371/journal.pbio.1000557.t001

Several residues belonging to the end of the α 1-helix and the beginning of the α 2-helix, i.e. “above” the F pocket of classical MHC class I molecules, are characterized by double conformations (Figure 3B). This suggests the presence of conformational dynamics that may aid in binding structurally distinct ligands to the YF1*7.1 binding groove. The fact that the positive charges of Arg82 and Arg142 are compensated by binding an acetate molecule derived from the crystallization solution (Figure 3) suggests, in addition, that a YF1*7.1 ligand might interact with these two residues.

Possible Structural Consequences of Allelic Variations

Although the sequence of the first 27 amino acids has not been determined for other YF1 alleles (Figure S1), the available information permits us to predict that several exchanges might have an impact on the shape and the electrostatic properties of the binding groove (Figure 4). By modeling the altered residues onto the YF1*7.1 structure, the three exchanges between YF1*7.1 and YF1*15 (Asn75Gly, Met92Leu, and Phe119Tyr) will probably lead to an enlargement of the binding groove ($\sim 1,070 \text{ \AA}^3$ versus $\sim 1,030 \text{ \AA}^3$), predominantly in the region of the F pocket. In contrast, YF1*16 not only possesses three replacements involving the same residues as in case of YF1*15 but also exhibits three additional exchanges (Arg82Cys, Met94Arg, and Phe96Ile).

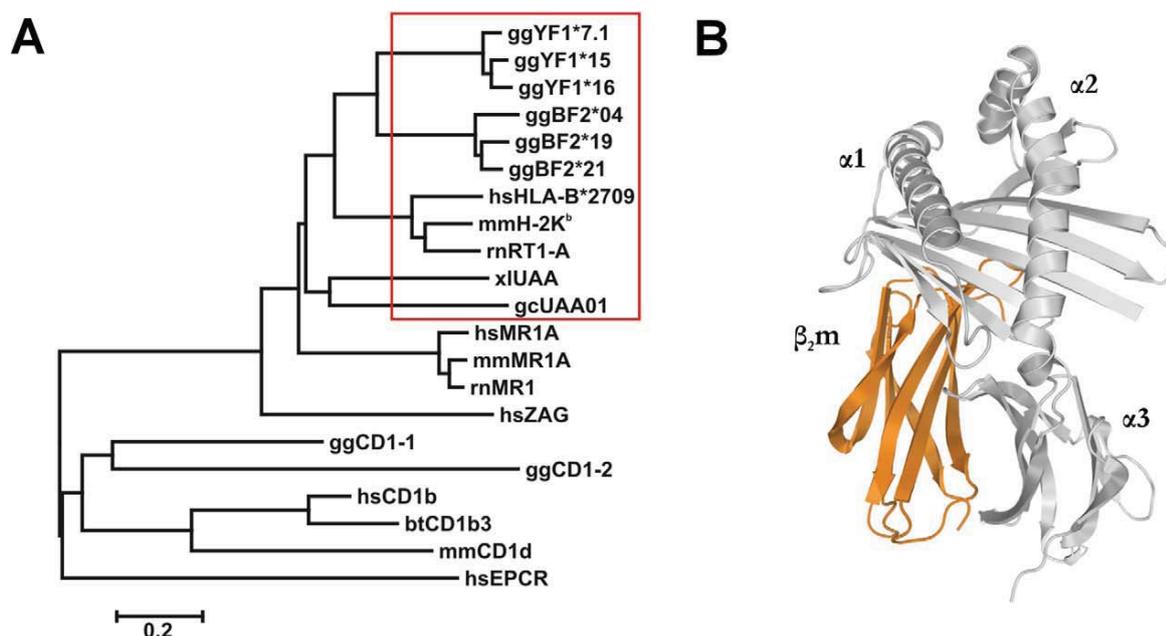


Figure 1. Evolutionary and structural characteristics of YF1*7.1. (A) The evolutionary tree reveals that YF1 isoforms are closely related to chicken MHC-BF2 variants and classical mammalian as well as non-mammalian (frog, nurse shark) class I heavy chains (red box) but are also similar to mammalian MR1 chains and human ZAG. YF1 heavy chains are, however, only distantly related to chicken and mammalian CD1 molecules as well as to EPCR. The designations of the molecules are given in the Accession Numbers section. The tree is drawn to scale, with branch lengths equivalent to evolutionary distances in the units of the number of amino acid substitutions per site. (B) Ribbon diagram of YF1*7.1 non-covalently associated with β_2m (orange), as seen along the binding groove. A ligand has been omitted for clarity. doi:10.1371/journal.pbio.1000557.g001

YF1*16 is thus characterized by pronounced alterations of the groove: the novel Phe at HC position 75 is expected to intersect the cleft, thereby separating the A pocket from the F pocket. This bears some resemblance to amino acid exchanges in mammalian CD1 molecules, where the long T' tunnel of human CD1b is blocked in mouse CD1d molecules due to the replacement of two Gly residues by Leu and Val, respectively [18]. This alteration has consequences for the type of ligands that can be bound by the two CD1 molecules (reviewed in [4]). The volumes of the remaining A ($\sim 490 \text{ \AA}^3$) and F ($\sim 430 \text{ \AA}^3$) pockets of YF1*16 will most likely lead to exposure of the middle section of ligands with two hydrophobic segments, as e.g. in case of phosphatidylcholines (PC) [4]. In addition, the Arg82Cys and the Met94Arg exchanges can be expected to alter the electrostatic properties of the groove. In particular, the novel Arg94 residue at the floor of the binding cleft of YF1*16 might allow the interaction with ligands possessing acidic groups. These allele-specific changes are reminiscent of sequence-dependent alterations that give particular classical MHC class I molecules the opportunity to bind defined sets of ligands [3]. In contrast, the binding grooves of avian and mammalian CD1 molecules do not exhibit such polymorphisms [4,16,19–21]. These comparisons indicate that, other than in the case of mammalian species, where dissimilar non-polymorphic CD1 *genes* with distinct binding grooves serve to enlarge the repertoire of displayed lipids [4], YF1 *alleles* might be responsible for differential interaction with ligands.

A Structural Peculiarity Characterizes Chicken Classical Class I Molecules

A further distinct structural feature that YF1*7.1 shares with chicken BF2*2101 molecules [15], but not with classical or non-

classical MHC class I molecules from human, rhesus macaque, mouse, rat, cattle, as well as chicken CD1 molecules, is the particular location and conformation of the HC loop 1 (Loop1) (Figures 2A, 5A). This is due to a salt bridge that is formed between Loop1 (residue Asp14 of YF1*7.1 and BF2*2101 molecules) and β_2m (Lys34). Instead, Arg14 of classical mammalian MHC class I HC contacts Asp39 within HC loop 2 (Loop2) via a salt bridge. Sequence comparisons suggest that the existence of the Loop1- β_2m contact is probably present also in classical MHC class I molecules from other birds, certain amphibians, and possibly reptiles (Figure 5B). In contrast, the Loop1-Loop2 contact is expected to be restricted to mammals, including egg-laying mammals (monotremes) such as echidna and platypus, indicating that the replacement of the intermolecular contact found in non-mammalian vertebrates by an intramolecular salt bridge is likely to have preceded the development of monotremes, about 170 million years ago [22]. On the other hand, the Loop1- β_2m interaction probably constitutes an example of structure-dependent co-evolution between two genetically unlinked genes (classical class I HC and β_2m). Non-classical class I molecules (e.g. CD1 molecules from mammals and chicken, endothelial protein C receptor (EPCR) [23], or Zn- $\alpha 2$ -glycoprotein (ZAG) [24]) lack both the Loop1-Loop2 and the Loop1- β_2m contacts (see also Figures 2A, S1 and interactive Figure S2). This is in line with their evolutionary history, which suggests an early separation of the lineages leading to CD1 and EPCR, on one hand, and to ZAG as well as classical MHC class I molecules, on the other, about 300 million years ago (Figure 1A) [16,19,25]. Although no structure has yet been determined for mammalian MHC class I-related (MR)1 molecules [26], their predicted Loop1 and Loop2 do not appear to be connected (Asp14, Val39; Figure S1). Likewise, a salt

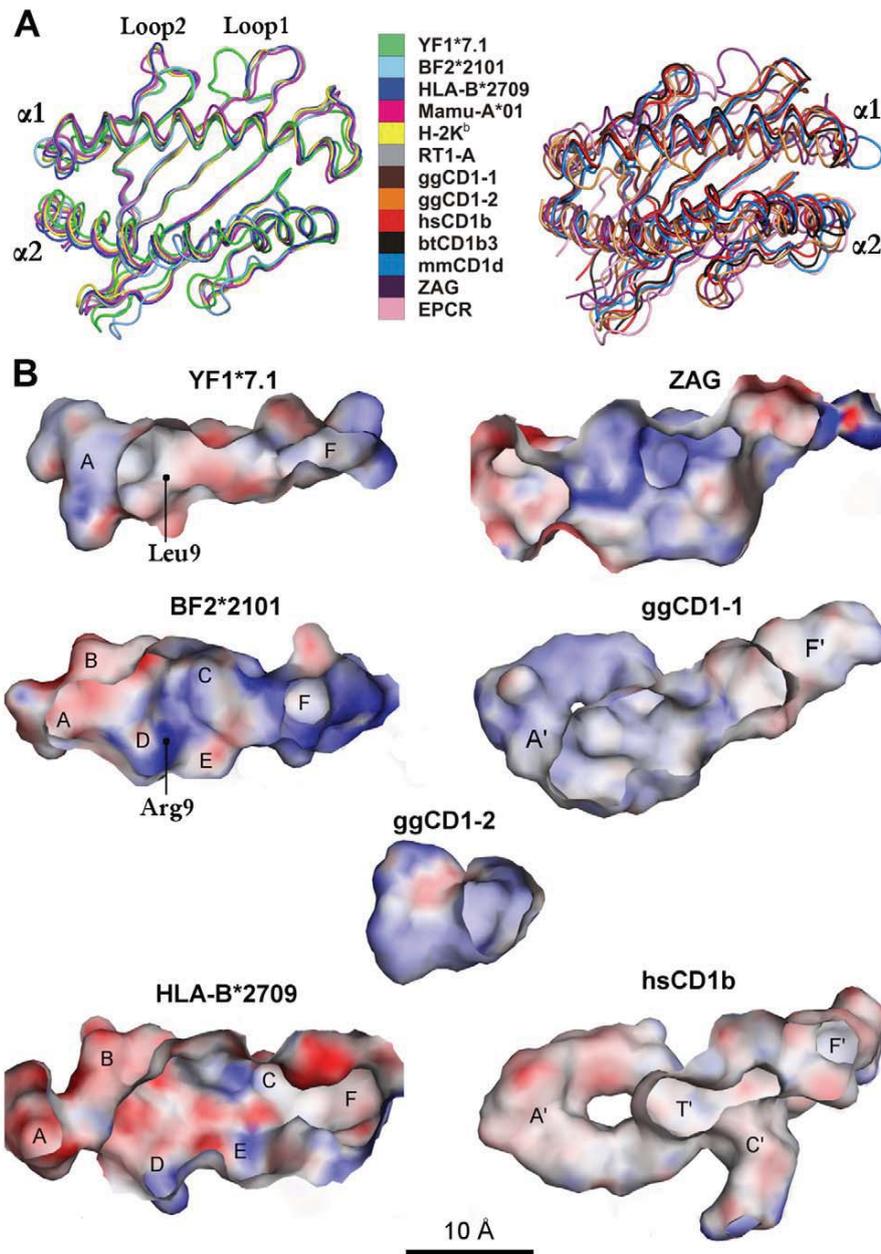


Figure 2. Binding grooves of YF1*7.1 and selected classical or non-classical class I molecules. (A) Overlay of $\alpha 1$ - and $\alpha 2$ -domains, viewed from above. Classical (YF1*7.1, BF2*2101, HLA-B*2709, Mamu-A*01, H-2K^b, RT1-A, left) and non-classical class I molecules (all others, right) (see Materials and Methods for details) are superimposed onto the α -backbone of the $\alpha 1$ -helix and the β -sheet platform, with selected interstrand loops (Loop1 and Loop2) designated. The Loop1 locations of YF1*7.1 and BF2*2101 are nearly indistinguishable but are distinct from those of classical mammalian MHC class I molecules. An interactive three-dimensional (3D) comparison of these molecules is available in Figure S2. (B) Interior molecular surfaces of ligand-devoid binding grooves. The binding pockets of classical (A–F) and non-classical (A', C', F', T') molecules are indicated. The approximate position of HC residues 9 (Leu in YF1*7.1, Arg in BF2*2101) is indicated (see main text for further explanation). Electrostatic potentials are mapped to the molecular surfaces with positive potential (≥ 20 mV) in blue, neutral potential (0 mV) in white, and negative potential (≤ -40 mV) in red. Although the ZAG groove is predicted to bind hydrophobic ligands [24,31] like CD1 molecules, it appears closely related to that of YF1*7.1 (see also Table S1).
 doi:10.1371/journal.pbio.1000557.g002

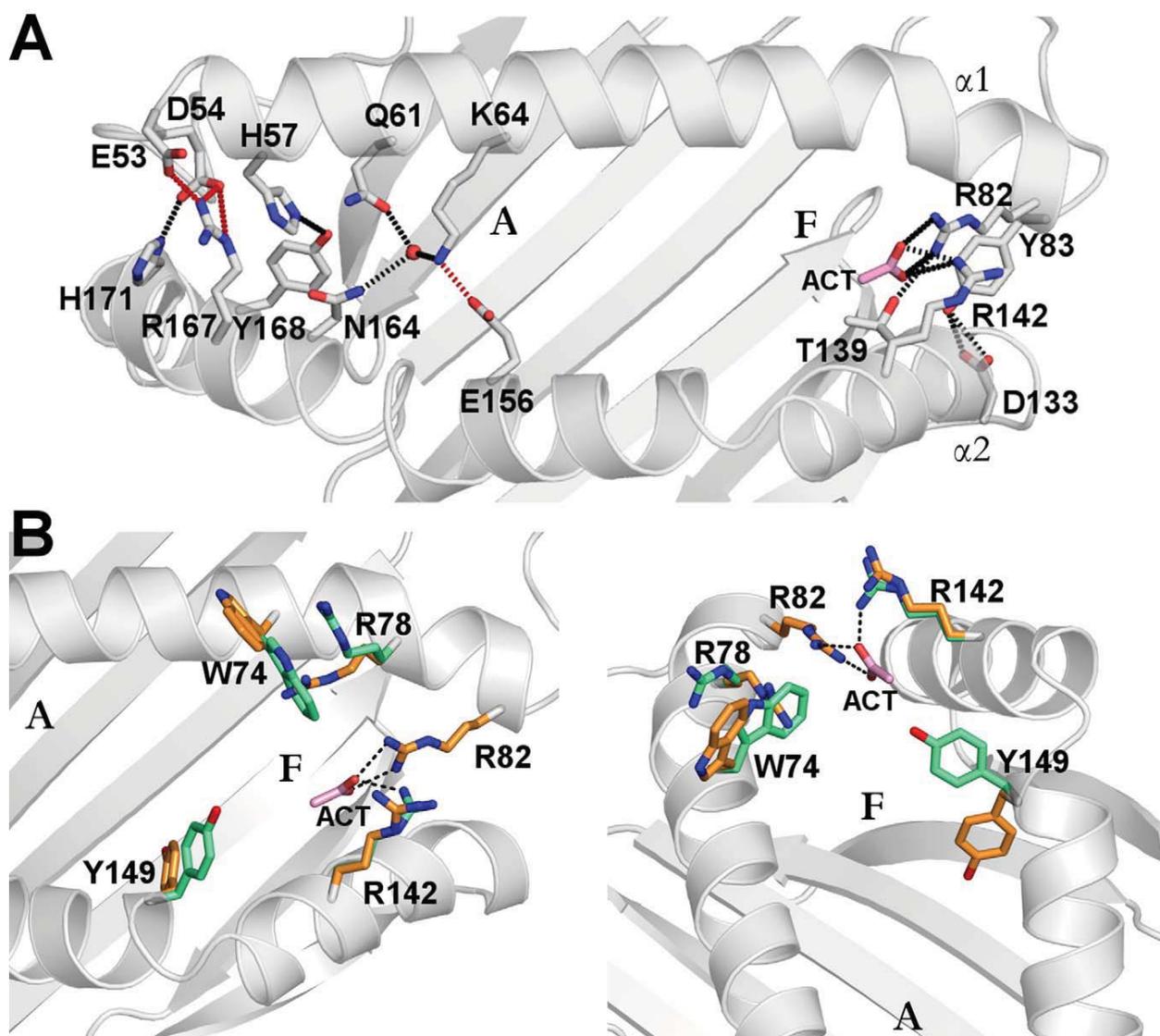


Figure 3. Side chain interactions in the vicinity of the YF1*7.1 binding groove. (A) Side chain interactions between α 1- and α 2-helices partially cover the A pocket and close the binding groove terminals. Side chains are shown as stick representation with hydrogen bonds and salt bridges indicated by dotted lines. An acetate molecule (ACT) “above” the F pocket is shown as pink stick representation. (B) Two views of the F pocket showing residues that might be involved in ligand binding due to their conformational dynamics. On the left is the view from “above” the binding groove and, on the right, the view from the A pocket along the binding groove towards the F pocket. Residues exhibiting dual conformations are distinguished by orange and green colors. The side chains of Trp74, Arg78, Arg142, and Tyr149 “above” the F pocket have poorly defined electron density compared to the surrounding residues, indicating that they might interact with ligands captured within the F pocket. Also near the F pocket, Arg82 interacts with an acetate molecule, indicating that this HC residue might also be involved in ligand binding.
doi:10.1371/journal.pbio.1000557.g003

bridge-mediated interaction between Loop1 and mammalian β_2m (Asp, His or Asn34; Figure 5B) is also not obvious.

Thermodynamic Behavior of the YF1*7.1 Complex

The thermodynamic properties of MHC class I molecules are crucially influenced by the presence of ligands within the binding groove [27]. Therefore, we sought to gain insight into the stability of the YF1*7.1 complex using differential scanning calorimetry (DSC). As a consequence of a higher degree of inter-experimental variability than in case of peptide-binding MHC class I complexes

(e.g. HLA-B27 molecules) [28], the thermodynamic behavior of the complex could not be determined reliably. We observed, however, that the “melting” temperature of the YF1*7.1 complex was considerably decreased in comparison with typical peptide-presenting mammalian MHC class I molecules (Figure S3), since its dissociation began already at $\sim 40^\circ\text{C}$. This value is lower than the body temperature of a chicken ($\sim 41.8^\circ\text{C}$) [29], indicating that the YF1*7.1 complex exhibits only a limited degree of structural integrity and might thus be prone to interaction with a ligand that could confer an improved stability *in vivo*.

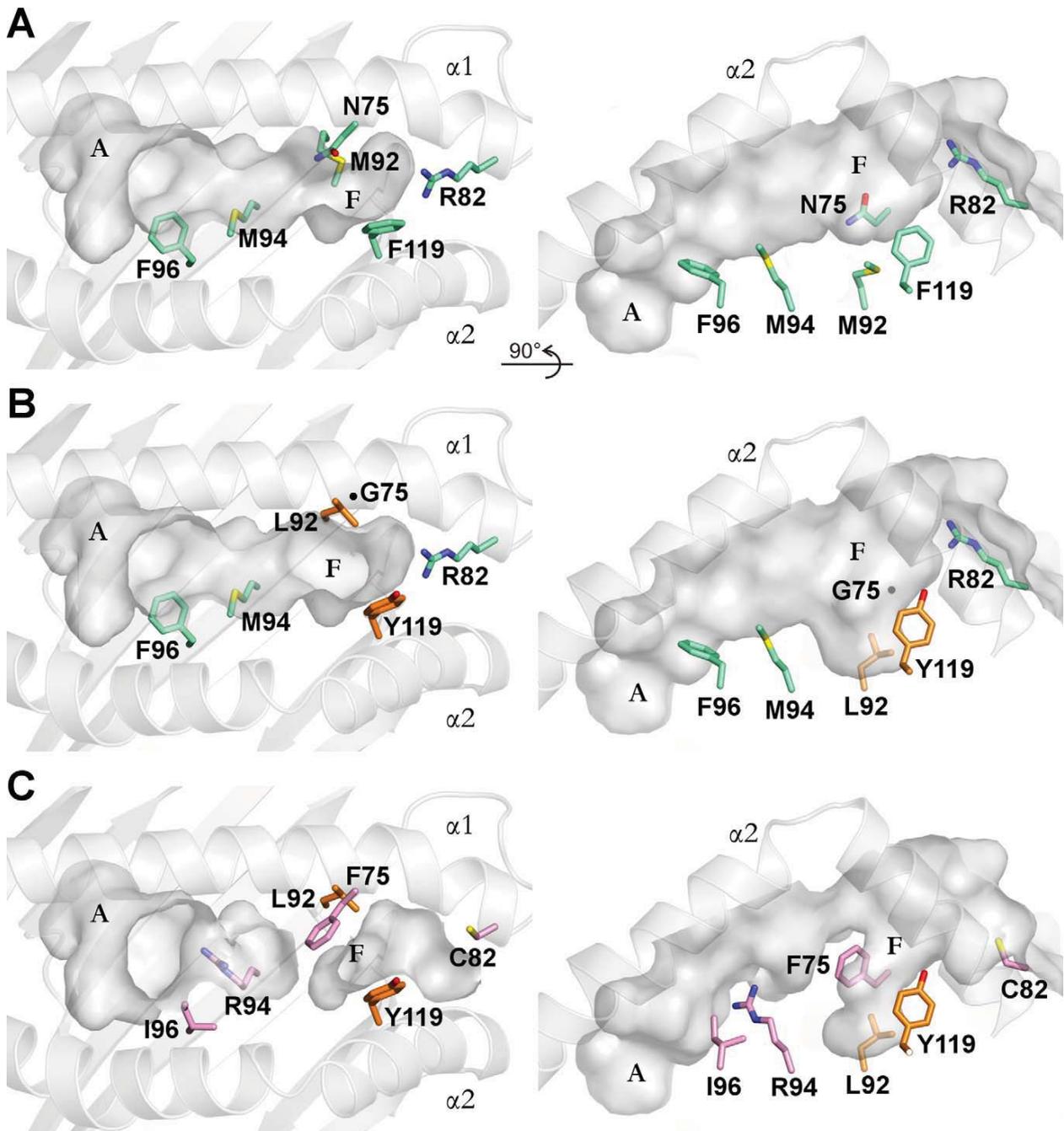


Figure 4. Polymorphic residues of YF1 alleles within the binding groove. Polymorphic residues of YF1*15 and YF1*16 are “mutated” in silico using the YF1*7.1 structure as model. Two views are shown for each allele: on the left, views from “above” the binding groove, and on the right, views “through” the α 2-helix. The A and F pockets as well as the α 1- and α 2-helices are labeled accordingly. (A) Six polymorphic residues that influence the binding groove architectures are shown in green stick representation in YF1*7.1. (B) Substitutions of Asn75Gly and Met92Leu (orange stick representation) in YF1*15 result in a wider groove entrance and deeper F pocket, respectively. The Phe119Tyr exchange slightly narrows the F pocket. The position of Gly75 is shown as a black dot. (C) Six substitutions in YF1*16 result in a division of the binding groove into two parts. The Phe96Ile exchange in YF1*16 slightly enlarges the A pocket, while the substitutions Asn75Phe and Met94Arg are expected to disrupt the middle part of the binding groove. The Asn75Phe exchange also narrows the F pocket together with the Phe119Tyr substitution. The Arg82Cys substitution leaves the F terminal part of the binding groove open, and the Met92Leu substitution results in a deeper F pocket similar to that observed in the case of YF1*15.

doi:10.1371/journal.pbio.1000557.g004

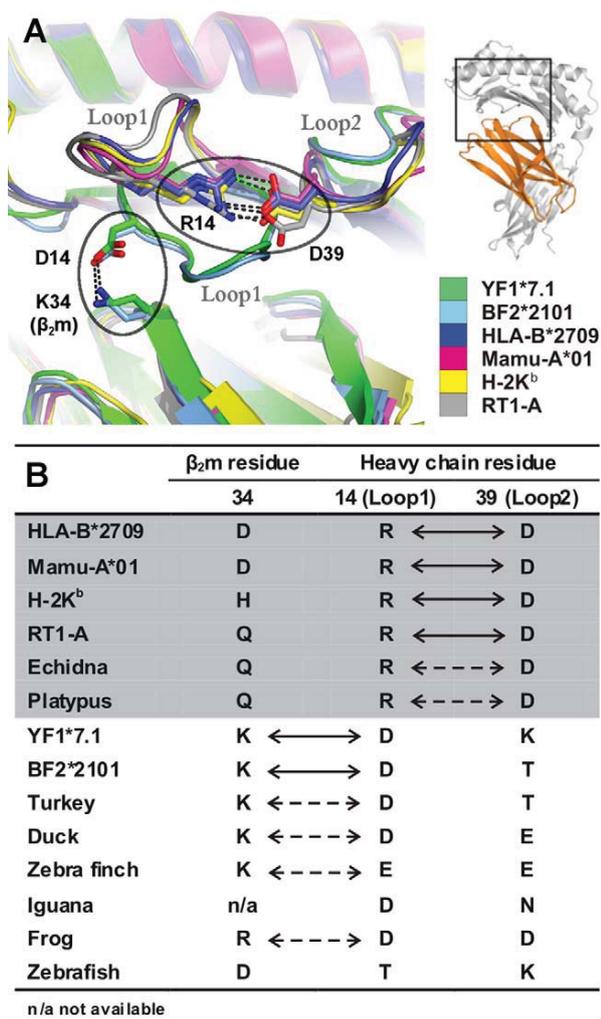


Figure 5. Distinct loop interactions in classical mammalian and chicken MHC class I molecules. (A) Overlay of chicken YF1*7.1 and selected class I molecules reveal that the Loop1 conformations in the YF1*7.1 and BF2*2101 molecules of the chicken deviate from those of mammalian classical class I antigens. This is due to different contacts made by residue 14 of the HC: Asp14 of YF1*7.1 and BF2*2101 interact with Lys34 of β_2m , whereas Arg14 of HLA-B*2709, Mamu-A*01 (rhesus macaque), H-2K^b, and RT1-A contact Asp39 of Loop2 (regions of interest indicated by ellipses). Salt bridges are indicated by black dotted lines. The location of the enlarged area within YF1*7.1 is shown on the right, together with a color legend. Carboxyl group atoms of Asp residues and nitrogen side chain atoms of Lys and Arg are colored in red and blue, respectively. (B) Summary of residues involved in the Loop1 - β_2m or Loop1-Loop2 interactions in various species. Contacts supported by molecular structures are indicated by arrows, and suggested interactions are shown by dotted arrows. The area shaded in grey indicates placental (human, rhesus macaque, mouse, rat) and egg-laying mammals (echidna, platypus).
doi:10.1371/journal.pbio.1000557.g005

Non-Peptidic Ligands in the YF1*7.1 Binding Groove

Due to the high-resolution density map of the YF1*7.1 structure, we were also able to model a ligand (L1) with a linear chain of 17 atoms and a tetragonal head group in the groove (Figure 6A). This ligand is buried within the binding groove, with the end of its tail (~6 backbone atoms) inserted deeply into the hydrophobic A pocket and a tetragonal head group located in the

middle of the groove entrance (Figure 6A). The binding of this ligand resembles the situation observed for CD1 molecules of the chicken, bovine, mouse, and human, where unidentified hydrophobic molecules, some of considerable length, have also been found (see e.g. interactive Figure S2, “UL-ggCD1-1” [19], “palmitate-ggCD1-2” [16], and “spacer-hsCD1b” [20]). These unidentified ligands are thought to stabilize CD1 molecules and may facilitate the interchange with other ligands both inside and outside of the cell [4,19,20]. Although we suspected the YF1*7.1 ligand to be cetrimonium (hexadecyltrimethylammonium), an antiseptic cationic surfactant, various approaches including mass spectrometric analyses did not allow us to identify L1 unambiguously.

Nevertheless, this detection of a hydrophobic molecule bound to YF1*7.1 as well as the hydrophobic character of the narrow binding groove suggest that the natural ligands for this protein might be lipids and not peptides. We therefore attempted to form complexes with lipids that are typical cellular components and that YF1*7.1 might encounter within the endoplasmic reticulum during or after assembly. These included palmitoylcholine (POPC), dioleoylphosphatidylcholine (DOPC), palmitic acid (PLM), oleic acid (OLA), and PC. Following reconstitution and chromatographic analyses, we obtained several YF1*7.1 complexes for crystallization trials. Polyethyleneglycol (PEG) 4000 was always necessary to obtain crystals, but we used either glycerol or PEG 200 as cryoprotectants. For nearly all complexes, diffraction data sets at resolutions of 1.5–1.7 Å were obtained (Table S2). The refined structures reveal that reconstitution of YF1*7.1 in the presence of these lipids does not result in additional electron density within the binding groove that can unequivocally be assigned to any of these ligands, suggesting that the examined lipids are not bound by this protein.

In contrast, the use of PEG 200 for cryoprotection (Tables 1, S2) invariably leads to the presence of two additional stretches of electron density within the binding groove, which can best be modeled as two PEG 200 molecules, in addition to a longer fragment exhibiting similarity to the unidentified ligand of YF1*7.1:L1 (YF1*7.1:L2; Figure 6B). In the absence of another plausible explanation, we modeled a fragment of PEG 4000 into this electron density, with the hydrophobic terminus extending into the depth of the A pocket. The identification of the two short ligands as PEG 200 is supported by the characteristic horseshoe-like shape [30,31], which is particularly obvious in case of the molecule within the F pocket. The conformation of this ligand is maintained by multiple predicted contacts to Asn75. In turn, these lead to a reorientation of the larger ligand, which can now be contacted directly by Tyr112. The second small ligand is located above the larger molecule but is still buried completely within the YF1*7.1 complex. No detectable electron density towards the F pocket is present in the YF1*7.1:L1 structure, suggesting that disordered water molecules occupy this section of the binding groove. The bound hydrophobic chains indicate that natural ligand(s) might occupy comparable positions.

A schematic representation of the types of natural ligands that could possibly bind to the YF1*7.1 groove is provided in Figure S4. A ligand with a long, hydrophobic tail would fit ideally into the highly hydrophobic environment of the A pocket (Figure S4A). Further hydrophobic ligands as seen in the YF1*7.1:L2 structure might bind in addition, e.g. to the F pocket (Figure S4B). However, this would leave Arg82 and Arg142 without charge compensation. Our detection of an acetate molecule that is bound to these residues provides evidence that such salt bridge-mediated interactions could be favored (Figure S4A,B), indicating that a negatively charged head group of a lipid might be accommodated

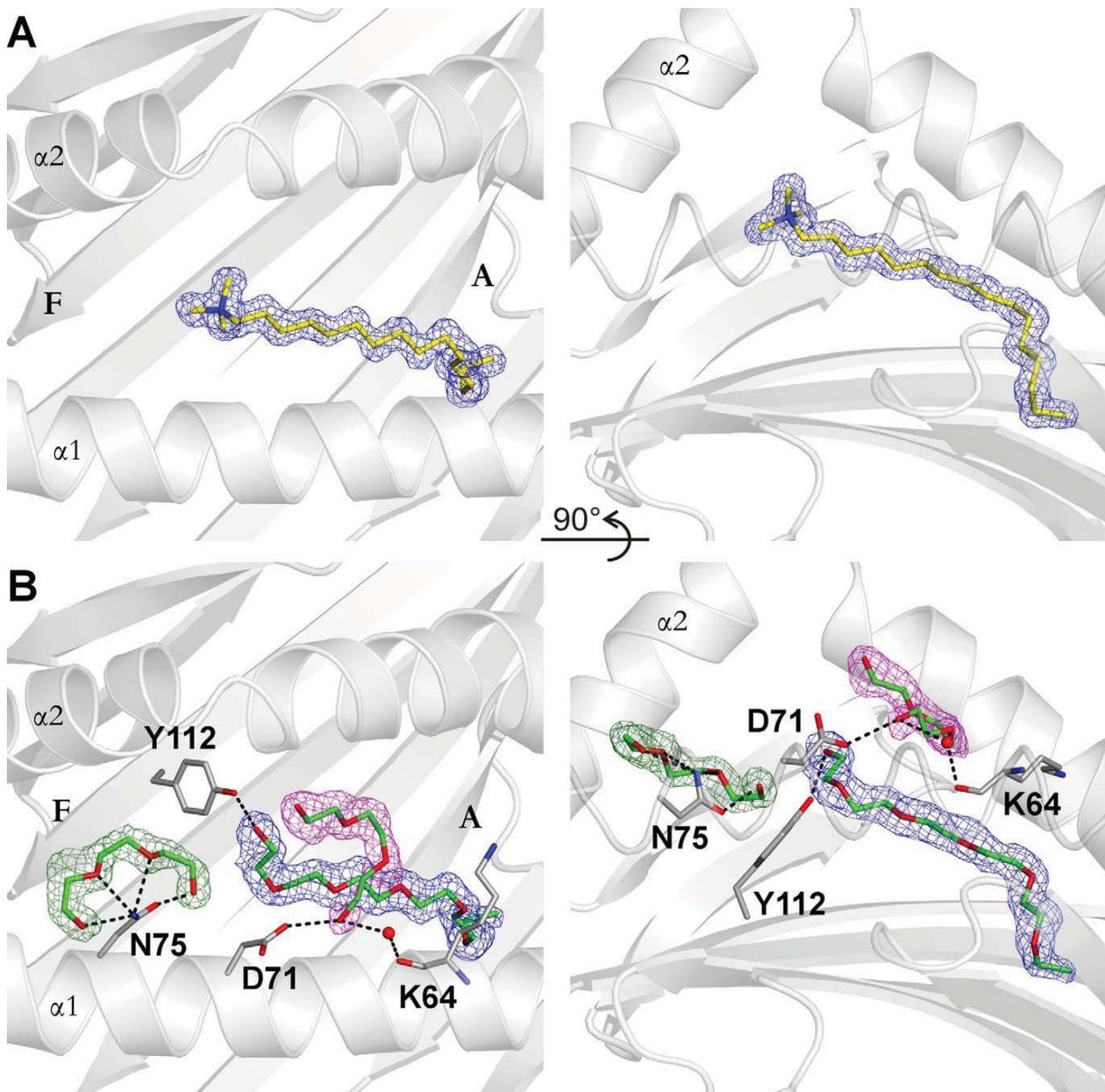


Figure 6. Electron densities observed in YF1*7.1 structures. The electron densities derived from $2F_o - F_c$ maps after refinement are shown as blue, magenta, and green meshes with a contour level of 1σ . Two different types of electron densities (resembling those depicted in A or B) can be observed in eight data sets collected under different cryo-conditions. Side chains interacting with ligands are shown as grey stick representation, with oxygen and nitrogen atoms indicated with red and blue color, respectively. Hydrogen bonds are shown as black dashed lines. Two views are displayed for each type of electron density: on the left, views from "above" the binding groove, and on the right, views from the $\alpha 1$ -helix (visible as loop in the foreground). (A) Electron density of the YF1*7.1 complex without an added ligand cryo-protected with glycerol (YF1*7.1:L1). (B) Electron density of the YF1*7.1:L2 complex, using PEG 200 for cryoprotection. Three $\alpha 1$ -helical residues are involved in indirect (Lys64) or direct (Asp71, Asn75) contacts to the PEG 200 molecules.
doi:10.1371/journal.pbio.1000557.g006

in or "above" the F pocket (Figure S4C). Comparable charge compensatory interactions are occasionally observed between negatively charged lipidic head groups and positively charged membrane protein residues such as arginine and lysine [32,33], and contacts involving the side chain of $\alpha 1$ -helical arginine residues (Arg73, Arg79) can also be found in case of CD1:ligand

interactions [21,34–36]. We consider it unlikely that lipids would bind with hydrophobic portions into the F pocket, unless a negatively charged head group of a lipid would be positioned above the groove and permit Arg78, Arg82, or Arg142 to form contacts with the ligand (Figure S4D). The conformational flexibility of arginine residues might facilitate such interactions.

The existing ambiguities in modeling ligands into the observed electron density and in identifying them are likely to be also a reflection of considerable dynamics exhibited by YF1*7.1-bound molecules. Wang and co-workers detected a comparable phenomenon in the case of two lipids bound to mouse CD1d molecules [21]. The difficulties in assigning ligands for YF1*7.1 molecules may be compared to those encountered by Bjorkman and co-workers with the human ZAG protein [31]. Despite extensive attempts and the prediction that the ZAG binding groove might accommodate hydrophobic molecules (compare Table S1), no natural ligand for the binding cleft of this protein has so far been identified. Although ZAG and YF1*7.1 are evolutionarily related (Figure 1A), there are also pronounced differences between them: ZAG lacks an association with β_2m and instead interacts with prolactin-inducible protein [37], possesses a larger binding groove (Figure 2B), is non-polymorphic, and exists as a secreted molecule [24,31]. The closest mammalian relatives of YF1*7.1 molecules could be the MR1 antigens of human, mouse, and rat (Figure 1A). Although there is evidence that these molecules might bind lipids and are recognized by a specialized subpopulation of T cells [26,38,39], it is currently difficult to judge to what extent these comparisons are valid, since structures for MR1 antigens have so far not been reported.

Ligand Search for YF1*7.1 Molecules

Since our structural studies did not yield support for an endogenous ligand within the YF1*7.1 binding groove, we attempted to bind also a number of nonself lipids to this molecule, employing native isoelectric focusing (IEF). Instead of reconstituting YF1*7.1 with various lipid preparations, we incubated these potential ligands with the reconstituted, purified HC: β_2m complex and performed IEF (Figure 7). The unloaded YF1*7.1 complex migrates as one major species and four minor components of which one appears to be due to free β_2m . Incubation with oleic acid did not alter this pattern, but incubation with a mixture of lipopolysaccharides from *Salmonella enterica* or *Escherichia coli* revealed a novel band with a pI of ~ 5.6 . In addition, a mycolic

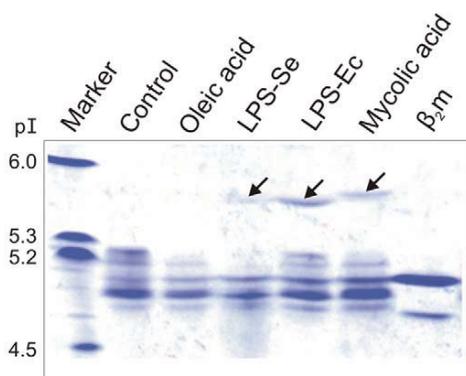


Figure 7. Isoelectric focusing analysis of YF1*7.1 complexes incubated with lipid preparations. The pI of marker protein (lane left) is indicated. The analysis comprised (I) YF1*7.1 without added lipid, (II) YF1*7.1 with oleic acid, (III) YF1*7.1 with lipopolysaccharide from *S. enterica*, (IV) YF1*7.1 with lipopolysaccharide from *E. coli*, (V) YF1*7.1 with mycolic acid, and (VI) monomeric β_2m . The arrows indicate the positions of novel bands obtained following incubation of YF1*7.1 complexes with selected lipid preparations.
doi:10.1371/journal.pbio.1000557.g007

acid preparation from *Mycobacterium tuberculosis* yielded a novel species with a pI of ~ 5.7 . These results indicate that in some of the YF1*7.1 molecules, a charged species has replaced the molecule(s) that may have been loaded into the binding groove. Similarly, as in the case of CD1 molecules, charged lipids do not always bind to all proteins within the preparation, leading to a pI shift only in a subpopulation of the molecules [20,40].

