PEPTIDES NATURALLY SELECTED BY MAJOR HISTOCOMPATIBILITY COMPLEX CLASS II MOLECULES ASSOCIATED WITH RHEUMATOID ARTHRITIS

by

James Bernard Hayden

A DISSERTATION

Presented to the Department of Cell Biology and Anatomy
and the Oregon Health Sciences University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

July 1994

Approved:

[redacted]

(Professor in Charge of Thesis)

[redacted]

(Chairman, Graduate Council)

Table of Contents

List of Tables v

List of Figures vi

Abstract viii

Introduction 1

General Description of Rheumatoid Arthritis 2
Epidemiology of Rheumatoid Arthritis 3
Genetics of Rheumatoid Arthritis 4
Major Histocompatibility Complex 6
Organization of the Class I Locus 7
Expression of HLA Molecules 11
Antigen Processing 12
Antigen Processing for Class I 12
Antigen Processing for Class II 15
Kinetics of Class II/Peptide Binding 19
Crystal Structure of MHC Class I 20
Peptide Binding to MHC 22
Peptide Binding to Class I 22
Crystal Structure of Class II 26
Allele Specific Motifs for Class I 31
Allele Specific Motifs for Class II 33
Antigen Presentation 34
Role of T cells in RA 35
Rationale 35

Materials and Methods 37

Cell Lines 37
Flow Cytometry 38
HLA Typing 38
Antibody Purification and Column Production 39
Immunoaffinity Purification 40
Acid Elution 41
HPLC Fractionation 41
Capillary Electrophoresis Separation 43
Peptide Sequence Determination 43

Results 45

Comparison of Peptides Derived from Different DR Alleles 45
Comparison of Peptides Derived from a Single DR Allele 47
Reproducibility of HPLC Separations 49
Complexity of Peptide Mixtures 51
Purity of Fractions by CE 54

Peptides Derived from DRB1*0404 58
Peptides Derived from DRB1*0402 59
Comparison of Nested Clusters 60

Discussion 62

Class I in Class II 62
Class II Allele Specific Motifs 66
Different Peptides Derived from Homologous Proteins 71
Implications for Rheumatoid Arthritis 73

Summary and Conclusions 75

References 77

Appendix 99

Properties of Class I and Class II MHC Molecules 99
Amino Acid Codes 100

List of Tables

Table 1 1987 Revised Criteria for Rheumatoid Arthritis (1) 2
Table 2 HLA-DR Molecules and Ethnic Relationship with RA 5
Table 3 Peptides Derived from HLA-B27 and HLA-B7 32
Table 4 A Nested Cluster Derived from DRB1*0101 33
Table 5 Peptides Eluted from DRB1*0404 58
Table 6 Peptides Eluted from DRB1*0402 59
Table 7 Comparison of the Class I Donor Protein Amino Acid Sequences 61

List of Figures

Figure 1 HLA Locus 7
Figure 2 HLA Class I Gene Map 7
Figure 3 HLA class II Gene Map 9
Figure 4 DR Haplotype Groups 10
Figure 5 Class I Antigen Processing Pathway 13
Figure 6 Class II Antigen Processing Pathway 18
Figure 7 Class I Crystal Structure and Binding Cleft 21
Figure 9 Model of Pockets in the Class I Binding Groove 23
Figure 8 Conserved Hydrogen Bonds of the A and F Pockets 25
Figure 10 Predicted Structure of the DRB1*0404 Antigen Binding Groove 27
Figure 11 Amino Acid Sequences of DRB1*0401, *0402, *0404 30
Figure 12 HPLC Chromatogram of Peptides Derived from Different DR Alleles 45
Figure 13 HPLC Chromatogram of Peptides Derived for DRB1*0401 47
Figure 14 HPLC Chromatogram of Peptides Derived from DRB1*0404 49
Figure 15 HPLC Chromatogram of Peptides Eluted from DRB1*0402 51
Figure 16 Chromatogram of a Secondary Separation of DRB1*0402 Peptides 54
Figure 17 CE Chromatograms of Fractions from DRB1*0402 55

Table of Abbreviations

AChR	Acetyl Choline Receptor
ACN	Acetonitrile
APC	Antigen Presenting Cell
ASO	Allele Specific Oligonucleotide hybridization
C	Constant region
C proteins	Compliment proteins
CE	Capillary Electrophoresis
CPL	Compartment for Peptide Loading
D	Diversity region
DNA	Deoxyribose Nucleic Acid
EBV	Epstein Barr Virus
EDTA	Ethylenediaminetetraacetic Acid
ER	Endoplasmic Reticulum
Fc	Disulfide linked Constant region of Immunoglobulin heavy chain
FITC	Fluorescene IsoThioCynate
HEPES	N-(2-Hydroxyethly) Piperazine-N-(2-ethenesulfonic) acid
HLA	Human Leukocyte Antigen (MHC in Humans)
HPLC	High Pressure Liquid Chromatography
Ig	Immunoglobulin
Ii	Invariant Chain
J	Joining region
MHC	Major Histocompatibility Complex
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulfonyl Flouride
RA	Rheumatoid Arthritis
SDS PAGE	Sodium Dodecyl Sulfate PolyAcrilamide Electrophoresis
TAP	Transporter Associated with Antigen Processing
TFA	TriFluoroAcetic acid
V	Variable region

Abstract

Rheumatoid arthritis is a chronic inflammatory disease of unknown etiology. It is associated with the major histocompatibility complex allele DRB1*0404 and is not associated with the closely related allele DRB1*0402. These molecules differ by only three residues which are strategically placed to interact with two of the proposed binding pockets in the antigen binding cleft. One possible model to explain the variation in disease susceptibility between these alleles is that the small sequence alteration causes different processed peptides to be presented to T cells. The DRB1*0404 allele may present a processed peptide that stimulates or maintains the inflammatory process while DRB1*0402 may select a different processed peptide. To test this hypothesis, naturally processed peptides from homozygous cell lines expressing DRB1*0404 or DRB1*0402 were analyzed. The MHC/peptide complexes were purified by affinity chromatography. The peptides were dissociated from the MHC molecule by acid elution, and then the peptide pool was separated from the MHC chains by filtration. The peptide pool was fractionated by two reversed-phase HPLC gradients, and the purity of each fraction was checked by capillary electrophoresis. Pure peptides were sequenced by Edman degradation, and the sequence was confirmed by tandem mass spectrometry sequencing. Ten unique peptides were sequenced from DRB1*0404, and eight unique peptides were sequenced from DRB1*0402. The peptides eluted from both alleles were dominated by a common core sequence with variations at both the N and C termini called a nested cluster. The donor proteins for the nested clusters in both alleles were HLA class I molecules. Class I molecules have hypervariable regions that are unique to the allele, and framework regions that are relatively conserved between alleles. The nested clusters presented by the DRB1*0404 and DRB1*0402 alleles were derived from different conserved regions even though the donor proteins contained sequences very similar to the other molecule's nested cluster.

The core of the peptides eluted from DRB1*0404 was very similar to a region of the protein that donated the nested cluster eluted from DRB1*0402. However, the peptide found in DRB1*0404 was never found in DRB1*0402. The sequence presented in DRB1*0402 was homologous to the class I molecule that donated a different region for the DRB1*0404 peptides, but the peptide eluted from DRB1*0402 was never found in DRB1*0404. Therefore, even when the donor proteins that have primary, secondary and tertiary structure homology, these two alleles, DRB1*0404 and DRB1*0402, select different peptides for presentation. This variation may explain the different disease susceptibility found with these two molecules.

Introduction

Rheumatoid arthritis is a chronic inflammatory disease of unknown etiology. It affects approximately one percent of the world's population with little variation for race or geography. The most common model for the etiology of rheumatoid arthritis (RA)[1] proposes that the disease develops in a genetically susceptible individual who is exposed to an antigenic challenge which precipitates the disease. Currently, the only gene associated with RA is a major histocompatibility complex (MHC) class II gene, HLA-DRB1. The particular MHC class II allele associated with RA varies across races as the frequency of the class II alleles change. All of the associated alleles have a small, but critical, common amino acid sequence at position 70-74 of the HLA-DRB1 gene product, QKRAA or QRRAA[2]. MHC class II alleles that are homologous except at this short region show no association with RA.

The normal function of the MHC class II molecule is to present processed antigens in the form of short peptides on antigen presenting cells (APC). The MHC class II molecule/peptide complex is then recognized by a $CD4^+$ T cell leading to a cellular immune response. If the RA associated sequence is important for determining which peptides are selected for presentation, then this could account for the differences in disease susceptibility. An APC with an RA associated allele might present an antigen that could stimulate particular T cells leading to the development of RA, but an APC with a non-RA associated allele would select a different portion of the antigen or perhaps a completely different antigen leading to a non-pathological T cell response. The hypothesis of this thesis was that an RA associated MHC class II molecule would select different processed antigens than a highly homologous, non-RA associated MHC class II molecule.

---

[1]See page vii for a complete list of abbreviations

[2]Single letter abbreviations for amino acid residues. See appendix for a list of amino acid codes

General Description of Rheumatoid Arthritis

In 1987 a revised criteria for the diagnosis of RA was released by the predecessor to the American College of Rheumatology (see table 1) (1). To be diagnosed with RA, a patient must demonstrate 4 of 7 possible criteria. The disease can occur at any time, but the peak incidence occurs in the fourth to sixth decade (2). Disease onset can be acute, but usually RA develops over weeks or months. A prominent feature of RA is its involvement of diarticular joints. The general disease course involves intermittent symptoms over years or decades that gradually become worse leading to progressive joint deformity (2). Joint involvement is usually symmetrical between left and right sides, with the hand, wrist, knee, and foot joints the most commonly affected. Joint destruction occurs by an inflammatory process marked by swelling, redness, and pain and leading to joint laxity. It is the laxity that allows the characteristic deformities such as the ulnar deviation and palmar subluxation of metacarpophalangeal joints to develop.

Table 1 1987 Revised Criteria for Rheumatoid Arthritis (1)	
1.	Morning stiffness in and around joints lasting at least 1 hour before maximal improvement.
2.	Soft tissue swelling of 3 or more joint areas observed by a physician.
3.	Swelling of the proximal interphalangeal, metacarpophalangeal, or wrist joints.
4.	Symmetric swelling on the right and left sides.
5.	Rheumatoid nodules.
6.	Presence of rheumatoid factor.
7.	Radiographic erosions and /or peri articular osteopenia in hand and/or wrist joints
Criteria 1 to 4 must be present for at least 6 weeks. RA is defined by the presence of any 4 of the 7 criteria.	

Histologically, an involved joint is characterized by a thickening of the synovial lining from the normal 1-3 cells to multiple cell layers, an infiltration of mononuclear cells and the development of a pannus or granulation tissue. The synovialcyte

proliferation involves both the type A (macrophage like) and type B (fibrocyte like) synovialcytes (2) and leads to the synovium becoming edematous and developing villous projections into the joint cavity. The infiltrating mononuclear cells include macrophages, B cells, dendritic cells, and T cells; these cells may collect into follicles although rarely into true germinal centers. The pannus is composed of fibroblasts and inflammatory cells that are supported by rapid angiogenesis. The development of the pannus is the most destructive change in the rheumatoid joint as it gradually erodes the periarticular bone and the articular cartilage.

Even though RA is usually associated with joint inflammation, it is actually a systemic disease. Extraarticular features can affect nearly any organ system. Rheumatoid nodules are one of the most common features occurring in 20-35 percent of RA patients (3). A nodule is composed of a central necrotic area surrounded by palisading fibroblasts contained within a collagenous capsule. A nodule also contains inflammatory cells often collected perivascularly (3). Another possible complication of RA is vasculitis, usually seen only in the most progressive disease (3). The lesion is usually a pan arteritis, with fibrinoid necrosis and infiltrating mononuclear cells (3). RA can also have a number of effects in the lung and the heart (3).

Epidemiology of Rheumatoid Arthritis

RA is common throughout the world with an average prevalence of approximately 1 percent (2, 4), and the prevalence increases with age reaching 2-5 percent by age 65 (2). The prevalence of RA is considered to be relatively uniform, but some racial groups seem to differ, such as an increased rate in some American Indians (2, 4) and a lower than usual rate in rural South African Blacks (4). There may also be some differences based on socioeconomic groups with a higher prevalence and greater

disease mortality in those with less than 5 years of education and in the lower economic strata (3). In the United States, RA is estimated to afflict 5 million people with 150,000 new cases each year (3). With an estimated per patient average lifetime cost which includes treatment and lost work time of nearly $29,000, RA rivals stroke and coronary artery disease for health care expenditures (3).

RA affects women 2-3 times more often than men (2, 4). However, when considering cases of early probable RA, the ratio of men to women may approach 1.0 (3). Interestingly, an English study reported that the use of oral contraceptives modified the course of the disease by decreasing the percentage of women that develop severe disease when compared to women who have never used oral contraceptives (5). Parity status is also an independent risk factor with pregnancy being associated with remission (5) and nulliparity having an odds ratio for disease development of about 2.0 (5). However, a study in Rochester, MN, did not show an association between RA and hormonal therapy (3). If there is an association in women or a modification of the disease by reproductive hormones and parity status, the reason remains unknown, but it has been suggested that long-term hormonally mediated immunosuppression may be responsible (5).

Genetics of Rheumatoid Arthritis

Despite the uniformity across nearly all races, RA does have a genetic component other than sex. A hospital based twin study report a concordance rate of 30 percent in identical twins and 10 percent in dizygotic twins (4, 6) while more clinically based study showed a concordance rate of 12.3 percent for monozygotic and 11.1 percent dizygotic twins (5). The variations in the concordance rate is thought to be caused by the hospital based study relying on patients with more severe disease and

the clinical study was using patients with milder RA. RA has been shown cluster in some families (4, 6). While the average prevalence is about 1 percent, the prevalence in first degree relatives is 2-3 percent (7). The concordance between same sex siblings shows a stronger relationship with males having a rate of 3.0 percent and females a rate of 9.3 percent (6). This rate increases to an average of 20.5 percent when considering same sex siblings that are serologically matched for MHC (6).