Conclusions

The results presented here attest to the unusual versatility of classical MHC class I antigens. Not only do they display peptides [3,15,28] or post-translationally modified protein fragments [41,42], but they may also bind hydrophobic ligands, as shown here. The prerequisites for the latter characteristic are so far met only by the YF1*7.1 molecule, which can be regarded as a “hybrid” structure, favoring the interaction with “non-classical” ligands through the combination of a hydrophobic binding groove with a classical scaffold. As there is only a single fully functional, polymorphic classical MHC class I gene, BF2, in most MHC-B haplotypes [43,44], and just one non-polymorphic CD1 molecule is available to display complex lipids (CD1-1) [19], the chicken’s antigen-presenting capabilities might be limited. Together with the unusual, massive expansion of highly polymorphic immunoglobulin-like loci (“CHIR”) within the leukocyte receptor complex [45,46], lipid-binding YF1 molecules might be part of a strategy to overcome inadequacies in the repertoire of displayed ligands and thus improve the ability of this species to fight successfully against infections [15,16,19,44,45].

The analysis of MHC-encoded class I molecules of the chicken lags far behind that of mammals, so that many questions, e.g. with regard to the nature of the cellular interaction partners of YF1 and CD1 molecules, are currently unsolved. However, our analyses demonstrate that YF1 molecules deserve to be studied in more detail, because they bridge, at least in structural terms, the traditional gap between peptide-presenting classical and lipid-displaying non-classical class I molecules.

Materials and Methods

Protein Purification and Crystallization

Procedures for protein preparation and crystallization of YF1*7.1 HC and β_2m have been reported previously [47]. A similar procedure was applied to produce the other YF1*7.1 complexes. In the reconstitution experiments, however, the respective lipid was added in 10-fold molar excess to YF1*7.1 HC in a buffer pre-warmed to 37°C, followed by incubation at room temperature for 2 d. All lipids were obtained from Sigma-Aldrich with the following product numbers: palmitic acid (PLM), P5585; oleic acid (OLA), O1008; phosphatidylcholine (PC), P2772; dioleoylphosphatidylcholine (DOPC), P6354; and palmitoyloleoylphosphatidylcholine (POPC), P3017. Lipids were dissolved in dimethylsulfoxide (DMSO) and the concentration adjusted to 10 mg/ml and preheated to $\sim 40^\circ\text{C}$ prior to adding them to the reconstitution experiments of YF1*7.1 HC and β_2m . Protein complexes were purified and crystallization experiments were performed as described before [47]. Crystals were cryoprotected with the respective reservoir solution supplemented with either 19% (v/v) glycerol or 19% (v/v) PEG 200.

Data Collection and Structure Determination

X-ray diffraction data sets were collected at BESSY II, Berlin, Germany, at beamline 14.1 or 14.2. Structure solution of the YF1*7.1 complex has been described by us [47]. Molecular

replacement for the other structures was performed by employing a search model of ligand- and water-depleted YF1*7.1:L1. Crystals of the YF1*7.1:L2 complex were isomorphous to the initially determined YF1*7.1:L1 complex. Restrained and TLS refinement with 3 TLS groups designated for α 1- α 2 domain, α 3 domain, and β _{2m} were then carried out with Refmac5 [48] and the model building was performed with COOT [49]. The refined models show excellent steric and geometric quality and have no residue in the disallowed region of the Ramachandran plot, as assessed with MolProbity [50].

Structure Presentation and Computational Analyses

Figures depicting structures were prepared with PyMOL [51]. In silico mutagenesis for YF1*15 and YF1*16 alleles (Figure 4) was performed using PyMOL. The polymorphic residues were substituted in the YF1*7.1 structure and the side chain conformations were chosen based on the frequencies calculated by the program, taking into consideration that the side chains do not clash and have least van der Waals overlap with the surrounding residues. Electrostatics potentials of binding grooves in Figure 2B were calculated with ABPS tools [52] embedded in PyMOL. Binding groove volumes were calculated using the web-based program CASTp with a probe radius of 1.4 Å [53]. The procedure to create interactively accessible 3D images and to integrate them into a PDF document using Adobe Acrobat 9 Pro Extended as well as Adobe 3D Reviewer (as in the interactive Figure S2) has been described [54]. For viewing, the latest version of the freely available Adobe Reader 9 should be installed. The sequence alignments (Figure S1) were generated with Clustal W [55]. Evolutionary analyses (Figure 1A) were conducted using the Neighbor-Joining method [56] in program MEGA4 [57]. The evolutionary distances were computed using the Poisson correction method [58]. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option).

Differential Scanning Calorimetry (DSC)

For DSC, the YF1*7.1:L1 complex and the separately reconstituted β _{2m} samples were prepared in a buffer containing 10 mM phosphate (pH 7.5) and 150 mM NaCl at a protein concentration of 0.2 mg/ml as determined by UV absorption at 280 nm. Molecular masses and extinction coefficients were calculated from the amino acid composition using the ProtParam tool on the ExPASy-server (www.expasy.ch/tools/protparam.html). DSC measurements and the determination of melting temperature (T_m) values were carried out as previously described [59]. The data were analyzed using the "ORIGIN for DSC" software package.

Isoelectric Focusing (IEF)

YF1*7.1 complexes and β _{2m} used for IEF were purified as described previously [47]. Briefly, the protein complexes in a buffer composed of 20 mM Tris pH 7.5 and 150 mM NaCl were incubated at 37°C for 2 d with the respective lipid dissolved in DMSO at a 1:10 (YF1*7.1:lipid) molar ratio. Lipids were purchased from Sigma-Aldrich with the following product numbers: oleic acid (O1008), lipopolysaccharides from *Salmonella enterica* (L2525), lipopolysaccharides from *Escherichia coli* (L5293), and mycolic acid from *Mycobacterium tuberculosis* (M4537). Five μ g of protein sample and 10 μ g IEF Marker 3–10 (Invitrogen) were applied to the native IEF gel at a pH range of 3–7 (Invitrogen). Electrophoresis was performed at 4°C using the XCell SureLock Mini-Cell system (Invitrogen) according to manufacturer's proto-

col. Proteins were detected by staining with Coomassie Brilliant Blue R 250 (SERVA).

Accession Numbers

The National Center for Biotechnology Information (NCBI) accession numbers (<http://www.ncbi.nlm.nih.gov>) for proteins discussed in this article are as follows: chicken YF1*7.1 (AF218783), YF1*15 (AY257165), and YF1*16 (AY257166); chicken BF2*04 (Z54323), BF2*19 (Z54360), and BF2*21 (AF013493); human HLA-B*2709 (Z33453); mouse H-2K^b (P01901); rat RT1-A (M31018); frog UAA (L20733); nurse shark UAA01 (AF220063); human MR1A (AJ249778); mouse MR1A (AF010448); rat MR1 (Y13972); human ZAG (M76707); chicken CD1-1 (AY874074) and CD1-2 (AY375530); human CD1b (AL121986); bovine CD1b3 (Q1L1H6); mouse CD1d (AK002582); and human EPCR (AF106202).

The Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank accession numbers (<http://www.pdb.org>) for the YF1*7.1:L1 and YF1*7.1:L2 structures are 3P73 and 3P77, respectively. The accession numbers for other proteins discussed in this article are as follows: BF2*2101 (3BEV), HLA-B*2709 (1OF2), Mamu-A*01 (1ZVS), H-2K^b (1S7Q), RT1-A (1KJM), ZAG (1ZAG), ggCD1-1 (3JVG), ggCD1-2 (3DBX), hsCD1b (2H26), btCD1b3 (3L9R), mmCD1d (2FIK), and EPCR (1LQV).

Supporting Information

Figure S1 Amino acid sequence alignment of the α 1- and α 2-domains of YF1 alleles with selected classical and non-classical MHC class I molecules. Numbering refers to YF1*7.1. Secondary structure is shown at the top of the alignment: known or predicted α -helices as pink bars and β -sheets as blue bars. Residues with side chains contributing to the binding grooves (crystallographic evidence) are colored according to their biochemical properties: acidic as red, basic as blue, polar as green, and hydrophobic as yellow. Residues contributing to pockets A and F of classical class I molecules are marked with "A" and "F" above the alignment, respectively. The sequences of YF1*15 and YF1*16 are only partial and lack the first 27 amino acids (indicated by dots).

Found at: doi:10.1371/journal.pbio.1000557.s001 (1.15 MB EPS)

Figure S2 Binding grooves of YF1*7.1 and selected classical or non-classical class I molecules, together with an embedded interactive three-dimensional figure.

The three-dimensional (3D) comparison of the molecules in (A) can be activated by clicking on the image in (B). Each individual structural component (with its designation shown on the left panel) can be selected or removed by checking the boxes in the model tree, using the mouse buttons. A tree of all available models is available through clicking onto the respective icons to the right of the "Views" drop-down menu. Each model can be manipulated individually (the tools to rotate, pan, or zoom can be selected through the toolbar or the contextual menu). Preset views (shown below the model tree) can be selected in the form of a "tour" by clicking the green arrows in the middle of the opened model tree menu. Termination of the interactive session can be accomplished by right-clicking anywhere onto the model and choosing "Disable 3D."

Found at: doi:10.1371/journal.pbio.1000557.s002 (2.22 MB PDF)

Figure S3 Thermodynamic stabilities of YF1*7.1 complexes and β_2m measured by differential scanning calorimetry.

Examples for experimental excessive heat capacity curves (black curved lines) and deconvolution results (red curved lines) of (A) a YF1*7.1:L1 complex and (B) free β_2m . The experimental curve of the YF1*7.1:L1 complex can be deconvoluted into three two-state transitions with $T_m^1 = 47.7^\circ\text{C}$, $T_m^2 = 57.9^\circ\text{C}$, and $T_m^3 = 64.1^\circ\text{C}$, while only one two-state transition can be deconvoluted for β_2m ($T_m = 59.1^\circ\text{C}$).

Found at: doi:10.1371/journal.pbio.1000557.s003 (0.27 MB TIF)

Figure S4 Schematic representation of YF1*7.1 ligand binding modes.

The area around the A pocket is marked in green, that around the F pocket in blue; the lengths of the ligands are approximations. (A) The binding of a hydrophobic ligand with a long aliphatic chain within the A pocket is depicted (compare Figure 6A); an acetate molecule forms salt bridges with Arg82 and Arg142. (B) A hydrophobic ligand within the A pocket is shown, together with two short aliphatic molecules (compare Figure 6B); an acetate molecule compensates the charges of Arg82 and Arg142. (C) A large hydrophobic ligand with a negatively charged head group occupies most of the binding groove and interacts with the positively charged amino acids in the vicinity of the F pocket. (D) A ligand with branched hydrophobic segments rests within the binding groove; its exposed head group might interact with positively charged residues at the surface of the YF1*7.1 complex.

Found at: doi:10.1371/journal.pbio.1000557.s004 (1.07 MB TIF)

References

- Pink JRL, Koch C, Ziegler A (1981) Immuno-ornithological conversation. In: Steinberg CM, Lefkowitz I, eds. Festschrift in honor of Niels Kaj Jerne on the occasion of his 70th birthday: the immune system. Basel: Karger. pp 69–75.
- Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexes. *Immunogenetics* 56: 683–695.
- Madden DR (1995) The 3-dimensional structure of peptide-MHC complexes. *Ann Rev Immunol* 13: 587–622.
- Silk JD, Salio M, Brown J, Jones EY, Cerundolo V (2008) Structural and functional aspects of lipid binding by CD1 molecules. *Ann Rev Cell Dev Biol* 24: 369–395.
- Bennett MJ, Lebron JA, Bjorkman PJ (2000) Crystal structure of the hereditary haemochromatosis protein HFE complexed with transferrin receptor. *Nature* 403: 46–53.
- Delany ME, Robinson CM, Goto RM, Miller MM (2009) Architecture and organization of chicken microchromosome 16: order of the NOR, MHC-Y, and MHC-B subregions. *J Hered* 100: 507–514.
- Schierman L, Nordskog AW (1961) Relationship of blood type to histocompatibility in chickens. *Science* 134: 1008–1009.
- Kaufman J, Milne S, Göbel TW, Walker BA, Jacob JP, et al. (1999) The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 401: 923–925.
- Bacon LD, Hunt HD, Cheng HH (2001) Genetic resistance to Marek's disease. *Curr Top Microbiol Immunol* 255: 121–141.
- Miller MM, Wang C, Parisini E, Coletta RD, Goto RM, 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.
- Salomonsen J, Sørensen MR, Marston DA, Rogers SL, Collen T, et al. (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.
- Pharr GT, Gwynn AV, Bacon LD (1996) Histocompatibility antigen(s) linked to Rfp-Y (Mhc-like) genes in the chicken. *Immunogenetics* 45: 52–58.
- LePage KT, Miller MM, Briles WE, Taylor RL (2000) Rfp-Y genotype affects the fate of Rous sarcomas in B-2 B-5 chickens. *Immunogenetics* 51: 751–754.
- Afanassieff M, Goto RM, Ha J, Sherman MA, Zhong L, et al. (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.
- Koch M, Camp S, Collen T, Avila D, Salomonsen J, et al. (2007) Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* 27: 885–899.
- Zajonc DM, Striegl H, Dascher CC, Wilson IA (2008) The crystal structure of avian CD1 reveals a smaller, more primordial antigen-binding pocket compared to mammalian CD1. *Proc Natl Acad Sci U S A* 105: 17925–17930.

Table S1 Comparison of binding groove residues of YF1*7.1 as well as selected classical and non-classical class I molecules.

Found at: doi:10.1371/journal.pbio.1000557.s005 (0.03 MB DOC)

Table S2 Crystallization and cryo-protectant conditions.

Found at: doi:10.1371/journal.pbio.1000557.s006 (0.03 MB DOC)

Acknowledgments

We thank Dr. R. Misselwitz for help with differential scanning calorimetric measurements and three anonymous reviewers for critical but very constructive comments on an earlier version of this manuscript. We acknowledge access to beamline BL14.1 and 14.2 of the BESSY II storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie.

Author Contributions

The author(s) have made the following declarations about their contributions: Conceived and designed the experiments: MMM BUZ OD AZ. Performed the experiments: CSH SG. Analyzed the data: CSH SG BL BUZ OD. Contributed reagents/materials/analysis tools: MMM. Wrote the paper: CSH BUZ OD AZ.

2. The Structure of the Chicken YF1*7.1 Molecule and its Ligands

Crystal Structures of the Chicken YF1*7.1 Molecule

33. Loll B, Kern J, Saenger W, Zouni A, Biesiadka J (2007) Lipids in photosystem II: interactions with proteins and cofactors. *Biochem Biophys Acta* 1767: 509–519.
34. Zajonc DM, Elsiger MA, Teyton L, Wilson IA (2003) Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 Å. *Nat Immunol* 4: 808–815.
35. Batuwangala T, Shepherd D, Gadola SD, Gibson KJ, Zaccari NR, et al. (2004) The crystal structure of human CD1b with a bound bacterial glycolipid. *J Immunol* 172: 2382–2388.
36. Zajonc DM, Crispin MD, Bowden TA, Young DC, Cheng TY, et al. (2005) Molecular mechanism of lipopeptide presentation by CD1a. *Immunity* 22: 209–219.
37. Hassan MI, Bilgrami S, Kumar V, Singh N, Yadav S, et al. (2008) Crystal structure of the novel complex formed between Zinc α 2-glycoprotein (ZAG) and prolactin-inducible protein (PIP) from human seminal plasma. *J Mol Biol* 384: 663–672.
38. Le Bourhis L, Martin E, Péguillet I, Guihot A, Froux N, et al. (2010) Antimicrobial activity of mucosal-associated invariant T cells. *Nat Immunol* 11: 701–708.
39. Gold MC, Cerri S, Smyk-Pearson S, Cansler ME, Vogt TM, et al. (2010) Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biol* 8: e1000407. doi:10.1371/journal.pbio.1000407.
40. Girardi E, Wang J, Mac TT, Vershuis C, Bhowruth V, et al. (2010) Crystal structure of bovine CD1b3 with endogenously bound ligands. *J Immunol* 185: 376–386.
41. Beltrami A, Rossmann M, Fiorillo MT, Paladini F, Sorrentino R, et al. (2008) Citrullination-dependent differential presentation of a self-peptide by HLA-B27 subtypes. *J Biol Chem* 283: 27189–27199.
42. Mohammed F, Cobbold M, Zarling AL, Salim M, Barrett-Wilt GA, et al. (2008) Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. *Nat Immunol* 9: 1236–1243.
43. Ziegler A, Pink JRL (1976) Chemical properties of two antigens controlled by the major histocompatibility complex of the chicken. *J Biol Chem* 251: 5391–5396.
44. Wallny HJ, Avila D, Hunt LG, Powell TJ, Riegert P, et al. (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.
45. Laun K, Coggill P, Palmer S, Sims S, Ning Z, et al. (2006) The leukocyte receptor complex in chicken is characterized by massive expansion and diversification of immunoglobulin-like loci. *PLoS Genet* 2: e73.
46. Viertlboeck BC, Gick CM, Schmitt R, Du Pasquier L, Göbel TW (2010) Complexity of expressed CHIR genes. *Dev Comp Immunol* 34: 866–873.
47. Hee CS, Gao S, Miller MM, Goto RM, Ziegler A, et al. (2009) Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1. *Acta Crystallogr F65*: 422–425.
48. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D53*: 240–255.
49. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D66*: 486–501.
50. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D66*: 12–21.
51. Delano WL (2002) The PyMOL Molecular Graphic System. San CarlosCA: DeLano Scientific.
52. Lerner MG, Carlson HA (2006) APBS plugin for PyMOL. Ann Arbor: University of Michigan.
53. Liang J, Edelsbrunner H, Woodward C (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Prot Sci* 7: 1884–1897.
54. Kumar P, Ziegler A, Grahn A, Hee CS, Ziegler A (2010) Leaving the structural ivory tower, assisted by interactive 3D PDF. *Trends Biochem Sci* 35: 419–422.
55. Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673–4680.
56. Saitou N, Nei M (1987) The neighbor-joining method - a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425.
57. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
58. Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, eds. *Evolving genes and proteins*. New York: Academic Press. pp 97–166.
59. Hillig RC, Hülsmeier M, Wellle K, Misselwitz R, Wellle H, et al. (2004) Thermodynamic and structural analysis of peptide- and allele-dependent properties of two HLA-B27 subtypes exhibiting differential disease association. *J Biol Chem* 279: 652–663.

Supporting information

Figure S1

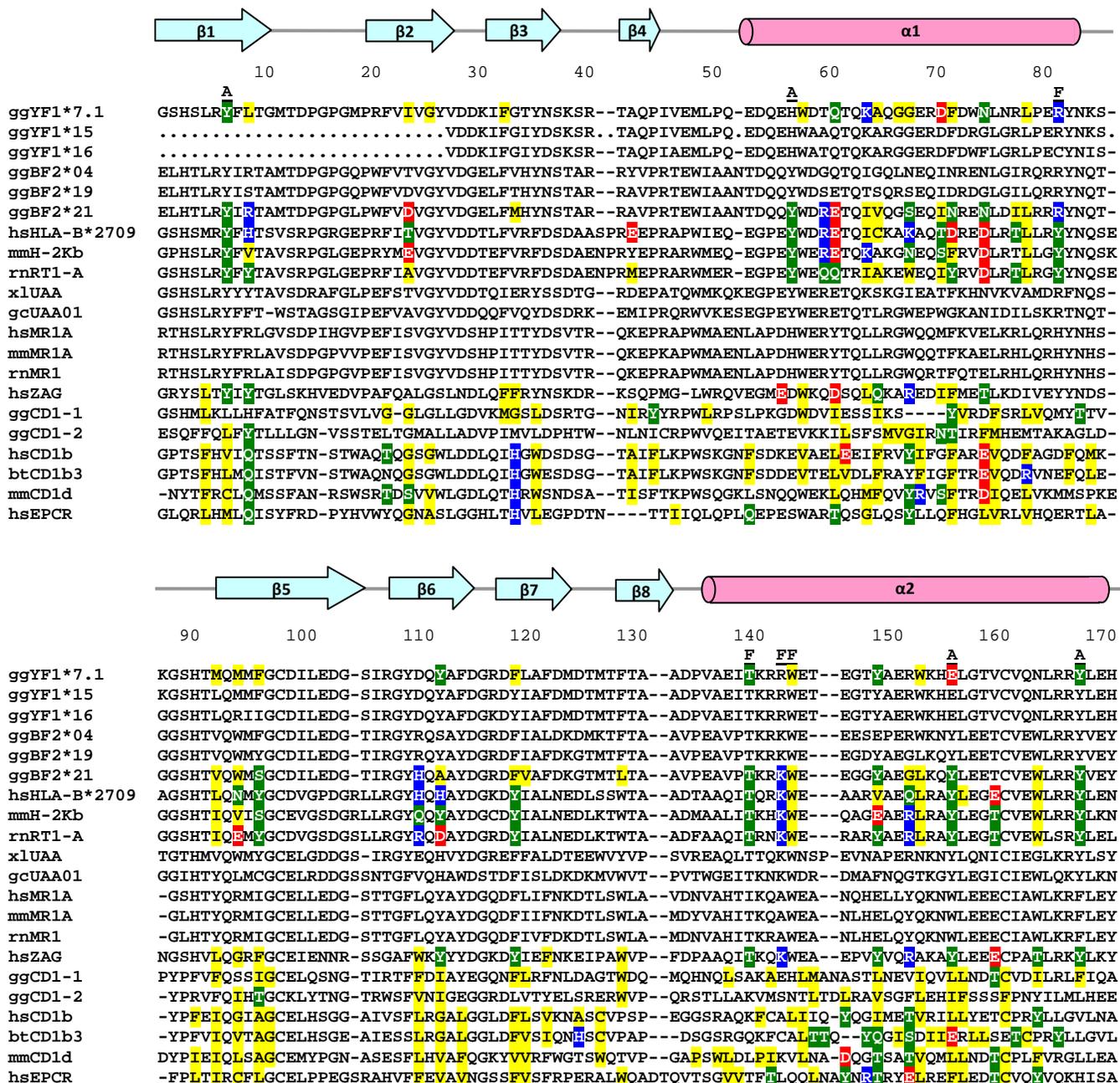


Figure S2

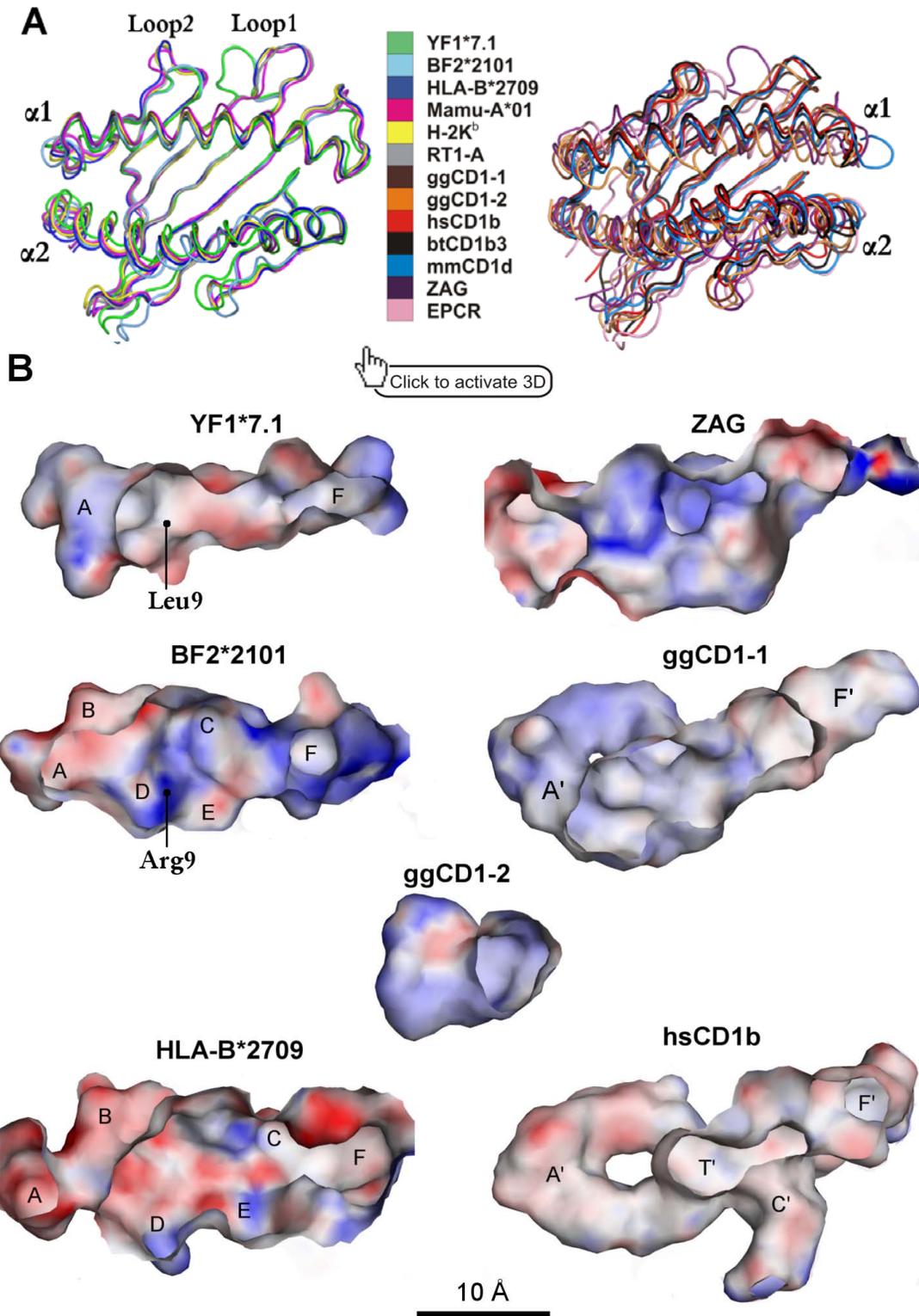


Figure S3

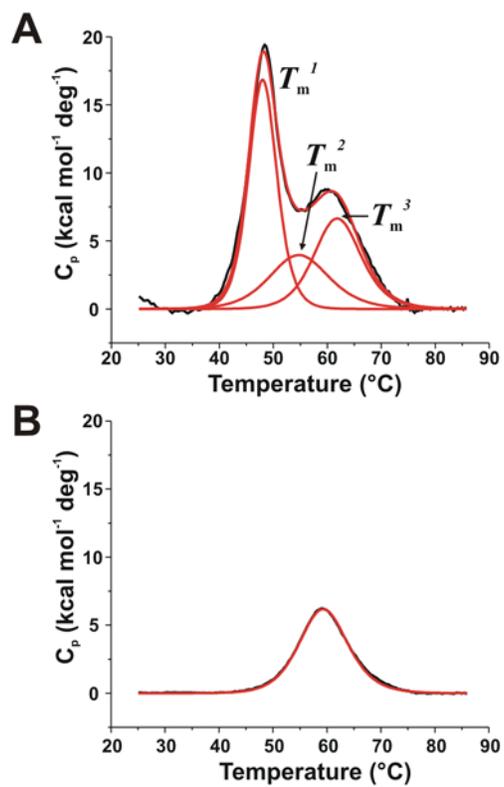


Figure S4

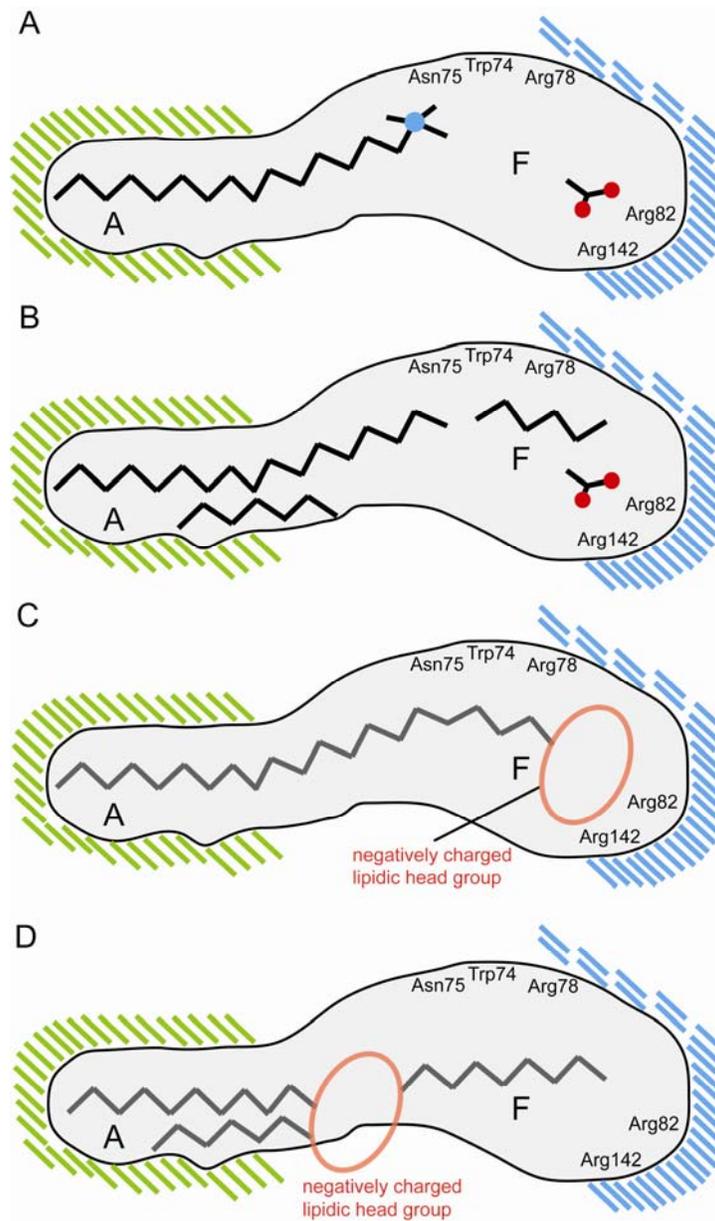


Table S1. Comparison of binding groove residues of YF1*7.1 as well as selected classical and non-classical class I molecules.

	YF1*7.1	BF2*2101	HLA-B*2709	H-2K ^b	RT1-A	ZAG	EPCR	ggCD1-1	ggCD1-2	hsCD1b	btCD1b3	mmCD1d
PDB entries	3P73	3BEV	1OF2	1S7Q	1KJM	1ZAG	1LQV	3JVG	3DBX	2H26	3L9R	2FIK
Total	27*	32	33	29	32	33	43	38	23	49	47	46
Hydrophobic	16	16	9	8	11	16	31	34	19	40	34	34
Polar	7	9	13	13	15	11	9	4	4	6	8	8
Basic	2	5	6	4	3	3	2	0	0	1	3	2
Acidic	2	2	5	4	3	3	1	0	0	2	2	2

* Number of amino acids contributing to the binding groove

Table S2. Crystallization and cryo-protectant conditions

Designation	Ligand*	Resolution	Crystallization	Cryoprotectant
L1	-	1.3	0.20 M AmAc, 0.1 M NaAc pH 5.0, 20% PEG 4000	glycerol
L2	-	1.6	0.24 M AmAc, 0.1 M NaAc pH 5.0, 20% PEG 4000	PEG 200
L3	PC	1.65	0.25 M AmAc, 0.1 M NaAc pH 5.0, 19% PEG 4000	glycerol
L4	PLM	1.55	0.32 M AmAc, 0.1 M NaAc pH 5.0, 16% PEG 4000	glycerol
L5	PLM	2.6	0.28 M AmAc, 0.1 M NaAc pH 5.0, 14% PEG 4000	PEG 200
L6	POPC	1.55	0.25 M AmAc, 0.1 M NaAc pH 5.9, 19% PEG 4000	PEG 200
L7	DOPC	1.55	0.25 M AmAc, 0.1 M NaAc pH 5.9, 19% PEG 4000	PEG 200
L8	OLA	1.6	0.23 M AmAc, 0.1 M NaAc pH 5.0, 19% PEG 4000	PEG 200

*Ligand present in the reconstitution

3. Comparative Biophysical Characterizations of Human and Chicken β_2 -microglobulin and the Evolution of Interfaces in MHC Class I Molecules

3.1 Summary

β_2 m is a non-polymorphic component that forms complexes with a broad range of MHC class I and class I-like molecules. This chapter deals with two aspects of β_2 m whose details are described in the two articles below. The first article deals with the structural and biophysical characterization of free chicken β_2 m by X-ray crystallography, differential scanning calorimetry (DSC) and infrared-spectroscopy. The results reveal that the structure of chicken β_2 m is virtually identical to those of free β_2 m of other vertebrates studied (human, bovine and fish), but thermodynamically different. This can be explained by the fewer intra-molecular contacts found in this molecule. In addition, an unprecedentedly tight crystal packing of chicken β_2 m molecules was observed. The second part describes results dealing with the dynamics of free and HLA-B*27:05-complexed human β_2 m as well as the interface interactions between β_2 m and HC of selected MHC class I and class I-like complexes. Using nuclear magnetic resonance (NMR) spectroscopy, we detected several β_2 m residues involved in interface contacts that exhibit considerable dynamics in the unbound form, but gain stabilization upon binding to the HC. A number of these residues engage in evolutionarily retained inter-polypeptide chain contacts that are detectable in species as diverse as mammals and chicken. In addition, we identified interface interactions that are characteristic for a particular group of MHC class I complexes such as peptide- or lipid-binding molecules or are limited to a specific vertebrate taxon. Taken together, the analysis of interfaces demonstrates that β_2 m is an unusually versatile molecule which fulfils its role of stabilizing a wide range of MHC class I and class I-like complexes by engaging in intermolecular contacts that may be absolutely conserved, semi-conserved or characteristic for a group of complexes or particular vertebrate taxa.

3.2 Manuscripts

3.2.1 Hee, C.S., Fabian, H., Uchanska-Ziegler, B., Ziegler, A., Loll, B. (2011) Comparative biophysical characterization of chicken β_2 -microglobulin. (Submitted)

3.2.2 Hee, C.S., Beerbaum, M., Diehl, A., Loll, B., Schmieder, P., Uchanska-Ziegler, B., Ziegler, A. (2011) Dynamics of free versus complexed β_2 -microglobulin and the evolution of interfaces in MHC class I molecules. (Submitted)

Comparative biophysical characterization of chicken β_2 -microglobulin

Running title: Biophysical properties of chicken β_2 -microglobulin

Chee-Seng Hee¹, Heinz Fabian², Barbara Uchanska-Ziegler¹, Andreas Ziegler^{1,*}, and
Bernhard Loll^{3,*}

¹ Institut für Immungenetik, Charité - Universitätsmedizin Berlin, Campus Benjamin
Franklin, Freie Universität Berlin, Thielallee 73, 14195 Berlin, Germany

² Robert Koch Institut, P 25, Nordufer 20, 13353 Berlin, Germany

³ Institut für Chemie und Biochemie, Abteilung Strukturbiochemie, Freie Universität Berlin,
Takustrasse 6, 14195 Berlin, Germany

* Prof. Dr. Andreas Ziegler: Tel: +49-30-450 564731, Fax: +49-30-450 564920,
e-mail: andreas.ziegler@charite.de

* Dr. Bernhard Loll: Tel: +49-30-838-57348, Fax: +49-30-838-54936
e-mail: loll@chemie.fu-berlin.de

Keywords: β_2 -microglobulin, immune system, major histocompatibility complex, IR
spectroscopy, crystal structure

Summary

β_2 -microglobulin (β_2 m) is not only the smallest building block of molecules belonging to the immunoglobulin superfamily, but serves also essential functions as a non-polymorphic component of major histocompatibility complex (MHC) class I molecules. By comparing thermodynamic and structural characteristics of chicken β_2 m with those of human, bovine, and carp, we seek to elucidate whether it is possible to pinpoint features of β_2 m that set the avian protein apart from those of other vertebrates. The thermodynamic assays revealed that monomeric chicken β_2 m exhibits a lower melting temperature than human β_2 m, and the H/D exchange behavior observed by infrared spectroscopy indicates a more flexible structure of this protein in comparison to its human counterpart. To understand these differences at a molecular level, we determined the structure of free chicken β_2 m by X-ray crystallography to a resolution of 2.0 Å. Chicken β_2 m is highly similar to previously determined structures from other species, whether free or in the complex of an MHC class I molecule, but its crystal lattice is unprecedented, as a particularly tight packing was observed. The crystal structure also revealed that chicken β_2 m has fewer intra-molecular contacts compared to β_2 m from other vertebrates, which might explain the distinct thermodynamic behaviors of the chicken β_2 m. Our comparisons indicate that certain biophysical and structural characteristics of the chicken protein diverge considerably from those of the other β_2 m analyzed, although basic features of β_2 m have been retained through evolution.

Introduction

β_2 -microglobulin (β_2m) is a protein composed of 99 residues which folds into a seven-stranded β -sandwich fold.⁽¹⁾ The strands are termed A to G and the loops are named according to the strands which they connect, e.g. AB, BC, CD, and so on. The typical immunoglobulin fold of β_2m comprises two β -sheets (strands A, B, D, E on one hand and strands C, F, G on the other) mediated by a highly conserved intra-molecular S-S bridge with a spacing of ~ 54 amino acids between the two Cys residues of the protein. β_2m constitutes the basic building unit of the immunoglobulin superfamily.⁽²⁾ In addition, the protein serves as a light chain that is non-covalently bound to the heavy chain (HC) of major histocompatibility complex (MHC) class I or class I-like molecules such as CD1.^(3,4)

β_2m is an important prerequisite for ligand binding by MHC class I molecules⁽⁵⁾. A ligand is anchored within a groove that is formed by the HC, but also β_2m is crucial for maintaining complex stability,^(6,7) and this protein is necessary as well for the intracellular transport of the complex to the surface of nearly all nucleated cells.⁽⁸⁻¹¹⁾ Free β_2m , on the other hand, is found in most body fluids,⁽¹²⁾ including serum, urine, milk and colostrums.⁽¹³⁻¹⁵⁾ As the protein is normally eliminated by the kidney, renal failure leads to an increase of the concentration of β_2m in the serum by up to 60-fold.^(15,16) Greatly elevated concentrations of free β_2m result in the formation of amyloid fibrils that may be deposited at various sites in the body of patients who undergo dialysis due to kidney failure.^(17,18)

In this study, we employ a comparative biophysical approach to identify features of chicken β_2m that distinguish it from those of other vertebrates. In particular, we investigate the suitability of this polypeptide as a model for amyloidosis because calorimetric measurements had revealed that this molecule is less stable than its human counterpart⁽¹⁹⁾ and might therefore be expected to exhibit considerable dynamics. To gain further insights into the thermodynamic properties of chicken β_2m and its dynamic behavior, we performed differential scanning calorimetry (DSC) and infrared (IR) spectroscopy. Since structural information about chicken β_2m was limited to complexes with chicken HC, i.e. the BF2*2101,⁽²⁰⁾ YF1*7.1,⁽¹⁹⁾ and CD1-1⁽²¹⁾ complexes, we determined also the crystal structure of the free protein by X-ray crystallography. The results allow us not only to compare free β_2m from two mammalian, an avian, and a fish species, but permit us to suggest structure-based explanations for a number of biophysical properties of these proteins.