The only specific gene found to be associated with the disease is an MHC gene, human leukocyte antigen (HLA) DR. Early studies showed an association with HLA-DR4 which was defined serologically. HLA-DR4 was found to be public specificity that included may subtypes recognized by T cells in mixed lymphocyte cultures. The HLA-DR4 family currently contains 12 members, and they are classified by genotype (8). In the DR4 family, the alleles defined by DRB1*0401 and DRB1*0404 are the most strongly associated with RA in Caucasians (9-11) This association can be further refined to a precise region of the DR molecule, amino acids 70-74 of the β chain, that has the sequence QKRAA or QRRAA (9, 11, 12). The allele frequency of DRB1*0401 and DRB1*0404 varies between races. Races that have a low frequency of these alleles often have another HLA-DR allele associated with RA as demonstrated in table 2. However, all the associated alleles contain a homologous sequence from 70-74 of the β chain which has been called the shared epitope (9)

Table 2 HLA-DR Molecules and Ethnic Relationship with RA

Race	DR allele	70-74	Source
Caucasian	*0401 *0404	QKRAA QRRAA	(9-11)
Indian	*0101	QRRAA	(13)
Yakima Indian	*1402	QRRAA	(14)
Japanese	*0405	QRRAA	(15)

RA is a disease with many possible manifestations and degrees of severity which can be correlated with the different HLA-DR allele a patient expresses. Patients that express an allele in the DRB1*01 serological family, DRB1*0101 or DRB1*0102, tend to have a milder form of the disease than those expressing one from the DRB1*04 family, DRB1*0401, DRB1*0404 (16-18). A DR4 expressing patient is more likely to have earlier disease onset, be rheumatoid factor positive, have more joints involved, and more aggressive joint erosion than the DR1 patient (16). As the maternal and paternal HLA-DR genes are codominantly expressed, the association of RA and HLA-DR is also subject to a gene dosage effect (19, 20). More severe forms of RA develop extraarticular involvement including rheumatoid nodules and rheumatoid organ disease. In one study, nodular disease was present in 100 percent of patients expressing two DRB1*04 alleles and only 59 percent of patients expressing only one DRB1*04 allele. Major organ involvement was present in 61 percent and 11 percent, respectively, of these two groups of patients (20). Clearly, the HLA-DR gene plays an important role in the development and progression of RA.

Major Histocompatibility Complex

To understand the possible roles of HLA-DR in RA, it is necessary to understand the normal functions of the molecule. The HLA molecules are divided into class I and class II based on their different structure and function. However, there remains enough similarity between these two classes that information learned about one can often be applied to the other. Since more is known about the class I system, both systems will be described with the class I system being used as an analogy to predict currently unknown characteristics of the class II system.

Organization of the Class I Locus

The HLA locus is located on the short arm of chromosome 6 and is divided into three regions designated class I, class II and class III as shown in fig 1. The class III region genes have no known function in antigen presentation so will not be discussed further. The class I region consist of loci for HLA-A, B, C, E, F, G, and J along with 11 pseudogenes (21). All of these genes are located within 1600 kilobases forming a tightly linked group as shown in fig 2(21). The class I genes are divided into the classical, A, B, C; and the non classical, E, F, G. The function and tissue distribution of the non classical gene products is not completely understood, so this discussion will be limited to the classical HLA genes. The protein products of these genes, called a heavy chain, pairs with another protein, β2 microglobulin, to make a functional dimer. The

Figure 1 HLA Locus

The HLA locus on the short arm of chromosome 6 has distinct regions for Class I, Class II and Class III genes. The entire region only spans 4000 kilobases making this important region very closely linked (21).

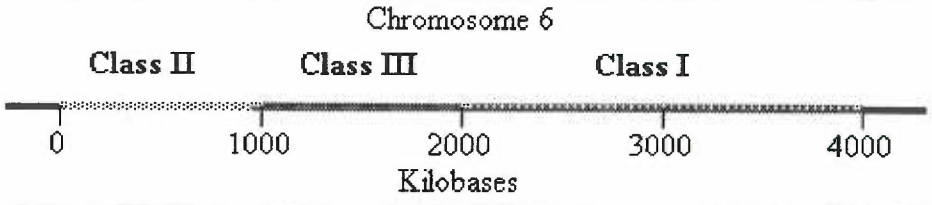
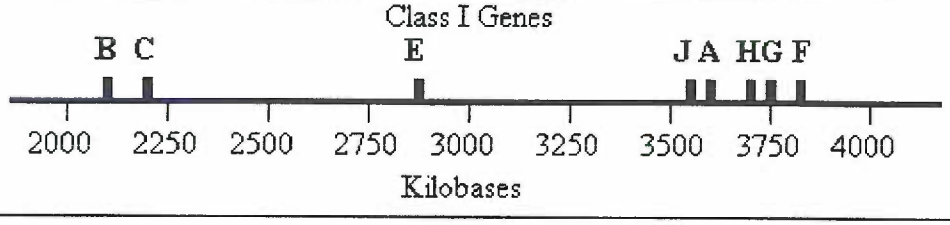


Figure 2 HLA Class I Gene Map

An enlargement of the Class I region of the HLA locus showing the arrangement of the specific genes (21).



classical class I molecules are highly polymorphic with 41 alleles in the A locus, 61 in the B and 18 in the C (22). β2 microglobulin is not encoded in the MHC locus. It is non-polymorphic and dimerizes with all heavy chains. Identification of each molecule is based on the World Health Organization nomenclature. The actual assignment for a particular HLA molecule is given at periodical international workshops. All identifications are according to a standard format which identifies the gene first followed by the number of the serological classification, and then the number of the allele based on amino acid sequence within the serological family (22). As an example HLA-B*2705 defines the fifth allele of the 27th serological family from the B gene. A completely heterozygous individual would express 2 unique A, B and C genes for a total of 6 class I molecules.

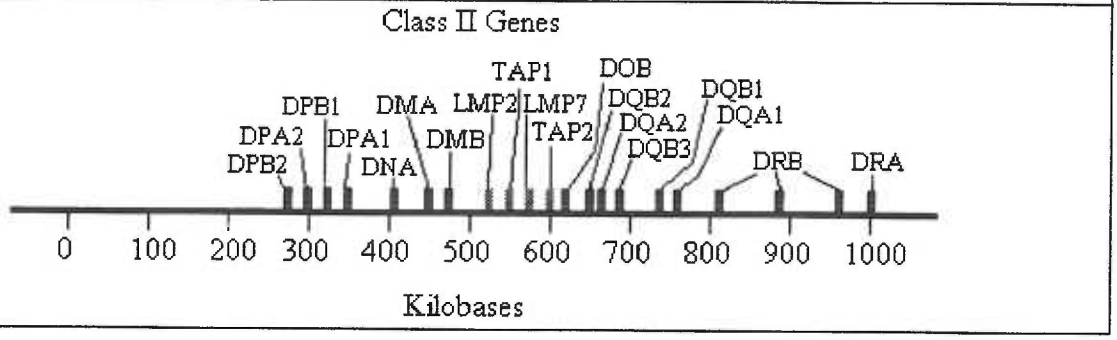
Organization of the Class II Locus

The organization of the class II locus is more complicated than the class I. The class II locus stretches over only 900 kilobases, making its genes very tightly linked as demonstrated in fig 3. All of the genes for the class II molecules are identified as D followed by a letter designating the subregion. Like the class I molecule, the class II protein is a heterodimer. However, the dimer is formed by equally contributing α and β chains that are both unique to a subregion. The three main subregions are defined as DR, DQ, and DP, and the minor subregions are DO, DM, and DN. Of the minor subregions, only the DM genes have been shown to be transcribed into RNA (21), and they will be discussed later with antigen processing. The DP region has one functional α and β gene and one α and β pseudogene (21). The DQ region contains one functional α and β gene and includes one α and two β pseudogenes (21). All DR regions contain one functional α gene that is not polymorphic. The product of the α

gene, the α chain, can pair with all expressed β chains to make a functional DR molecule.