Material and Methods

β_2 m purification

The coding sequence for the extra-cellular domain of β_2 m (residues 21-120 of the signal peptide-containing protein) were cloned into the expression vector pMAL-p4x and consequently expressed as maltose-binding protein fusion (MBP- β_2 m) and purified using the pMAL purification system (New England Biolabs, Germany). *Escherichia coli* TB1 cells transformed with the MBP- β_2 m construct were grown at 310 K in LB-medium containing 100 mg/l ampicillin to an OD₆₀₀ of 0.5. Protein expression was induced by adding 0.4 mM isopropyl β -D-1-thiogalactopyranoside and cells were incubated for 20 h at 298 K. Cells were harvested by centrifugation at 4500 g for 10 min at 277 K. The MBP- β_2 m fusion protein was extracted from the periplasm according to the pMAL purification protocol. The cell pellet obtained by centrifugation was resuspended and incubated in 100 ml buffer containing 30 mM Tris-HCl pH 8.0, 20% (w/v) sucrose, 1 mM EDTA and 0.3 mM PMSF on a shaker for 10 min. After incubation, the sample was centrifuged at 8000g for 10 min at 277 K and the pellet was resuspended in 100 ml chilled 5 mM MgSO₄ followed by incubation for 10 min in an ice bath. The sample was then centrifuged at 8000 g for 10 min at 277 K; the supernatant was collected and filtered through a 0.22 μ m filter membrane (Millipore). The filtered sample was applied onto a 10 ml self-packed amylose resin (New England Biolabs, Germany) column, washed with 100 ml column buffer (20 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM EDTA) and eluted with 30 ml column buffer containing 10 mM maltose. Fractions containing MBP- β_2 m fusion protein were collected and concentrated using an Amicon Ultra-15 protein centrifugal filter with 10 kDa molecular weight cutoff (Millipore) before purification by size exclusion chromatography on a SuperdexTM200 16/60 column (GE Healthcare) connected to an ÄKTA FPLC instrument (GE Healthcare, Germany). The MBP-tag was cleaved with protease factor Xa (New England Biolabs, Germany) by incubation at 296 K for 2 d. β_2 m was separated from MBP using a 10 ml self-packed amylose resin column followed by purification on a SuperdexTM75 column (GE Healthcare, Germany) in buffer composed of 20 mM Tris-HCl pH 7.5, 150 mM NaCl, and 0.01% (w/v) sodium azide. For crystallization experiments β_2 m was concentrated to 17 mg ml⁻¹. The purity of the protein was assessed by SDS-PAGE. The protein concentration was determined by measuring the absorption at a wavelength of 280 nm, assuming an absorption coefficient of 19200 M⁻¹cm⁻¹. Human β_2 m was refolded from inclusion bodies.⁽²²⁾

*Preparation of YF1*7.1, BF2*1301:pMDV and BF2*2101:p1 complexes*

The cDNA sequences of the extracellular domains of BF2*1301 and BF2*2101 (residues 22-294) were cloned into the pMAL-p4x vector. Both clones were verified by DNA sequencing. Preparation and purification of β_2m and the YF1*7.1 complex have been reported previously.⁽²³⁾ For the BF2*1301 and BF2*2101 complexes, a similar protocol was established with slight modifications. The BF2*1301 and BF2*2101 HC were expressed as MBP-fused proteins as inclusion bodies in *E. coli* TB1 cells. The HC were refolded in a reconstitution buffer together with solubilized β_2m and a peptide. For the reconstitution of BF2*1301, the pMDV peptide (¹⁵³VDAYDRDE¹⁶⁰) derived from *Meleagrid herpesvirus* glycoprotein B was used, whereas the pC1 peptide (³⁵³SELKDFQKL³⁶²) is derived from *Gallus gallus* Ras-GTPase-activating protein and was employed for reconstitution of BF2*2101.⁽²⁴⁾ Both peptides were purchased from Alta Bioscience, UK. Refolding was performed according to the established protocol for HLA complexes.⁽²⁵⁾ Reconstituted protein complexes were subjected to Factor Xa digestion to remove the MBP-tags followed by affinity chromatography step on an amylose column and gel filtration chromatography as described.⁽²³⁾

Differential scanning calorimetry

Excessive heat capacity curves were recorded using an ultrasensitive scanning microcalorimeter (VP-DSC, MicroCal Inc., Northampton, MA, USA) with a heating rate of 1 K/min and a sample cell volume of 0.5 ml. For refolding experiments, after the initial heating from 20 °C to 90 °C, the samples were cooled to 20 °C and subsequently heated again to 90 °C. All DSC samples were prepared in 10 mM phosphate buffer (pH 7.5), 150 mM NaCl and the protein concentrations were adjusted to the range of 0.18-0.22 mg/ml. DSC measurements were conducted as described previously.⁽²⁵⁾ The experimental data were analyzed by standard procedures using the “ORIGIN for DSC” software package supplied by the manufacturer.

Infrared spectroscopy

The protein solutions were always freshly prepared and placed into demountable calcium fluoride IR-cells with an optical pathlength of 50 μ m for measurements in D₂O-buffer or 8 μ m for samples in H₂O-buffer. IR-spectra were recorded with IFS-28B and IFS-66 Fourier transform infrared (FTIR) spectrometers (Bruker Optics, Ettlingen, Germany)

equipped with DTGS (deuterated triglycine sulphate) detectors and continuously purged with dry air. For each sample, 128 interferograms were co-added and Fourier-transformed to yield spectra with a nominal resolution of 4 cm^{-1} . The sample temperature was controlled by means of thermostated cell jackets. Buffer spectra were recorded under identical conditions and subtracted from the spectra of the proteins in the relevant buffer (10 mM sodium phosphate, pH 7.5, 150 mM NaCl). Spectral contributions from residual water vapor, if present, were eliminated using a set of water vapor spectra. The final unsmoothed protein spectra were used for further analysis. Band intensities were determined by standard functions of the Bruker OPUS software, implemented into home-built macros for data analysis. A detailed description of the entire procedure has been described previously.⁽²⁶⁾

Crystallization and crystal cooling

Initial crystals were obtained at 291 K in a sitting drop setup in 96-well plates even though the crystals were insufficient in size for X-ray data collection. Crystal size could be improved in a hanging drop setup with a reservoir solution with a volume of 700 μl composed of 22% (w/v) polyethylene glycol 3350 and 200 mM MgCl_2 . The crystallization drop was mixed from 1.5 μl chicken $\beta_2\text{m}$ and 1.5 μl reservoir solution. The needle-shaped, inter-grown crystals appeared after three weeks. Inter-grown needles were carefully separated with an acupuncture needle yielding single thin specimen. Prior to cooling the crystals were transferred to a cryo-solution containing 25% (w/v) polyethylene glycol 3350, 200 mM MgCl_2 , and 20% (v/v) glycerol and subsequently flash cooled in liquid nitrogen.

X-ray data collection, structure determination and refinement

Synchrotron diffraction data were collected at the beamline 14.2 of the MX Joint Berlin laboratory at the BESSY in Berlin, Germany. X-ray data collection was performed at 100 K at a wavelength of $\lambda=0.9184\text{ \AA}$ and diffraction data were processed with the XDS package.⁽²⁷⁾ For calculation of the free R-factor, a randomly generated set of 5% of the reflections from the diffraction data set was used and excluded from the refinement. Initial phases were determined by molecular replacement with the program PHASER⁽²⁸⁾ employing the coordinates of $\beta_2\text{m}$ from the structure of BF2*2101 (PDB entry 3BEW).⁽²⁰⁾ The structure was refined by maximum-likelihood restrained refinement implemented in REFMAC5⁽²⁹⁾ followed by iterative model building cycles with COOT.⁽³⁰⁾ Water molecules were picked with COOT⁽³⁰⁾ and manually inspected. At final stages of refinement, TLS refinement⁽³¹⁾ was applied with one TLS-groups for each polypeptide chain. Model quality was evaluated with

PROCHECK⁽³²⁾ and MolProbity⁽³³⁾. Secondary structure elements have been assigned with DSSP.⁽³⁴⁾ Figures were prepared using PyMOL⁽³⁵⁾ and 3D figures were produced as described previously.⁽³⁶⁾ The 3D figure can be activated by clicking on the image. Each model (structure) can be selected or removed by checking the boxes in the model tree, which is available upon clicking on the icon to the right of the “Views” drop-down menu. Preset views are shown below the model tree and can be displayed in the form of a “tour” by clicking the green arrows in the middle of the opened model tree menu. 3D session can be terminated by right-clicking anywhere onto the model and choosing “Disable 3D”. The atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under the accession code 3O81.

Results and discussion

Thermodynamic properties of chicken β_2 m and class I complexes

The presence of β_2 m has been described as a prerequisite for the stabilization of mammalian MHC class I complexes.^(6, 7) While the same is suspected for avian MHC class I molecules, systematic studies of these proteins have not been carried out so far. Our reconstitution experiments with three different classical chicken subtypes, BF2*2101, BF2*1301 and YF1*7.1 together with β_2 m clearly revealed the absolute requirement of the light chain for complex reconstitution (data not shown). We then employed DSC to assess the thermodynamic stability of free β_2 m as well as of MHC class I complexes. Free chicken β_2 m exhibits a single two-state transition upon temperature increase, with a melting temperature (T_m) of ~ 59.1 °C (**Figure 1A and E**), about 5 °C lower than that of human β_2 m.⁽²⁵⁾ The unfolding enthalpies of the two proteins, however, were practically indistinguishable. Furthermore, the folding reversibility of chicken β_2 m was found to be akin to that of human β_2 m (results not shown).

The results from the DSC analysis of free chicken β_2 m are supported by denaturation experiments performed by IR spectroscopy. These confirm a decreased T_m value of ~ 5 °C when compared to human β_2 m (data not shown). In addition, we studied the hydrogen/deuterium (H/D) exchange behavior of chicken and human β_2 m samples by monitoring the disappearance of the amide II band near 1550 cm^{-1} in the IR-spectra. This approach can provide valuable information on the structure and flexibility of proteins.⁽³⁷⁾ To achieve this, the corresponding lyophilized samples were dissolved in D₂O buffer and the time course and extent of exchange of the amide protons with deuterons in the samples, kept

at 25 °C and 37 °C, respectively, were measured over a period of 24 h. The IR-data reveal that already shortly after dissolution (first data point measured 3 min after starting the experiment), 70-75% (at 25 °C) and 75-80% (at 37 °C), respectively, of the N-H groups of both proteins are exchanged by deuterons (**Figure 1F and G**). The degree of exchange is initially slightly lower for human β_2m than for chicken β_2m , but these differences in H/D exchange become more pronounced over time. At room temperature, it takes ~10 h to fully exchange all N-H groups of chicken β_2m , while ~10% remain unexchanged in human β_2m even after one day at 25 °C (**Figure 1F**). As expected, an increase in temperature makes the proteins more prone to H/D exchange, but also at 37 °C the loss of residual N-H groups is much faster in chicken than in human β_2m (**Figure 1G**). Generally, protons exposed to the solvent and involved in unstable regions of a protein will exchange rapidly, whereas the protons which are part of the hydrophobic core of a protein are typically much more resistant to exchange. Therefore, the observed differences in H/D exchange between human and chicken β_2m are most likely due to conformational differences.

We then analyzed whether these properties of free chicken β_2m could also be observed in the HC-complexed form by using DSC. As expected, and other than in the case of free β_2m , the experimental excessive heat capacity curves of chicken BF2*2101:pC1 and BF2*1301:pMDV were more complex than that of free β_2m . They could each be deconvoluted into two two-state transitions (**Figure 1B, D, and E**), while the YF1*7.1 complex was clearly less stable (**Figure 1C and E**). Unfolding of this complex started at 40 °C, a value that is lower than the body temperature (41.8 °C) of a chicken⁽³⁸⁾ and the experimental excessive heat capacity curve could be deconvoluted into three two-state transitions with the T_m peaks at 48 °C, 58 °C and 64 °C (**Figure 1C**). We have already argued that these characteristics are connected to the lack of an appropriate ligand within the binding groove of YF1*7.1.⁽¹⁹⁾ In addition to being distinct from the latter, the unfolding behaviors of the BF2 complexes were not identical either. Two-state transitions corresponding to two T_m peaks at ~58 °C and ~63 °C (**Figure 1B and E**) characterize the BF2*2101:pC1 complex, whereas the BF2*1301:pMDV complex was more stable with a first T_m of ~66 °C and a second at T_m ~69 °C (**Figure 1D and E**). This could be due to the unusual anchor amino acid requirement of the BF2*1301 protein, which binds only peptides with three appropriately spaced acidic residues.⁽³⁹⁾ All three chicken complexes did not refold when cooled down to 20 °C after the first heating step (data not shown), as observed for HLA-B27 complexes.⁽²⁵⁾

An interpretation of these distinct melting behaviors found by DSC experiments is difficult, although it may be aided by considering the T_m values in conjunction with the corresponding unfolding enthalpies for a given complex and a comparison with the values obtained with free β_2 m. For example, the first transition (T_{m1} ' of the YF1*7.1 complex is very likely not due to unfolding of β_2 m, because the temperature is much too low and the enthalpy does not correspond to that of free β_2 m (**Figure 1C**). This transition should therefore reflect the partial melting of the HC. Whether one of the next two transitions (T_{m1} and T_{m2}) is due to unfolding of β_2 m, as suggested by the enthalpies observed, to further unfolding of the YF1*7.1 HC or a combination of both, is impossible to deduce from these DSC experiments. Similarly, from a comparison with the values obtained for free β_2 m, it seems plausible to assume that T_{m1} reflects the melting of β_2 m from the BF2*2101:pC1 complex. In the case of BF2*1301:pMDV, the HC must contribute to both melting temperatures observed for the complex, but it remains unknown whether β_2 m melts as part of the first or the second transition. The DSC experiments do show, however, that chicken MHC class I molecules are stabilized in an allele-dependent fashion when bound to appropriate ligands. Melting temperatures higher than that of free β_2 m can thus be obtained. This has also been observed in the case of selected human HLA class I complexes.^(25, 40)

Overall structure

To obtain further information on the detected flexibility of chicken β_2 m, we decided to crystallize the protein and to determine its structure. Free chicken β_2 m (residues 21-120 of the signal peptide-containing protein) was purified⁽²³⁾ and successfully crystallized. Using synchrotron radiation, an X-ray diffraction data set was collected at 2.0 Å resolution (**Table 1**). The structure (**Figure 2A**) was solved by molecular replacement, employing the structure of chicken β_2 m in complex with BF2*2101 (protein data bank (PDB) entry 3BEW)⁽²⁰⁾ as a model. Refinement of the structure converged with an R/R_{free}-factor of 19.5/24.3%. The geometry of the structure is excellent with no Ramachandran outliers. Detailed refinement statistics are given in **Table 1**. The electron density was of excellent quality, allowing the unambiguous tracing of the polypeptide chain, even in the case of the few residues that were found to display multiple conformations (**Figure 2B**). These residues are stabilized upon binding of the molecule to an MHC class I HC (**Figure 2C**). The model is complete except for the two carboxy-terminal amino acid residues.

Chicken β_2 m adopts the classical fold of the immunoglobulin superfamily constant domain with a seven-stranded β -sandwich fold (**Figure 2A**) that is composed of two β -sheets arranged face-to-face. These are connected and stabilized by a disulfide bond between Cys25 and Cys80 on β -strands B and F, respectively. The β -strand assignment is according to Saper and colleagues⁽³⁾ and resembles that used for human β_2 m. The first β -sheet is composed of β -strands A, comprising the residues 6-11 (A⁶⁻¹¹), B²¹⁻³⁰, D₁⁵⁰⁻⁵¹ and D₂⁵⁵⁻⁵⁶, as well as E⁶²⁻⁷⁰, and the second β -sheet is formed by C³⁶⁻⁴¹, F⁷⁸⁻⁸³, and G⁹¹⁻⁹⁵ (**Figure 2A**). Additionally, a short β -strand is located between the β -strands C and D which is termed C'. It encompasses residues 44 and 45. The β -strands D₁ and D₂ are separated by a three residue long β -bulge that is twisting the β -strand (**Figure 2A**). When complexed with the HC, such an arrangement allows for efficient interaction. The crystallographic asymmetric unit harbors two chicken β_2 m polypeptide chains that are practically indistinguishable with a root mean square deviation (r.m.s.d.) of 0.3 Å for 97 pairs of C α -atoms.

Monomer-monomer interface

The monomer-monomer interface within the asymmetric unit is exclusively hydrophobic, leading to a head-to-head arrangement of the loop connecting strands B and C (BC loop) and the D₂E loop of both β_2 m molecules (**Figure 3A and B**). Further analyses revealed additional crystal contacts between adjacent symmetry-related molecules. The hydrophobic interactions between the two chicken β_2 m molecules appear to be unique, since such an interface has not been observed in other structures of free β_2 m. By formation of this hydrophobic interface *in crystallo*, a number of hydrophobic amino acid residues (**Figure 3B**) are shielded from the solvent, reminiscent of the situation when β_2 m is part of an MHC class I complex. For example, these interactions include Phe56, Trp60 and Phe62. All of these residues are highly conserved from fish to human (**Figure 4**) and may exhibit, despite their involvement in multiple hydrophobic contacts, considerable dynamics (high B-factor) in both polypeptide chains within the asymmetric unit, as in the case of Phe56 and Trp60 (**Figure 5**). We also noticed that several residues located in this segment (from D₁ to E strand) are flexible in free β_2 m, exhibiting either poorly defined electron density (Phe56) or double conformations (Ser55, Asp58 and also Trp60) (**Figure 2**). In the NMR structure of monomeric human β_2 m, an increased conformational flexibility of the same segment was observed.⁽⁴¹⁾ This segment becomes stabilized upon binding of β_2 m to an HC, as observed in

the structure of chicken β_2 m in complex with the HC of YF1*7.1 (**Figure 2B, and Figure 5**).⁽¹⁹⁾ In the complex, the side chain of Trp60 adopts a single conformation that intercalates between HC residues that reside on the floor of the peptide binding groove.

Structural comparison of monomeric β_2 m structures

Compared to the vast number of β_2 m structures solved as part of MHC class I molecules deposited in the PDB, a structure of free β_2 m is only available from four species, namely chicken β_2 m (PDB entry 3O81, this study), bovine β_2 m (1BMG),⁽¹⁾ human β_2 m (1LDS),⁽⁴²⁾ and (2YXF)⁽⁴³⁾ as well as carp β_2 m (3GBL).⁽⁴⁴⁾ In addition, there is structural information available for the human protein obtained using NMR studies.^(41, 42, 45)

The overall fold of all free β_2 m structures is very similar. Superposition of 97 pairs of C α -atoms of chicken and human β_2 m results in an r.m.s.d. of 3.5 Å, chicken and bovine β_2 m exhibit an r.m.s.d. of 3.1 Å, while chicken and carp β_2 m are characterized by an r.m.s.d. of 2.6 Å. These relatively small structural differences can be attributed to different sites within the β -sandwich structure of these proteins (**Figure 6**). The most drastic conformational difference is the AB loop (residues 12-20) between the human and β_2 m of the other species. The human AB loop points away from the β -sandwich, while the chicken, bovine and carp AB loops lean onto the β -sandwich (**Figure 6A**). A further difference is the presence of a β -bulge that is flanked by strands D₁⁵⁰⁻⁵¹ and D₂⁵⁵⁻⁵⁶ in the chicken, bovine and carp β_2 m structures, but not in human β_2 m (**Figure 6a**). Instead of possessing the D₁ – β -bulge – D₂ conformation, human β_2 m is characterized by a continuous β -strand D.⁽⁴²⁾ These authors wondered whether the β -strand D persists only under the conditions of crystallization and performed ¹H-NMR experiments that revealed the presence of the β -bulge in solution, in agreement with earlier NMR studies.^(41, 45) Crystal packing must be regarded as the most likely reason for an induction of the continuous β -strand D in the crystal structure of free human β_2 m. The NMR studies make it very probable that this structural peculiarity of the human protein cannot be found *in vivo*. Apart from this, the differences in H/D exchange between human and chicken β_2 m observed by IR spectroscopy (**Figure 1F and G**) demonstrate conformational differences between the two molecules in solution. IR spectroscopy cannot localize the regions where the two proteins differ, but the comparison of the crystallographic temperature factors suggest that free chicken β_2 m is significantly more flexible than free human β_2 m in the regions between the BC, D₁E and FG strands (**Figure 5**).

Such conformational differences could be the cause of the experimentally observed differences in H/D exchange behavior.

The C'⁴⁵⁻⁴⁶ β -strand and the downstream loop region (C'D₁ loop) adopt different conformations in all structures (**Figure 6A**). This is most likely due to the low level of sequence conservation in the primary sequence of the four β_2 m molecules. The bovine primary sequence, for instance, possesses even an amino acid deletion in this region (**Figure 4**). Further differences can be detected in the overlay of the β_2 m structures, as the DE loop (connecting β -strand D₂⁵⁵⁻⁵⁶ and E⁶²⁻⁷⁰) adopts different conformations in all four species, although the conformational difference between this loop of chicken and carp is minor when compared to the bovine and human DE loops that point into opposite directions (**Figure 6A**). Further structural differences are found in the amino- and carboxy-terminal residues, reflected also by higher B factors (**Figure 5**).

Structural comparison of free and complexed β_2 m

Superposition of the C α backbones of free and complexed chicken β_2 m in the structures BF2*2101 (PDB entry (3BEW)⁽²⁰⁾, CD1-1 (3JVG)⁽²¹⁾ and YF1*7.1 (3P77)⁽¹⁹⁾) reveal that the four β_2 m structures are very similar, with r.m.s.d. of maximally 0.6 Å (**Figure 6B**). The largest differences between free and bound β_2 m are found at both polypeptide termini due to different contacts between β_2 m and HC residues. The structures of free bovine β_2 m and this protein in complex with the HC of bovine BoLA-N*01301 (2XFX)⁽⁴⁶⁾ or CD1b3 (3L9R)⁽⁴⁷⁾ are also nearly identical, with r.s.m.d. of 0.5 Å (**Figure 6C**). In contrast, the crystal structure of free human β_2 m differs significantly (r.m.s.d. of 3.1 Å) from that of the HC-complexed human protein. For example, the HLA-B*2705-complexed β_2 m (2A83)⁽⁴⁸⁾ is distinct at two particular segments: the β -bulge region between the D₁-D₂ strands and the AB loop as described in the previous section. Upon binding to the HC, the two segments of human β_2 m “revert” back to the typical conformations seen in the free forms of chicken, bovine and carp β_2 m (**Figure 6B, C, D**).

We also noticed that the residues in the chicken monomeric β_2 m exhibiting dynamics, such as Asp53, Ser55, Phe56 and Trp60, are stabilized when associated with the HC of the three complexes referred to previously.⁽¹⁹⁻²¹⁾ However, the β_2 m residue Asp58 in the BF2*2101 and YF1*7.1 complexes is an exception, apparently due to the lack of an interacting partner. In contrast, Asp58 of β_2 m in the chicken CD1-1 complex is stabilized by forming a salt bridge with HC residue Lys6. All of the β_2 m residues mentioned above are located beneath the

binding groove. They are those residues mainly responsible for interacting with the HC. It thus seems that these β_2 m residues are crucial for the stabilization of binding grooves of MHC class I molecules. In line with this assumption, Asp53, Ser55, Phe56, and Trp60 are nearly totally conserved among β_2 m molecules (**Figure 4**). Only Ser55 deviates somewhat, since bony fish with the exception of cod, possess alanine at this position. However, it might be that serine is in fact originally found at position 55, since not only all other higher vertebrates, but also cartilaginous fish exhibit this amino acid. It is remarkable that there seem to be in general, only very few residues that are characteristic of one of the large vertebrate taxa. We have previously shown that β_2 m residue 34 provides an example: while this residue is a lysine in birds which forms an inter-chain salt bridge with Asp14 of classical chicken MHC class I molecules, Arg14 of classical mammalian HC cannot engage in a comparable contact and interacts intra-molecular with HC Asp39.⁽¹⁹⁾

The comparison provided in **Figure 4** demonstrates also that proteins sharing ~48% amino acid sequence similarity (chicken and mammalian β_2 m) or only ~35% similarity (chicken and carp β_2 m) can have practically identical structures. It is very likely that a distinct set of highly conserved residues that engage in intra-molecular contacts is responsible for this finding. As hydrophobic intra-chain contacts are too numerous to be listed here, **Table 2** provides all such contacts that are due to salt bridges or hydrogen bonds. Four highly conserved contacts are seen in β_2 m molecules of all four species (R/H12 – N21; N21 – F/Y/V/T22; N/D42 – D/S76/75; N/D42 – E/Q/T76/77; colored yellow in **Table 2**), although different amino acids or atoms can be involved. In addition, there are further contacts that are conserved in three of the four molecules (colored green in **Table 2**). We also notice that the number of salt bridges (coloured pink in **Table 2**) differs drastically between the four⁽⁴⁹⁾ proteins. While the chicken polypeptide exhibits only a single salt bridge near its N-terminus, carp β_2 m possesses three such contacts, in widely distinct locations of the molecule. Furthermore, there are five (human) or even six (bovine) intra-molecular salt bridges, several of them clustered, in the mammalian proteins. Since thermodynamic analyses have only been carried out on the human and the chicken β_2 m, we cannot be sure whether the relative instability of the avian protein is also a consequence of the distinct endorsement of the two proteins with salt bridges. A unifying feature, however, is the presence and topographical conservation of the S-S bridge which is a hallmark of immunoglobulin domains.⁽⁴⁹⁾

Conclusions

It has been estimated that proteins belonging to the immunoglobulin superfamily started to emerge about 500 million years ago in cartilaginous fish,⁽⁵⁰⁾ although their origin might be traced back even to very distantly related polypeptides such as the T cell receptor-like receptor of lamprey⁽⁵¹⁾ or the variable (V) and constant (C) domain-bearing protein of *Amphioxus*⁽⁵¹⁾ PMID: 15749885). Structurally-conserved, but thermodynamically and sequence-divergent β ₂m molecules might be the products of “restrained” evolution, since the monomorphic proteins are under the pressure to bind to a vast variety of HC ranging from polymorphic classical MHC class I to a variety of distantly related MHC class I-like molecules such as CD1. This is also consistent with the fact that β ₂m loci in all species examined to date are unlinked to MHC loci, a feature that has been postulated to help in protecting the β ₂m gene from the polymorphism, duplications and deletions that are so characteristic of many MHC-encoded loci.⁽⁵²⁻⁵⁴⁾

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (grants Na226/12-3, UC8/1-2, and SFB 449/B6). A. Ziegler acknowledges support by the Volkswagen Stiftung (grant I/79 989). B. Loll is grateful for support by the Forschungskommission of the Freie Universität Berlin as well as the Fonds der Chemischen Industrie. C.-S. Hee was supported by the Berliner Krebsgesellschaft and the Deutsche Forschungsgemeinschaft. The cDNA was a kind gift of Dr. M. M. Miller and R. Goto. We are thankful to M. Wahl for continuous encouragement and support. We acknowledge access to beamline BL14.2 of the BESSY II storage ring (Berlin, Germany) *via* the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum and the Leibniz-Institut für Molekulare Pharmakologie.

References

1. Becker JW, Reeke GN, Jr. Three-dimensional structure of β ₂-microglobulin. *Proc Natl Acad Sci U S A* 1985;82(12):4225-9.
2. Hunkapiller T, Hood L. Diversity of the immunoglobulin gene superfamily. *Adv Immunol* 1989;44:1-63.
3. Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J Mol Biol* 1991;219(2):277-319.
4. Zeng Z, Castano AR, Segelke BW, Stura EA, Peterson PA, Wilson IA. Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science* 1997;277(5324):339-45.
5. Otten GR, Bikoff E, Ribaldo RK, Kozlowski S, Margulies DH, Germain RN. Peptide and beta 2-microglobulin regulation of cell surface MHC class I conformation and expression. *J Immunol* 1992;148(12):3723-32.
6. Lancet D, Parham P, Strominger JL. Heavy chain of HLA-A and HLA-B antigens is conformationally labile: a possible role for beta 2-microglobulin. *Proc Natl Acad Sci U S A* 1979;76(8):3844-8.
7. Krangel MS, Orr HT, Strominger JL. Assembly and maturation of HLA-A and HLA-B antigens in vivo. *Cell* 1979;18(4):979-91.
8. Fellous M, Gerbal A, Nobillot G, Weils J. Studies on the biosynthetic pathway of human P erythrocyte antigen using genetic complementation tests between fibroblasts from rare p and Pk phenotype donors. *Vox sanguinis* 1977;32(5):262-8.
9. Dobberstein B, Kvist S, Roberts L. Structure and biosynthesis of histocompatibility antigens (H-2, HLA). *Philosophical transactions of the Royal Society of London* 1982;300(1099):161-72.
10. Williams DB, Barber BH, Flavell RA, Allen H. Role of Beta-2-Microglobulin in the Intracellular-Transport and Surface Expression of Murine Class-I Histocompatibility Molecules. *Journal of Immunology* 1989;142(8):2796-806.
11. Zijlstra M, Bix M, Simister NE, Loring JM, Raulet DH, Jaenisch R. Beta 2-microglobulin deficient mice lack CD4-8+ cytolytic T cells. *Nature* 1990;344(6268):742-6.
12. Peterson PA, Rask L, Ostberg L. beta2-microglobulin and the major histocompatibility complex. *Advances in cancer research* 1977;24:115-63.
13. Vincent C, Revillard JP, Charpentier B. Presence of glycoproteins binding human beta 2 microglobulin in rat serum. *Comptes rendus des seances de l'Academie des sciences* 1981;292(12):755-8.
14. Cejka J, Van Nieuwkoo JA, Mood DW, Kithier K, Radl J. Beta2-microglobulin in human colostrum and milk: effect of breast feeding and physico-chemical characterization. *Clin Chim Acta* 1976;67(1):71-8.
15. Winchester JF, Salsberg JA, Levin NW. Beta-2 microglobulin in ESRD: an in-depth review. *Adv Ren Replace Ther* 2003;10(4):279-309.
16. Floege J, Ehlerding G. Beta-2-microglobulin-associated amyloidosis. *Nephron* 1996;72(1):9-26.
17. Heegaard NH. beta(2)-microglobulin: from physiology to amyloidosis. *Amyloid* 2009;16(3):151-73.
18. Platt GW, Radford SE. Glimpses of the molecular mechanisms of beta2-microglobulin fibril formation in vitro: aggregation on a complex energy landscape. *FEBS letters* 2009;583(16):2623-9.
19. Hee CS, Gao S, Loll B, Miller MM, Uchanska-Ziegler B, Daumke O, Ziegler A. Structure of a Classical MHC Class I Molecule That Binds "Non-Classical" Ligands. *PLoS biology* 2010;8(12):e1000557.
20. Koch M, Camp S, Collen T, *et al.* Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* 2007;27(6):885-99.
21. Dvir H, Wang J, Ly N, Dascher CC, Zajonc DM. Structural basis for lipid-antigen recognition in avian immunity. *J Immunol* 2010;184(5):2504-11.

22. Garboczi DN, Hung DT, Wiley DC. Hla-A2-Peptide Complexes - Refolding and Crystallization of Molecules Expressed in Escherichia-Coli and Complexed with Single Antigenic Peptides. *P Natl Acad Sci USA* 1992;89(8):3429-33.
23. Hee CS, Gao S, Miller MM, Goto RM, Ziegler A, Daumke O, Uchanska-Ziegler B. Expression, purification and preliminary X-ray crystallographic analysis of the chicken MHC class I molecule YF1*7.1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2009;65(Pt 4):422-5.
24. Sherman MA, Goto RM, Moore RE, Hunt HD, Lee TD, Miller MM. Mass spectral data for 64 eluted peptides and structural modeling define peptide binding preferences for class I alleles in two chicken MHC-B haplotypes associated with opposite responses to Marek's disease. *Immunogenetics* 2008;60(9):527-41.
25. Hillig RC, Hülsmeier M, Saenger W, *et al.* Thermodynamic and structural analysis of peptide- and allele-dependent properties of two HLA-B27 subtypes exhibiting differential disease association. *J Biol Chem* 2004;279(1):652-63.
26. Fabian H, Gast K, Laue M, Misselwitz R, Uchanska-Ziegler B, Ziegler A, Naumann D. Early stages of misfolding and association of beta2-microglobulin: insights from infrared spectroscopy and dynamic light scattering. *Biochemistry* 2008;47(26):6895-906.
27. Kabsch W. XDS. *Acta Crystallographica Section D* 2010;66(2):125-32.
28. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *Journal of applied crystallography* 2007;40(Pt 4):658-74.
29. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta crystallographica* 1999;55 (Pt 1):247-55.
30. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta crystallographica* 2010;66(Pt 4):486-501.
31. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta crystallographica* 2001;57(Pt 1):122-33.
32. Laskowski RA, McArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallog* 1993;26:283-91.
33. Chen VB, Arendall WB, III, Headd JJ, *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D* 2010;66(1):12-21.
34. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983;22(12):2577-637.
35. DeLano WL. The PyMOL Molecular Graphics System on World Wide Web <http://www.pymol.org> 2002.
36. Kumar P, Ziegler A, Grahn A, Hee CS, Ziegler A. Leaving the structural ivory tower, assisted by interactive 3D PDF. *Trends Biochem Sci* 2010;35(8):419-22.
37. Englander SW, Kallenbach NR. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Quarterly reviews of biophysics* 1983;16(4):521-655.
38. Bolzani R, Ruggeri F, Olivo OM. Average normal temperature of the chicken in the morning and after 1-2 days of fasting. *Bollettino della Societa italiana di biologia sperimentale* 1979;55(16):1618-22.
39. Wallny HJ, Avila D, Hunt LG, *et al.* Peptide motifs of the single dominantly expressed class I molecule explain the striking MHC-determined response to Rous sarcoma virus in chickens. *P Natl Acad Sci USA* 2006;103(5):1434-9.
40. Ziegler A, Loll B, Misselwitz R, Uchanska-Ziegler B. Implications of structural and thermodynamic studies of HLA-B27 subtypes exhibiting differential association with ankylosing spondylitis. *Adv Exp Med Biol* 2009;649:177-95.
41. Verdone G, Corazza A, Viglino P, *et al.* The solution structure of human β 2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci* 2002;11(3):487-99.
42. Trinh CH, Smith DP, Kalverda AP, Phillips SE, Radford SE. Crystal structure of monomeric human β -2-microglobulin reveals clues to its amyloidogenic properties. *Proc Natl Acad Sci U S A* 2002;99(15):9771-6.

43. Iwata K, Matsuura T, Sakurai K, Nakagawa A, Goto Y. High-resolution crystal structure of β 2-microglobulin formed at pH 7.0. *Journal of biochemistry* 2007;142(3):413-9.
44. Chen W, Gao F, Chu F, Zhang J, Gao GF, Xia C. Crystal structure of a bony fish β 2-microglobulin: insights into the evolutionary origin of immunoglobulin superfamily constant molecules. *J Biol Chem* 2010;285(29):22505-12.
45. Okon M, Bray P, Vucelic D. ^1H NMR assignments and secondary structure of human β 2-microglobulin in solution. *Biochemistry* 1992;31(37):8906-15.
46. Macdonald IK, Harkiolaki M, Hunt L, *et al.* MHC class I bound to an immunodominant *Theileria parva* epitope demonstrates unconventional presentation to T cell receptors. *PLoS pathogens* 2010;6(10):e1001149.
47. Girardi E, Wang J, Mac TT, *et al.* Crystal structure of bovine CD1b3 with endogenously bound ligands. *J Immunol* 2010;185(1):376-86.
48. Rückert C, Fiorillo MT, Loll B, *et al.* Conformational dimorphism of self-peptides and molecular mimicry in a disease-associated HLA-B27 subtype. *J Biol Chem* 2006;281(4):2306-16.
49. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 1987;329(6139):506-12.
50. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet* 2010;11(1):47-59.
51. Pancer Z, Mayer WE, Klein J, Cooper MD. Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. *P Natl Acad Sci USA* 2004;101(36):13273-8.
52. Ono H, Figueroa F, O'HUigin C, Klein J. Cloning of the beta 2-microglobulin gene in the zebrafish. *Immunogenetics* 1993;38(1):1-10.
53. Horton R, Wilming L, Rand V, *et al.* Gene map of the extended human MHC. *Nat Rev Genet* 2004;5(12):889-99.
54. Kelley J, Walter L, Trowsdale J. Comparative genomics of major histocompatibility complexes. *Immunogenetics* 2005;56(10):683-95.

Table 1: Data collection and refinement statistics

Data collection	
PDB entry	3O81
Space group	$P2_1$
Unit cell a; b; c [Å]	36.9; 30.5; 87.4
α ; β ; γ [°]	90; 87.4; 90
Resolution [Å] ^a	30.0 - 2.0 (2.1 - 2.0)
Completeness ^a	96.0 (92.2)
$\langle I/\sigma(I) \rangle$ ^a	12.5 (6.8)
R_{meas} ^b	8.8 (19.3)
Redundancy ^a	3.6 (3.5)
Refinement	
Non-hydrogen atoms	1765
R_{work} ^{a, c}	19.5 (19.3)
R_{free} ^{a, d}	24.2 (22.0)
$\beta_2\text{m}$ (chain A), no. of atoms / average B factor [Å ²]	810 / 18.0
$\beta_2\text{m}$ (chain B), no. of atoms / average B factor [Å ²]	801 / 19.0
Water, no. of molecules / average B factor [Å ²]	160 / 23.8
Rmsd ^e from ideal geometry, bond length [Å]	0.011
bond angles [°]	1.271

^a values in parentheses refer to the highest resolution shell.

^b $R_{\text{meas}} = \sum_h [n/(n-1)]^{1/2} \sum_i |I_h - I_{h,i}| / \sum_h \sum_i I_{h,i}$. where I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.

^c $R_{\text{work}} = \sum_h |F_o - F_c| / \sum F_o$ (working set, no σ cut-off applied).

^d R_{free} is the same as R_{cryst} , but calculated on 5% of the data excluded from refinement.

^e Root-mean-square deviation (Rmsd) from target geometries.

3. Comparative Biophysical Characterizations of Human and Chicken β_2 -microglobulin and the Evolution of Interfaces in MHC Class I Molecules

Table 2. Intra-molecular contacts in human (PDB entry 1LDS), bovine (1BMG), chicken (3O81) and carp (3GBL) β_2 m. Only hydrogen bonds and salt bridges (shaded in pink) are listed in the table. Completely conserved interactions in the β_2 m of the four species are shaded in yellow and semi-conserved interactions are shaded in green.