Figure 3 HLA class II Gene Map

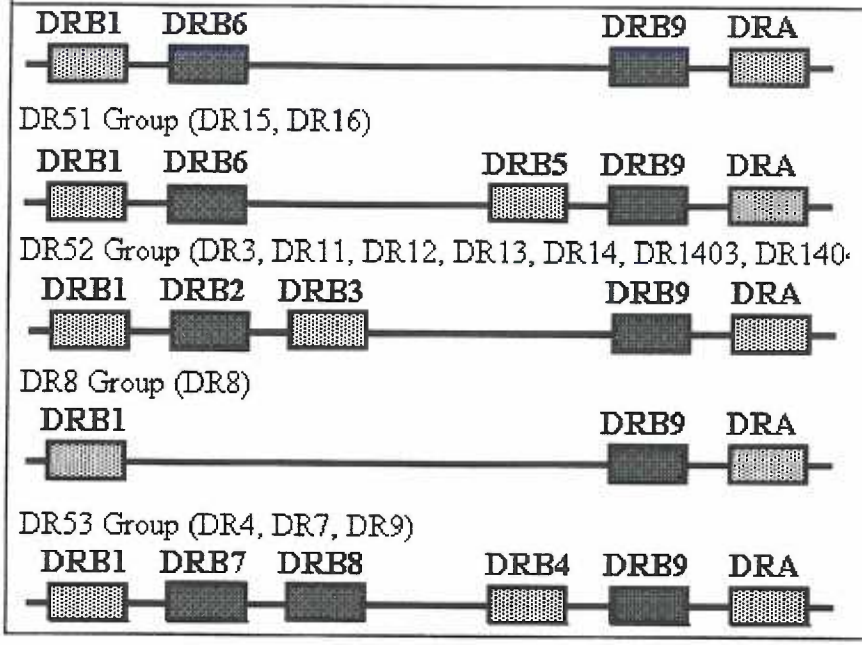
The HLA class II genes are very tightly linked with the entire locus contained within 1000 kilobases. Each major subregion, DP, DQ, and DR contains both α and β genes. The class II region also contains the LMP and TAP genes discussed with antigen processing (21).



Since the region for the DR genes is very small, all the genes tend to be very tightly linked. Therefore, all the β genes segregate as a consistent unit. Different haplotypes are defined as a group depending on the complement of functional β genes and pseudogenes present. Several groups have been defined containing one or two functional β genes and several β pseudogenes (8). The haplotype groups are defined in fig 4. For instance, the DR1 group contains DR1, DR10, and DR15. All of these haplotypes express only the β1 gene and carry the β6 and β9 pseudogenes.
In contrast, the DR53 group, which includes DR4, DR7 and DR9, expresses mainly the β1 gene with a small amount from the β4 gene and carry the β7, β8 and β9 pseudogenes (8).

Figure 4 DR Haplotype Groups

Expression of DR β genes can be divided into groups that are consistent within a haplotype. One the serological family of the DRB1 gene product is known, the haplotype can be assigned. The haplotype will determine the number of different DR molecules that are expressed by the individual (21).



Like the class I genes, the class II genes are highly polymorphic. As an example the DRB1 gene has 60 possible alleles (8). The nomenclature to define the class II alleles is similar to the class I system with the gene first, followed by the serological family, then the allele variant based on amino acid sequence within that family (22). For instance DPB1*0402 defines the second variation in the fourth serological family from the β1 gene of the DP subregion. In the DP and DQ subregions, both the α and β genes are polymorphic, so both alleles must be defined, such as DPA1*0401/DPB1*2201, to completely identify a class II molecule. Like the class I genes, the maternal and paternal class II genes are codominantly expressed, allowing

the maternal α to dimerize with the paternal β and vice versa. This allows a heterozygous individual to express four unique DP and DQ molecules. In the DR subregion, the α is considered nonpolymorphic, so only the polymorphic β must be identified. The nomenclature follows the same guidelines of subregion, gene, serological family, and amino acid sequence variant. DRB1*0404 would define the fourth variant in the fourth serological family from the β1 gene in the DR subregion, and DRB4*0101 would identify the first variant in the first serological family from the β4 gene. It is understood that both these β gene products would pair with the nonpolymorphic α molecule. Because of the variations in the number of functional genes in a haplotype, a completely heterozygous individual may express between 2 and 4 unique DR molecules. Over the entire class II region, a completely heterozygous individual would express 10 to 12 unique class II molecules.

Expression of HLA Molecules

The class I molecules have a larger tissue distribution in humans than the class II molecules. Class I is present on all nucleated cells. Constitutive class II expression is usually limited to cells of the immune system, monocyte/macrophages, B cells, thymic epithelium, and dendritic cell lineage. In their normal distribution, both class I and class II expression levels can be increase under the influence of cytokines particularly interferon-γ (23). Class II can also be induced on other cell types such as endothelial cells, fibroblasts, synovialcytes by cytokines and on T cells upon activation. This is thought to be an important point for understanding the etiology of RA. Most of the non-T infiltrating mononuclear cells isolated from the synovial tissue which include specifically, macrophages, B cells, and dendritic cells demonstrate high levels of DR expression (24, 25). The T cells isolated from inflamed synovial tissue or synovial fluid

have been induced to express DR (25-27). Importantly, the other cell types in the inflamed tissue, fibroblasts (26), synovialcytes (25, 26, 28) and endothelial cells (26) also express DR. DR expression in RA tissue has been measured both pre and post intraarticular corticosteroid injection and found to correlate well with the inflammatory score (29). Therefore, the DR molecule is not only associated with the disease, but its expression has been induced or enhanced on many of the cell types in the affected tissue, and expression correlates with disease activity. It is possible that the enhanced DR expression found within joints from patients with RA plays a role in initiating or maintaining the chronic inflammatory state.