Human β_2 m		Distance (Å)	Bovine β_2 m		Distance (Å)	Chicken β_2 m		Distance (Å)	Carp β_2 m		Distance (Å)
I1 ^O	H31 ^{ND1}	2.7									
Q2 ^O	R3 ^{NH1}	3.0				D2 ^{OD1}	L3 ^N	3.4			
						D2 ^{OD1}	T4 ^{OG1}	3.2			
						D2 ^{OD1/2}	K6 ^{NZ}	3.3/3.1			
			R3 ^{NH1}	S60 ^{OG}	3.1				S3 ^{OG1}	G29 ^O	2.8
									S3 ^{OG1}	H31 ^{N/ND1}	3.3/3.5
T4 ^{OG1}	T86 ^{OG1}	3.5				T4 ^O	K6 ^{NZ}	2.8			
T4 ^{OG1}	P5 ^N	3.3				T4 ^{OG1}	K6 ^{NZ}	3.4			
			P5 ^O	R90 ^{NH1}	3.4						
Y10 ^O	N24 ^{ND2}	3.6				Y10 ^O	N24 ^{ND2}	3.3			
S11 ^O	R12 ^{NH1}	2.9				S11 ^N	W95 ^{NE1}	3.5			
S11 ^{OG}	W95 ^{NE1}	3.7				S11 ^{OG}	W95 ^{NE1}	3.6			
						S11 ^{OG}	F13 ^O	2.9	S11 ^{OG}	Y13 ^O	2.7
R12 ^O	N21 ^{OD1}	3.5	R12 ^N	N21 ^{ND2}	3.4	R12 ^N	N21 ^{OD1}	3.3	H12 ^N	N21 ^{OD1}	3.5
			H13 ^N	N21 ^{ND2}	3.1	F13 ^N	N21 ^{OD1}	3.1	Y13 ^N	N21 ^{OD1}	3.1
			H13 ^O	N21 ^{OD1}	2.6	F13 ^O	N21 ^{ND2}	3.0	Y13 ^O	N21 ^{ND2}	2.9
H13 ^{ND1}	N21 ^{OD1}	3.8									
H13 ^{ND1}	G18 ^O	3.2									
H13 ^{ND1}	S20 ^O	2.9									
									P14 ^O	N21 ^{ND2}	3.4
			E16 ^{OE1}	K19 ^{NZ}	3.1	S16 ^O	T19 ^{OG1}	3.5	E16 ^{OE2}	K19 ^{NZ}	3.7
						S16 ^{OG}	T19 ^{OG1}	3.0			
N17 ^{OD1}	S20 ^{OG}	3.0	D17 ^{OD1}	R96 ^{NH1/2}	3.2/3.4						
			D17 ^{OD2}	S73 ^N	3.2						
									G18 ^N	E73 ^{OE2}	3.5
S20 ^{OG}	E69 ^{OE2}	3.1							E20 ^{OE1}	S69 ^{OG}	3.2
N21 ^{OD1}	F22 ^N	3.1	N21 ^{ND2}	Y22 ^{N/O}	2.6/3.3	N21 ^{OD1}	V22 ^N	2.8	N21 ^{OD1}	T22 ^{N/O}	2.7/3.5
									N21 ^O	T22 ^{OG1}	3.5
									N21 ^O	S69 ^{OG}	3.4
F22 ^O	N24 ^{ND2}	3.0				V22 ^O	N21 ^{ND1}	3.4			
			H31 ^{ND1}	P32 ^N	3.3	H31 ^O	H84 ^{NE2}	3.1	H31 ^O	H84 ^{NE2}	3.6
P32 ^O	H84 ^{NE2}	2.8	P32 ^O	H83 ^{NE2}	2.7	P32 ^N	H31 ^{ND1}	3.3	P32 ^O	H84 ^{NE2}	2.9
S33 ^{OG}	D34 ^N	3.2									
D34 ^O	H84 ^{NE2}	3.7	Q34 ^O	H83 ^{NE2}	3.7	K34 ^O	H84 ^{ND1}	3.1	D34 ^{OD1}	I35 ^N	3.4
E36 ^{OE1}	N83 ^{ND2}	3.4	E36 ^{OE1}	K82 ^{NZ}	3.6						
V37 ^O	D38 ^{OD1}	3.5							I37 ^O	D38 ^{OE1}	3.6
									I37 ^O	K66 ^{NZ}	2.8
D38 ^{OD1}	R45 ^{NE}	2.7	D38 ^{OD1}	K45 ^{NZ}	3.8				E38 ^{OE1}	K66 ^{NZ}	3.5
D38 ^{OD2}	R81 ^{NH2}	3.6	D38 ^{OD2}	R80 ^{NH2}	3.0						
N42 ^{ND2}	D76 ^{OD1}	2.7	N42 ^{ND2}	D75 ^{OD1}	3.4	D42 ^{OD2}	S76 ^{OG}	2.7	N42 ^{ND2}	D76 ^{OD1}	3.1
N42 ^{OD1}	E77 ^N	2.9	N42 ^{OD1}	Q76 ^N	3.2	D42 ^{OD1}	T77 ^N	3.0	N42 ^{OD1}	E77 ^N	2.7
R45 ^{NH1}	E47 ^{OE1/2}	2.7/3.1									
R45 ^{NH2}	E47 ^{OE2}	3.4									
K48 ^{NZ}	E50 ^{OE1}	2.6	K47 ^{NZ}	E68 ^{N/OE1}	3.4/3.6				D48 ^{OD1}	Q50 ^{NE2}	3.2
			E49 ^O	H66 ^{ND1}	3.4						
E50 ^N	T68 ^{OG1}	3.1				Q50 ^{OE1}	H67 ^{NE2}	3.6			
E50 ^O	H51 ^{ND1}	3.6									
S52 ^O	D53 ^{OD1}	3.5	S51 ^{OG}	L64 ^{N/O}	3.0/3.2	S52 ^O	R64 ^{NE/NH2}	2.8/3.3			

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

						S52 ^{OG}	D53 ^{NO}	3.3/2.7	T52 ^{OG1}	D53 ^O	3.6
						S52 ^{OG}	L65 ^{NO}	3.0/3.5	T52 ^{OG1}	T65 ^{NO}	2.8/3.5
			D52 ^N	S51 ^{OG}	3.4						
			D52 ^O	S54 ^{OG}	3.4						
			L53 ^O	S54 ^{OG}	3.6	M54 ^O	S55 ^{OG}	3.2			
S55 ^{OG}	F56 ^N	3.5				S55 ^{OG}	Q63 ^{NE2}	3.5			
S57 ^{OG}	K58 ^N	2.6	S56 ^{OG}	K57 ^N	2.9	N57 ^{ND2}	Q63 ^{OE1}	2.9	E57 ^{OE1/2}	H63 ^{NE2}	3.4/2.9
S57 ^{OG}	Y63 ^{OH}	3.3	S56 ^{OG}	D58 ^N	3.4				E57 ^{OE1}	Q61 ^{NE2}	3.5
						D58 ^{OD1}	D59 ^N	2.7			
						D59 ^{OD1/2}	T61 ^{OG1}	2.7/3.5			
						D59 ^{OD1}	W60 ^N	3.3			
						D59 ^{OD1}	T61 ^N	3.2			
S61 ^{OG}	Y63 ^{OH}	2.8									
						F62 ^N	T61 ^{OG1}	3.4			
			S65 ^{OG}	H66 ^N	3.6						
T68 ^{OG1}	E69 ^N	3.1									
									S69 ^{OG}	F70 ^N	3.3
P70 ^O	T71 ^{OG1}	3.5	F69 ^O	T70 ^{OG1}	3.8						
			P71 ^O	N72 ^{ND2}	2.6						
T73 ^{OG1}	K75 ^N	3.6				S73 ^N	Y78 ^{OH}	3.1	E73 ^{OE2}	G18 ^N	3.5
T73 ^{OG1}	D76 ^{N/OD2}	3.5/2.7				S73 ^{OG}	S76 ^{OG}	2.9			
						S73 ^{OG}	Y78 ^{OH}	3.7			
			Q76 ^{OE1/NE2}	K93 ^{NZ}	3.0/3.0				S75 ^{OG}	D76 ^N	3.4
									E77 ^{OE2}	K90 ^{NZ}	2.8
N83 ^{OD1}	P90 ^N	3.6							R83 ^{NH1}	T88 ^{OG1}	2.8
									R83 ^{NH1/2}	S86 ^O	3.5/3.2
			H83 ^{ND1}	T85 ^{OG1}	2.8						
H84 ^{ND1}	T86 ^{OG1}	2.8				H84 ^{NE2}	T86 ^{OG1}	3.5			
T86 ^{OG1}	L87 ^N	3.4	T85 ^{OG1}	L86 ^N	3.4	T86 ^{OG1}	L87 ^N	3.3			
S88 ^O	Q89 ^{NE2}	3.0									
S88 ^{OG}	Q89 ^N	3.2									
						E89 ^O	Q91 ^{NE2}	3.0			
						P90 ^O	Q91 ^{NE2}	3.1			
									E94 ^{OE2}	N96 ^{ND2}	2.8
									S95 ^{OG}	N96 ^N	3.3
			D95 ^{OD1}	R96 ^N	3.0						

Figure captions

Figure 1. Thermodynamic stabilities of β_2 m and class I complexes of the chicken as determined by differential scanning calorimetry (DSC) and infrared (IR) spectroscopy. Experimental excessive heat capacity curves (black dashed lines) and deconvolution curves (blue lines) as well as the calculated fit (red lines) obtained from DSC of (A) free β_2 m, (B) BF2*2101:pC1 complex, (C) YF1*7.1 complex, and (D) BF2*13:pMDV complex. (E) The table summarizes the deconvoluted melting temperatures (T_m) and enthalpy changes (ΔH_m) from DSC experiments. ΔH_m^{tot} is the total enthalpy change that accompanies a given melting process. Loss of residual unexchanged amide groups in free chicken β_2 m (filled circles) and free human β_2 m (open circles) at 25 °C (f) and 37 °C (g) as monitored by IR spectroscopy. X is the fraction of the unexchanged amide protons as calculated by setting the difference in peak intensity at 1550 cm^{-1} (amide II band) between the spectra in H_2O -buffer and after 24 h in D_2O -buffer at 37 °C to 100%.

Figure 2. Structural details of chicken β_2 m. (A) Cartoon representation of chicken β_2 m. The rainbow chain is colored from N-terminal (blue) to C-terminal (red) with the β -strands labeled from A to G. The first β -sheet is formed by β -strands A, B, and E while the second β -sheet comprises β -strands C, D₁, D₂, F and G. The two β -sheets are connected by a disulfide bond which is shown as black ball-and-stick representation. (B and C) The electron densities derived from $2F_o-F_c$ maps after refinement are contoured at a level of 1.5σ and shown as blue meshes. (B) Residues of free β_2 m with multiple conformations or poorly defined electron density are displayed as yellow stick representation. Double conformations are indicated with (i) and (ii) following residue numbering. (C) The flexible residues of chicken β_2 m are stabilized (orange stick representation) upon binding to YF1*7.1 HC (grey cartoon representation).

Figure 3. Crystal lattice and the interface contacts of chicken β_2 m. (A) The two β_2 m molecules in the asymmetric unit are colored as orange and blue. The residues involved in the head-to-head arrangement of the two β_2 m molecules are exclusively hydrophobic. Further interface contacts are observed with the neighboring symmetry-related molecules which are colored as green and yellow. Hydrogen bonds and salt bridges are indicated with black and

red dashed lines, respectively. (B) The interface contact details as shown in (A). β_2 m residues are color-shaded according to the β_2 m molecules in (A). Hydrogen bonds and salt bridges are listed with a distance cut-off of 3.5 Å and hydrophobic contacts with a distance cut-off of 4.0 Å. HB = hydrogen bonds, SB = salt bridges and vdW = van der Waals interactions.

Figure 4. Amino acid sequence alignment of β_2 m from various species. Numbering refers to the human β_2 m sequence. Completely conserved residues are highlighted with red boxes and residues with more 70% sequence similarity are shown as bold letters and boxed in yellow. Black arrows indicate with β -strands A to G and are displayed above the alignment according to human β_2 m (PDB entry 1LDS).

Figure 5. Average B-factor plots of free (blue line) and YF1*7.1-bound chicken β_2 m (red lines) as well as free human β_2 m (green line). The average B-factor values of free chicken β_2 m was calculated from the B-factor of the two monomers in the asymmetric unit (PDB entry 3O81) and the values for YF1*7.1-bound β_2 m was an average between the B-factor of β_2 m bound to two YF1*7.1 structures (PDB entries 3P73 and 3P33).⁽¹⁹⁾ The average values for free human β_2 m was determined from the structure with PDB entry 2YXF.⁽⁴³⁾ β -strands A to G are labeled and indicated above the plots with black arrows. Residues located in the loop regions have generally higher B-factor values than the residues assigned as β -strand. Free chicken β_2 m residues are significantly more flexible than free human β_2 m (in regions between BC, D₁E and FG strands) and YF1*7.1-bound chicken β_2 m (in regions between BC and D₁E strands).

Figure 6. Comparison of β_2 m structures. (A) Overlay of monomeric β_2 m structures from chicken, human, bovine and carp. Segments with drastic conformational changes between molecules are highlighted with grey box or ellipses. Human β_2 m crystal structure differs considerably from the other β_2 m structures at two parts: the AB loop and the absence of a β -bulge between D₁-D₂ strands. Other significant differences between the β_2 m molecules compared are the C'D₁ and DE loops that is due to high flexibility of these segments. (B) Overlay of monomeric chicken β_2 m (gg β_2 m) and gg β_2 m bound to chicken classical class I BF2*2101, YF1*7.1 as well as non-classical CD1-1 molecule. (C) Overlay of monomeric

bovine β_2m (bt β_2m) and bt β_2m associated with BoLA-N*01301 and bovine CD1b3. Like the chicken β_2m structures, bovine β_2m are structurally identical whether in the monomeric or bound form. (D) Overlay of monomeric human β_2m (hs β_2m) obtained from X-ray crystallization (green) or NMR (blue) together with hs β_2m bound to HLA-B*2705 and human CD1b molecules. Areas bordered with grey lines indicate the drastic difference in the crystal structure of hs β_2m (green) compared to the other hs β_2m . For the ease of comparison, a three-dimensional figure is embedded here; see “Material and Methods” for instructions.

Figure 1

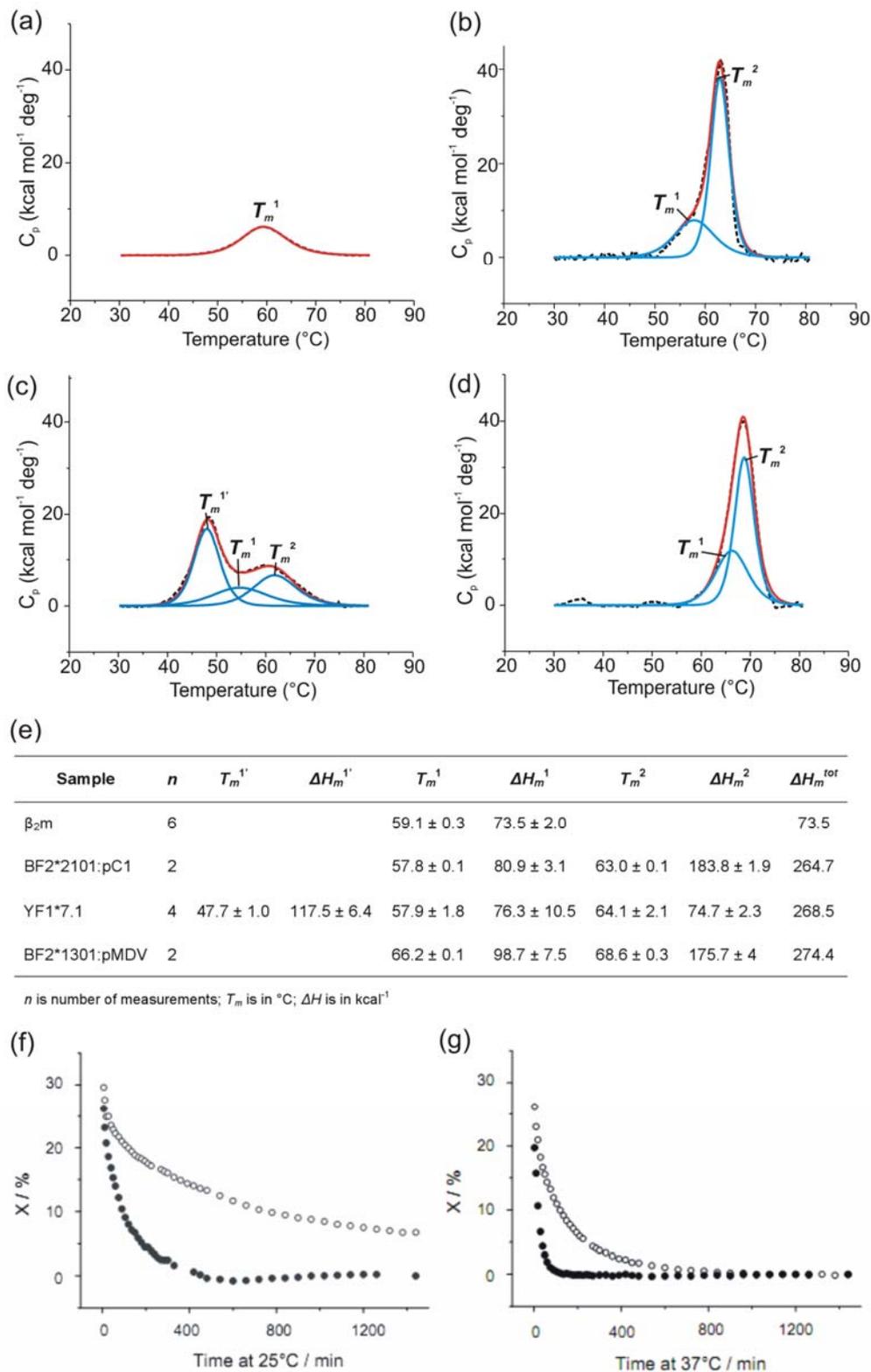


Figure 2

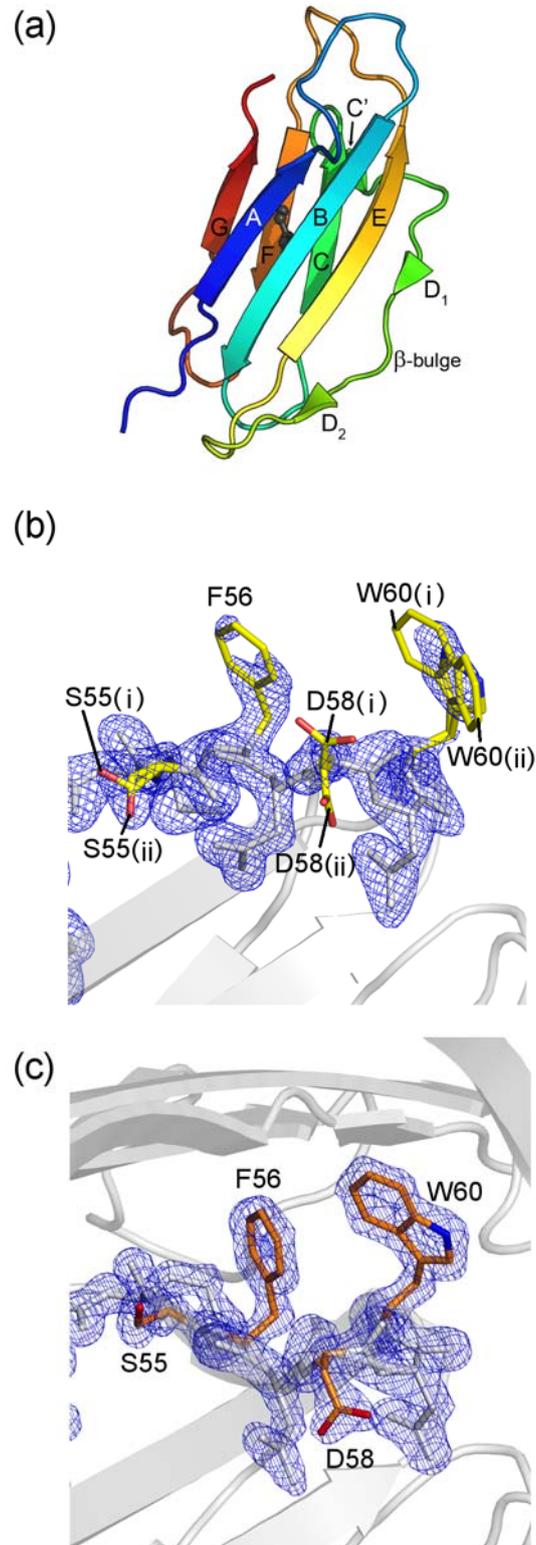
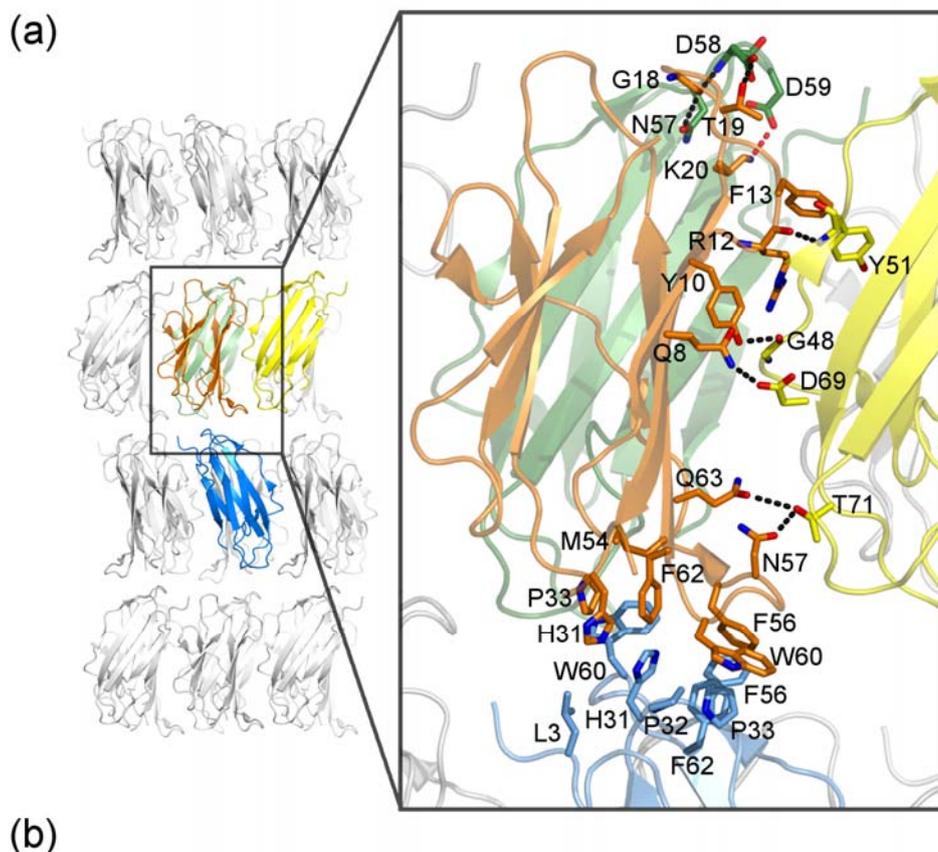


Figure 3



		Distance (Å)	Interaction
Q8 ^{NE2}	D69 ^{OD2}	2.77	HB
Y10 ^{OH}	G48 ^O	2.71	HB
R12 ^O	Y51 ^N	2.76	HB
F13	Y51	3.54-3.96	π -stack
G18 ^O	N57 ^{OD1}	3.46	HB
G18 ^O	D58 ^N	2.68	HB
T19 ^{OG1}	D58 ^{OD2}	2.39	HB
K20 ^{NZ}	D59 ^{OD2}	3.51	SB
H31	L3	3.80	vdW
H31	H31	3.31-3.72	vdW
P33	W60	3.42-3.98	vdW
M54	W60	3.49-3.80	vdW
F56	F62	3.95-3.98	vdW
F56	F56	3.32-3.94	vdW
N57 ^{OD1}	T71 ^{OG1}	2.80	HB
W60	P33	3.42-3.98	vdW
W60	F62	3.26-3.99	vdW
F62	W60	3.26-3.99	vdW
Q63 ^{OE1}	T71 ^{OG1}	3.41	HB

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

Figure 4

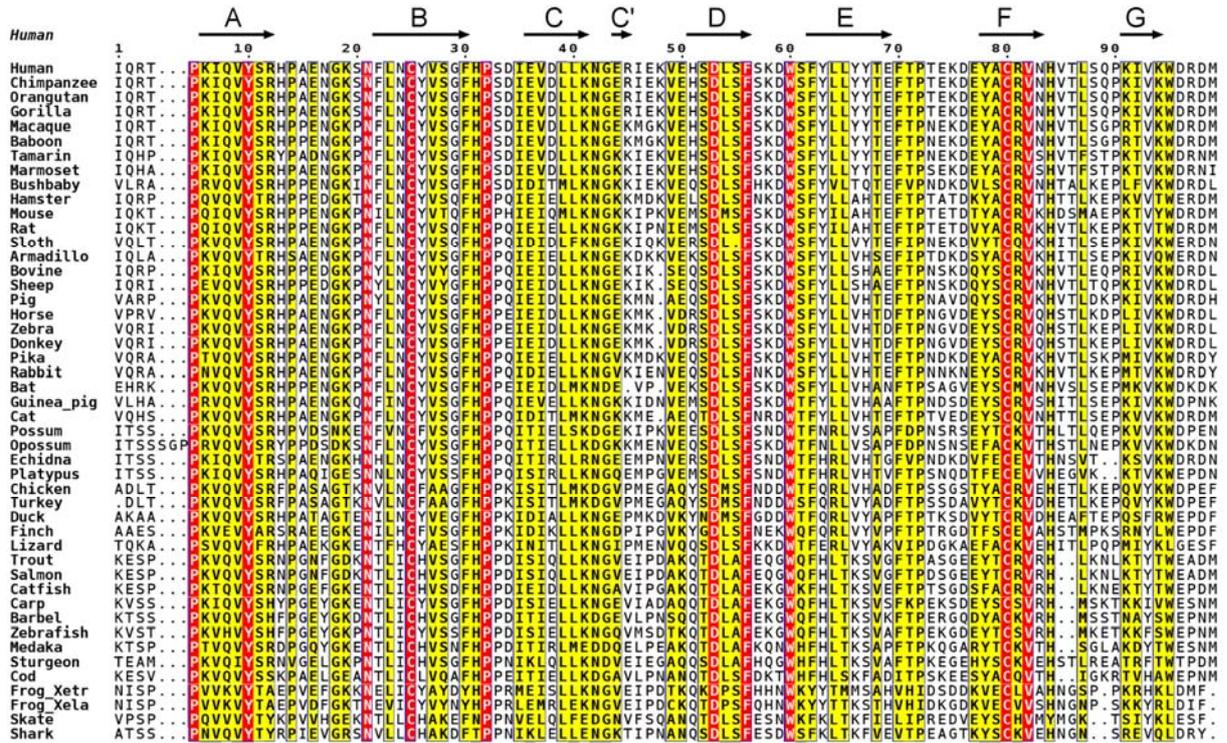


Figure 5

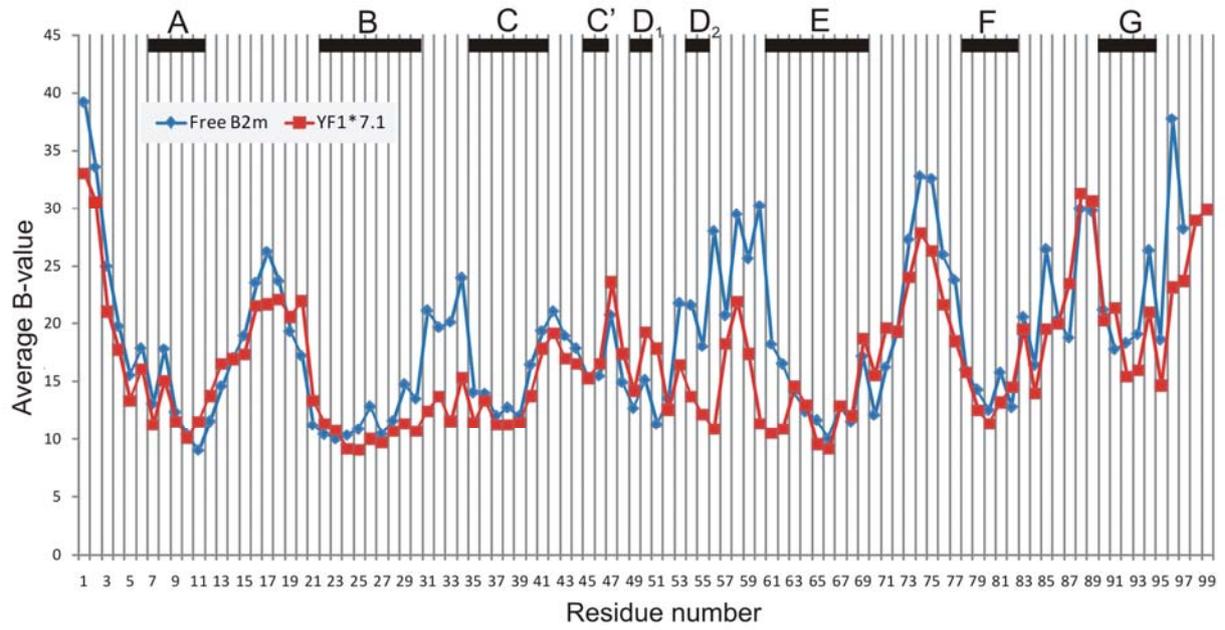
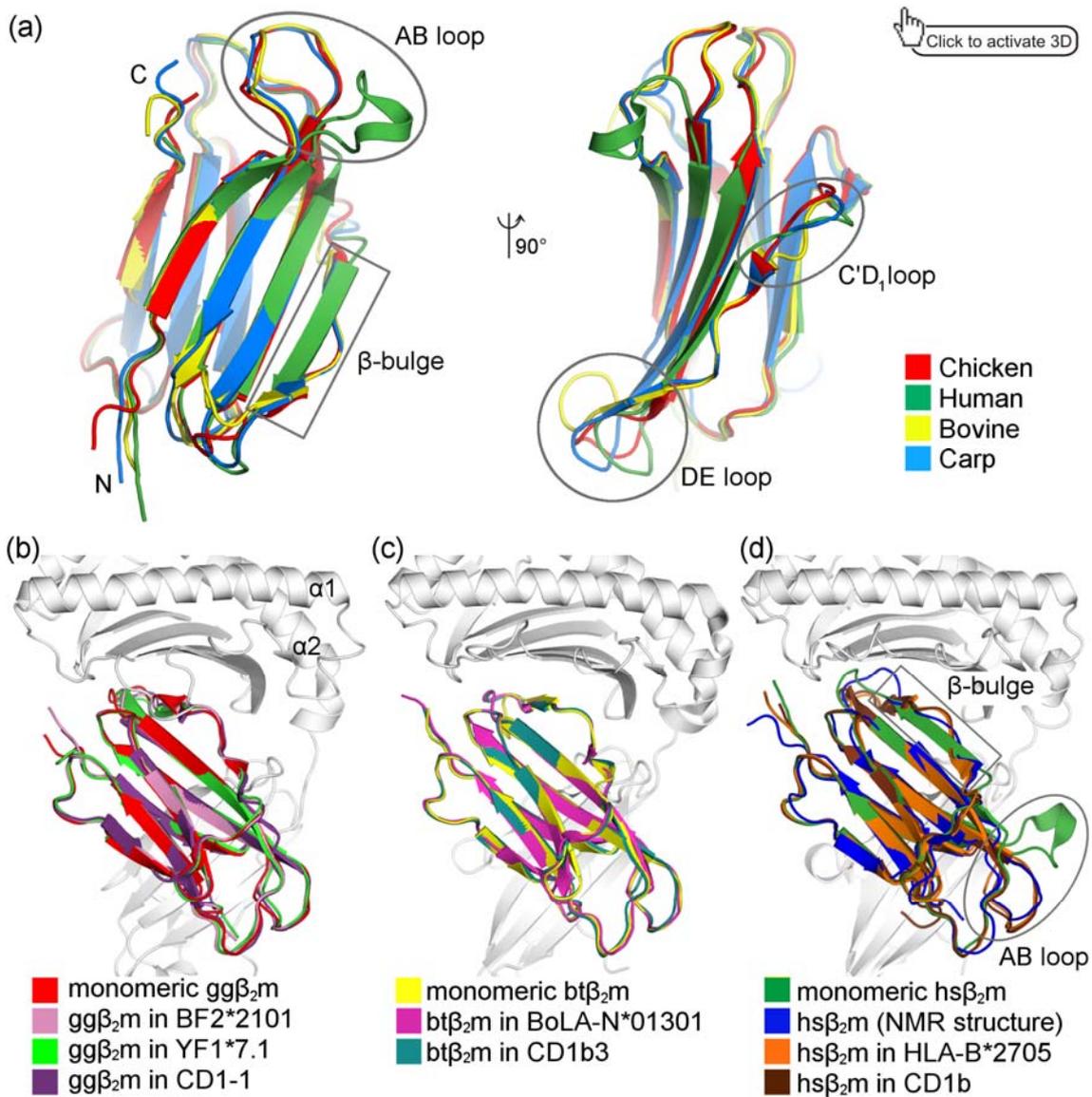


Figure 6



Dynamics of free versus complexed β_2 -microglobulin and the evolution of interfaces in MHC class I molecules

Chee-Seng Hee^a, Monika Beerbaum^b, Annette Diehl^b, Bernhard Loll^c, Peter
Schmieder^b, Barbara Uchanska-Ziegler^a, and Andreas Ziegler^{a,&}

^a Institut für Immungenetik, Charité - Universitätsmedizin Berlin, Campus Benjamin
Franklin, Freie Universität Berlin, Thielallee 73, 14195 Berlin, Germany

^b Leibniz-Institut für Molekulare Pharmakologie, Robert Rössle-Strasse 10, 13125
Berlin, Germany

^c Institut für Chemie und Biochemie, Abteilung Strukturbiochemie, Freie Universität
Berlin, Takustrasse 6, 14195 Berlin, Germany

& Correspondence:

andreas.ziegler@charite.de

Tel: +49-30-450 564731

Fax: +49-30-450 564920

Abstract

In major histocompatibility complex (MHC) class I molecules, monomorphic β_2 -microglobulin (β_2m) is non-covalently bound to a heavy chain (HC) exhibiting a variable degree of polymorphism. β_2m can stabilize a wide variety of HC ranging from classical peptide-binding to non-classical lipid-presenting MHC class I molecules as well as to MHC-class I-like molecules that do not bind small ligands. In this study, we aim to understand the evolution of the interface between β_2m and different HC by scrutinizing their interactions in MHC class I complexes across species. Using human β_2m and the HLA-B*27:05 complex as a model system, a comparison of free and HC-bound β_2m by nuclear magnetic resonance spectroscopy was initially carried out, revealing that several β_2m residues involved in interface contacts with the exterior of the peptide binding groove exhibit considerable dynamics in the unbound form, but gain stabilization upon binding to the HC. A number of these residues engage in evolutionarily retained inter-polypeptide chain contacts that are detectable in species as diverse as mammals and chicken. In addition, sequence analyses demonstrate that these contacts may be present in cold-blooded vertebrates as well. We also identified interface interactions that are characteristic for a particular group of MHC class I complexes such as peptide- or lipid-binding molecules or are limited to a specific vertebrate taxon. Taken together, this study demonstrates that β_2m is an unusually versatile molecule which fulfils its role of stabilizing a wide range of MHC class I and class I-like complexes by engaging in intermolecular contacts that may be absolutely conserved, semi-conserved or characteristic for a group of complexes or particular vertebrate taxa.

Introduction

β_2 -microglobulin (β_2m) is a molecule that can associate non-covalently with a wide range of Major Histocompatibility Complex (MHC) class I heavy chains (HC) (Bjorkman et al., 1987). For example, the human HLA-A, -B, -C, -E and -G HC all interact with β_2m (Bjorkman et al., 1987; O'Callaghan et al., 1998; Boyington et al., 2000; Hülsmeier et al., 2002; Clements et al., 2005). Furthermore, β_2m is able to bind also to a variety of non-MHC encoded class I-like HC such as CD1 (Zeng et al., 1997; Zajonc et al., 2003). and exerts an impact on the stability and conformation of MHC class I HC that strongly influences peptide presentation (Pedersen et al., 1994; Shields et al., 1998). Several studies have also shown an involvement of β_2m in the function of class I molecules in the context of recognition by receptors on T cells, natural killer (NK) cells, and other cell types (Michaelsson et al., 2001; Wang et al., 2002; Mitsuki et al., 2004). Although free β_2m molecules of mammals share less than 50% of amino acid sequence similarity with chicken or fish β_2m , their overall structures are virtually identical (Becker and Reeke, Jr., 1985; Trinh et al., 2002; Chen et al., 2010). Due to this structural similarity, interactions between β_2m and HC from two highly unrelated species are possible. For instance, human β_2m has been shown to associate with two very distinct HC, the classical MHC class I mouse H-2D^b and the non-classical chicken CD1-2, with an enhanced stability (Achour et al., 2006; Zajonc et al., 2008).

Until a few years ago, the structures of MHC class I complexes had been determined exclusively from human, mouse, and rat (Maenaka and Jones, 1999). However, these studies were recently extended to a non-human primate (Chu et al., 2007), bovine (Li et al., 2011) and even to the chicken. In case of the latter, the structures of a peptide-binding classical (Koch et al., 2007) and two lipid-binding non-classical CD1 molecules (Dvir et al., 2010; Zajonc et al., 2008) are available. In addition, the structure of an unusual, lipid-binding classical complex (YF1*7.1) has recently been determined (Hee et al., 2010).

The obvious versatility of β_2m prompted us to investigate two aspects using a combined experimental and *in silico* approach. By employing Nuclear Magnetic Resonance (NMR) spectroscopy, we investigated whether all β_2m residues that engage in interactions with the HC exhibit exceptional dynamic properties when compared to other parts of the molecule. In addition, we analyzed the interface

interactions established between β_2 m and classical as well as non-classical class I HC in order to identify any general features that might be shared between these diverse complexes, but also differences characterizing particular groups of HC. Our study shows (i) how the exceptional dynamics exhibited by selected residues of free β_2 m can be harnessed following the interaction with an HC, and it reveals (ii) that particular groups of HC can be distinguished by the presence or absence of distinct β_2 m-HC interactions. Furthermore, we show (iii) that certain interchain contacts appear to be restricted to particular vertebrate taxa.

Results and Discussions

Flexible free β_2 m residues are stabilized by HC

In order to investigate differences between free and MHC-bound β_2 m, we used heteronuclear NMR spectroscopy, in particular ^{15}N -relaxation data, which are capable of providing a picture of the mobility of a molecule at atomic resolution. Free human β_2 m has been extensively studied using solution state NMR before (Verdone et al., 2002; Esposito et al., 2005; Platt et al., 2005), and an assignment of the ^1H , ^{15}N -Heteronuclear Single Quantum Coherence (HSQC) spectrum has already been obtained (McParland et al., 2002). Significantly fewer than the expected 94 signals (there are 5 Pro residues within the protein) are observable in a spectrum, indicating the loss of the other signals due to exchange broadening. We re-analyzed the ^1H , ^{15}N -HSQC of free β_2 m and were able to assign 75 of the 94 signals using triple resonance NMR experiments (**Fig. 1A**) (Cavanagh et al., 1996). The signals that are missing are those belonging to the amino acids 30-33, 53-63, 88-89 as well as residues at the N-terminus. These residues correspond to that side of β_2 m that contacts the bottom of the peptide binding groove (BC-loop, DE-loop, FG-loop, N-terminus) when bound to the HC in an MHC molecule (**Fig. 1B**). It should be noted, however, that they also include the second half of the proposed D-strand. This is a first indication that this strand is not present in solution as is further confirmed by relaxation data (see below).

We also investigated β_2 m in an MHC class I molecule, i.e. bound to the HC of HLA-B*27:05 (in short, B*27:05) presenting pVIPR, a vasoactive intestinal peptide type 1 receptor-derived peptide with the sequence RRKWRRWHR (**Fig. 1C**) (Hülsmeier et

al., 2004). Because of the size of the complex (in which only β_2 m was labeled for NMR studies), the assignment of the resonances had to be performed using Transverse relaxation optimized spectroscopy (TROSY)-based triple resonance experiments and employing deuteration of the protein (Pervushin et al., 1997; Tugarinov et al., 2004), although the number of expected signals stayed the same. Interestingly, a more extensive ^1H , ^{15}N -HSQC spectrum was obtained and all but four residues of those undetectable in the case of free β_2 m could be unambiguously assigned. A comparison between the two spectra is shown in **Fig. 1D**. These results demonstrate that the dynamics of uncomplexed β_2 m changes significantly upon binding to the HC.

Based on these assignments, we measured ^{15}N -heteronuclear Nuclear Overhauser Effects (NOEs) that are sensitive to mobility (Farrow et al., 1994; Palmer, III, 2004). Values can, of course, only be obtained for residues that show signals in an ^1H , ^{15}N -correlation and in the absence of overlap in such spectra. In case of free β_2 m, an average value of 0.8 was obtained for the residues in the β -sheets. Not unexpectedly, slightly reduced values were found in those loop regions that show detectable signals as well as the N- and C-termini, but a significant reduction of the values could be detected for residues 48-52. These are the residues preceding those undetectable in the putative D-strand and the DE loop, indicating that not only this loop is mobile, but that the whole region downstream of the C' strand is more flexible than the remainder of the protein. This would explain the different outcome of the structure determinations performed so far (Trinh et al., 2002; Verdone et al., 2002; Iwata et al., 2007), since the structure of such a flexible region would strongly depend on the experimental conditions.

However, the situation is different for β_2 m bound to B*27:05. In this case, the whole protein exhibits fairly uniform values for the heteronuclear NOE, indicating a less mobile polypeptide chain, in line with the uniform linewidth of nearly all residues. Small effects are still visible for the BC and the FG loops and in particular for the region preceding the putative D-strand, which still appears to be more flexible than the rest of the protein. We expect that this region will also exhibit differences when distinct MHC class I complexes would be compared. Interestingly, the region of the EF loop appears to become more flexible than it is in free β_2 m (**Fig. 1D**).

To understand how β_2 m residues exhibiting dynamics are stabilized or destabilized by the interaction with an HC, we scrutinized the β_2 m-HC interfaces for which structural information is available.

Analyzing β_2 m-HC interface interactions

An alignment of β_2 m protein sequences reveals that the sequences are highly conserved between species (**Fig. S1**). Human β_2 m shares between 55-100% of its sequence with that of β_2 m from other mammals, slightly less than 50% with avian species, about 31% with frogs, about 44% with fish, and 30-40% with β_2 m from cartilaginous fishes. Ten residues are identical in all species compared in this study, including the two characteristically spaced cysteine residues which form a disulfide bridge that constitutes a typical feature of the immunoglobulin superfamily (IgSF) domain (Becker and Reeke 1985; Trinh, Smith et al. 2002; Chen, Gao et al. 2010). Another 50 positions share more than 70% sequence similarity comprising residues with comparable side chain characteristics. This similarity between the β_2 m sequences prompted us to analyze in detail if also these highly conserved residues are involved in interface interactions with HC residues. Furthermore, we intended to understand, whether differences could be found among these amino acid residues when classical, non-classical and MHC class I-like complexes are compared. Finally, we investigated whether the dynamics exhibited by individual β_2 m residues could be correlated with their involvement in inter-molecular interfaces.

To date, more than two hundred MHC class I or class I-like structures, mostly from human and mouse, are available from the Protein Data Bank (PDB). The recent determination of classical class I structures from chicken, like BF2*2101 and YF1*7.1 as well as the non-classical CD1-1 and CD1-2 complexes permit now a comprehensive cross-species study of the interactions between β_2 m and HC to be carried out. However, our study had to be confined to mammals and chicken, because MHC class I structures of lower vertebrates have not been determined to date. To identify the conserved interactions, we chose a “bottom-up” approach by which conserved β_2 m-HC interactions were first identified in the four chicken structures. This information was then used for a comparison with MHC class I and class I-like molecules of mammalian species.

β_2 m-HC interface interactions in chicken molecules

First we evaluated two classical MHC class I molecules of the chicken, the peptide-presenting BF2*2101 molecule and the lipid-displaying YF1*7.1 complex. In the two BF2*2101 complexes determined (Koch et al., 2007), 15 β_2 m residues are involved in interface interactions with the HC. Three of these (Tyr10, Asp53, Trp60) are conserved in all species studied (**Fig. S1** and **Table 1**). The YF1*7.1 molecule (Hee et al., 2010), shares most of the β_2 m-BF2*2101 interactions (**Table 1**). However, only about half of these are shared by the non-classical CD1 molecules of the chicken (**Table 1**). Although the structure of chicken CD1-2 was solved in complex with *human* β_2 m, most of the β_2 m-HC interactions are identical to the CD1-1 complex, whose structure was determined with *chicken* β_2 m (**Table 1**). It is obvious that many of the conserved interactions in the structures of the four distinct HC either involve the same β_2 m and HC residues or amino acids with closely related properties. For example, the van der Waals contact between Trp60 and Ala109, 113, or 114 (HC residues of CD1-1, YF1*7.1, or BF2*2101) involves Trp60 of human β_2 m and Gly113 of the chicken CD1-2 HC.

Highly conserved β_2 m-HC interface interactions

Based on the results from the chicken molecules, we extended this comparison to mammalian MHC class I and class I-like molecules (**Table 1**). For the sake of simplicity, we summarize the most important findings of **Table 1** in **Table 2**. MHC class I HC can be divided into five groups: (i) classical, β_2 m- and peptide-binding HC, (ii) classical, β_2 m- and lipid-binding HC, (iii) non-classical, β_2 m- and lipid-binding HC, (iv) non-classical, β_2 m- and protein-binding HC without the need for a small ligand in the binding groove, and (v) non-classical HC that do not bind β_2 m. In **Table 2**, only HC residues of representative molecules for each of these groups are displayed.

Two highly conserved hydrogen bonds are present in all β_2 m-HC complexes in all molecules studied: β_2 m-Tyr10 interacts with the backbone carbonyl of HC-Pro235 (HC residue numbering refers to the human HLA-A2 HC) and the backbone carbonyl

of β_2 m-Trp60 contacts the carboxamide of HC-Gln96 (**Table 1**). The residues β_2 m-Tyr10, β_2 m-Trp60, HC-Gln96 and HC-Pro235 are highly conserved across all species (**Fig. S1 and S2**), indicating that these interactions are very likely also retained in complexes for which structural information is not yet available. Additionally, the side chain of β_2 m-Tyr10 is involved in van der Waals interactions with HC-Val230 (BF2*2101 and YF1*7.1) or HC-Leu232 of all CD1 (Met231 in bovine CD1b3), but not in the mammalian peptide-binding molecules, possibly because the homologous positions for the Val/Leu residues are exclusively Arg in these mammalian HC. However, the basic side chain of Arg stack hydrophobically with β_2 m-Tyr10. The β_2 m-Tyr10 – HC-Leu232 interaction is also detected in human HFE, but not in the human or rat FcRn complexes, because the latter possess Gly at the homologous position. (**Table 1**).

We can also identify two highly conserved hydrophobic interactions encompassing conserved or homologous residues: β_2 m-Tyr26 (mammalian complexes) or β_2 m-Phe26 (chicken complexes) contacts HC-Pro235 and β_2 m-Trp60 interacts with HC-Ala117 except for chicken CD1-2 and HFE which possess Gly at an equivalent position (**Table 1 and Fig. 2, S2**).

These sequence- and structure-based analyses reveal β_2 m-HC interactions that are prerequisites for complex formation and explain why MHC class I-like HC such as zinc- α 2-glycoprotein (ZAG) (Sanchez et al., 1999) and MHC class I chain-related (MIC) (Li et al., 1999; Holmes et al., 2002) molecules cannot bind β_2 m. The ZAG and MIC HC possess an Ala \rightarrow Tyr substitution at residue 113 or 111, respectively (**Table 2**), which prevents β_2 m from binding because the large aromatic Tyr side chain would sterically clash with β_2 m-Trp60. β_2 m-Trp60 is also involved in several other conserved interface interactions that are apparently essential for complex stability (**Tables 1 and 2**). Instead, ZAG has been shown to bind prolactin inducible protein (PIP) (Hassan et al., 2008), in an overall similar, yet in detail completely different manner as compared to that of β_2 m. Analyses of the ZAG-PIP interactions and sequence comparison to MIC molecules indicate that the latter might also be able to bind PIP (data not shown), although this has not been demonstrated so far. The absolutely conserved residue β_2 m-Trp60 also forms a hydrogen bond to HC-Asp122 (Asn in chicken CD1-1 and Glu in human FcRn) in all complexes except for murine CD1d (Giabbai et al., 2005) and rat FcRn (Vaughn and Bjorkman, 1998). The

homologous HC-Gln116 of murine CD1d and HC-Glu115 of rat FcRn do not form hydrogen bonds with β_2 m-Trp60 although this appears in principle possible.

Semi-conserved β_2 m-HC interface interactions

Apart from these conserved interactions, β_2 m-HC interfaces are also characterized by several semi-conserved contacts. The salt bridge between β_2 m-Arg12 and the side chain of HC-Asp238 (HC-Asp234 in chicken HC) has been found in all molecules examined except for classical mammalian peptide-binding complexes. Intriguingly, HC-Asp238 is highly conserved in mammalian HC and the salt bridge could in principle be formed if the side chain conformation of β_2 m-Arg12 would resemble that of the chicken or non-classical molecules. However, the side chain of β_2 m-Arg12 adopts a different conformer and thereby establishes an interaction with the main chain carbonyls of HC-Ala236 and/or HC-Gly237 and/or HC-Ser236 (**Table 1**).

Similarly, the hydrogen bond between β_2 m-Asn24 and HC-Asn232 (equivalent to HC-236 in mammals) is conserved in all molecules examined except for mammalian peptide-binding molecules, HFE and human CD1a. This is because HC-Asn232 is only found in chicken and mammalian non-classical molecules (**Fig. S2**). Instead, in mammalian peptide binding molecules, β_2 m-Asn24 contacts the main chain oxygen of HC-Ala236 or HC-Ser236 (**Tables 1 and 2**).

In most of the complexes studied except for chicken BF2*2101 and YF1*7.1, rat RT1-A, mouse H-2D^b and CD1d, β_2 m-His31 and HC-Gln89 are two highly conserved, hydrogen bonded residues (**Tables 1 and 2, Figs. S1 and S2**). The reason for the difference between the complexes lies in a distinct conformation of β_2 m-His31 which may be tilted away from HC-Gln89 (**Fig. 2**). In the BF2*2101, YF1*7.1, and mouse CD1d complexes, β_2 m-His31 establishes intrachain contacts to β_2 m-Trp60 via water molecules while β_2 m-His31 of H-2Db and RT1-A interacts with the main chain oxygen atoms of HC-Glu119 and HC-Asp119, respectively (**Tables 1 and 2**).

Group- or species-specific HC- β_2 m interface interactions

In addition to conserved and semi-conserved interface interactions, we identified also group- or species-specific contacts. For example, the residue β_2 m-Gln8 is highly conserved in all species, but the interacting residues of the HC differ between complexes. In the mammalian peptide-binding molecules, β_2 m-Gln8 contacts the backbone carbonyl of HC-Glu232 within the α 3-domain through hydrogen bonding, while in the BF2*2101 and YF1*7.1 complexes of the chicken, an interaction with the non-homologous backbone carbonyl of HC-Gly227 is established (**Table 1**). A comparable contact is also found in HFE, to HC-Asp230, but these interactions are not found in FcRn or any CD1 molecules. Additionally, β_2 m-Gln8 hydrogen bonds with HC-Arg234 in all mammalian peptide-binding molecules except for mouse H-2M3 which binds exclusively N-terminally formylated peptides (Wang et al., 1995). However, an analogous interaction is not found in other molecules because HC-Arg234 is only conserved in mammalian classical class I molecules (**Fig. S2**).

We further noticed a β_2 m-HC contact that is conserved in all mammalian peptide-binding molecules: hydrogen bonding between β_2 m-Asp98 and HC-Arg202 (HC-His192 in BoLA and H-2K^b) (**Table 1**). In chicken BF2*2101 and YF1*7.1, β_2 m-Glu98 forms a hydrogen bond with the non-homologous HC-Arg185, while β_2 m-Glu98 of the two chicken CD1 molecules (Asp98 in CD1-2, because its structure was solved with human β_2 m) contacts HC-Arg187 and HC-Arg190 through salt bridges, respectively (**Table 1**). The backbone carbonyl of β_2 m-Asp98 interacts with HC-Ser188 in mammalian CD1 molecules, except for human CD1a, and contacts HC-Ser197 in human, but not in rat FcRn because the latter possesses an Ala at this position. β_2 m-Glu98 of BF2*2101 and chicken CD1-1 establish an additional salt bridge to HC-Arg200 and HC-Arg199, respectively (**Table 1**).

Previously, we reported a distinctive β_2 m-HC interaction shared between classical chicken, but not classical mammalian molecules (Hee et al., 2010). In BF2*2101 and YF1*7.1, β_2 m-Lys34 establishes an intermolecular salt bridge to HC-Asp14 that is very likely the reason for the distinct conformation of HC Loop1 in classical class I molecules of the chicken, and possibly in classical class I complexes of other non-mammalian species as well (Hee et al., 2010). On the other hand, the mammalian HC residue 14 is exclusively Arg and it forms a bidentated, intramolecular salt bridge

with HC-Asp39. Non-classical molecules like CD1 do not possess any of these interactions. Based on the sequence comparisons (**Figs. S1 and S2**), the HC-Arg14 – HC-Asp39 interaction is likely to be restricted to mammalian classical molecules, where it might facilitate an interaction, possibly with a protein ligand, that is irrelevant in classical class I complexes of non-mammalian vertebrates.

Similarly, residue β_2 m-Asp53 is conserved in all β_2 m analyzed, but it interacts distinctly with HC residues in different molecules (**Table 1**). In all peptide-binding molecules and HFE, β_2 m-Asp53 forms salt bridge(s) to HC-Arg48 (Fig. 3A-D) (via a water molecule in HLA-B27, see Fig. 4). This contact is absent from all non-peptide-binding molecules. One explanation could be that the residue HC-48 of the non-peptide-binding molecules has been substituted by non-charged residues (**Figs. 3E, F and S2**). However, we regard another explanation as more plausible. It comprises a more comprehensive picture in which the β_2 m-Asp53 – HC-Arg48 salt bridge(s), the peptide-binding capability and conformation of the H1 helix appear to be interrelated structural features. The H1 helix is only present when the β_2 m-Asp53 – HC-Arg48 salt bridge(s) are established. These are, however, only found in peptide-binding molecules, except for HFE (**Fig. 3**). One can presume that HFE is in the evolutionary transition from a peptide- to a non-peptide-binding state, because it still possesses the salt bridges involving β_2 m-Asp53 and HC-Arg48, but has already lost the H1 helix. This helix is not directly involved in peptide-binding, but its presence and the concomitant binding of a protein-derived ligand suggest that this structure feature is imperative for the architecture of the N-terminal section of the peptide-binding groove. This part of the molecule is made up of the A pocket that accommodates the N-terminus of a bound peptide. Non-peptide-binding molecules like CD1 and the chicken YF1*7.1 complex with its otherwise classical scaffold (Hee et al., 2010), do not possess the H1 helix, further indicating that it is important for histocompatibility antigens that present peptides (**Fig. 3**).

In some of the HLA-B27 and HLA-B44 molecules, β_2 m-Asp53 – HC-Arg48 interactions are mediated by water molecules because of the substitutions of HC-Gln32 by Leu32 (**Fig. 4 and Table S2**). Different interaction networks are established in this part of the HLA-B27 and HLA-B44 complexes that are not found in other HLA-A and HLA-B molecules which possess Gln32 (see **Fig. 4 and Table S2**). In non-classical molecules, i.e. the mammalian CD1 and FcRn molecules, only the β_2 m-

Asp53 – HC-Gln32 contacts are conserved (**Table 1**). For the chicken molecules, only BF2*2101 retained the β_2 m-Asp53 – HC-Arg46 salt bridge in addition to the β_2 m-Asp53 – HC-His35 interactions, but YF1*7.1, CD1-1 and CD1-2 do not show this contact due to substitution of the corresponding HC residue.

Concluding remarks

In this study, we identified β_2 m residues that are highly dynamic in free form but are stabilized upon association with HC. Many of these residues interact with HC amino acids that form the underside of the peptide binding groove, suggesting an induced fit-type of interaction during complex formation. β_2 m-Trp60 occupies a central position within this β_2 m-HC interface, as it is involved in more universally conserved contacts than any other β_2 m residue. Furthermore, our β_2 m-HC interface analyses identify a number of evolutionarily conserved β_2 m and HC residues that engage in interface interactions in all complexes examined, indicating their critical role in complex stabilization. Based on these results, we derive also an explanation as to why non- β_2 m binders, e.g. ZAG and MIC, cannot associate with β_2 m. In addition, we pinpoint interface interactions that are only found in a particular group of molecules or species, at least according to our present knowledge. In summary, this study leads to advanced insights into the contacts between β_2 m and HC and demonstrates that the molecular dynamics of β_2 m should not be neglected when considering its interactions with MHC HC.

Materials and methods

Accession numbers and PDB entries.

Universal Protein Resource (UniProt) or Ensembl transcript accession numbers for the β 2m sequences used in this article: human (P61769), chimpanzee (P61770), orangutan (P16213), gorilla (P61771), macaque (Q6V7J5), baboon (Q9TS09), tamarin (P55079), marmoset (Q71UN5), hamster (Q9WV24), mouse (P01887), rat (P07151), sloth (ENSCHOP00000008882), armadillo (ENSDNOP00000006209), bovine (P01888), sheep (Q6QAT4), pig (Q07717), horse (P30441), zebra (Q863A9), donkey (Q861S3), pika (ENSOPRP00000001924), rabbit (P01885), bat (ENSMLUP00000004920), guinea pig (P01886), cat (Q5MGS7), possum (Q9GKM2), opossum (Q864T8), echidna (Q864T6), platypus (Q864T7), chicken (Q6L755), turkey (P21612), duck (Q14U75), finch (B5G1N3), trout (Q92004), salmon (Q9DG62), catfish (O42240), carp (Q03422), barbel (Q70XT2), zebrafish (B0UYS1), medaka (Q90ZJ6), sturgeon (Q9PRF8), cod (Q9YGI3), western clawed frog (D4AHU6), African clawed frog (Q9IA97), skate (Q8AXA0), sandbar shark (D0EP39).

Accession numbers for the MHC class I or class I-like sequences are as following: human HLA-A2 (P01892), HLA-B*27:05 (P30466), HLA-E (P13747), HLA-G (P17693); macaque Mamu-A*01 (Q30596); tamarin (P30515); orangutan (P16211); chimpanzee (P13750); gorilla (P30381); baboon (Q30869); dog (P18466); horse (Q30485); cat (Q95485); bovine (Q30291); sheep (Q30838); pig (Q8SPA6); rat (M31018); mouse H-2K^b (P01901), H-2M3 (Q31093), H-2D^b (P01899); opossum (Q9TPL0); possum (Q95IT0); bandicoot (A7YKN2); Tasmanian devil (B3F356); kangaroo (Q30768); platypus (Q8HX81); echidna (Q8HX83); chicken BF2*2101 (AF013493), YF1*7.1 (AF218783); quail (Q76LI6); turkey (B1N1D8); goose (Q6TU42); duck (Q2VQY9); crane (O62881); warbler (O98192); guineafowl (A6YQY7); lizard (Q31307); marine iguana, *Amblyrhynchus cristatus* (B3UZY4); common iguana, *Iguana iguana* (B3UZZ5); Galapagos iguana, *Conolophus subcristatus* (B3UZZ1); tuatara (Q1I0Q9); axolotl (P79458); leopard frog, *Rana pipiens* (Q9TP98); African clawed frog, *Xenopus laevis* (Q9TPA1); Western clawed frog, *Xenopus tropicalis* (Q6XXY4); Uganda clawed frog, *Xenopus ruwenzoriensis* (Q8SNP3); catfish (O62896); grass carp (Q76CD9); zebrafish (Q31365); trout

(Q9GJJ4); Atlantic salmon (Q31594); Atlantic cod (O98272); medaka (Q589Q0); fugu (Q9GJE4); leopard shark (O46817); nurse shark (AF220063); human HFE (Q30201); human FcRn (P55899); rat FcRn (P13599); human CD1a (P06126), CD1b (P29016), CD1d (P15813); bovine CD1b3 (Q1L1H6); mouse CD1d (P11609) and chicken CD1-1 (A5HUM9), CD1-2 (Q5GL29).

Structures used in this article were obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank with the following accession numbers: chicken BF2*2101 (3BEV), YF1*7.1 (3P73), CD1-1 (3JVG), CD1-2 (3DBX); human HLA-A2 (3HLA, 1I4F), HLA-B27 (1OGT, 3CZF), HLA-E (1KTL), HLA-G (1YDP); macaque Mamu-A*01 (1ZVS), Mamu-A*02 (3JTT); bovine BoLA (2XFX); mouse H-2Kb (2ZSV), H-2M3 (1MHC), H-2Db (3FTG, 3ZOK); rat RT1-A (1KJM); human HFE (1A6Z); human FcRn (1EXU); rat FcRn (3FRU); human CD1a (1ONQ), CD1b (2H26), CD1d (1ZT4, 2PO6); bovine CD1b3 (3L9R); mouse CD1d (1ZHN, 2FIK, 2GAZ); ZAG (1ZAG), ZAG-PIP (3ES6) and human MICA (1B3J).

Structural presentation and computational analyses

Figures depicting structures were generated with PyMOL (Delano, 2002) and labeled using the GNU Image Manipulation Program (GIMP; www.gimp.org). Atomic distances were calculated using CONTACT of the CCP4 Program Suite (Collaborative Computational Project, 1994). A distance cut-off of 4 Å was used to define hydrogen bonds, salt bridges and van der Waals interactions. Amino acid sequence alignments were generated using ClustalW (Thompson et al., 1994). Alignment with secondary structures on Figure S1 was constructed using ESPript (Gouet et al., 1999).

Samples for NMR-spectroscopy

Four different samples were used in the investigation using NMR spectroscopy. For the resonance assignment of free β ₂m, labeling with ¹³C and ¹⁵N was used while in addition deuteration of the protein was necessary for an assignment of MHC-bound β ₂m. Only β ₂m, however, was labeled with ²H, ¹³C and ¹⁵N; the B*27:05 HC and the

peptide were unlabeled. For the relaxation experiment the same types of samples were used, but the free and MHC-bound protein were only labeled with ^{15}N .

NMR spectroscopy

NMR experiments were performed at 37°C at 600 or 750 MHz on AVIII NMR spectrometers (Bruker, Karlsruhe) using 5 mm cryogenically cooled triple resonance probes with actively shielded gradients. Each sample had a volume of 600 μL of 90% H_2O /10% D_2O using a buffer of 10 mM NaPhosphate, 150 mM NaCl, pH 7.5. The concentrations were 13, 21, 8 and 18 mg/ml for ^{15}N - $\beta_2\text{m}$, ^{13}C , ^{15}N - $\beta_2\text{m}$, ^2H , ^{13}C , ^{15}N - $\beta_2\text{m}$ in B*27:05-pVIPR and ^{15}N - $\beta_2\text{m}$ in B*27:05-pVIPR, respectively.

For the assignment of free $\beta_2\text{m}$ two pairs of three-dimensional triple resonances experiments were performed, namely HNCACB/HN(CO)CACB as well as HNCO/HN(CA)CO (Cavanagh et al., 1996). For the assignment of MHC-bound $\beta_2\text{m}$ the same pairs of three-dimensional triple resonance experiments were performed. However, the sequences were based on the TROSY-sequence (Pervushin et al., 1997; Tugarinov et al., 2004). Heteronuclear NOE measurements (Farrow et al., 1994) were performed using an HSQC-based and a TROSY-based sequence for free and MHC bound $\beta_2\text{m}$, respectively. Data were recorded and processed using TOPSPIN (Bruker, Karlsruhe) and subsequently transferred to CCPN (Vranken et al., 2005) for resonance assignment and evaluation of the intensities in the heteronuclear NOE spectra.

Acknowledgement

We are grateful to Christina Schnick and Natalya Erdmann for excellent technical support. This work was supported by the Deutsche Forschungsgemeinschaft (grants Na226/12-3, UC8/1-2, UC8/2-1 and SFB 449/B6). A. Ziegler acknowledges support by the VolkswagenStiftung (grant I/79 989). B. Loll is grateful for support by the Forschungskommission of the Freie Universität Berlin and Fonds der Chemischen Industrie. C.-S. Hee thanks the Berliner Krebsgesellschaft (Ernst von Leyden Stipendium) and the Deutsche Forschungsgemeinschaft for support through UC8/2-1.

References

- Achour, A., Michaelsson, J., Harris, R.A., Ljunggren, H.G., Karre, K., et al. (2006). Structural basis of the differential stability and receptor specificity of H-2Db in complex with murine versus human beta2-microglobulin. *J. Mol. Biol.* *356*, 382-396.
- Becker, J.W. and Reeke, G.N., Jr. (1985). Three-dimensional structure of beta 2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 4225-4229.
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., et al. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* *329*, 506-512.
- Boyington, J.C., Motyka, S.A., Schuck, P., Brooks, A.G., and Sun, P.D. (2000). Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. *Nature* *405*, 537-543.
- Cavanagh, J., Fairbrother, J.W., Palmer III, A.G., and Skelton, N.J. (1996). *Protein NMR Spectroscopy Principles And Practice*. (San Diego: Academic Press).
- Chen, W., Gao, F., Chu, F., Zhang, J., Gao, G.F., et al. (2010). Crystal structure of a bony fish beta2-microglobulin: insights into the evolutionary origin of immunoglobulin superfamily constant molecules. *J. Biol. Chem.* *285*, 22505-22512.
- Chu, F., Lou, Z., Chen, Y.W., Liu, Y., Gao, B., et al. (2007). First glimpse of the peptide presentation by rhesus macaque MHC class I: crystal structures of Mamu-A*01 complexed with two immunogenic SIV epitopes and insights into CTL escape. *J. Immunol* *178*, 944-952.
- Clements, C.S., Kjer-Nielsen, L., Kostenko, L., Hoare, H.L., Dunstone, M.A., et al. (2005). Crystal structure of HLA-G: a nonclassical MHC class I molecule expressed at the fetal-maternal interface. *Proc. Natl. Acad. Sci. U. S. A* *102*, 3360-3365.
- Collaborative Computational Project, N.4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* *50*, 760-763.
- Delano, W. L. *The PyMOL Molecular Graphic System*. 2002. San Carlos, CA, DeLano Scientific.
- Ref Type: Generic
- Dvir, H., Wang, J., Ly, N., Dascher, C.C., and Zajonc, D.M. (2010). Structural basis for lipid-antigen recognition in avian immunity. *J. Immunol.* *184*, 2504-2511.
- Esposito, G., Corazza, A., Viglino, P., Verdone, G., Pettirossi, F., et al. (2005). Solution structure of beta(2)-microglobulin and insights into fibrillogenesis. *Biochim. Biophys. Acta* *1753*, 76-84.
- Farrow, N.A., Zhang, O., Forman-Kay, J.D., and Kay, L.E. (1994). A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* *4*, 727-734.
- Giabbai, B., Sidobre, S.P., Crispin, M.D.M., Sanchez-Ruiz, Y., Bachi, A., et al. (2005). Crystal structure of mouse CD1d bound to the self ligand phosphatidylcholine: A molecular basis for NKT cell activation. *Journal of Immunology* *175*, 977-984.
- Gouet, P., Courcelle, E., Stuart, D.I., and Metz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics.* *15*, 305-308.

- Hassan, M.I., Bilgrami, S., Kumar, V., Singh, N., Yadav, S., et al. (2008). Crystal structure of the novel complex formed between zinc alpha2-glycoprotein (ZAG) and prolactin-inducible protein (PIP) from human seminal plasma. *J. Mol. Biol.* *384*, 663-672.
- Hee, C.S., Gao, S., Loll, B., Miller, M.M., Uchanska-Ziegler, B., et al. (2010). Structure of a classical MHC class I molecule that binds "non-classical" ligands. *PLoS Biol.* *8*, e1000557.
- Holmes, M.A., Li, P., Petersdorf, E.W., and Strong, R.K. (2002). Structural studies of allelic diversity of the MHC class I homolog MIC-B, a stress-inducible ligand for the activating immunoreceptor NKG2D. *J. Immunol* *169*, 1395-1400.
- Hülsmeier, M., Fiorillo, M.T., Bettosini, F., Sorrentino, R., Saenger, W., et al. (2004). Dual HLA-B27 subtype-dependent conformation of a self-peptide. *Journal of Experimental Medicine* *199*, 271-281.
- Hülsmeier, M., Hillig, R.C., Volz, A., Ruhl, M., Schroder, W., et al. (2002). HLA-B27 subtypes differentially associated with disease exhibit subtle structural alterations. *Journal of Biological Chemistry* *277*, 47844-47853.
- Iwata, K., Matsuura, T., Sakurai, K., Nakagawa, A., and Goto, Y. (2007). High-resolution crystal structure of beta2-microglobulin formed at pH 7.0. *J. Biochem.* *142*, 413-419.
- Koch, M., Camp, S., Collen, T., Avila, D., Salomonsen, J., et al. (2007). Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* *27*, 885-899.
- Li, P., Willie, S.T., Bauer, S., Morris, D.L., Spies, T., et al. (1999). Crystal structure of the MHC class I homolog MIC-A, a gammadelta T cell ligand. *Immunity* *10*, 577-584.
- Li, X., Liu, J., Qi, J., Gao, F., Li, Q., et al. (2011). Two distinct conformations of a rinderpest virus epitope presented by bovine major histocompatibility complex class I n*01801: a host strategy to present featured peptides. *J. Virol.* *85*, 6038-6048.
- Maenaka, K. and Jones, E.Y. (1999). MHC superfamily structure and the immune system. *Curr. Opin. Struct. Biol.* *9*, 745-753.
- McParland, V.J., Kalverda, A.P., Homans, S.W., and Radford, S.E. (2002). Structural properties of an amyloid precursor of beta(2)-microglobulin. *Nat. Struct. Biol.* *9*, 326-331.
- Michaelsson, J., Achour, A., Rolle, A., and Karre, K. (2001). MHC class I recognition by NK receptors in the Ly49 family is strongly influenced by the beta(2)-microglobulin subunit. *J. Immunol.* *166*, 7327-7334.
- Mitsuki, M., Matsumoto, N., and Yamamoto, K. (2004). A species-specific determinant on beta2-microglobulin required for Ly49A recognition of its MHC class I ligand. *Int. Immunol* *16*, 197-204.
- O'Callaghan, C.A., Tormo, J., Willcox, B.E., Braud, V.M., Jakobsen, B.K., et al. (1998). Structural features impose tight peptide binding specificity in the nonclassical MHC molecule HLA-E. *Mol. Cell* *1*, 531-541.
- Palmer, A.G., III (2004). NMR characterization of the dynamics of biomacromolecules. *Chem. Rev* *104*, 3623-3640.

Pedersen, L.O., Hansen, A.S., Olsen, A.C., Gerwien, J., Nissen, M.H., et al. (1994). The interaction between beta 2-microglobulin (beta 2m) and purified class-I major histocompatibility (MHC) antigen. *Scand. J. Immunol* 39, 64-72.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A* 94, 12366-12371.

Platt, G.W., McParland, V.J., Kalverda, A.P., Homans, S.W., and Radford, S.E. (2005). Dynamics in the unfolded state of beta2-microglobulin studied by NMR. *J. Mol. Biol.* 346, 279-294.

Sanchez, L.M., Chirino, A.J., and Bjorkman, P. (1999). Crystal structure of human ZAG, a fat-depleting factor related to MHC molecules. *Science* 283, 1914-1919.

Shields, M.J., Kubota, R., Hodgson, W., Jacobson, S., Biddison, W.E., et al. (1998). The effect of human beta2-microglobulin on major histocompatibility complex I peptide loading and the engineering of a high affinity variant. Implications for peptide-based vaccines. *J. Biol. Chem.* 273, 28010-28018.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research* 22, 4673-4680.

Trinh, C.H., Smith, D.P., Kalverda, A.P., Phillips, S.E.V., and Radford, S.E. (2002). Crystal structure of monomeric human beta-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9771-9776.

Tugarinov, V., Hwang, P.M., and Kay, L.E. (2004). Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev Biochem.* 73, 107-146.

Vaughn, D.E. and Bjorkman, P.J. (1998). Structural basis of pH-dependent antibody binding by the neonatal Fc receptor. *Structure* 6, 63-73.

Verdone, G., Corazza, A., Viglino, P., Pettirossi, F., Giorgetti, S., et al. (2002). The solution structure of human beta2-microglobulin reveals the prodromes of its amyloid transition. *Protein Sci.* 11, 487-499.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., et al. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59, 687-696.

Wang, C.R., Castano, A.R., Peterson, P.A., Slaughter, C., Lindahl, K.F., et al. (1995). Nonclassical binding of formylated peptide in crystal structure of the Mhc class-Ib molecule H2-M3. *Cell* 82, 655-664.

Wang, J., Whitman, M.C., Natarajan, K., Tormo, J., Mariuzza, R.A., et al. (2002). Binding of the natural killer cell inhibitory receptor Ly49A to its major histocompatibility complex class I ligand - Crucial contacts include both H-2D(d) and beta(2)-microglobulin. *J. Biol. Chem.* 277, 1433-1442.

Zajonc, D.M., Crispin, M.D.M., Rudd, P.M., Dwek, R.A., Teyton, L., et al. (2003). Structural basis of glycolipid presentation by CD1. *Glycobiology* 13, 835.

Zajonc, D.M., Striegl, H., Dascher, C.C., and Wilson, I.A. (2008). The crystal structure of avian CD1 reveals a smaller, more primordial antigen-binding pocket compared to mammalian CD1. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 17925-17930.

Zeng, Z.H., Castano, A.R., Segelke, B.W., Stura, E.A., Peterson, P.A., et al. (1997). Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science* *277*, 339-345.

Table 1. β_2 m and HC residues involved in β_2 m-HC interactions. β_2 m numbering refers to chicken β_2 m, homologous β_2 m residues are indicated in the boxes showing individual contacts. For example, the F26→P231 contact within the YF1*7.1 complex resembles the Y26→P235 contact in the HLA-A2 molecule. Non-homologous HC residues in comparison to chicken BF2*2101 are highlighted with blue text. Cut-off distances are as follows: hydrogen bonds and salt bridges, 3.5 Å; van der Waals interactions and π -stacking, 4.0 Å. Hydrogen bonds or salt bridges between 3.5-4.0 Å are indicated with red text. Four highly conserved β_2 m-HC interactions are highlighted with a yellow background. The residue range for α_1 , α_2 and α_3 domain of HC are approximately 1-90, 91-180 and 181-270, respectively.