Antigen Processing

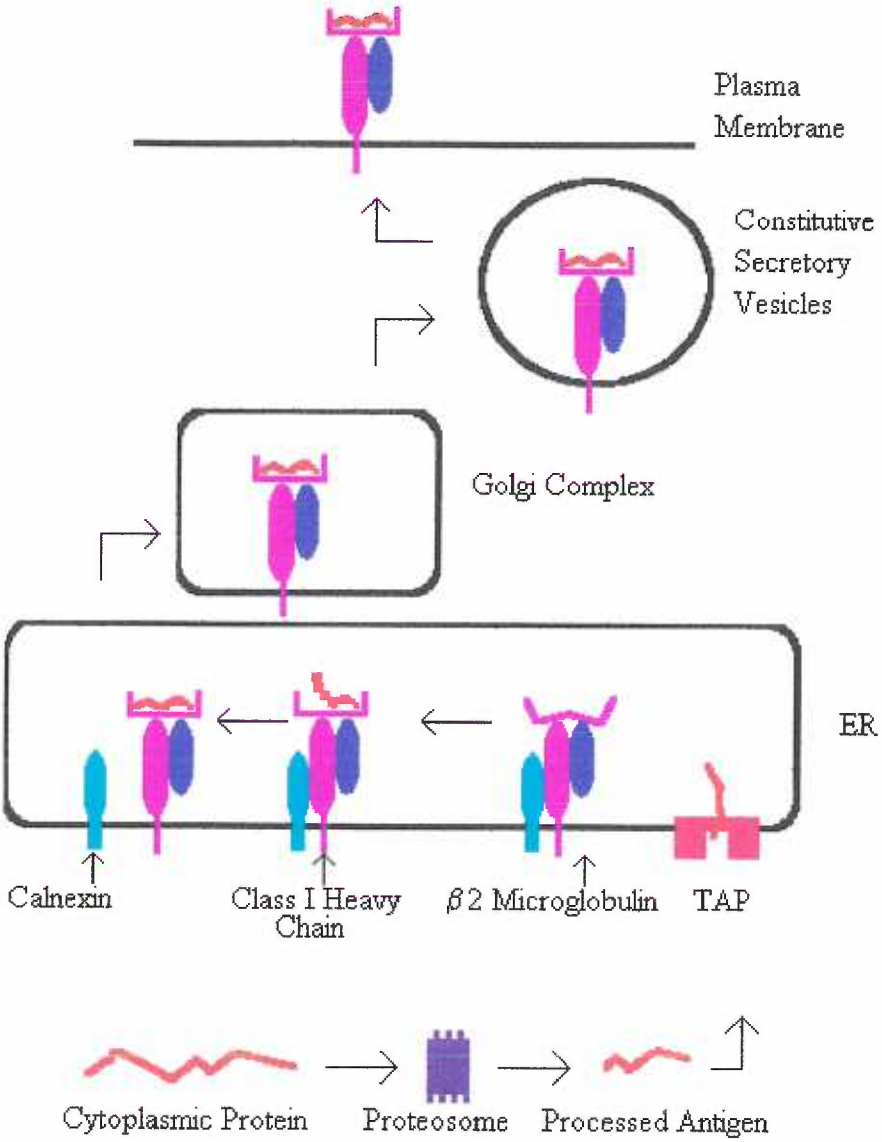
T cells do not recognize intact, soluble antigen. Rather, antigens must first be digested into short peptide fragments and placed in the antigen binding cleft of a class I or class II molecule. This process of digestion and binding to MHC is called antigen processing. The MHC/peptide complex is displayed on the surface of the cell for interaction with the T cell TCR, a process called antigen presentation. Class I molecules present processed antigen to $CD8^+$ T cells, and class II molecules present peptides to $CD4^+$ T cells. The method of antigen processing is one of the biggest differences between the two types of MHC molecules.

Antigen Processing for Class I

The class I synthesis and antigen processing pathway is diagrammed in fig 5. The class I heavy chain is synthesized on rough endoplasmic reticulum (ER). Once in

Figure 5 Class I Antigen Processing Pathway

The class I heavy chain and β2 microglobulin are synthesized on the ER. The heavy chain associates with calnexin in the ER. Peptides are degraded by the proteosome and transported in the ER by TAP. The peptide, β2 microglobulin and the heavy chain form a complex that is released from calnexin. The class I/peptide complex proceeds through the golgi and to the cell surface by the default pathway (30).



the ER, it quickly binds to an ER resident protein, calnexin or p88. This protein holds the heavy chain until it forms a complex with β2 microglobulin and a peptide. Both β2 microglobulin and peptide have been shown to associate weakly with the heavy chain in the absence of the other (30), so the order of association does not appear to be important. If either the β2 microglobulin or the peptide is missing, the complex fails to achieve a stable form, and it is degraded in the ER. Once a stable trimolecular complex is produced, it is released from calnexin, and the class I complex moves to the golgi and the cell surface by the default secretory pathway (31). However, in mutant cell lines that lack either β2 microglobulin or a functional peptide processing pathway, low levels of class I molecules can still be found on the cell surface, but they are unstable and are rapidly removed (31).

The protein donors normally utilized by the class I processing system are intracellularly derived. Cellular proteins are degraded by proteosomes, which are multicatalytic assemblies (30). Two proteins that may be part of proteosomes that generate peptides for presentation are LMP-2 and LMP-7 (32). The genes for these two proteins are located in the class II region of the MHC locus; (33, 34). However, peptides for class I presentation can be generated in mutant cell lines lacking functional LMP-2 and LMP-7 proteins, so their complete role is not understood. Once peptide fragments are produced in the cytoplasm, they are transported into the ER by TAP (transporter associated with antigen processing). TAP is an ATP dependent transporter that demonstrates both size and sequence specificity (30). The gene for TAP is also located in the class II region of the MHC locus (35), and it has some limited polymorphism (30). Since TAP is critical for supplying peptides to the class I molecule and it shows peptide sequence specificity, its polymorphism may influence the repertoire of peptides available for class I presentation (30). In the absence of a functional system to produce and transport peptides into the ER, signal sequences and

other cleaved proteins from within the ER have been shown to bind to the class I molecule and be presented (36).

Antigen Processing for Class II

The class II α and β chains are also synthesized on the rough endoplasmic reticulum. The α and β chains complex with invariant chain (Ii) immediately in the ER. Invariant chain is thought to have three roles: (1) it aids the folding of the α/β chain dimer which is necessary for the molecule to leave the ER (37); (2) it blocks proteins and peptides from binding in the antigen binding cleft (38); and (3) it provides the targeting signals necessary for the endocytic pathway (39). The complex of α/β/Ii then moves through the golgi to the early endosome and on to a recently discovered compartment, compartment for peptide loading (CPL) (40). The CPL has electrophoretic characteristics that are between early and late endosomes (37). However, two-dimensional electrophoresis has shown this compartment to have several proteins not found in either early or late endosomes (37) suggesting that this it is truly unique and not merely early endosomes maturing to late endosomes.

One model for class II processing suggests that the a/b/Ii complex moves through the early endosomes where MHC begins to be released from the Ii (40). Ii is known to be cleaved away from the class II molecule in a process that is acid dependent and can be inhibited by leupeptin (41). The majority of MHC/Ii complex would proceed on to the CPL for complete release of the MHC. MHC class II molecules that lack peptides are thought to have a "floppy" conformation and are known to aggregate. An MHC molecule could be released from an aggregate and assume a more stable conformation once it had bound a peptide (40). MHC/peptide

complex would then travel to the cell surface, and any MHC aggregates still remaining would be shuttled to the lysosome for degradation (40).

Any model for class II processing pathway needs to explain how antigen is encountered and processed. According to the model above, any protein that was phagocytosed or endocytosed might be available for processing. Phagocytosis would be expected to bring in larger particles that may contain multiple copies of an antigen at one time. This could deliver enough identical peptide fragments to the MHC molecules for presentation to lead eventually to T cell stimulation (31). Without this benefit, cells would need some type of mechanism to collect antigens that may be at very low concentrations. B cells accomplish this with surface Ig (42). B cells have been shown to be 100-10,000 times better APC when their surface Ig is specific for the antigen (31). Other APC also express receptors that aid them in collecting antigen. For example, macrophages express three different types of receptors for the Fc portion of IgG which enable them to harness the collecting power of immunoglobulins particularly when multiple antibodies have bound to the same antigen or complex (43). Receptors for the cleavage products of complement (C) proteins, particularly C3, aid the macrophage in collecting particles that have stimulated the complement cascade (43). A variety of receptors for polysaccharides, such as the mannose receptor, allows the macrophage to collect glycoproteins that express terminal mannose groups that are often found on microorganisms (43). Finally, macrophages secrete $\alpha 2$ macroglobulin that when proteolytically activated has been shown to covalently bind a diverse array of proteins. Once coupled, the complex binds to high affinity receptors on the macrophage for rapid endocytosis (44). Other cell types, such as Langerhans cells of the dendritic cell lineage, express many of these surface receptors while they reside in the epidermis where they are thought to collect antigen. However, when they migrate to the lymph nodes and function as presenting cells, they lose these receptors (43). In

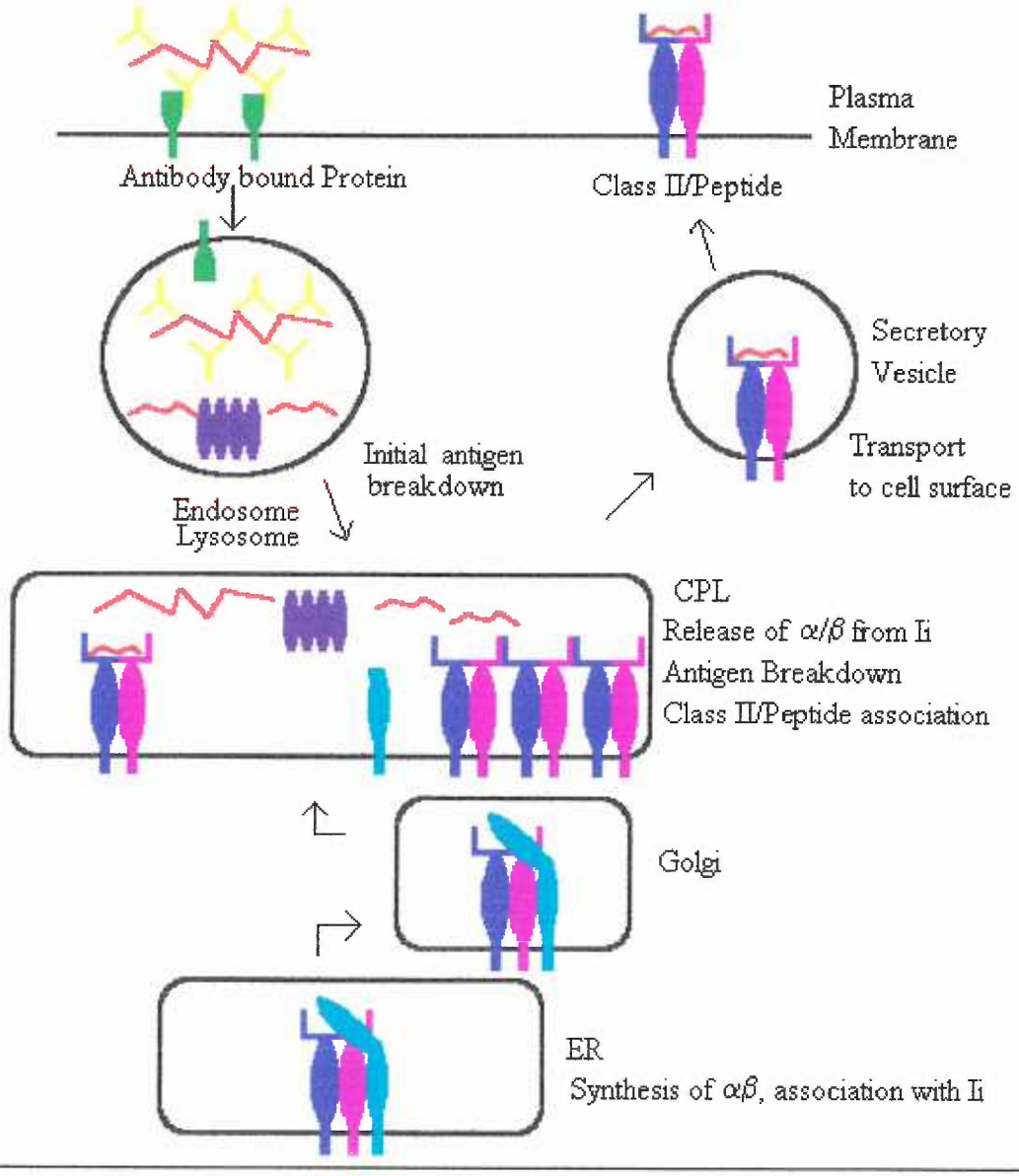
summary, APC may use a variety of receptors to help collect and direct potential antigens to the endosome and possibly the CPL for class II processing.

Despite an understanding of the pathway taken by class II from the ER through the CPL to the cell surface, very little is understood about the actual machinery for class II processing. Recently, the DM molecule, which is derived from the class II locus, was found to be critical in the processing pathway (45, 46). A normal amount of class II reaches the cell surface when DM is missing, but 60-70 percent of the molecules are associated with peptides derived from invariant chain (45), and the class II has an altered conformation that can be detected by antibodies for some DR alleles (45, 46). The complete function of this molecule is not known, but it is clearly important in process.

Unlike the proteosome complex for class I processing, specific proteolytic complexes or enzymes have yet to be identified. There are two models used to explain how the antigen is actually processed. The first suggests that an antigen is directed to the endocytic pathway then degraded under the influence of acid and proteases. MHC class II is released from Ii and the degraded peptides bind in the unoccupied binding cleft in a competitive fashion. The second model suggests that the antigen is brought into the endocytic pathway then denatured. MHC class II is released from Ii and captures the long, denatured protein. Proteases then attack the protein, leaving the peptide portion that is protected by the class II binding cleft. Thiol and aspartyl proteases are thought to be involved in the degradation based on inhibitor studies, but it is not known if their role is degradation of antigens or release of Ii from the class II molecules (43). It has been shown for some antigens that the type of "receptor" and the binding region that captured the antigen can influence which peptides are actually presented by the class II molecule (47). The selection of peptides may also be influenced by the three dimensional conformation of the parent protein (48, 49).

Figure 6 Class II Antigen Processing Pathway

Class II α and β chains are synthesized and transported into the ER. They rapidly complex with Ii. The complex of Ii/α/β is transported through the golgi to early endosomes then to the CPL. Ii may be released in the early endosomes or the CPL, and the class II lacking peptides aggregates. Antigen or self proteins are brought to the CPL for processing. The peptides generated bind to the class II molecules freeing them from the aggregate. The peptide/class II complex then moves to the cell surface (40).