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

Interaction	β 2m residues	Heavy chain residue																						
		Chicken molecule				Mammalian peptide-binding molecule										Mammalian non-classical/ class I-like molecule								
		BF2*21	YF1*7.1	CD1-1	CD1-2#	HLA-A2	HLA-B27	HLA-E	HLA-G	Mamu-A*01/*02	BoLA	H-2K ^b	H-2D ^b	RT1-A	H-2M3	HFE	hsFcRn	rnFcRn	hsCD1a	hsCD1b	hsCD1d	btD1b3	mmCD1d	
H	Q8 ^{NE2}	G227 ^O	G227 ^O	-	-	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	E232 ^O	D230 ^{OD1}	-	-	-	-	-	-	-	
H	Q8 ^{OE1}	W240 ^{NE1}	S240 ^{OG} (H2O)	-	-	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	R234 ^{NH1}	-	-	F224 ^O	-	-	-	D230 ^{OD2}	R242 ^{NH1}	-	
VdW	Y10	V230	V230	L231	L234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	π -stack R234	L232	-	-	L232	L232	L232	M232	L232	-	
H	Y10 ^{OH}	P231 ^O	P231 ^O	P232 ^O	P236 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P235 ^O	P233 ^O	P226 ^O	P228 ^N	P233 ^O	P233 ^O	P233 ^O	P233 ^O	P235 ^O	
H	S11 ^{OG}	-	-	R199 ^{NH2}	R202 ^{NH2}	-	-	-	-	-	-	-	-	-	-	R201 ^{NH2}	-	-	-	-	-	-	-	
H	S11 ^O	H202 ^{ND1}	H202 ^{ND1}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	S11 ^O	-	-	Q239 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	-	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q240 ^{NE2}	H233 ^{ND1}	H235 ^{ND1}	-	-	-	-	-	
H	R12 ^O	H202 ^{ND1}	H202 ^{ND1}	T201 ^{OG}	-	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q242 ^{NE2}	Q240 ^{NE2}	S200 ^{OG}	N229 ^{ND2}	S205 ^{OG}	S205 ^{OG}	S205 ^{OG}	S205 ^{OG}	S207 ^{OG}		
H	R12 ^{NH1/NH2}	-	-	S202 ^{OG}	S205 ^{OG}	-	-	-	-	-	-	-	-	-	-	-	-	S205 ^{OG}	S205 ^{OG}	-	-	-	-	
SB/H	R12 ^{NE/NH}	D234 ^{OD1/2}	D234 ^{OD2}	D235 ^{D1/2}	D238 ^{OD1/2}	A236 ^O G237 ^O	A236 ^O	A236 ^O G237 ^O	G237 ^O	A236 ^O G237 ^O	S236 ^O	A236 ^O G237 ^O	G237 ^O	A236 ^O G237 ^O	S236 ^O G237 ^O	D236 ^{OD2}	D229 ^{OD2}	D231 ^{OD2}	D236 ^{OD2}	N236 ^{OD1}	D236 ^{OD1}	D236 ^{OD2}	D238 ^{OD1/2}	
H	A15 ^O	R185 ^{NH1/2}	R185 ^{NH1/2}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	N24 ^{ND2}	N232 ^{OD1}	N232 ^{OD1}	N233 ^{OD1}	N236 ^{OD1}	A236 ^O	A236 ^O	A236 ^O	A236 ^O	A236 ^O	S236 ^O	A236 ^O	A236 ^O	A236 ^O	S236 ^O	-	N227 ^{OD1}	N229 ^{OD1}	-	-	N234 ^{OD1}	N234 ^{OD1}	N234 ^{OD1}	N236 ^{OD1}
VdW	F26	V230	V230	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
VdW	F26	P231	P231	P232	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P235	Y26-P233	Y26-P226	Y26-P228	Y26-P233	Y26-P233	Y26-P233	Y26-P233	Y26-P233	Y26-P233	
H	H31 ^{NE2}	-	-	Q89 ^{OE1}	Q92 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	Q96 ^{OE1}	E119 ^O	D119 ^O	Q96 ^{OE1}	Q94 ^{OE1}	Q89 ^{OE1}	Q91 ^{OE1}	Q93 ^{OE1}	Q93 ^{OE1}	Q93 ^{OE1}	Q93 ^{OE1}	
VdW	P33	M12	M12	-	-	-	-	-	-	-	-	-	V12	-	-	-	-	-	-	-	-	-	-	
SB	K34 ^{NZ}	D14 ^{OD2}	D14 ^{OD2}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SB/H	D53 ^{OD1/2}	H35 ^{NE2}	-	-	-	Q32 ^{NE2} & R35 ^{NH1}	R35 ^{NE/NH2}	Q32 ^{NE2} & R35 ^{NE}	Q32 ^{NE2} †	Q32 ^{NE2} & R35 ^{NE}	Q32 ^{OE1} & R35 ^{NE}	-	-	-	Q32 ^{NE2} & R35 ^{NH1}	-	Q32 ^{NE2} & S35 ^{OG}	Q32 ^{NE2} & T35 ^{OG1}	Q32 ^{NE2}	Q32 ^{NE2}	Q32 ^{NE2} & R35 ^{NE/NH2}	Q32 ^{NE2}	Q32 ^{NE2} & R35 ^{NE/NH2}	
SB/H	D53 ^{OD1/2}	R46 ^{NH1}	-	Y45 ^{OH}	-	R48 ^{NE}	R48 ^{NE} (H2O)	R48 ^{NE/NH1}	R48 ^{NE}	R48 ^{NE/NH2}	R48 ^{NE/NH2}	R48 ^{NH1/2}	R48 ^{NH1/2}	R48 ^{NH1/2}	R48 ^{NH2/E}	R46 ^{NH2/E}	-	-	-	-	-	-	-	
vdW	M54	M12	-	F10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
vdW	M54	V23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	S55 ^{OG}	Y27 ^{OH}	Y27 ^{OH}	-	-	-	-	-	-	-	-	-	-	-	-	Y27 ^{OH}	-	-	-	-	-	-	-	
VdW	F56	I8	F8	L8	F8	F8	F8	F8	F8	F8	F8	F8	F8	F8	F8	L8	-	-	-	-	-	-	-	
SB	D58 ^{OD1}	-	-	K6 ^{NZ}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	W60 ^O	Q94 ^{NE2}	Q93 ^{NE2}	Q89 ^{NE2}	Q92 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{OE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q96 ^{NE2}	Q94 ^{NE2}	Q89 ^{NE2}	Q91 ^{NE2}	Q93 ^{NE1}	Q93 ^{NE1}	Q93 ^{NE2}	Q93 ^{NE1}	Q95 ^{NE1}	
VdW	W60	A114	A113	A109	G113	A117	A117	A117	A117	A117	G117	A117	A117	A117	A117	G114	A109	A111	A113	A113	A113	A113	A115	
H	W60 ^{NE1}	D119 ^{OD1}	D118 ^{OD1}	N114 ^{OD1}	D117 ^{OD1}	D122 ^{OD2}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D122 ^{OD1}	D119 ^{OD1}	E114 ^{OE1}	-	D118 ^{OD1}	D118 ^{OD2}	D118 ^{OD1/2}	D118 ^{OD2}	-	
H	E98 ^O	R185 ^{NE/NH1}	R185 ^{NH1}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	E98 ^O	-	-	-	-	D98 ^O R202 ^{NH2}	D98 ^O R202 ^{NH1}	D98 ^O R202 ^{NE}	D98 ^O R202 ^{NH1}	D98 ^O R202 ^{NE/NH1}	D97 ^O H192 ^{NE2}	D98 ^O H192 ^{NE2}	D98 ^O R202 ^{NH1}	D98 ^O R202 ^{NE}	D98 ^O R202 ^{NE}	-	D98 ^O S197 ^{OG}	-	-	-	D98 ^O S188 ^{OG}	D98 ^O S188 ^{OG}	D97 ^O S188 ^{OG}	D98 ^O S190 ^{OG}
SB	E98 ^{OE1}	-	-	R187 ^{NE/NH2}	D98 ^{OD1/2} R190 ^{NE/NH2}	-	-	D98 ^{OD2} H192 ^{ND1}	-	-	-	-	-	-	-	-	-	-	-	D98 ^{OD2} H189 ^N	-	-	-	
SB	E98 ^{OE1/2}	R200 ^{NE}	-	R199 ^{NH1}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
H	F99 ^O	-	-	R199 ^{NH2}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Crystallized with human β 2m; † Crystallized with murine β 2m; ‡ Asparagine (N) or Glutamine (Q) rotamer (e.g. NE1/2 and OE1/2 could be at either end)

3. Comparative Biophysical Characterizations of Human and Chicken β_2 -microglobulin and the Evolution of Interfaces in MHC Class I Molecules

Table 2. Summary of β_2 m-HC interactions. For the ease of comparison, only HC residues of a representative molecule from a particular group are shown, for example, HLA-A2 is chosen to represent the group of mammalian classical peptide- and β_2 m-binding molecules whereas BF2*2101 represents non-mammalian classical peptide- and β_2 m-binding molecules. The representative HC residue is labelled as “variable” when there are ambiguous interactions in the group. For the non-classical molecules that do not bind β_2 m (ZAG and MICA), the homologous HC residues are shown for comparison.

Types of HC	β_2 m residue														
	Q8 ^{NE2}	Q8 ^{OE1}	Y10	Y10 ^{OH}	R12 ^{NE/NH}	N24 ^{ND2}	Y/F26	H31 ^{NE2}	D53 ^{OD1/2}	F56	W60 ^O	W60	W60 ^{NE1}	D/E98 ^O	
1. Classical, β_2 m-binding															
a) peptide-binding															
HLA-A2	E232 ^O	R234 ^{NH1}	R234 (π) [§]	P235 ^O	A236 ^O /G237 ^O	A236 ^O	P235	Q96 ^{OE1}	R48 ^{NE}	F8	Q96 ^{NE2}	A117	D122 ^{OD1}	R202 ^{NH2}	
BF2*2101	G227 ^O	W240 ^{NE1}	V230	P231 ^O	D234 ^{OD1/2}	N232 ^{OD1}	P231	-	R46 ^{NH1}	I8	Q94 ^{NE2}	A114	D119 ^{OD1}	R185 ^{NE/NH1}	
b) lipid-binding															
YF1*7.1	G227 ^O	S240 ^{OG} (H ₂ O) [§]	V230	P231 ^O	D234 ^{OD2}	N232 ^{OD1}	P231	-	-	F8	Q93 ^{NE2}	A113	D118 ^{OD1}	R185 ^{NH1}	
2. Non-classical															
a) β_2 m-binding															
i) lipid-binding															
e.g. hsCD1d	-	variable	L232	P233 ^O	D236 ^{OD1}	N234 ^{OD1}	P233	Q93 ^{OE1}	-	-	Q93 ^{NE2}	A113	D118 ^{OD1/2}	S188 ^{OG}	
ggCD1-1	-	-	L231	P232 ^O	D235 ^{OD1/2}	N233 ^{OD1}	P232	Q89 ^{OE1}	variable	L8	Q89 ^{NE2}	A109	N114 ^{OD1}	R187 ^{NE/NH2}	
ii) protein-binding															
e.g. hsFcRn	-	F224 ^O	-	P226 ^O	D229 ^{OD2}	N227 ^{OD1}	P226	Q89 ^{OE1}	-	-	Q89 ^{NE2}	A109	variable	Variable	
HFE	D230 ^{OD1}	-	L232	P233 ^O	D236 ^{OD2}	-	P233	Q94 ^{OE1}	R46 ^{NE/NH2}	L8	Q94 ^{NE2}	G114	D119 ^{OD1}	-	
b) not associated with β_2 m															
e.g. ZAG	D228	L230	L230	H231	N232/G233	N232	H231	Q93	M46	I8	Q93	Y113	D118	K199	
hsMICA	D232	L234	L234	P235	D236/G237	D236	P235	Q91	Q46	N8	Q91	Y111	L116	T201	

[§]H₂O, water-mediated contact; π , pi-stacking

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

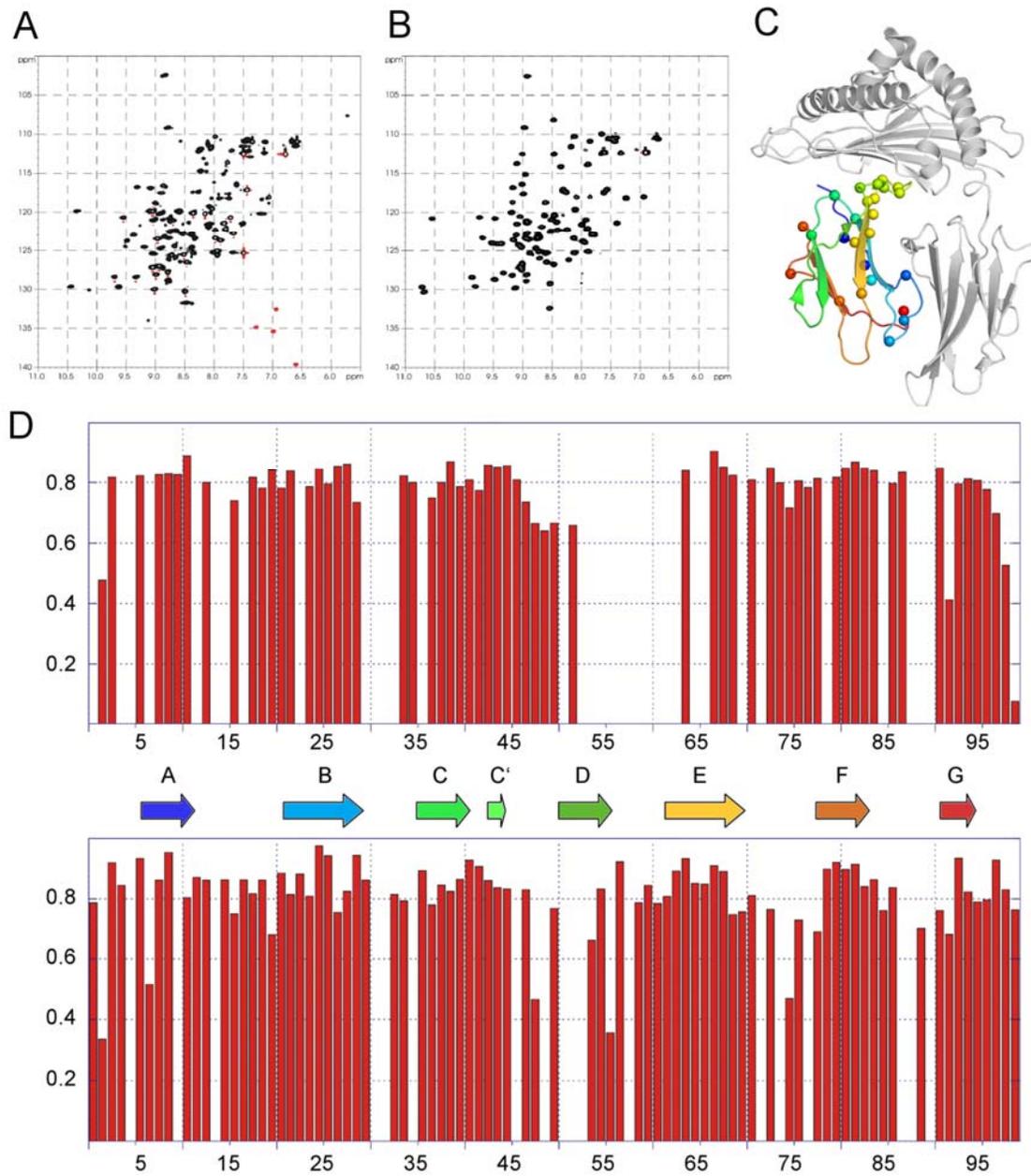


Figure 1. Comparison between the heteronuclear NMR data of human β_2 m free in solution and as part of the complex B*27:05-pVIPR.

(A) In the $^1\text{H}, ^{15}\text{N}$ HSQC of β_2 m free in solution, only 75 of the expected 94 backbone resonances can be observed, the others are missing due to exchange broadening. (B) In the $^1\text{H}, ^{15}\text{N}$ HSQC of the complexed β_2 m, a uniform linewidth can be observed and an almost complete assignment is possible. (C) Crystal structure of the B*27:05-pVIPR complex (PDB entry 1OGT). HC is in grey and β_2 m is in rainbow colour according to the β -strands A-G in (D). Unbound β_2 m residues rendered inflexible upon binding to HC are shown as spheres. (D) Comparison between the heteronuclear NOE of β_2 m free in solution (upper chart) and when complexed with the HC of B*27:05 and the pVIPR peptide (lower chart). The secondary structure of free β_2 m according to the X-ray structure (PDB entry 1LDS) is shown in between.

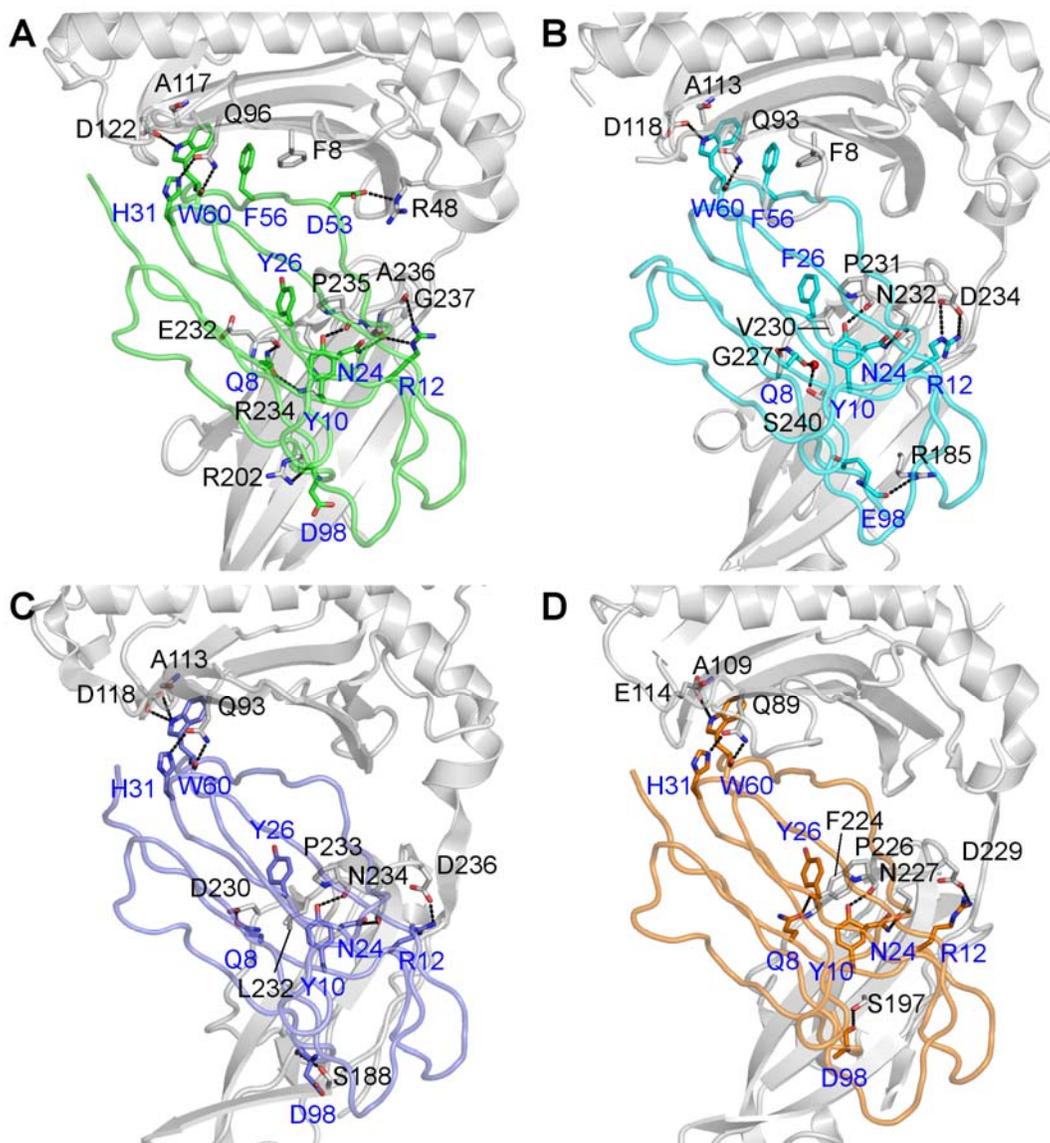


Figure 2. Comparison of β_2 m-HC interactions.

HC are shown as grey cartoon representation and β_2 m is depicted as loop representation. The β_2 m molecule of the HLA-A2 is shown in green (A), YF1*7.1 in light blue (B), human CD1d in violet (C), and human FcRn in orange (D). Only residues involve in interface interactions are shown. β_2 m and HC residues are labelled with blue and black text, respectively. Hydrogen bonds and salt bridges are denoted with black dashed lines. A water molecule in (B) is shown as red sphere.

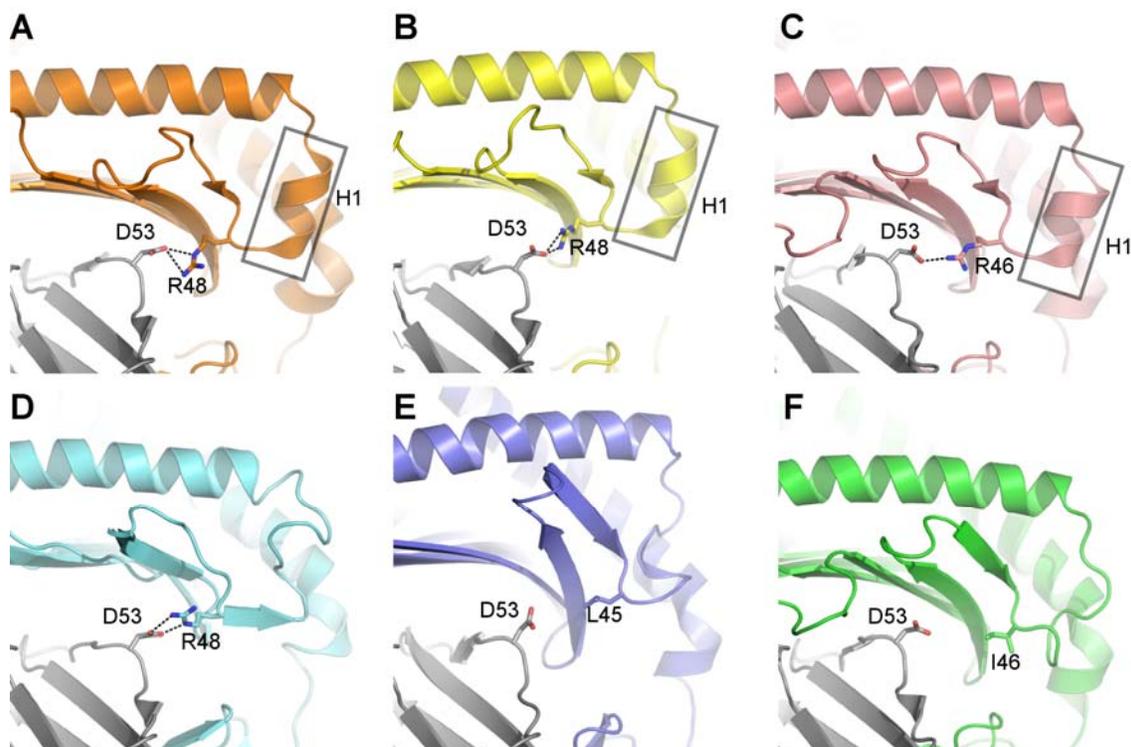


Figure 3. β_2 m-HC interface interactions and H1 helix conformations.

Molecules are viewed from the α 1 helix towards the α 2 helix and ligands are omitted for clarity. Salt bridges are shown as black dashed lines. The position of the H1 helices is labelled and indicated with a black rectangle. Structures used for the comparison are (A) HLA-A2, (B) mouse H-2Kb, (C) chicken BF2*2101, (D) HFE, (E) human CD1a and (F) chicken YF1*7.1 (see Methods for the PDB entries). β_2 m in all structures are shown in grey. The H1 helices are retained in classical peptide-binding MHC class I molecules that possess the highly conserved β_2 m-HC interactions between β_2 m-Asp53 and HC-Arg48 (HC-Arg46 in BF2*2101) (Figure A-C). Despite the presence of the β_2 m-Asp53 – HC-Arg48 contact, the H1 helix is not found in the HFE complex which does not bind a peptide (D). Molecules that do not bind a peptide such as CD1a (E) and YF1*7.1 (F) possess neither the H1 helix nor the Asp53-Arg48 contact due to amino acid substitutions of residues resembling the mammalian Arg48.

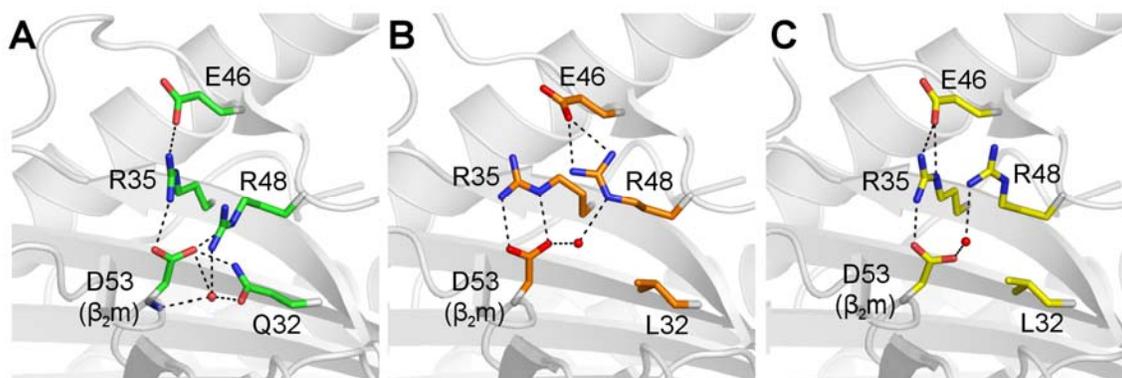


Figure 4. Influence of HC residue 32 on β_2 m-HC interface contacts.

Hydrogen bonds and salt bridges are shown as black dashed lines. Water molecules are displayed as red spheres. (A) A typical interacting network observed in most of the HLA-A and HLA-B complexes that possess HC residue Gln32. (B) Interactions observed exclusively in HLA-B27 complexes except for the water-mediated contacts which occur only in about half of the HLA-B27 structures studied (see **Table S2** for details). (C) Interactions in HLA-B44 structures, except for HLA-B*44:05. The Gln32Leu substitutions in HLA-B27 and HLA-B44 complexes result in different interface contacts made between β_2 m-Asp53 and HC residues. HLA-A2 (PDB entry 3D25), HLA-B*27:05 (PDB entry 1OGT) and HLA-B*44:02 (PDB entry 1M6O) were used to generate these figures.

3. Comparative Biophysical Characterizations of Human and Chicken β_2 -microglobulin and the Evolution of Interfaces in MHC Class I Molecules

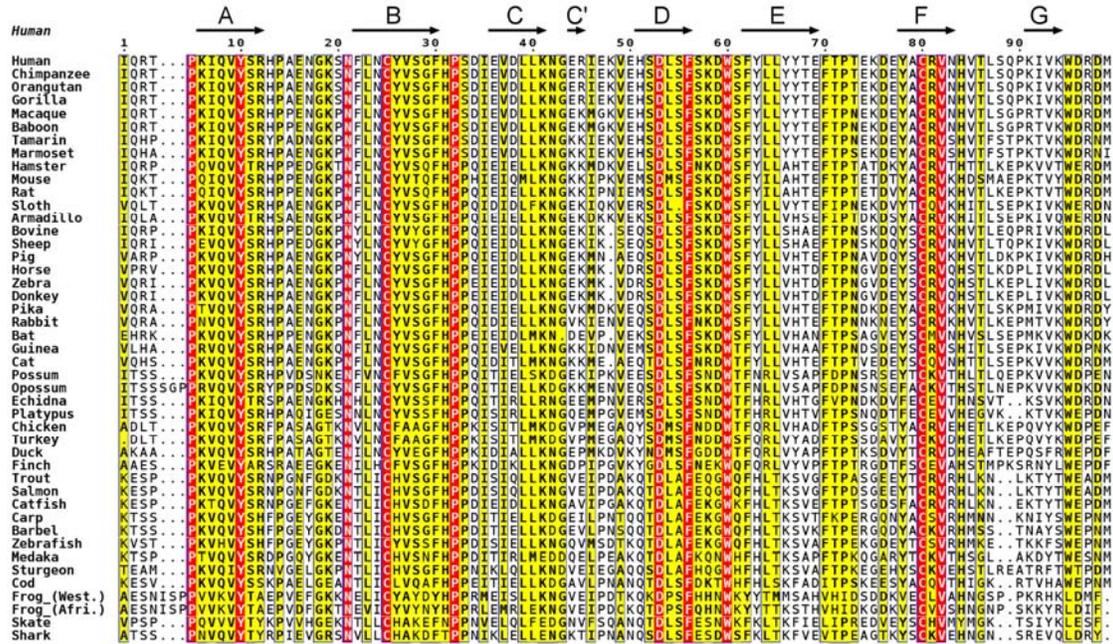


Figure S1. Amino acid sequence alignment of β_2m from various species. The numbering refers to the human β_2m sequence. Completely conserved residues are highlighted with red boxes and residues with more than 70% sequence similarity are shown as bold letters and boxed in yellow. Black arrows above the alignment indicate β -strands A to G of the human β_2m structure (PDB entry 1LDS). See accession numbers in “Materials and Methods” for sequences used in this alignment.

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

HLA-A2 numbering	8	10	12	14	23	27	32	35	48	96	117	122	188	202	204	206	231	232	234	235	236	237	238	242	244		
<i>Placental mammals</i>																											
HLA-A2	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
HLA-B*2705	F	T	V	R	I	Y	L	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
HLA-E	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
HLA-G	F	A	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Macaque	F	T	M	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Tamarin	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Orangutan	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Chimpanzee	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	E	G	D	Q	W		
Gorilla	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Baboon	F	T	M	R	F	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Dog	F	T	V	R	I	Y	Q	R	R	Q	A	D	R	R	W	L	A	E	R	P	A	G	D	Q	W		
Horse	F	T	V	R	I	Y	Q	R	R	Q	A	D	Q	R	W	L	A	E	R	P	A	G	D	Q	W		
Cat	F	T	V	R	I	Y	Q	R	R	Q	A	D	Q	R	W	L	A	E	R	P	A	G	D	Q	W		
Bovine BoLA	F	T	V	R	I	Y	Q	R	R	Q	G	D	H	R	W	L	E	E	R	P	S	G	D	Q	W		
Sheep	F	T	V	R	L	Y	Q	R	R	Q	G	D	H	R	W	L	E	E	R	P	S	G	D	Q	W		
Pig	F	T	V	R	L	Y	Q	R	R	Q	A	D	H	R	W	L	K	G	R	P	S	G	D	Q	W		
Rat RT1	F	T	V	R	I	Y	E	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Mouse H-2Db	F	T	V	R	M	Y	E	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
Mouse H-2M3	F	T	V	R	I	Y	Q	R	R	Q	A	D	H	R	W	L	A	E	R	P	S	G	D	Q	W		
Mouse H-2Db	F	T	V	R	I	Y	E	R	R	Q	A	D	H	R	W	L	A	E	R	P	A	G	D	Q	W		
<i>Marsupial mammals</i>																											
Opossum	F	T	M	R	I	Y	Q	R	R	Q	A	D	R	R	R	Q	A	E	R	P	G	G	D	Q	W		
Foosum	F	T	V	G	L	Y	Q	R	R	Q	A	D	R	R	R	Q	A	E	R	P	A	G	D	Q	W		
Bandicoot	F	T	W	R	V	Y	E	R	R	Q	A	D	R	R	R	Q	S	E	R	P	A	G	D	Q	W		
Tas. Devil	F	T	V	R	L	Y	Q	R	R	Q	A	D	R	Q	R	Q	S	E	R	P	A	G	D	Q	W		
Kangaroo	F	T	V	G	V	Y	P	S	R	Q	A	D	R	R	R	Q	K	E	R	P	A	G	D	Q	W		
<i>Egg-laying mammals</i>																											
Platypus	F	T	V	R	T	Y	Q	R	R	Q	G	D	E	R	R	L	A	G	R	P	S	G	D	Q	W		
Echidna	F	T	V	R	T	Y	P	R	R	Q	G	D	R	R	R	L	A	E	R	P	S	G	E	.	.		
<i>Birds</i>																											
Chicken BF2*21	I	T	M	D	V	Y	L	H	R	Q	A	D	R	S	R	H	R	G	V	P	N	G	D	H	W		
Chicken YF1*7.1	F	T	M	D	V	Y	I	T	I	Q	A	D	R	S	H	H	R	G	V	P	N	S	D	H	S		
Quail	F	T	M	D	V	Y	I	H	R	Q	A	D	R	S	R	H	R	G	V	P	N	S	D	H	W		
Turkey	F	T	M	D	V	Y	I	H	R	Q	A	D	R	S	R	H	R	G	A	P	N	S	D	H	L		
Goose	F	T	V	D	V	Y	V	R	Q	Q	G	D	R	S	R	H	R	S	V	P	N	S	D	H	W		
Duck	F	T	V	D	V	S	V	Y	R	Q	G	D	R	S	R	H	R	S	V	P	N	S	D	H	W		
Crane	F	I	V	E	V	Y	L	R	G	Q	A	D	R	H	R	Y	R	S	A	P	N	S	D	Y	W		
Warbler	L	V	V	E	M	F	P	R	L	L	G	D	H	S	H	Y	N	G	V	P	N	S	D	H	W		
Guinea fowl	V	T	M	D	V	Y	L	H	R	Q	A	D	R	S	R	H	R	G	V	P	N	G	D	H	W		
<i>Reptiles</i>																											
Lizard	F	T	V	E	I	Y	L	Q	R	Q	G	D	K	I	R	D	K	L	A	P	N	V	D	H	W		
Iguana (A.crist.)	F	T	V	D	F	Y	Q	S	T	Q	A	D	K	I	R	G	K	L	S	P	N	S	D	Y	W		
Iguana (I.iguana)	F	T	V	E	F	Y	P	S	K	Q	A	D	K	I	R	D	K	L	S	P	N	S	D	Y	W		
Iguana (C.Subcr.)	F	T	V	D	F	Y	P	S	K	Q	G	N	K	I	Q	G	K	L	S	P	N	S	D	Y	W		
Tuatara	F	T	V	E	M	Y	P	H	R	Q	G	D	L	S	R	H	R	G	V	P	S	G	D	Q	W		
<i>Amphibians</i>																											
Frog (R.pipi.)	Y	T	V	S	S	Y	E	N	R	Q	G	D	K	Q	L	Y	R	P	L	P	H	P	D	Q	R		
Frog (X.laevius)	Y	T	V	D	S	Y	Q	R	A	Q	G	E	R	R	Q	Y	R	E	L	P	N	P	D	Q	R		
Frog (X.tropi.)	Y	T	V	D	S	Y	Q	R	A	Q	G	E	K	R	Q	Y	R	E	L	P	N	P	D	Q	R		
Frog (X.ruwen.)	Y	T	V	D	S	Y	Q	R	A	Q	V	E	K	R	Q	Y	R	E	L	P	N	P	D	Q	R		
Axolotl	F	S	L	E	S	F	P	G	R	Q	A	D	M	T	Y	Y	R	Q	L	P	N	P	D	Q	K		
<i>Bony fish</i>																											
Catfish	F	T	V	P	T	Q	Q	Y	K	Q	S	D	S	V	H	T	K	E	L	P	N	Q	D	Q	R		
Grass carp	F	T	V	G	T	L	Q	Y	K	Q	G	D	S	T	H	T	S	E	L	P	N	G	D	Q	S		
Zebrafish	Y	T	T	G	V	L	L	Y	Q	Q	G	D	S	V	H	T	S	E	I	P	N	E	D	Q	T		
Trout	F	T	S	E	V	M	Q	H	K	Q	G	D	S	T	H	T	S	E	L	Q	N	D	D	Q	S		
Atlantic salmon	F	T	S	E	V	V	Q	H	K	Q	G	D	S	T	H	T	S	E	L	Q	N	D	D	Q	S		
Atlantic cod	F	T	S	G	V	M	Q	H	K	Q	G	D	S	V	H	T	D	E	L	P	N	H	N	Q	S		
Medaka	F	T	S	G	L	Y	Q	H	K	Q	G	D	S	S	H	T	D	Q	L	P	N	N	D	Q	S		
Fugu	F	T	S	G	V	M	Q	R	K	Q	G	D	S	S	H	T	D	E	L	L	N	P	D	Q	S		
<i>Cartilaginous fish</i>																											
Leopard shark	F	T	M	-	V	Y	E	H	R	Q	G	N	S	S	V	T	Q	G	R	P	N	H	D	Q	E		
Nurse shark	F	T	S	-	V	Y	Q	Q	R	Q	A	D	T	S	V	T	H	G	R	P	N	H	D	Q	H		
<i>Non-classical MHC-I</i>																											
Human HFE	L	M	A	E	E	Y	L	F	R	Q	G	D	K	R	R	L	Q	D	L	P	N	G	D	Q	W		
Human FcRn	H	T	V	S	N	W	Q	S	C	Q	A	E	R	T	S	F	P	D	G	P	N	S	D	H	S		
Rat FcRn	H	A	V	D	W	Q	T	C	Q	A	E	R	T	A	F	P	S	G	P	N	G	D	H	W			
Human CD1a	I	I	S	Y	L	W	Q	T	L	Q	A	D	W	V	H	S	K	D	L	P	S	A	D	Y	R		
Human CD1b	I	T	S	T	Q	W	Q	G	L	Q	A	D	W	V	H	S	K	D	L	P	N	A	N	Y	R		
Human CD1d	M	I	T	V	Q	W	Q	G	L	Q	A	D	W	V	H	S	K	D	M	P	N	A	D	Y	R		
Bovine CD1b3	L	I	S	A	D	W	Q	S	L	Q	A	D	W	V	H	S	K	D	L	P	N	A	D	Y	R		
Mouse CD1d	L	M	S	A	D	W	Q	R	T	Q	A	Y	W	V	H	S	K	D	L	P	N	A	D	Y	Q		
Chicken CD1-1	L	F	T	Q	G	L	K	S	Y	Q	A	N	V	V	R	T	R	T	L	P	N	A	D	Q	R		
Chicken CD1-2	F	T	L	G	T	L	P	V	C	Q	G	D	V	V	R	T	R	T	L	P	N	A	D	Q	R		

Figure S2. Amino acid sequence alignment of MHC class I and class I-like HC. Only positions involved in contacting β 2m are shown, see Table S1 for the interactions and Accession Numbers in Materials and Methods for the amino acid sequences used in this alignment.

(Next page) **Table S1.** Contact table between β ₂m residue Asp53 to HC residues Gln32, Arg35 and Arg48 within classical human HLA class I complexes as well as HC to HC residue interactions i.e. water-mediated Gln32-Arg48 contact and Asp46 to Arg35 and/or Arg48. PDB entry is shown for molecule without properly designated peptide (in parenthesis) and water-mediated contacts between Asp53 and Arg48 are indicated with (H₂O).

3. Comparative Biophysical Characterizations of Human and Chicken β 2-microglobulin and the Evolution of Interfaces in MHC Class I Molecules

β m residue	HC residues			HC to HC residue contacts			Substitution	
	D53	Q32	R35	R48	Q32-H ₂ O-R48	E46 – R35		E46 - R48
A*02:01-MAGE-A4	✓	✓	✓		✓	✓	✗	
A*02:01 (3HLA)	✓	✓	✗		✗	✓	✓	
A*02:01-pHIV	✓	✓	✓		✓	✓	✗	
A*02:01-pHA1	✓	✓	✓		✓	✓	✗	
A*11:01-HBV	✓	✓	✓		✓	✓	✗	
A*68:01 (IHSB)	✓	✓	✗		✗	✓	✗	
A*68:01 (ITMC)	✓	✓	✓		✓	✓	✗	
A*68:01 (2HLA)	✓	✓	✓		✗	✓	✗	
B*14:02-pLMP2	✓	✓	✓		✓	✓	✗	
B*14:02-pCatA	✓	✓	✓		✓	✓	✗	
B*15:01-pEBV1	✓	✓	✓		✓	✗	✗	E46A
B*15:01-pEBV2	✓	✓	✓		✓	✗	✗	E46A
B*35:01-pCYP	✓	✓	✓		✓	✓	✗	
B*35:01-pEBV	✓	✓	✓		✗	✓	✗	
B*51:01-KM1	✓	✓	✓		✓	✓	✗	
B*51:01-KM2	✓	✓	✓		✗	✓	✗	
B*53:01-pHIV	✓	✓	✓		✓	✓	✗	
B*53:01-pHIV2	✓	✓	✓		✓	✓	✗	
B*53:01-pLS6	✓	✗	✓		✗	✓	✓	
B*57:03-pHIV1	✓	✓	✓		✓	✗	✗	E46A
B*57:03-pHIV2	✓	✓	✓		✓	✗	✗	E46A
B*27:03-pLMP2	✗	✓	✗		✗	✗	✓	Q32L
B*27:03-pVIPR	✗	✓	✗		✗	✗	✓	Q32L
B*27:04-pB27	✗	✓	✗		✗	✗	✓	Q32L
B*27:04-pLMP2	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:04-pVIPR	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:05-pCatA	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:05-pCAC	✗	✓	✗		✗	✗	✓	Q32L
B*27:05-pGR	✗	✓	✗		✗	✗	✓	Q32L
B*27:05-M9	✗	✓	✗		✗	✗	✓	Q32L
B*27:05-TIS	✗	✓	✗		✗	✗	✓	Q32L
B*27:05-pVIPR	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:06-pLMP2	✗	✓	✗		✗	✗	✓	Q32L
B*27:06-pVIPR	✗	✓	✗		✗	✗	✓	Q32L
B*27:09-pCatA	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:09-pGR	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:09-m9	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:09-TIS	✗	✓	✓ (H ₂ O)		✗	✗	✓	Q32L
B*27:09-pVIPR	✗	✓	✗		✗	✗	✓	Q32L
B*44:02-pDPA	✗	✓	✓ (H ₂ O)		✗	✓	✗	Q32L
B*44:03-pDPA	✗	✓	✓ (H ₂ O)		✗	✓	✗	Q32L
B*44:05-pABC	✗	✓	✓		✗	✓	✓	Q32L
B*44:05-pMP	✗	✓	✗		✗	✓	✗	Q32L

4. PDF-embedded Interactive 3D Images

4.1 Summary

The conventional means of publishing scientific information involves printed media or the current form of electronic PDF files that are available through the internet. These are, however, inadequate to convey complex 3D structural information. It is especially hard for readers unfamiliar with the subject to “visualize” 3D details from 2D mediums. Although molecular visualization programs like Jmol are available through the Protein Data Bank (PDB) and Proteopedia for those who are interested in a particular 3D structure, these programs are not able to communicate subtle structural details like multiple conformations of a side chain or the dynamics of a short peptide. More advanced programs like PyMOL and COOT can perform more complicated tasks, but they require a prior in depth knowledge of the programs in order to manipulate a 3D structure as one wishes. The technology and tools to embed 3D models into PDF files were available since 2008, but the process was far from being straightforward. Together with two colleagues, Drs. Pravin Kumar and Alexander Ziegler, I was able to establish a protocol which requires as much reduced number of step and programs in order to embed not only one but even multiple models into a PDF file. To further enhance the clarity of 3D structural information, valuable features like labelling and cross-sectioning can now also be added to the 3D models. Several articles containing embedded 3D models have already been published since the establishment of this procedure.

4.2 Publications

4.2.1 Kumar, P., Ziegler, A., Grahn, A., **Hee, C.S.**, Ziegler, A. (2010) Leaving the structural ivory tower, assisted by interactive 3D PDF. *Trends Biochem. Sci.* 35, 419-422.

4.2.2 Loll, B., Rückert, C., **Hee, C.S.**, Saenger, W., Uchanska-Ziegler, B., Ziegler, A. (2011) Loss of recognition by cross-reactive T cells and its relation to a C-terminus-induced conformational reorientation of an HLA-B*2705-bound peptide. *Prot. Sci.* 20, 278-290.

4.2.3 Uchanska-Ziegler, B., Loll, B., Fabian, H., **Hee, C.S.**, Saenger, W., Ziegler, A. (2011) HLA class I-associated diseases with a suspected autoimmune etiology: HLA-B27 subtypes as a model system. *Eur. J. Cell Biol.* (in press)

4.2.1 Kumar, P., Ziegler, A., Grahn, A., **Hee, C.S.**, Ziegler, A. (2010) Leaving the structural ivory tower, assisted by interactive 3D PDF. Trends Biochem. Sci. 35, 419-422.

The original article is online available at <http://dx.doi.org/10.1016/j.tibs.2010.03.008>

4.2.2 Loll, B., Rückert, C., **Hee, C.S.**, Saenger, W., Uchanska-Ziegler, B., Ziegler, A. (2011)
Loss of recognition by cross-reactive T cells and its relation to a C-terminus-induced conformational reorientation of an HLA-B*2705-bound peptide. *Prot. Sci.* 20, 278-290.

The original article is online available at <http://dx.doi.org/10.1002/pro.559>

4.2.3 Uchanska-Ziegler, B., Loll, B., Fabian, H., **Hee, C.S.**, Saenger, W., Ziegler, A. (2011) HLA class I-associated diseases with a suspected autoimmune etiology: HLA-B27 subtypes as a model system. Eur. J. Cell Biol. (in press)

The original article is online available at <http://dx.doi.org/10.1016/j.ejcb.2011.03.003>

5. Discussion

The results of the chapters 2, 3 and 4 have already been discussed in the relevant articles or manuscripts and will not be repeated here. In this section, I aim to outline the broader aspects related to the results of this dissertation that have not been discussed previously.

5.1 *MHC-Y* and Class I *YF* of Other Avian Species

The major finding of this doctoral study is the identification of the lipid-binding feature of the chicken classical class I molecule, YF1*7.1, whose gene maps to the *MHC-Y* region. Several avian species belonging to the Phasianidae family have also been found to contain MHC gene clusters located within two separate chromosomal segments, like the chicken *MHC-B* and *MHC-Y* regions. The turkey (*Meleagris gallopavo*) is another galliform species with an MHC that has been the subject of genetic and immunological research. Similar to the chicken, the turkey *MHC-B* and *MHC-Y* regions are located on the same chromosome that contains also a *NOR* (Chaves et al., 2007; Delany et al., 2009). The arrangement of the three regions in this species is similar to that of chicken, but the direction is distinct: *MHC-B* in turkey is located on the p-arm instead of the q-arm as observed in chicken (Fig. 1.3). Interestingly, MHC class I sequences similar to the chicken *YFI* have also been detected in some turkey BAC clones, suggesting that these genes are orthologous to chicken *YFI* loci (Chaves et al., 2007).

The ring-necked pheasant (*Phasianus colchicus*) (Jarvi et al., 1996; Wittzell et al., 1995), the black grouse (*Tetrao tetrix*) (Strand et al., 2007) and the greater prairie chicken (*Tympanuchus cupido*) (Eimes et al., 2010) are further three galliform species with their *MHC* genes organized into two genetically independent clusters such as the *MHC-B/MHC-Y* system in chicken. In these studies, the authors characterized the class IIB genes (*BLB*) in the *MHC-B* region, using a combination of restriction fragment length polymorphisms (RFLP), polymerase chain reaction (PCR) and DNA sequencing. They concluded that the class IIB genes in these species assort independently, similar to the chicken. Strand (2007), Eimes (2010) and their co-workers took a step further to investigate the class IIB loci (*YLB*) in the *MHC-Y* of black grouse and greater prairie chicken and demonstrated that the *YLB* sequences of these two species are orthologous to those of the chicken, because they cluster on a separate evolutionary branch with strong bootstrap support next to the *BLB* sequences of other galliform species (Fig. 5.1). Eimes and co-workers (2010) also included an unpublished *YLB* sequence from common quail (*Coturnix coturnix*) (Nishibori et al., 2007, unpublished data, Genbank accession number AB020333) in the phylogenetic tree and this quail *YLB*

grouped together with the other *YLB* sequences (Fig. 5.1). The phylogenetic tree shown in Fig. 5.1 also supports the independent evolution of the *BLB* and the *YLB* loci, which have been shown to be genetically unlinked in chicken (Briles et al., 1993) and turkey (Chaves et al., 2007).

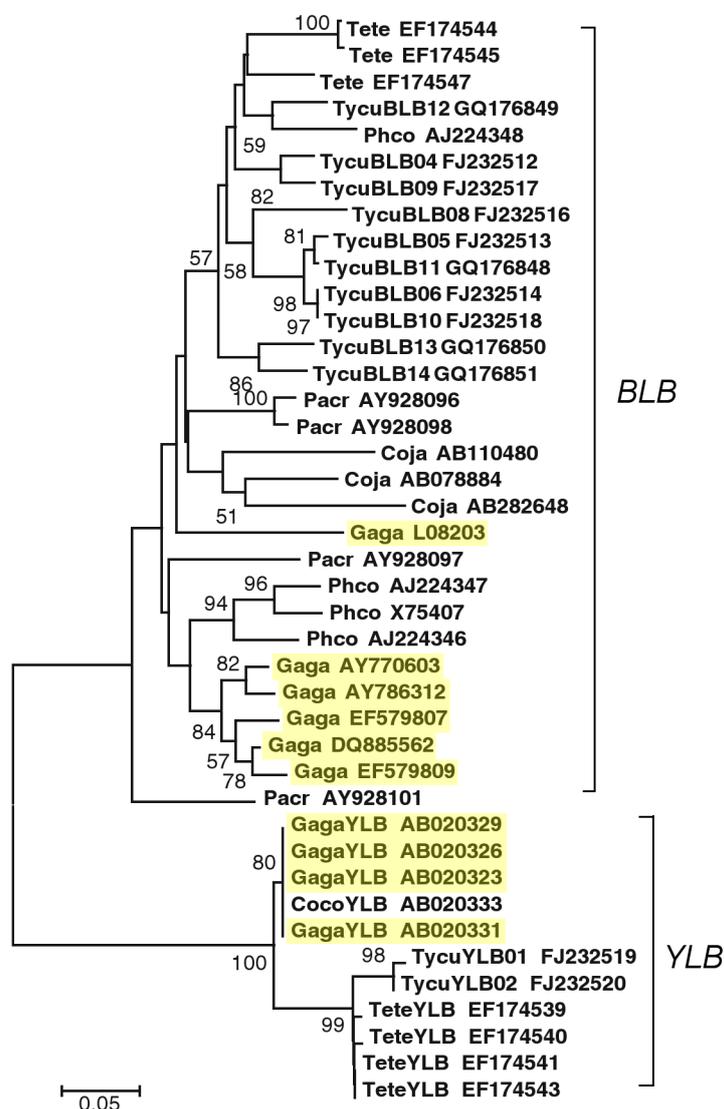


Figure 5.1. Phylogenetic tree based on the 192 bp fragment of exon 2 of *BLB* and *YLB* sequences from seven galliform species. Chicken *BLB* and *YLB* sequences are highlighted. The scale bar at the bottom left indicates nucleotide substitutions per site. Bootstrap values below 50 are not shown. *Tycu*: *Tympanuchus cupido* (greater prairie chicken); *Tete*: *Tetrao tetrix* (black grouse); *Phco*: *Phasianus colchicus* (ring-necked pheasant); *Pacr*: *Pavo cristatis* (Indian peafowl); *Coco*: *Coturnix coturnix* (common quail); *Coja*: *Coturnix japonica* (Japanese quail); *Gaga*: *Gallus gallus* (chicken). (Adapted and modified from Eimes et al., 2010).

Although these data provide evidence for the existence of *MHC-Y* in these galliform species, it remains to be determined whether these gene clusters are fully comparable to that of chicken. The MHC class I loci (*YF*) in the *MHC-Y* region of galliform species other than chicken and turkey have not been reported as most of the avian MHC studies focus on class II genes, while class I genes are often disregarded. This is even more evident for the class I loci located outside of the core MHC region like the *YFI* genes. It will be interesting to determine if these *MHC-Y*-possessing galliform species also contain class I loci orthologous to chicken *YFI* genes and whether these loci also share the characteristics of chicken *YFI* region genes.

Based on these data, we can now postulate that the segregation of MHC genes into two separate gene clusters can be traced back, at least, to the common ancestor of the galliform family Phasianidae about 37 MYA (Fig. 5.2) (Dimcheff et al., 2002). It is currently unclear if the two gene families of *MHC-B* and *MHC-Y* arose after the divergence of the Odontophoridae/Numididae and Phasianidae lineages or earlier because *MHC-Y* loci of non-galliform species have not been reported. There is a recent study on the MHC class IIB loci in guinea fowl (*Numida meleagris*), but the authors only focused on the *BLB* genes in the *MHC-B* region (Singh et al., 2010). Therefore, at the moment, we can only assume that *MHC-Y* is a common feature of Phasianidae species. The detection of *YLB* genes in several Phasianidae species (Fig. 5.1) allows us to postulate that the *MHC-Y* of birds belonging to this family may also contain the chicken MHC class I *YF* orthologs since the class I and class II loci are usually physically linked, as in the *MHC-B* region and mammalian MHC (Kelley et al., 2005).

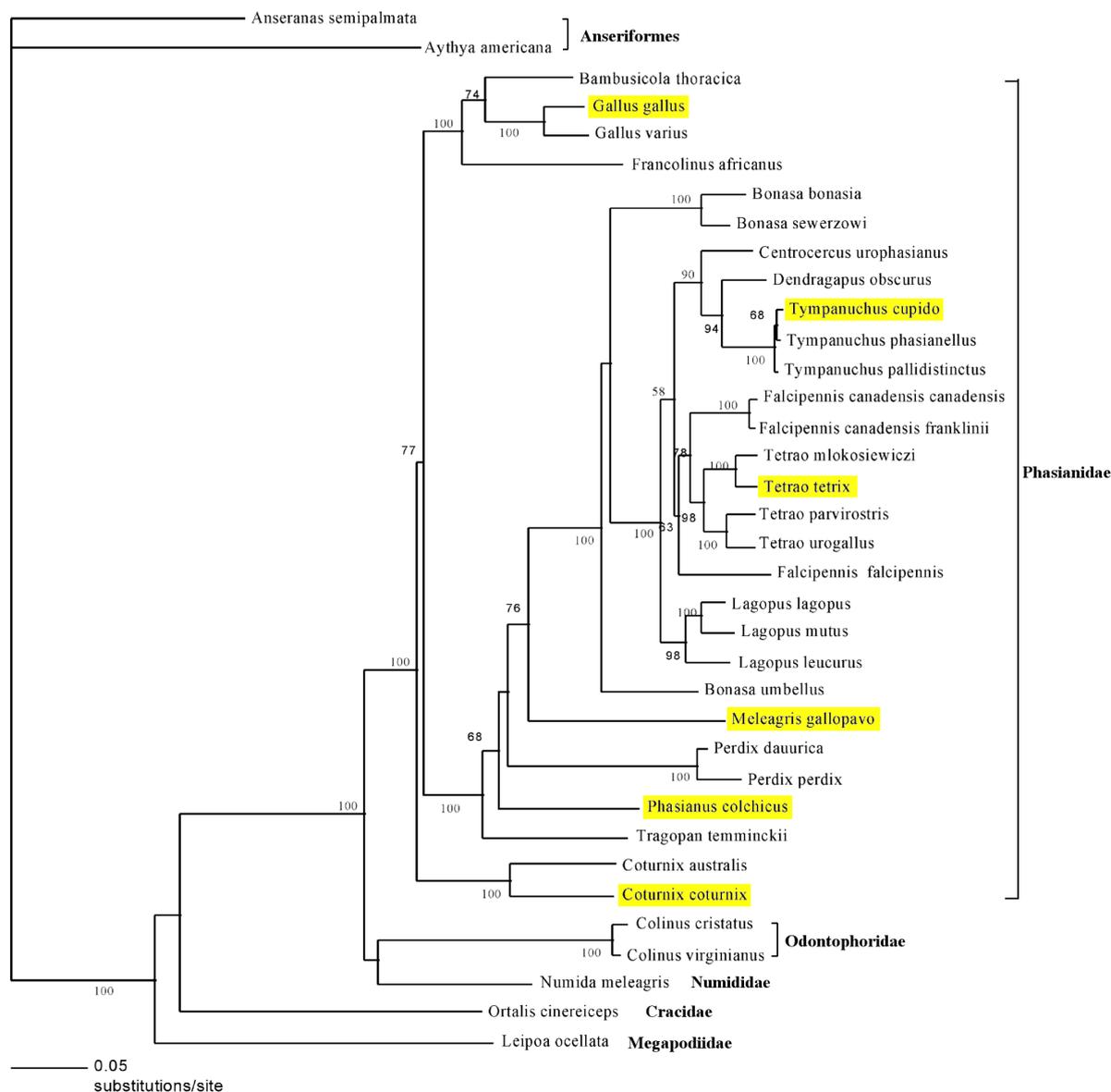


Figure 5.2. Phylogenetic relationships among Galliform and selected distantly related avian species. Species with evidence for the presence of *MHC-Y* are highlighted with yellow. Branch lengths are proportional to the number of inferred changes as shown at the bottom left corner of the tree. Bootstrap values greater than 50 are displayed next to the nodes. Common names for the bird families: Anseriformes (waterfowl), Phasianidae (pheasants and partridges), Odontophoridae (New World quails), Numididae (guineafowl), Cracidae (chachalacas, guans and curassows) and Megapodiidae (incubator birds or mound-builders). (Adapted and modified from Dimcheff et al., 2002).

Another interesting feature of the Phasianidae family is the small size and reduced complexity of their “minimal essential” *MHC-B*. Based on the number of *BLB* loci, turkey, black grouse, ring-necked pheasant and greater prairie chicken have been predicted to also having the “minimal essential” *MHC-B* (Fig. 1.2), similar to the chicken. An exception to this miniscule MHC in Phasianidae birds is the Japanese quail (*Coturnix japonica*) which appears

to have an expanded number of MHC class I (Shiina et al., 1999b) and class II loci (Shiina et al., 2004). Seven MHC class I loci map to the core 180 kb *Coja-MHC* of quail and two of these have been classified as classical class I loci based on their expression levels and tissue distribution (Shiina et al., 1995; Shiina et al., 1999a; Shiina et al., 1999b; Shiina et al., 2004). Outside of the *Coja-MHC*, at least six MHC class I loci have been detected, but the exact locations of these genes remain to be established. It is also not known whether some of these loci resemble the *YF* genes in the chicken *MHC-Y* (Shiina et al., 2004). It would be interesting to know if Japanese quail also contains an *MHC-Y* region because *MHC-Y* has been found (through the detection of a *YLB* gene) in a closely related bird, the common quail (Nishibori et al., 2007, unpublished data, Genbank accession number AB020333). The differences in the number of MHC class I loci in quail and chicken have been explained by the migrating behaviour of quail as opposed to the domestication of chicken. The quail immune system may have been stimulated by a larger variety of pathogens, subsequently promoting the expansion of MHC class I loci (Shiina et al., 2004), while chickens, on the other hand, have been domesticated and raised under artificial selection, possibly in more “pathogen-free” environments for at least the last 7500 years (Fumihito et al., 1994). This may have resulted in the possible contraction of the *MHC-B* and the expansion of the *MHC-Y* regions in this species (Shiina et al., 2004).

Outside of the Galliformes, several passerine birds have been predicted to possess the two *MHC-B* and *MHC-Y* clusters based on the number of RFLP bands and the phylogenetic analyses of MHC class I sequences. These studies included great reed warblers (*Acrocephalus arundinaceus*) (Westerdahl et al., 1999) and house sparrows (*Passer domesticus*) (Bonneaud et al., 2004). However, it is not yet clear if some of these MHC class I loci are really related to chicken *YF1* loci. Other studies focusing on MHC class II loci identified a second MHC cluster in a number of passerine birds including Hawaiian honeycreepers (*Hemignathus virens* and *Vestiaria coccinea*) (Jarvi et al., 2004), Savannah sparrows (*Passerculus sandwichensis*) (Freeman-Gallant et al., 2002), and possibly some but not all continental and Darwin’s finches, e.g., the St. Lucia Black finch (*Melanospiza richardsoni*), the Lesser Antillean bullfinch (*Loxigilla noctis*), the Cuban Grassquit (*Tiaris canorus*), the Woodpecker finch (*Cactospiza pallid*) and the Large Cactus finch (*Geospiza conirostris*) (Sato et al., 2011). Whether these putative second MHC clusters do really resemble the chicken *MHC-Y* region and also contain MHC class I loci like chicken *YF1* will be subject to further studies. In addition to a second MHC cluster, some passerine birds have

also been shown to possess a much more expanded MHC. For instance, zebra finch (*Taeniopygia guttata*) MHC class I genes map to four chromosomes based on cytogenetic analyses and genome assembly, in addition to the 10 MHC class I gene-containing contigs that are presently unincorporated into the MHC genome assembly (Balakrishnan et al., 2010).

MHC class I loci in other avian species outside of the orders of Galliformes and Passeriformes have also been studied including duck (*Anas platyrhynchos*) (Xia et al., 2004) and goose (*Anser cygnoides*) (Xia et al., 2005) from the Anseriformes, Sandhill cranes (*Grus canadensis pratensis*) (Jarvi et al., 1999) from the Gruiformes and Red-billed gull (*Larus scopulinus*) (Cloutier et al., 2011) from the Chadriiformes. However, class I loci orthologous to the chicken *YF1* gene were not reported. These studies focused mainly on the MHC class I loci located in the core MHC by using PCR primers designed to correspond to conserved sites of classical MHC class I exons. Therefore, it is likely that these primers missed other MHC class I or class I-like sequences since these differ often considerably from classical MHC class I sequences.

The identification of the lipid-binding YF1*7.1 molecule described in this dissertation demonstrates that class I genes located outside of the core MHC region deserve to be studied in more detail. These non-MHC encoded class I genes may have evolved specialized functions in order to complement more “conventional” genes of the immune system or other functions not directly related to the immune system. The functions of class I genes residing outside of the classical MHC are often difficult to determine. This study provides an example on how these genes can be examined using a combination of structural and biochemical approaches. The major limitation of a structural study of MHC class I molecules is the identification of their ligands which is the prerequisite for complex formation. This work is, however, laborious and time-consuming. The methods described here are suitable at least for these non-MHC encoded class I molecules because many of them will have developed binding grooves that do not require a ligand stable complex formation. YF1*7.1 molecules provide the first example for a classical MHC class I protein that does not require a distinct peptide for complex formation.

5.2 Evolution of MHC Class I and β_2m Genes

This dissertation focused on two related molecules, the chicken MHC class I heavy chains and β_2m . Their molecular structures, as well as their interchain interactions were studied in detail. Furthermore, a comparison of interchain interactions with those of other MHC class I and class I-like molecules was carried out. Hence, it is rational to delve into the evolutionary relationship between these two types of genes.

β_2m shares a distinctive structure called constant-1 (C1) Ig superfamily (IgSF) domain with other adaptive immune system-related molecules including MHC class I and class II chains (Williams and Barclay, 1988). Despite their structural similarity, however, the β_2m locus is located outside of the MHC in all species so far examined from bony fish to mammals (Michaelson, 1981; Ono et al., 1993; Riegert et al., 1996). It has been speculated that the translocation of the β_2m gene out of the MHC has the advantage of preventing it being subjected to accumulation of polymorphisms, duplications and deletions, which affect class I genes in many vertebrates (Ono et al., 1993). However, β_2m had long been predicted to be tightly linked to the primordial MHC and this assumption has only recently been demonstrated in the MHC of nurse shark (*Ginglymostoma cirratum*) (Ohta et al., 2011). The same group also showed the close linkage between MHC class I and class II genes in the genome of this fish species (Ohta et al., 2000), which is in agreement with the genome of another shark species – the elephant shark (*Callorhinchus milii*) (Venkatesh et al., 2007). The close proximity of the β_2m , MHC class I and class II genes is consistent with the theory that they derived from a common ancestor by tandem (*cis*) duplication (Hughes and Nei, 1993; Klein and O'huigin, 1993).

To understand the evolutionary relationship between β_2m and MHC class I genes, it is essential to include MHC class II genes into the phylogenetic analyses. There has been a long debate on whether MHC class I or class II emerged first, dating back as long as 20 years (Flajnik et al., 1991; Hughes and Nei, 1993; Klein and O'huigin, 1993).

The “class I-first” hypothesis postulates that the ancestral class I HC was a merger between a C1-IgSF domain and a peptide-binding domain derived from heat shock protein (HSP) 70 (Fig. 5.3A) (Flajnik et al., 1991). This assumption is based on the fact that both the C1-IgSF domain as well as HSP are more ancient than MHC proteins (Lindquist, 1986; Benian et al., 1989), with HSP70 family genes present in all organisms. In eukaryotes, HSP occur in all

major cellular compartments (Hendrick and Hartl, 1993). The genes encoding a peptide binding domain of HSP are thought to have been transferred en bloc to a gene segment encoding an Ig-like domain, resulting in the proto-MHC class I gene (Fig. 5.3A) (Flajnik, 2000). Another fact that supports this hypothesis is that at least one of the members of the HSP70 family (in human, all three *HSP70* loci) is encoded in the MHC class III region (Sargent et al., 1989; Horton et al., 2004). The linkage of HSP and MHC genes may not be coincidental, since there is evidence for an essential interaction between HSP70 and other MHC-encoded molecules in reproduction (Ziegler et al., 2010). After duplication of the HSP70-derived class I HC and evolution of β_2m genes from a C1-IgSF domain, a class II α chain gene is thought to have been generated by fusion of a class I HC-domain exon with a β_2m gene, leaving a truncated HC behind that may have evolved into a class II β chain gene (Fig. 5.3A). The two-domain proteins encoded by these new class II genes were ready to associate during biosynthesis through their peptide-binding domains (Flajnik et al., 1991).

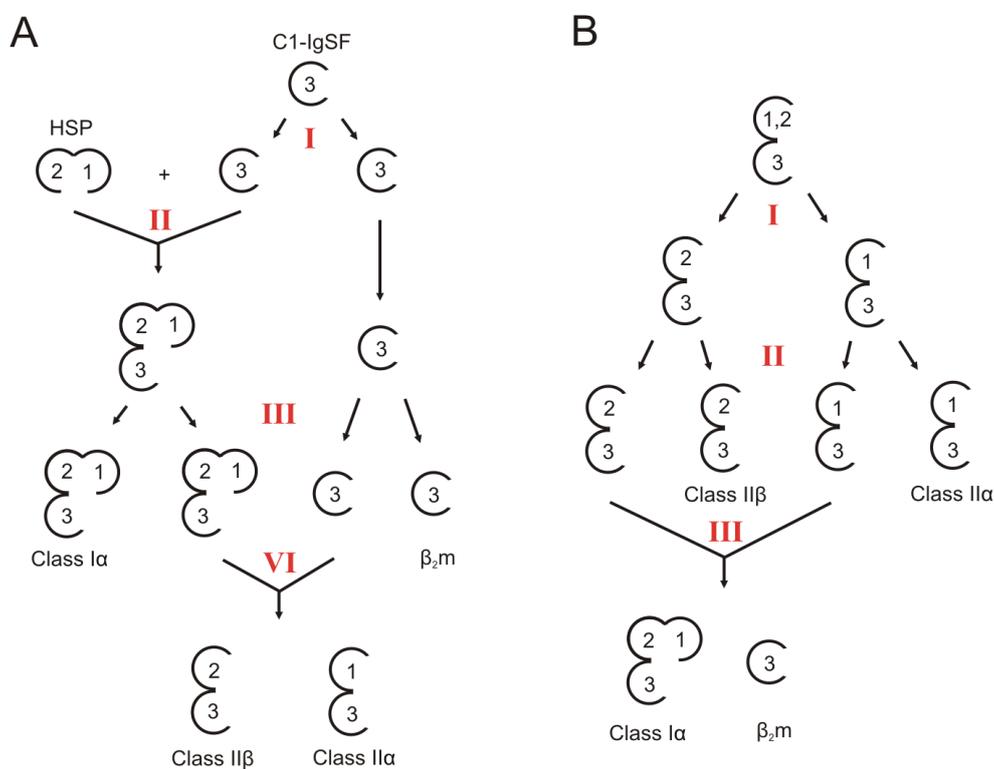


Figure 5.3. Models of the evolution of MHC molecules and β_2m . Models based on the “MHC class I-first” (A) and “MHC class II-first” (B) hypotheses. The domains are designated as follows: “1” is homologous to class I $\alpha 1$ domains, “2” is homologous to class I $\alpha 2$ domains and “3” is homologous to class I $\alpha 3$ domains. Genetic events that are believed to have taken place in the evolution of the molecules are labelled I to IV. (Adapted and modified from Hughes and Nei, 1993).

The “MHC class I-first” hypothesis is strengthened by the more “plastic” evolution of the MHC class I gene family than the class II evolution (Klein and Figueroa, 1986). Class I molecules have been implicated in a variety of phenomena including mate selection and reproduction (Edidin, 1988; Ziegler et al., 2005) and class I-like molecules have been demonstrated to perform other functions besides antigen presentation (Simister and Mostov, 1989; Feder et al., 1998; Oganesyanyan et al., 2002). Hence, the function of the primordial class I molecule might have been unrelated to the immune system (Klein and Figueroa, 1986), contrary to class II molecules, which appear to have no other function besides antigen presentation. Based on this “higher plasticity”, the proponents of the “MHC class I-first” hypothesis claim that a class I molecule is more likely to be the ancestor that evolved and gave rise to the class II and β_2m molecules (Flajnik, 2000).

The second hypothesis supposes that MHC class II loci originated first (Hood et al., 1985; Lawlor et al., 1990; Hughes and Nei, 1993; Klein and O’huigin, 1993). Hughes and Nei (1993) analyzed the phylogenetic relationships of C1-domains in class I α , class II α , class II β together with β_2m and discovered that class I α and class II β are closely related to each other while class II α and β_2m cluster in the phylogenetic tree, in agreement with a previous phylogenetic analysis (Hood et al., 1985). Based on this fact, it is postulated that an ancestral class II gene, either a monomer or a homodimer, duplicated and gave rise to the α/β heterodimer (Fig. 5.3B). The class II α and class II β chain genes then duplicated in tandem and eventually a recombinational event translocated the $\alpha 1$ domain of a class II α chain gene into a position between the leader and the $\beta 1$ exon of a class II β chain gene, thereby giving rise to the class I HC (Fig. 5.3B) (Hughes and Nei, 1993). The loss of the $\alpha 1$ domain exon left a class II α chain gene with the leader and the $\alpha 2$ exon, which eventually became the ancestral β_2m gene (Lawlor et al., 1990; Hughes and Nei, 1993).

The “class II-first” supporters oppose the idea that peptide-binding domains derive from HSP with the argument that MHC peptide-binding domains are more similar to the variable domains of members of the IgSF, e.g. CD4 and CD8, than to HSP70 (Hughes and Nei, 1993). In addition, the class II $\alpha 1$ and $\beta 1$ domains show much greater similarity to each other than to HSP domains. It is argued that if class II $\alpha 1$ and $\beta 1$ domains derive from class I $\alpha 1$ and $\alpha 2$ domains, which in turn derive from two separate domains of HSP, they must have diverged to a great extent over 1.5 billion years, making them unlikely to share high sequence similarity. Thus, the similarity between class II $\alpha 1$ and $\beta 1$ appears to disprove the “class I-first”

hypothesis which claims that the peptide binding domains are derived from HSP. Alternatively, if class II α and class II β chain genes diverged from each other by gene duplication, as proposed by the “class II-first” hypothesis, this event may have taken place only about 446-521 MYA (Hughes and Nei, 1990), thus explaining the similarity between the two genes. This is also in line with the observation that amphibians, which diverged from lobe-finned fish about 370 MYA, have both class II α and II β genes (Kaufman et al., 1990) and with the physical linkage between class I, II and III genes in the amphibian *Xenopus* (Nonaka et al., 1997). The “class I-first” hypothesis was first proposed by Flajnik and co-workers (1991), based on the homology modelling similarity between class I HC and the HSP70 peptide-binding domain (Rippmann et al., 1991). However, crystal structures of HSP70 later revealed that the HSP peptide-binding domain is very different compared to that of a MHC class I or class II peptide-binding groove (Zhu et al., 1996). The latter binding grooves are much better “defined” than the HSP70 peptide-binding domain, which does not possess an apparent peptide binding groove (Fig. 5.4). The peptides are also bound in a totally distinct manner by the HSP70 molecule (Zhu et al., 1996).

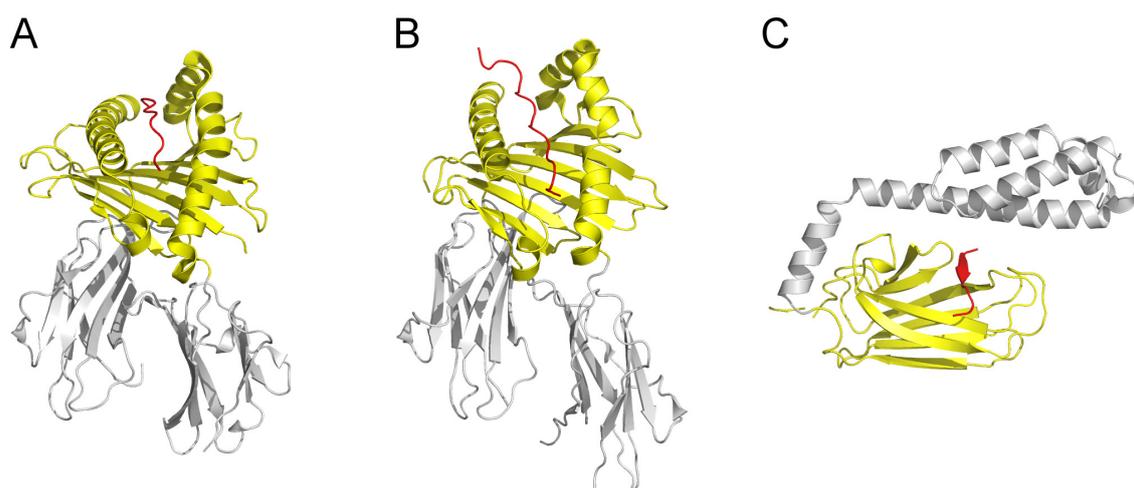


Figure 5.4. Comparison of peptide-binding domains. Peptide-binding domains of an MHC class I molecule (A), an MHC class II molecule (B), and HSP70 (C), are coloured in yellow, other parts of the molecules in grey. Peptides are shown as red lines. PDB entries: MHC class I (1W0V), MHC class II (3DP0), and HSP70 (1DKX).

The “class II-first” hypothesis is also more plausible than the “class I-first” theory because only three genetic events are required to explain the former hypothesis while four events are needed to account for the presence of class II, class I, and β_2m genes in the latter (Fig. 5.3). The “class II-first” scenario is thus more parsimonious, requiring fewer steps and the intermediate molecules it assumes might be more plausible (Hughes and Nei, 1990). It is also difficult to imagine the function of the ancestral single domain molecule as proposed by the “class I-first” and the role of the succeeding molecules over a long period of time until the emergence of class I α chain. Moreover, this scenario requires one C domain (β_2m) to immediately associate with the newly recombined class I α chain (derived from HSP and C domain) to form a more stable four-domain heterodimer, despite the fact that there was a long, separated evolutionary history between the two molecules (Fig. 5.3) (Hughes and Nei, 1993).

Note that certain evolutionary considerations favour the peptide-binding domain derived from the IgSF variable (V) domain as proposed by the “class II-first” hypothesis, despite the structural difference between MHC peptide-binding domains and V domains (Hunkapiller and Hood, 1989). First, all MHC molecules are members of the IgSF and it may therefore be more likely that the peptide-binding domain evolved from a domain type found in this family. Second, the peptide-binding domains display higher sequence similarity to IgSF V domains than to HSP70, further indicating that they evolved from V domains (Hughes and Nei, 1993). Finally, the length of exons encoding MHC peptide-binding domains is similar to that of V domains, contrary to HSP70 whose $\alpha 1$ and $\alpha 2$ domains are not separated by exon boundaries like those in MHC class I HC.

The two hypotheses place the emergence of the ancestral β_2m at different times in the course of MHC evolution (Fig. 5.3). The “class I-first” hypothesis assumes that the ancestral β_2m appeared long before the birth of class I and class II genes, while the “class II-first” hypothesis postulates the emergence of the ancestral β_2m gene as a result of class II chain genes losing a domain-encoding exon that gave rise to class I α genes. The origin of β_2m , like the origin of class I and class II genes, must currently remain an unsolved question.

However, also I regard it as more likely that class II genes emerged before the class I and β_2m genes because the arguments stated by “class II-first” hypothesis appear more plausible to me. The strongest evidence comes from phylogenetic analyses, which is an accurate and reliable method to study the evolution of a group of closely related genes. As pointed out in the

discussion above, the phylogenetic results may be more in favour of the “class II-first” hypothesis. In addition to the “class II-first” arguments, the evolutionary “route” provided by this hypothesis supports also the structural complementarity between β_2m and class I molecules since their predecessor, the class II α and II β chains, appear already to have interacted with each other to form a complex. Furthermore, as shown in our analyses of HC- β_2m interface interactions, several of these are highly conserved from chicken to mammals and in different classes of class I molecules. This indicates that these contacts are crucial for the formation of complexes. If the “class I-first” hypothesis would be right, one would probably have to assume that the class II α and II β chains that emerged, according to this hypothesis, after class I genes (Fig. 5.3), should have retained the conserved interactions. However, these are not found in MHC class II molecules. On the other hand, as the “class II-first” hypothesis proposes that class I and β_2m genes derive from class II genes, the conserved contacts may have appeared later in the evolution of class I HC and β_2m genes, perhaps to strengthen the interaction of the two newly emerged molecules.

5.3 Concluding remarks

The results presented in this dissertation primarily contribute to the understanding of the chicken immune system. Two related molecules, the chicken YF1*7.1 and β_2m , were biophysically characterized. The results show that YF1*7.1 is a classical MHC class I molecule with the capability to bind lipids. Structural and sequence comparisons of YF1*7.1 to other classical MHC class I molecules reveal a unique structural feature in chickens that appears to be shared by classical class I molecules of non-mammalian but not of mammalian vertebrates. On the other hand, chicken β_2m is structurally similar, but thermodynamically different to β_2m of other species so far investigated. This fact can be explained by the fewer intra-molecular salt-bridges found in chicken β_2m . The crystal packing of free chicken β_2m is also unprecedentedly tight compared to that of β_2m from other species. Analyses of HC- β_2m interface interactions in different classes of class I molecules across species identify several conserved interactions involving evolutionarily retained HC and β_2m residues. Finally, a simplified method to embed 3D models into a PDF document was optimized and developed into a powerful tool to communicate sophisticated 3D structures within scientific publications.

6. References

- Achour, A., Michaelsson, J., Harris, R.A., Ljunggren, H.G., Karre, K., et al. (2006). Structural basis of the differential stability and receptor specificity of H-2Db in complex with murine versus human beta2-microglobulin. *J. Mol. Biol.* *356*, 382-396.
- Afanassieff, M., Goto, R.M., Ha, J., Sherman, M.A., Zhong, L., et al. (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.
- Bacon, L.D., Hunt, H.D., and Cheng, H.H. (2001). Genetic resistance to Marek's disease. *Marek's Disease* *255*, 121-141.
- Bacon, L.D. and Witter, R.L. (1992). Influence of turkey herpesvirus vaccination on the B-haplotype effect on Marek's disease resistance in 15.B-congenic chickens. *Avian Dis.* *36*, 378-385.
- Bacon, L.D. and Witter, R.L. (1995). Efficacy of Marek's disease vaccines in Mhc heterozygous chickens: Mhc congenic x inbred line F1 matings. *J. Hered.* *86*, 269-273.
- Bahram, S., Bresnahan, M., Geraghty, D.E., and Spies, T. (1994). A second lineage of mammalian major histocompatibility complex class I genes. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 6259-6263.
- Balakrishnan, C.N., Ekblom, R., Volker, M., Westerdahl, H., Godinez, R., et al. (2010). Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biol.* *8*, 29.
- Bartl, S., Baish, M.A., Flajnik, M.F., and Ohta, Y. (1997). Identification of class I genes in cartilaginous fish, the most ancient group of vertebrates displaying an adaptive immune response. *J. Immunol.* *159*, 6097-6104.
- Bauer, M.M. and Reed, K.M. (2011). Extended sequence of the turkey MHC B-locus and sequence variation in the highly polymorphic B-G loci. *Immunogenetics* *63*, 209-221.
- Bauer, S., Groh, V., Wu, J., Steinle, A., Phillips, J.H., et al. (1999). Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. *Science* *285*, 727-729.
- Beck, S., Geraghty, D., Inoko, H., Rowen, L., Aguado, B., et al. (1999). Complete sequence and gene map of a human major histocompatibility complex. *Nature* *401*, 921-923.
- Beck, S. and Trowsdale, J. (2000). The human major histocompatibility complex: lessons from the DNA sequence. *Annu. Rev. Genomics Hum. Genet.* *1*, 117-137.
- Becker, J.W. and Reeke, G.N., Jr. (1985). Three-dimensional structure of beta 2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A.* *82*, 4225-4229.
- Bell, M.J., Burrows, J.M., Brennan, R., Miles, J.J., Tellam, J., et al. (2009). The peptide length specificity of some HLA class I alleles is very broad and includes peptides of up to 25 amino acids in length. *Mol. Immunol.* *46*, 1911-1917.
- Benian, G.M., Kiff, J.E., Neckelmann, N., Moerman, D.G., and Waterston, R.H. (1989). Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. *Nature* *342*, 45-50.
- Berggard, I. and Bearn, A.G. (1968). Isolation and properties of a low molecular weight beta-2-globulin occurring in human biological fluids. *J. Biol. Chem.* *243*, 4095-4103.