Kinetics of Class II/Peptide Binding

Once expressed on the cell surface, the stability of class II α/β heterodimers is heavily dependent on the presence of a peptide in the binding groove. Accordingly, peptide occupancy is thought to be nearly 100% for all the class II molecules on the cell surface. Binding of peptides to the MHC class II molecule is rather unusual with slow association and slow dissociation rates. Studies using purified MHC class II molecules have reported binding affinities of 10^{-6} to 10^{-7} M with association rates of 10^{0}-10^{2} $M^{-1}sec^{-1}$ (43). Once bound, the complex becomes very stable and dissociation rates are as slow as the association. Class II molecules without a peptide are in a an altered conformation described as floppy. When the peptide binds, the conformation changes to a more stable, compact form. This process can be followed by SDS PAGE where the floppy form dissociates into α and β chains and the compact form retains the integrity of the heterodimer. However, MHC class II molecules come to the surface of the cell with a peptide bound, and the *in vitro* analysis of peptide binding kinetics has utilized class II molecules purified from cells. Peptide exchange experiments have shown that peptide binding to purified class II molecules occurs at the same rate as peptide dissociation (50) which suggests that binding of an experimental peptide is limited by dissociation of endogenously bound peptides. Class II molecules produced in insect cells have an empty binding cleft (51). These molecules show association kinetics approximately 10 times faster than the same class II molecules purified from human cells. However, the rate is still very slow at 12 M^{-1} sec^{-1}. The dissociation rate for class II molecules produced in insect cells remained nearly the same as those purified with an endogenous peptide bound. The unusual kinetics, slow on and slow off, suggests the binding may require some secondary changes in the peptide, the class II molecule, or both (43). In contrast to the *in vitro* binding data requiring hours or

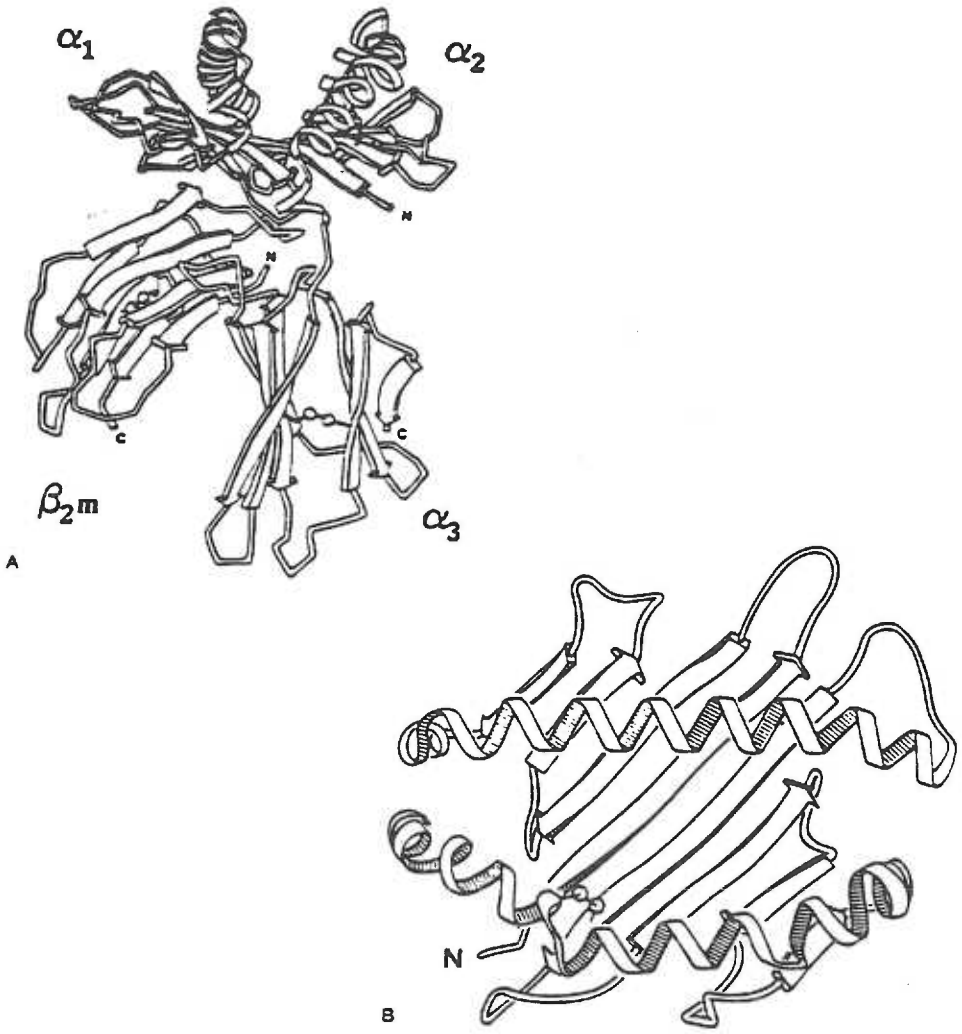
days to bind peptides, an APC can process an antigen, bind an appropriate peptide, and present it to T cells all within 30 minutes of exposure to that antigen.

Crystal Structure of MHC Class I

Once the peptide has been generated and bound to the MHC molecule, the complex assumes a stable conformation. The crystal structures for several human and murine class I molecules have been solved (52-56). All of the different alleles have conformations that are very similar. The heavy chain is divided into three external domains, and β2 microglobulin makes up a fourth domain. The α3 domain contains the CD8 binding region. The α1 and α2 domains contribute equally to the antigen binding cleft. The cleft has a floor of antiparallel β pleated sheets and the sides are bound by two α helices one from α1 and one from the α2 domain. This structure results in a groove that in each of the crystal structures has been filled with an ambiguous electron density (53). The electron density has been interpreted as a diverse mixture of processed peptides arranged linearly within the binding cleft (52, 55). One class I molecule has been crystallized after it was loaded with a single peptide, confirming the location and orientation of the peptide in the binding groove (56).

Figure 7 Class I Crystal Structure and Binding Cleft

A. Class I ribbon structure from the side demonstrating the α1 and α2 domains forming the antigen binding cleft. The α3 domain pairs with β2 microglobulin to support the binding cleft on the cell surface. B. Class I ribbon structure from the top showing the antigen binding cleft. The floor is made of β pleated sheets contributed by both the α1 and α2 domains. The cleft is bound by helices, one from the α1 domain and one from the α2 domain.