- Bernabeu, C., Maziarz, R., Murre, C., and Terhorst, C. (1985). Beta 2-microglobulin from serum associates with several class I antigens expressed on the surface of mouse L-cells. *Mol. Immunol.* *22*, 955-960.
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., et al. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* *329*, 506-512.
- Bonneaud, C., Sorci, G., Morin, V., Westerdahl, H., Zoorob, R., et al. (2004). Diversity of Mhc class I and IIB genes in house sparrows (*Passer domesticus*). *Immunogenetics* *55*, 855-865.
- Borg, N.A., Wun, K.S., Kjer-Nielsen, L., Wilce, M.C.J., Pellicci, D.G., et al. (2007). CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. *Nature* *448*, 44-49.
- Borrego, F., Ulbrecht, M., Weiss, E.H., Coligan, J.E., and Brooks, A.G. (1998). Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. *J. Exp. Med.* *187*, 813-818.
- Boyington, J.C., Brooks, A.G., and Sun, P.D. (2001). Structure of killer cell immunoglobulin-like receptors and their recognition of the class I MHC molecules. *Immunol. Rev.* *181*, 66-78.
- Brigl, M. and Brenner, M.B. (2004). CD1: antigen presentation and T cell function. *Annu. Rev. Immunol.* *22*, 817-890.
- 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., Stone, H.A., and Cole, R.K. (1977). Marek's disease: effects of B histocompatibility alloalleles in resistant and susceptible chicken lines. *Science* *195*, 193-195.
- Brossay, L., Burdin, N., Tangri, S., and Kronenberg, M. (1998). Antigen-presenting function of mouse CD1: one molecule with two different kinds of antigenic ligands. *Immunol. Rev.* *163*, 139-150.
- Burmeister, W.P., Huber, A.H., and Bjorkman, P.J. (1994). Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* *372*, 379-383.
- Cao, W., Xi, X., Hao, Z., Li, W., Kong, Y., et al. (2007). RAET1E2, a soluble isoform of the UL16-binding protein RAET1E produced by tumor cells, inhibits NKG2D-mediated NK cytotoxicity. *J. Biol. Chem.* *282*, 18922-18928.
- Carrington, M. and Bontrop, R.E. (2002). Effects of MHC class I on HIV/SIV disease in primates. *AIDS* *16 Suppl 4*, S105-S114.
- Castano, A.R., Tangri, S., Miller, J.E., Holcombe, H.R., Jackson, M.R., et al. (1995). Peptide binding and presentation by mouse CD1. *Science* *269*, 223-226.
- Cejka, J., Van Nieuwkoop, J.A., Mood, D.W., Kithier, K., and Radl, J. (1976). Beta2-microglobulin in human colostrum and milk: effect of breast feeding and physico-chemical characterization. *Clin. Chim. Acta* *67*, 71-78.
- Chaves, L.D., Krueth, S.B., and Reed, K.M. (2007). Characterization of the turkey MHC chromosome through genetic and physical mapping. *Cytogenet. Genome Res.* *117*, 213-220.

- Chaves, L.D., Krueth, S.B., and Reed, K.M. (2009). Defining the turkey MHC: sequence and genes of the B locus. *J. Immunol.* *183*, 6530-6537.
- Chen, W., Gao, F., Chu, F., Zhang, J., Gao, G.F., et al. (2010). Crystal structure of a bony fish beta2-microglobulin: insights into the evolutionary origin of immunoglobulin superfamily constant molecules. *J. Biol. Chem.* *285*, 22505-22512.
- Cloutier, A., Mills, J.A., and Baker, A.J. (2011). Characterization and locus-specific typing of MHC class I genes in the red-billed gull (*Larus scopulinus*) provides evidence for major, minor, and nonclassical loci. *Immunogenetics* *63*, 377-394.
- Comiskey, M., Goldstein, C.Y., De Fazio, S.R., Mammolenti, M., Newmark, J.A., et al. (2003). Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product. *Hum. Immunol.* *64*, 999-1004.
- Cosman, D., Mullberg, J., Sutherland, C.L., Chin, W., Armitage, R., et al. (2001). ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity* *14*, 123-133.
- D'Orazio, S.E., Halme, D.G., Ploegh, H.L., and Starnbach, M.N. (2003). Class Ia MHC-deficient BALB/c mice generate CD8⁺ T cell-mediated protective immunity against *Listeria monocytogenes* infection. *J. Immunol.* *171*, 291-298.
- Dausset, J. (1958). [Iso-leuko-antibodies]. *Acta Haematol.* *20*, 156-166.
- Davis, M.M. and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* *334*, 395-402.
- Delany, M.E., Robinson, C.M., Goto, R.M., and Miller, M.M. (2009). Architecture and organization of chicken microchromosome 16: order of the NOR, MHC-Y, and MHC-B subregions. *J. Hered.* *100*, 507-514.
- Deres, K., Beck, W., Faath, S., Jung, G., and Rammensee, H.G. (1993). MHC/peptide binding studies indicate hierarchy of anchor residues. *Cell Immunol.* *151*, 158-167.
- Dimcheff, D.E., Drovetski, S.V., and Mindell, D.P. (2002). Phylogeny of Tetraoninae and other galliform birds using mitochondrial 12S and ND2 genes. *Mol. Phylogenet. Evol.* *24*, 203-215.
- Dvir, H., Wang, J., Ly, N., Dascher, C.C., and Zajonc, D.M. (2010). Structural basis for lipid-antigen recognition in avian immunity. *J. Immunol.* *184*, 2504-2511.
- Edidin, M. (1988). Function by association? MHC antigens and membrane receptor complexes. *Immunol Today* *9*, 218-219.
- Eichner, T., Kalverda, A.P., Thompson, G.S., Homans, S.W., and Radford, S.E. (2011). Conformational conversion during amyloid formation at atomic resolution. *Mol. Cell* *41*, 161-172.
- Eimes, J.A., Bollmer, J.L., Dunn, P.O., Whittingham, L.A., and Wimpee, C. (2010). Mhc class II diversity and balancing selection in greater prairie-chickens. *Genetica* *138*, 265-271.
- Feder, J.N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D.A., et al. (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.* *13*, 399-408.

- Feder, J.N., Penny, D.M., Irrinki, A., Lee, V.K., Lebron, J.A., et al. (1998). The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. *Proc. Natl. Acad. Sci. U. S. A.* *95*, 1472-1477.
- Fernando, M.M., Stevens, C.R., Walsh, E.C., De Jager, P.L., Goyette, P., et al. (2008). Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* *4*, e1000024.
- Figuerola, F., Mayer, W.E., Sato, A., Zaleska-Rutczynska, Z., Hess, B., et al. (2001). Mhc class I genes of swordtail fishes, *Xiphophorus*: variation in the number of loci and existence of ancient gene families. *Immunogenetics* *53*, 695-708.
- Flajnik, M.F. (2000). Origins of the MHC. *J. Leukocyte Biol.* *34*.
- Flajnik, M.F., Canel, C., Kramer, J., and Kasahara, M. (1991). Which came first, MHC class I or class II? *Immunogenetics* *33*, 295-300.
- 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.
- Flajnik, M.F., Kasahara, M., Shum, B.P., Salter-Cid, L., Taylor, E., et al. (1993). A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian *Xenopus*. *EMBO J.* *12*, 4385-4396.
- Flajnik, M.F., Ohta, Y., Namikawa-Yamada, C., and Nonaka, M. (1999). Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunol. Rev.* *167*, 59-67.
- Floege, J. and Ehlerding, G. (1996). Beta-2-microglobulin-associated amyloidosis - Discussion. *Nephron* *72*, 9-26.
- Foster, T.L., Goto, R.M., Yokoyama, W.M., and Miller, M.M. (2005). Reporter cell assays reveal interaction between a putative NK cell receptor encoded within chicken MHC B and a non-classical class I allele encoded within chicken MHC Y. *FASEB J.* *19*, A335.
- Frangoulis, B., Park, I., Guillemot, F., Severac, V., Auffray, C., et al. (1999). Identification of the Tapasin gene in the chicken major histocompatibility complex. *Immunogenetics* *49*, 328-337.
- Freeman-Gallant, C.R., Johnson, E.M., Saponara, F., and Stanger, M. (2002). Variation at the major histocompatibility complex in Savannah sparrows. *Mol. Ecol.* *11*, 1125-1130.
- Fumihito, A., Miyake, T., Sumi, S., Takada, M., Ohno, S., et al. (1994). One subspecies of the red junglefowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 12505-12509.
- Garcia-Alles, L.F., Versluis, K., Maveyraud, L., Vallina, A.T., Sansano, S., et al. (2006). Endogenous phosphatidylcholine and a long spacer ligand stabilize the lipid-binding groove of CD1b. *EMBO J.* *25*, 3684-3692.
- Gasser, D.L., Klein, K.A., Choi, E., and Seidman, J.G. (1985). A new beta-2 microglobulin allele in mice defined by DNA sequencing. *Immunogenetics* *22*, 413-416.
- Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., et al. (1985). A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem. Biophys. Res. Commun.* *129*, 701-706.

- Girardi, E., Wang, J., Mac, T.T., Versluis, C., Bhowruth, V., et al. (2010). Crystal structure of bovine CD1b3 with endogenously bound ligands. *J. Immunol.* *185*, 376-386.
- Gorer, P.A. (1938). The antigenic basis of tumour transplantation. *J. Pathol.* *47*, 231-252.
- Goto, R.M., Wang, Y., Taylor, R.L., Jr., Wakenell, P.S., Hosomichi, K., et al. (2009). BG1 has a major role in MHC-linked resistance to malignant lymphoma in the chicken. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 16740-16745.
- Groh, V., Steinle, A., Bauer, S., and Spies, T. (1998). Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. *Science* *279*, 1737-1740.
- Groh, V., Wu, J., Yee, C., and Spies, T. (2002). Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature* *419*, 734-738.
- Hansen, J.D. and Kaattari, S.L. (1996). The recombination activating gene 2 (RAG2) of the rainbow trout *Oncorhynchus mykiss*. *Immunogenetics* *44*, 203-211.
- Hashimoto, K., Hirai, M., and Kurosawa, Y. (1995). A gene outside the human MHC related to classical HLA class I genes. *Science* *269*, 693-695.
- Hassan, M.I., Kumar, V., Singh, T.P., and Yadav, S. (2008a). Purification and characterization of zinc alpha2-glycoprotein-prolactin inducible protein complex from human seminal plasma. *J. Sep. Sci.* *31*, 2318-2324.
- Hassan, M.I., Waheed, A., Yadav, S., Singh, T.P., and Ahmad, F. (2008b). Zinc alpha 2-glycoprotein: a multidisciplinary protein. *Mol. Cancer Res.* *6*, 892-906.
- He, X., Tabaczewski, P., Ho, J., Stroynowski, I., and Garcia, K.C. (2001). Promiscuous antigen presentation by the nonclassical MHC Ib Qa-2 is enabled by a shallow, hydrophobic groove and self-stabilized peptide conformation. *Structure* *9*, 1213-1224.
- Hee, C.S., Gao, S., Loll, B., Miller, M.M., Uchanska-Ziegler, B., et al. (2010). Structure of a classical MHC class I molecule that binds "non-classical" ligands. *PLoS Biol.* *8*, e1000557.
- Heegaard, N.H. (2009). Beta(2)-microglobulin: from physiology to amyloidosis. *Amyloid.* *16*, 151-173.
- Hendrick, J.P. and Hartl, F.U. (1993). Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.* *62*, 349-384.
- Herberg, J.A., Beck, S., and Trowsdale, J. (1998). TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING 1, 3 to 5) form a dense cluster at the centromeric end of the MHC. *J. Mol. Biol.* *277*, 839-857.
- Hermel, E., Robinson, P.J., She, J.X., and Lindahl, K.F. (1993). Sequence divergence of B2m alleles of wild *Mus musculus* and *Mus spretus* implies positive selection. *Immunogenetics* *38*, 106-116.
- Hood, L., Kronenberg, M., and Hunkapiller, T. (1985). T cell antigen receptors and the immunoglobulin supergene family. *Cell* *40*, 225-229.
- Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., et al. (2004). Gene map of the extended human MHC. *Nat. Rev. Genet.* *5*, 889-899.
- Hughes, A.L. and Nei, M. (1990). Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol. Biol. Evol.* *7*, 491-514.

- Hughes, A.L. and Nei, M. (1993). Evolutionary relationships of the classes of major histocompatibility complex genes. *Immunogenetics* 37, 337-346.
- Huh, G.S., Boulanger, L.M., Du, H.P., Riquelme, P.A., Brotz, T.M., et al. (2000). Functional requirement for class I MHC in CNS development and plasticity. *Science* 290, 2155-2159.
- Hunkapiller, T. and Hood, L. (1989). Diversity of the immunoglobulin gene superfamily. *Adv. Immunol.* 44, 1-63.
- 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.
- Hunt, H.D., Goto, R.M., Foster, D.N., Bacon, L.D., and Miller, M.M. (2006). At least one YMHCI molecule in the chicken is alloimmunogenic and dynamically expressed on spleen cells during development. *Immunogenetics* 58, 297-307.
- Ishitani, A., Sageshima, N., Lee, N., Dorofeeva, N., Hatake, K., et al. (2003). Protein expression and peptide binding suggest unique and interacting functional roles for HLA-E, F, and G in maternal-placental immune recognition. *J. Immunol.* 171, 1376-1384.
- Ivanova, M.I., Sawaya, M.R., Gingery, M., Attinger, A., and Eisenberg, D. (2004). An amyloid-forming segment of beta2-microglobulin suggests a molecular model for the fibril. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10584-10589.
- Jarvi, S.I., Goto, R.M., Briles, W.E., and Miller, M.M. (1996). Characterization of Mhc genes in a multigenerational family of ring-necked pheasants. *Immunogenetics* 43, 125-135.
- Jarvi, S.I., Goto, R.M., Gee, G.F., Briles, W.E., and Miller, M.M. (1999). Identification, inheritance, and linkage of B-G-like and MHC class I genes in cranes. *J. Hered.* 90, 152-159.
- Jarvi, S.I., Tarr, C.L., McIntosh, C.E., Atkinson, C.T., and Fleischer, R.C. (2004). Natural selection of the major histocompatibility complex (Mhc) in Hawaiian honeycreepers (Drepanidinae). *Mol. Ecol.* 13, 2157-2168.
- Junghans, R.P. and Anderson, C.L. (1996). The protection receptor for IgG catabolism is the beta2-microglobulin-containing neonatal intestinal transport receptor. *Proc. Natl. Acad. Sci. U. S. A.* 93, 5512-5516.
- Juul-Madsen, H.R., Dalgaard, T.S., Guldbandsen, B., and Salomonsen, J. (2000). A polymorphic major histocompatibility complex class II-like locus maps outside of both the chicken B-system and Rfp-Y-system. *Eur. J. Immunogenet.* 27, 63-71.
- Kajikawa, M., Baba, T., Tomaru, U., Watanabe, Y., Koganei, S., et al. (2006). MHC class I-like MILL molecules are beta2-microglobulin-associated, GPI-anchored glycoproteins that do not require TAP for cell surface expression. *J. Immunol.* 177, 3108-3115.
- Kaufman, J., Jacob, J., Shaw, I., Walker, B., Milne, S., et al. (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., et al. (1999b). The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 401, 923-925.
- Kaufman, J. and Salomonsen, J. (1993). What in the dickens is with these chickens? An only slightly silly response to the first draft of Langman and Cohn. *Res. Immunol.* 144, 495-502.

- Kaufman, J., Skjoedt, K., and Salomonsen, J. (1990). The MHC molecules of nonmammalian vertebrates. *Immunol. Rev.* *113*, 83-117.
- 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.
- Kelley, J., Walter, L., and Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. *Immunogenetics* *56*, 683-695.
- Klein, J. (1976). Evolution and function of the major histocompatibility complex: facts and speculations. In *The major histocompatibility system in man and animals*. Gotze, D., ed. (New York, Berlin, Heidelberg: Springer), pp. 339-378.
- Klein, J. (1986). *Natural history of the major histocompatibility complex*. (New York: Wiley).
- Klein, J. and Figueroa, F. (1986). Evolution of the major histocompatibility complex. *Crit. Rev. Immunol.* *6*, 295-386.
- Klein, J. and O'huigin, C. (1993). Composite origin of major histocompatibility complex genes. *Curr. Opin. Genet. Dev.* *3*, 923-930.
- Koch, M., Camp, S., Collen, T., Avila, D., Salomonsen, J., et al. (2007). Structures of an MHC class I molecule from B21 chickens illustrate promiscuous peptide binding. *Immunity* *27*, 885-899.
- Koch, M., Stronge, V.S., Shepherd, D., Gadola, S.D., Mathew, B., et al. (2005). The crystal structure of human CD1d with and without alpha-galactosylceramide. *Nat. Immunol.* *6*, 819-826.
- Krangel, M.S., Orr, H.T., and Strominger, J.L. (1979). Assembly and maturation of HLA-A and HLA-B antigens in vivo. *Cell* *18*, 979-991.
- Kroemer, G., Zoorob, R., and Auffray, C. (1990). Structure and expression of a chicken MHC class I gene. *Immunogenetics* *31*, 405-409.
- Kubin, M., Cassiano, L., Chalupny, J., Chin, W., Cosman, D., et al. (2001). ULBP1, 2, 3: novel MHC class I-related molecules that bind to human cytomegalovirus glycoprotein UL16, activate NK cells. *Eur. J. Immunol.* *31*, 1428-1437.
- Kubota, K. (1984). Association of serum beta 2-microglobulin with H-2 class I heavy chains on the surface of mouse cells in culture. *J. Immunol* *133*, 3203-3210.
- Lancet, D., Parham, P., and Strominger, J.L. (1979). Heavy chain of HLA-A and HLA-B antigens is conformationally labile: a possible role for beta 2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A.* *76*, 3844-3848.
- Lawlor, D.A., Zemmour, J., Ennis, P.D., and Parham, P. (1990). Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. *Annu. Rev. Immunol.* *8*, 23-63.
- Lebron, J.A., Bennett, M.J., Vaughn, D.E., Chirino, A.J., Snow, P.M., et al. (1998). Crystal structure of the hemochromatosis protein HFE and characterization of its interaction with transferrin receptor. *Cell* *93*, 111-123.
- Lee, L.F., Bacon, L.D., Yoshida, S., Yanagida, N., Zhang, H.M., et al. (2004). The efficacy of recombinant fowlpox vaccine protection against Marek's disease: its dependence on chicken line and B haplotype. *Avian Dis.* *48*, 129-137.

- LePage, K.T., Miller, M.M., Briles, W.E., and Taylor, R.L., Jr. (2000). Rfp-Y genotype affects the fate of Rous sarcomas in B2B5 chickens. *Immunogenetics* *51*, 751-754.
- Li, P., McDermott, G., and Strong, R.K. (2002). Crystal structures of RAE-1beta and its complex with the activating immunoreceptor NKG2D. *Immunity* *16*, 77-86.
- Li, P., Morris, D.L., Willcox, B.E., Steinle, A., Spies, T., et al. (2001). Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nat. Immunol.* *2*, 443-451.
- Li, P., Willie, S.T., Bauer, S., Morris, D.L., Spies, T., et al. (1999). Crystal structure of the MHC class I homolog MIC-A, a gammadelta T cell ligand. *Immunity* *10*, 577-584.
- Lindquist, S. (1986). The heat-shock response. *Annu. Rev. Biochem.* *55*, 1151-1191.
- Liu, C., Sawaya, M.R., and Eisenberg, D. (2011). Beta-microglobulin forms three-dimensional domain-swapped amyloid fibrils with disulfide linkages. *Nat. Struct. Mol. Biol.* *18*, 49-55.
- Liu, Y., Xiong, Y., Naidenko, O.V., Liu, J.H., Zhang, R., et al. (2003). The crystal structure of a TL/CD8alphaalpha complex at 2.1 Å resolution: implications for modulation of T cell activation and memory. *Immunity* *18*, 205-215.
- Loconto, J., Papes, F., Chang, E., Stowers, L., Jones, E.P., et al. (2003). Functional expression of murine V213 pheromone receptors involves selective association with the M10 and M1 families of MHC class Ib molecules. *Cell* *112*, 607-618.
- 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.
- Matsui, M., Hioe, C.E., and Frelinger, J.A. (1993). Roles of the six peptide-binding pockets of the HLA-A2 molecule in allorecognition by human cytotoxic T-cell clones. *Proc. Natl. Acad. Sci. U. S. A.* *90*, 674-678.
- Michaelson, J. (1981). Genetic polymorphism of beta 2-microglobulin (B2m) maps to the H-3 region of chromosome 2. *Immunogenetics* *13*, 167-171.
- Michaelsson, J., Achour, A., Rolle, A., and Karre, K. (2001). MHC class I recognition by NK receptors in the Ly49 family is strongly influenced by the beta(2)-microglobulin subunit. *J. Immunol.* *166*, 7327-7334.
- Miller, K.M. and Withler, R.E. (1997). Mhc diversity in Pacific salmon: Population structure and trans-species allelism. *Hereditas* *127*, 83-95.
- Miller, M.M., Bacon, L.D., Hala, K., Hunt, H.D., Ewald, S.J., et al. (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., et al. (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., Young, S., Chirivella, J., Hawke, D., et al. (1991). Immunoglobulin variable-region-like domains of diverse sequence within the major histocompatibility complex of the chicken. *Proc. Natl. Acad. Sci. U. S. A.* *88*, 4377-4381.

- Miller, M.M., Goto, R., Young, S., Liu, J., and Hardy, J. (1990). Antigens similar to major histocompatibility complex B-G are expressed in the intestinal epithelium in the chicken. *Immunogenetics* 32, 45-50.
- Miller, M.M., Goto, R.M., Taylor, R.L., Jr., Zoorob, R., Auffray, C., et al. (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.
- 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.
- Munz, C., Stevanovic, S., and Rammensee, H.G. (1999). Peptide presentation and NK inhibition by HLA-G. *J. Reprod. Immunol.* 43, 139-155.
- Murphy, K., Travers, P., and Walport, M. (2008). *Janeway's Immunobiology*. (London: Garland Science).
- Murray, B.W., Nilsson, P., Zaleska-Rutczynska, Z., Sultmann, H., and Klein, J. (2000). Linkage relationships and haplotype variation of the major histocompatibility complex class I A genes in the cichlid fish *oreochromis niloticus*. *Mar. Biotechnol. (NY)* 2, 437-448.
- Nonaka, M., Namikawa, C., Kato, Y., Sasaki, M., SalterCid, L., et al. (1997). Major histocompatibility complex gene mapping in the amphibian *Xenopus* implies a primordial organization. *Proc. Natl. Acad. Sci. U. S. A.* 94, 5789-5791.
- Oganesyan, V., Oganesyan, N., Terzyan, S., Qu, D., Dauter, Z., et al. (2002). The crystal structure of the endothelial protein C receptor and a bound phospholipid. *J. Biol. Chem.* 277, 24851-24854.
- Ohta, Y., Okamura, K., McKinney, E.C., Bartl, S., Hashimoto, K., et al. (2000). Primitive synteny of vertebrate major histocompatibility complex class I and class II genes. *Proc. Natl. Acad. Sci. U. S. A.* 97, 4712-4717.
- Ohta, Y., Shiina, T., Lohr, R.L., Hosomichi, K., Pollin, T.I., et al. (2011). Primordial Linkage of {beta}2-Microglobulin to the MHC. *J. Immunol.* 186, 3563-3571.
- Ono, H., Figueroa, F., O'huigin, C., and Klein, J. (1993). Cloning of the beta 2-microglobulin gene in the zebrafish. *Immunogenetics* 38, 1-10.
- Persson, A.C., Stet, R.J.M., and Pilstrom, L. (1999). Characterization of MHC class I and beta(2)-microglobulin sequences in Atlantic cod reveals an unusually high number of expressed class I genes. *Immunogenetics* 50, 49-59.
- Peterson, P.A., Rask, L., and Ostberg, L. (1977). Beta2-microglobulin and the major histocompatibility complex. *Adv. Cancer Res.* 24, 115-163.
- Petrie, E.J., Clements, C.S., Lin, J., Sullivan, L.C., Johnson, D., et al. (2008). CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. *J. Exp. Med.* 205, 725-735.
- 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.
- Pink, J.R., Miggiano, V.C., and Ziegler, A. (1977). Antigens of the chicken major histocompatibility (B) complex. *Folia Biol. (Praha)* 23, 404-405.

- Plachy, J. and Benda, V. (1981). Location of the gene responsible for Rous sarcoma regression in the B-F region of the B complex (MHC) of the chicken. *Folia Biol. (Praha)* 27, 363-368.
- Praharaj, N., Beaumont, C., Dambrine, G., Soubieux, D., Merat, L., et al. (2004). Genetic analysis of the growth curve of Rous sarcoma virus-induced tumors in chickens. *Poult. Sci.* 83, 1479-1488.
- Radosavljevic, M., Cuillerier, B., Wilson, M.J., Clement, O., Wicker, S., et al. (2002). A cluster of ten novel MHC class I related genes on human chromosome 6q24.2-q25.3. *Genomics* 79, 114-123.
- Rammensee, H.G., Friede, T., and Stevanović, S. (1995). MHC ligands and peptide motifs: first listing. *Immunogenetics* 41, 178-228.
- Ribic, A., Zhang, M., Schlumbohm, C., Matz-Rensing, K., Uchanska-Ziegler, B., et al. (2010). Neuronal MHC class I molecules are involved in excitatory synaptic transmission at the hippocampal mossy fiber synapses of marmoset monkeys. *Cell Mol. Neurobiol.* 30, 827-839.
- Riegert, P., Andersen, R., Bumstead, N., Dohring, C., Dominguez-Steglich, M., et al. (1996). The chicken beta 2-microglobulin gene is located on a non-major histocompatibility complex microchromosome: a small, G+C-rich gene with X and Y boxes in the promoter. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1243-1248.
- Riewald, M., Petrovan, R.J., Donner, A., Mueller, B.M., and Ruf, W. (2002). Activation of endothelial cell protease activated receptor 1 by the protein C pathway. *Science* 296, 1880-1882.
- Rippmann, F., Taylor, W.R., Rothbard, J.B., and Green, N.M. (1991). A hypothetical model for the peptide binding domain of hsp70 based on the peptide binding domain of HLA. *EMBO J.* 10, 1053-1059.
- Rodgers, J.R. and Cook, R.G. (2005). MHC class IB molecules bridge innate and acquired immunity. *Nat. Rev. Immunol.* 5, 459-471.
- Rohrlich, P.S., Fazilleau, N., Ginhoux, F., Firat, H., Michel, F., et al. (2005). Direct recognition by alphabeta cytolytic T cells of Hfe, a MHC class Ib molecule without antigen-presenting function. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12855-12860.
- Rudolph, M.G., Stanfield, R.L., and Wilson, I.A. (2006). How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* 24, 419-466.
- Salomonsen, J., Marston, D., Avila, D., Bumstead, N., Johansson, B., 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., et al. (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.
- Sammut, B., Du Pasquier, L., Ducoroy, P., Laurens, V., Marcuz, A., et al. (1999). Axolotl MHC architecture and polymorphism. *Eur. J. Immunol.* 29, 2897-2907.
- Sanchez, L.M., Chirino, A.J., and Bjorkman, P. (1999). Crystal structure of human ZAG, a fat-depleting factor related to MHC molecules. *Science* 283, 1914-1919.
- Saper, M.A., Bjorkman, P.J., and Wiley, D.C. (1991). Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* 219, 277-319.

- Sargent, C.A., Dunham, I., Trowsdale, J., and Campbell, R.D. (1989). Human major histocompatibility complex contains genes for the major heat shock protein HSP70. *Proc. Natl. Acad. Sci. U. S. A.* *86*, 1968-1972.
- Sato, A., Tichy, H., Grant, P.R., Grant, B.R., Sato, T., et al. (2011). Spectrum of MHC class II variability in Darwin's finches and their close relatives. *Mol. Biol. Evol.* *28*, 1943-1956.
- Scharf, L., Li, N.S., Hawk, A.J., Garzon, D., Zhang, T., et al. (2010). The 2.5 Å structure of CD1c in complex with a mycobacterial lipid reveals an open groove ideally suited for diverse antigen presentation. *Immunity* *33*, 853-862.
- Shawar, S.M., Vyas, J.M., Rodgers, J.R., and Rich, R.R. (1994). Antigen presentation by major histocompatibility complex class I-B molecules. *Annu. Rev. Immunol.* *12*, 839-880.
- Sherman, M.A., Goto, R.M., Moore, R.E., Hunt, H.D., Lee, T.D., et al. (2008). Mass spectral data for 64 eluted peptides and structural modeling define peptide binding preferences for class I alleles in two chicken MHC-B haplotypes associated with opposite responses to Marek's disease. *Immunogenetics* *60*, 527-541.
- Shiina, T., Ando, A., Imanishi, T., Kawata, H., Hanzawa, K., et al. (1995). Isolation and characterization of cDNA clones for Japanese quail (*Coturnix japonica*) major histocompatibility complex (MhcCoja) class I molecules. *Immunogenetics* *42*, 213-216.
- Shiina, T., Briles, W.E., Goto, R.M., Hosomichi, K., Yanagiya, K., et al. (2007). Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J. Immunol.* *178*, 7162-7172.
- Shiina, T., Oka, A., Imanishi, T., Hanzawa, K., Gojobori, T., et al. (1999a). Multiple class I loci expressed by the quail Mhc. *Immunogenetics* *49*, 456-460.
- Shiina, T., Shimizu, C., Oka, A., Teraoka, Y., Imanishi, T., et al. (1999b). Gene organization of the quail major histocompatibility complex (MhcCoja) class I gene region. *Immunogenetics* *49*, 384-394.
- Shiina, T., Shimizu, S., Hosomichi, K., Kohara, S., Watanabe, S., et al. (2004). Comparative genomic analysis of two avian (quail and chicken) MHC regions. *J. Immunol.* *172*, 6751-6763.
- Shiina, T., Tamiya, G., Oka, A., Takishima, N., Yamagata, T., et al. (1999c). Molecular dynamics of MHC genesis unraveled by sequence analysis of the 1,796,938-bp HLA class I region. *Proc. Natl. Acad. Sci. U. S. A.* *96*, 13282-13287.
- Shum, B.P., Avila, D., Du, P.L., Kasahara, M., and Flajnik, M.F. (1993). Isolation of a classical MHC class I cDNA from an amphibian. Evidence for only one class I locus in the *Xenopus* MHC. *J. Immunol.* *151*, 5376-5386.
- Shum, B.P., Guethlein, L., Flodin, L.R., Adkison, M.A., Hedrick, R.P., et al. (2001). Modes of salmonid MHC class I and II evolution differ from the primate paradigm. *J. Immunol.* *166*, 3297-3308.
- Silk, J.D., Salio, M., Brown, J., Jones, E.Y., and Cerundolo, V. (2008). Structural and functional aspects of lipid binding by CD1 molecules. *Annu. Rev. Cell Dev. Biol.* *24*, 369-395.
- Simister, N.E. and Mostov, K.E. (1989). An Fc receptor structurally related to MHC class I antigens. *Nature* *337*, 184-187.
- Singer, D.S., Mozes, E., Kirshner, S., and Kohn, L.D. (1997). Role of MHC class I molecules in autoimmune disease. *Crit. Rev. Immunol.* *17*, 463-468.

- Singh, S.K., Mathew, J., Gupta, J., Mehra, S., Goyal, G., et al. (2010). Molecular characterization of MHC class II region in guinea fowl. *Br. Poultry Sci.* *51*, 769-775.
- Snell, G.D. and Higgins, G.F. (1951). Alleles at the histocompatibility-2 locus in the mouse as determined by tumor transplantation. *Genetics* *36*, 306-310.
- Stephens, R., Horton, R., Humphray, S., Rowen, L., Trowsdale, J., et al. (1999). Gene organisation, sequence variation and isochore structure at the centromeric boundary of the human MHC. *J. Mol. Biol.* *291*, 789-799.
- Story, C.M., Mikulska, J.E., and Simister, N.E. (1994). A major histocompatibility complex class I-like Fc receptor cloned from human placenta: possible role in transfer of immunoglobulin G from mother to fetus. *J. Exp. Med.* *180*, 2377-2381.
- Strand, T., Westerdahl, H., Høglund, J., Alatalo, V., and Siitari, H. (2007). The Mhc class II of the Black grouse (*Tetrao tetrix*) consists of low numbers of B and Y genes with variable diversity and expression. *Immunogenetics* *59*, 725-734.
- Sullivan, L.C., Hoare, H.L., McCluskey, J., Rossjohn, J., and Brooks, A.G. (2006). A structural perspective on MHC class Ib molecules in adaptive immunity. *Trends Immunol.* *27*, 413-420.
- Sutherland, C.L., Chalupny, N.J., and Cosman, D. (2001). The UL16-binding proteins, a novel family of MHC class I-related ligands for NKG2D, activate natural killer cell functions. *Immunol. Rev.* *181*, 185-192.
- Thoraval, P., Afanassieff, M., Bouret, D., Luneau, G., Esnault, E., et al. (2003). Role of nonclassical class I genes of the chicken major histocompatibility complex Rfp-Y locus in transplantation immunity. *Immunogenetics* *55*, 647-651.
- Totaro, A., Rommens, J.M., Grifa, A., Lunardi, C., Carella, M., et al. (1996). Hereditary hemochromatosis: Generation of a transcription map within a refined and extended map of the HLA class I region. *Genomics* *31*, 319-326.
- Treiner, E., Duban, L., Bahram, S., Radosavljevic, M., Wanner, V., et al. (2003). Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature* *422*, 164-169.
- Trinh, C.H., Smith, D.P., Kalverda, A.P., Phillips, S.E.V., and Radford, S.E. (2002). Crystal structure of monomeric human beta-2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci. U. S. A.* *99*, 9771-9776.
- Urban, R.G., Chicz, R.M., Lane, W.S., Strominger, J.L., Rehm, A., et al. (1994). A subset of HLA-B27 molecules contains peptides much longer than nonamers. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 1534-1538.
- Urdahl, K.B., Liggitt, D., and Bevan, M.J. (2003). CD8(+) T cells accumulate in the lungs of *Mycobacterium tuberculosis*-infected Kb⁻/Db⁻ mice, but provide minimal protection. *J. Immunol.* *170*, 1987-1994.
- Vaughn, D.E. and Bjorkman, P.J. (1998). Structural basis of pH-dependent antibody binding by the neonatal Fc receptor. *Structure* *6*, 63-73.
- Venkatesh, B., Kirkness, E.F., Loh, Y.H., Halpern, A.L., Lee, A.P., et al. (2007). Survey sequencing and comparative analysis of the elephant shark (*Callorhynchus milii*) genome. *PLoS Biol.* *5*, e101.

- Vilches, C. and Parham, P. (2002). KIR: diverse, rapidly evolving receptors of innate and adaptive immunity. *Annu. Rev. Immunol.* *20*, 217-251.
- Wakenell, P.S., Miller, M.M., Goto, R.M., Gauderman, W.J., and Briles, W.E. (1996). Association between the Rfp-Y haplotype and the incidence of Marek's disease in chickens. *Immunogenetics* *44*, 242-245.
- Waldmann, T.A. and Strober, W. (1969). Metabolism of immunoglobulins. *Prog. Allergy* *13*, 1-110.
- Wallny, H.J., Avila, D., Hunt, L.G., Powell, T.J., Riegert, P., et al. (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.
- Walpole, N.G., Kjer-Nielsen, L., Kostenko, L., McCluskey, J., Brooks, A.G., et al. (2010). The structure and stability of the monomorphic HLA-G are influenced by the nature of the bound peptide. *J. Mol. Biol.* *397*, 467-480.
- Wang, C.R., Castano, A.R., Peterson, P.A., Slaughter, C., Lindahl, K.F., et al. (1995). Nonclassical binding of formylated peptide in crystal structure of the Mhc class-Ib molecule H2-M3. *Cell* *82*, 655-664.
- Wang, C.R., Lindahl, K.F., and Deisenhofer, J. (1996). Crystal structure of the MHC class Ib molecule H2-M3. *Res. Immunol.* *147*, 313-321.
- Wang, J., Whitman, M.C., Natarajan, K., Tormo, J., Mariuzza, R.A., et al. (2002). Binding of the natural killer cell inhibitory receptor Ly49A to its major histocompatibility complex class I ligand - Crucial contacts include both H-2D(d) and beta(2)-microglobulin. *J. Biol. Chem.* *277*, 1433-1442.
- Watanabe, Y., Maruoka, T., Walter, L., and Kasahara, M. (2004). Comparative genomics of the Mill family: a rapidly evolving MHC class I gene family. *Eur. J. Immunol.* *34*, 1597-1607.
- Westerdahl, H., Wittzell, H., and von Schantz, T. (1999). Polymorphism and transcription of Mhc class I genes in a passerine bird, the great reed warbler. *Immunogenetics* *49*, 158-170.
- White, E.C., Briles, W.E., Briles, R.W., and Taylor, R.L., Jr. (1994). Response of six major histocompatibility (B) complex recombinant haplotypes to Rous sarcomas. *Poult. Sci.* *73*, 836-842.
- Williams, A.F. and Barclay, A.N. (1988). The immunoglobulin superfamily--domains for cell surface recognition. *Annu. Rev. Immunol.* *6*, 381-405.
- Wingren, C., Crowley, M.P., Degano, M., Chien, Y.H., and Wilson, I.A. (2000). Crystal structure of a gamma delta T cell receptor ligand T22: A truncated MHC-like fold. *Science* *287*, 310-314.
- Wittzell, H., von Schantz, T., Zoorob, R., and Auffray, C. (1995). Rfp-Y-like sequences assort independently of pheasant Mhc genes. *Immunogenetics* *42*, 68-71.
- Xia, C., Hu, T., Yang, T., Wang, L., Xu, G., et al. (2005). cDNA cloning, genomic structure and expression analysis of the goose (*Anser cygnoides*) MHC class I gene. *Vet. Immunol. Immunopathol.* *107*, 291-302.
- Xia, C., Lin, C.Y., Xu, G.X., Hu, T.J., and Yang, T.Y. (2004). cDNA cloning and genomic structure of the duck (*Anas platyrhynchos*) MHC class I gene. *Immunogenetics* *56*, 304-309.

- Ye, X., Fukudome, K., Tsuneyoshi, N., Satoh, T., Tokunaga, O., et al. (1999). The endothelial cell protein C receptor (EPCR) functions as a primary receptor for protein C activation on endothelial cells in arteries, veins, and capillaries. *Biochem. Biophys. Res. Commun.* *259*, 671-677.
- Zajonc, D.M., Elsliger, M.A., Teyton, L., and Wilson, I.A. (2003). Crystal structure of CD1a in complex with a sulfatide self antigen at a resolution of 2.15 angstrom. *Nat. Immunol.* *4*, 808-815.
- Zajonc, D.M., Striegl, H., Dascher, C.C., and Wilson, I.A. (2008). The crystal structure of avian CD1 reveals a smaller, more primordial antigen-binding pocket compared to mammalian CD1. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 17925-17930.
- Zappacosta, F., Tabaczewski, P., Parker, K.C., Coligan, J.E., and Stroynowski, I. (2000). The murine liver-specific nonclassical MHC class I molecule Q10 binds a classical peptide repertoire. *J. Immunol.* *164*, 1906-1915.
- Zekarias, B., Ter Huurne, A.A., Landman, W.J., Rebel, J.M., Pol, J.M., et al. (2002). Immunological basis of differences in disease resistance in the chicken. *Vet. Res.* *33*, 109-125.
- Zeng, Z.H., Castano, A.R., Segelke, B.W., Stura, E.A., Peterson, P.A., et al. (1997). Crystal structure of mouse CD1: An MHC-like fold with a large hydrophobic binding groove. *Science* *277*, 339-345.
- Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., et al. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* *272*, 1606-1614.
- Ziegler, A. (1997). Biology of chromosome 6. *DNA Seq.* *8*, 189-201.
- Ziegler, A., Kentenich, H., and Uchanska-Ziegler, B. (2005). Female choice and the MHC. *Trends Immunol.* *26*, 496-502.
- 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.
- Ziegler, A., Santos, P.S.C., Kellermann, T., and Uchanska-Ziegler, B. (2010). Self/nonsel perception, reproduction, and the extended MHC. *Self/Nonsel 1*, 176-191.