Peptide Binding to MHC

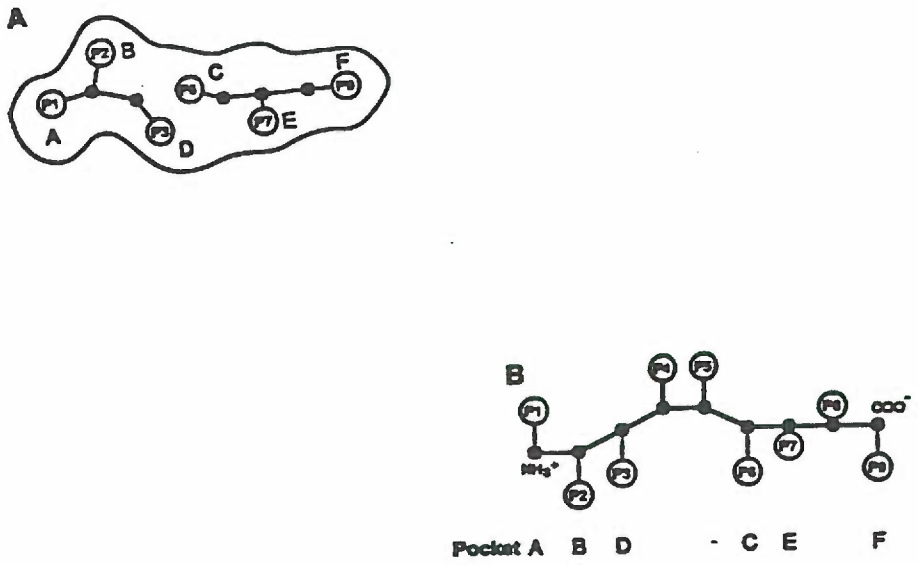
An individual who is heterozygous at the HLA class I locus will have only six unique, functional molecules that must be able to present the entire world of foreign antigens derived from intracellular infections such as viruses or intracellular parasites. A person heterzygous at the class II locus will have up to 12 unique molecules to present the entire world of extracellular antigens. This requires that the MHC binding cleft be able to bind and hold a virtually limitless number of unique peptides. To accomplish this task, the MHC system evolved to allow 2 different types of interactions between peptides and MHC, broad and specific peptide binding. Broad peptide binding refers to how the MHC molecule uses characteristics present in all peptides to provide a significant proportion of the binding energy. The backbone of the peptide itself is involved in the formation hydrogen bonds. Many of these interactions are governed by the non-polymorphic residues of the MHC molecule that are found in and around the antigen binding cleft. Specific peptide binding refers to how the individual side chains of the peptide's amino acids interact with the regions of the MHC binding cleft formed by the hypervariable residues.

Peptide Binding to Class I

The binding cleft can be divided into six pockets, labeled A through F, that can potentially interact with a peptide derived from a processed antigen. These pockets contribute to both the broad and specific peptide binding, Most of the polymorphism found in these molecules maps to residues that effect the shape and chemical nature of the different pockets (55). The A and F pockets have some conserved characteristics that are critical for the broad peptide binding as they aid in binding the N and C termini, respectively. Near the A pocket, conserved tyrosines at 59 and 171 are positioned to hydrogen bond to the amino terminal. Other conserved tyrosines, 7 and 159, may

Figure 8 Model of Pockets in the Class I Binding Groove

Model of the binding pockets in the class I antigen binding cleft. A. A map of the pockets on a schematic of the class I binding cleft. The pockets are labeled A through F and include the predicted residue of the peptide, P1-P9, that fits in the pocket. B. Model of a peptide in the cleft. Residues on the upper side are exposed to the TCR, and residues on the lower side interact with the class I molecule. The length of the line from the peptide backbone suggests a stronger interaction with TCR or MHC i.e. P2 is longer that P7 and would play a more important role in the binding to the MHC.



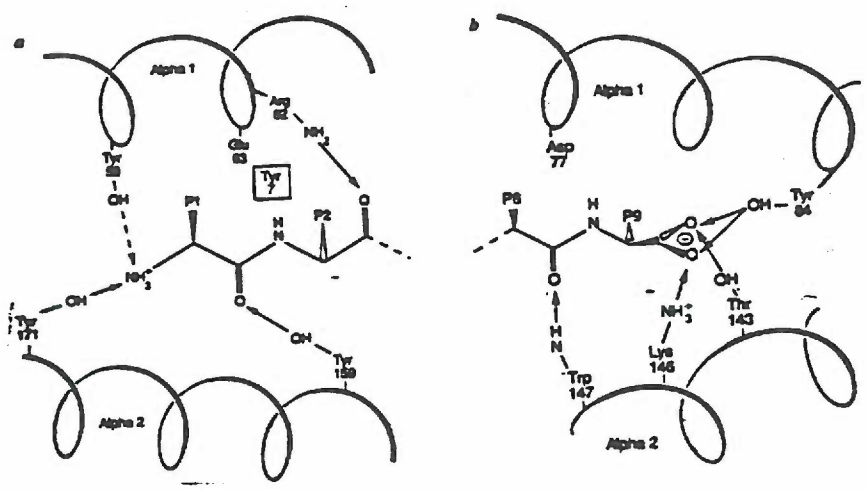
hydrogen bond to carbonyl oxygen of the first peptide (P1). The glutamate at position 63 is also conserved and may interact with the P2 amide nitrogen (57). High resolution

analysis of HLA-B*2705 showed that 5 of 8 possible non-side chain hydrogen binding sites on P1 and P2 would be utilized (52) Conserved residues near the F pocket interact with the C termini of the peptide. The epsilon amino group of lysine 146 can interact with the carboxylate group, and the hydroxyl groups of tyrosine 84 and threonine or serine at 143 are position to donate hydrogen bonds. The asparagine at 77 may interact with the amide nitrogen of P9 (57). Of the 8 potential sites for hydrogen bonding on the peptide backbone of P8 and P9, HLA-B*2705 successfully completes 6 (52). Due to the requirement of the amino- and carboxy- termini interacting with the A and F pockets, the optimum length of the peptide is limited to the distance between these pockets which corresponds to a peptide approximately 9 residues in length. Longer peptides have been found in class I molecules. These probably loop out of the groove to conserve the interactions at the termini. However, the predominant length of peptides found in class I binding groove is 9 residues.

The remaining pockets tend to vary more with the different polymorphic alleles as a comparison between HLA-A2 and HLA-Aw68 demonstrates (54). There are 13 amino acids that differ between A2 and Aw68, and 10 of these are in the solvent exposed areas of the antigen binding cleft. Five of the residues are located around the C pocket. In A2, residues 9F, 70H, 74H, 95V, and 97R tend to fill the C pocket and give it a strong positive charge. However, in Aw68 these residues are 9Y, 70Q, 74D, 95I, and 97M which leave the pocket open and give it net negative charge with hydrogen bonding and hydrophobic potential. In the A2 molecule the filled C pocket would influence the final conformation of the class I/peptide complex. However, in the Aw68 molecule, the deeper C pocket might precisely fit the side chain of an amino acid and form a stronger interaction and, thereby, act to anchor the peptide within the binding cleft.

Figure 9 Conserved Hydrogen Bonds of the A and F Pockets

A. From the α1 helix, the Y59 hydrogen bonds with the amino terminus, and the R62 hydrogen bonds with the carbonyl oxygen of residue 2. E63 from the α1 helix or Y7 from the β pleated sheet form hydrogen bonds with the amide of P2. Conserved residues Y159 and Y171 from the α2 domain hydrogen bond to the carbonyl oxygen of P1 and the amino termini, respectively. B. Y84 from the α1 helix along with T143 and K146 form hydrogen bonds and a potential salt bridge to stabilize the carboxylate group of P9. W147 from the helix on the α2 domain hydrogen bonds to the carbonyl group of P8.



Pockets that anchor peptides in the binding groove are best demonstrated in an analysis done with the B*27 family (58). Across the seven members of the B27 family analyzed, pockets A, C, D, E, and F all had variations. Only pocket B was completely conserved which is unusual since the B pocket is highly polymorphic in class I

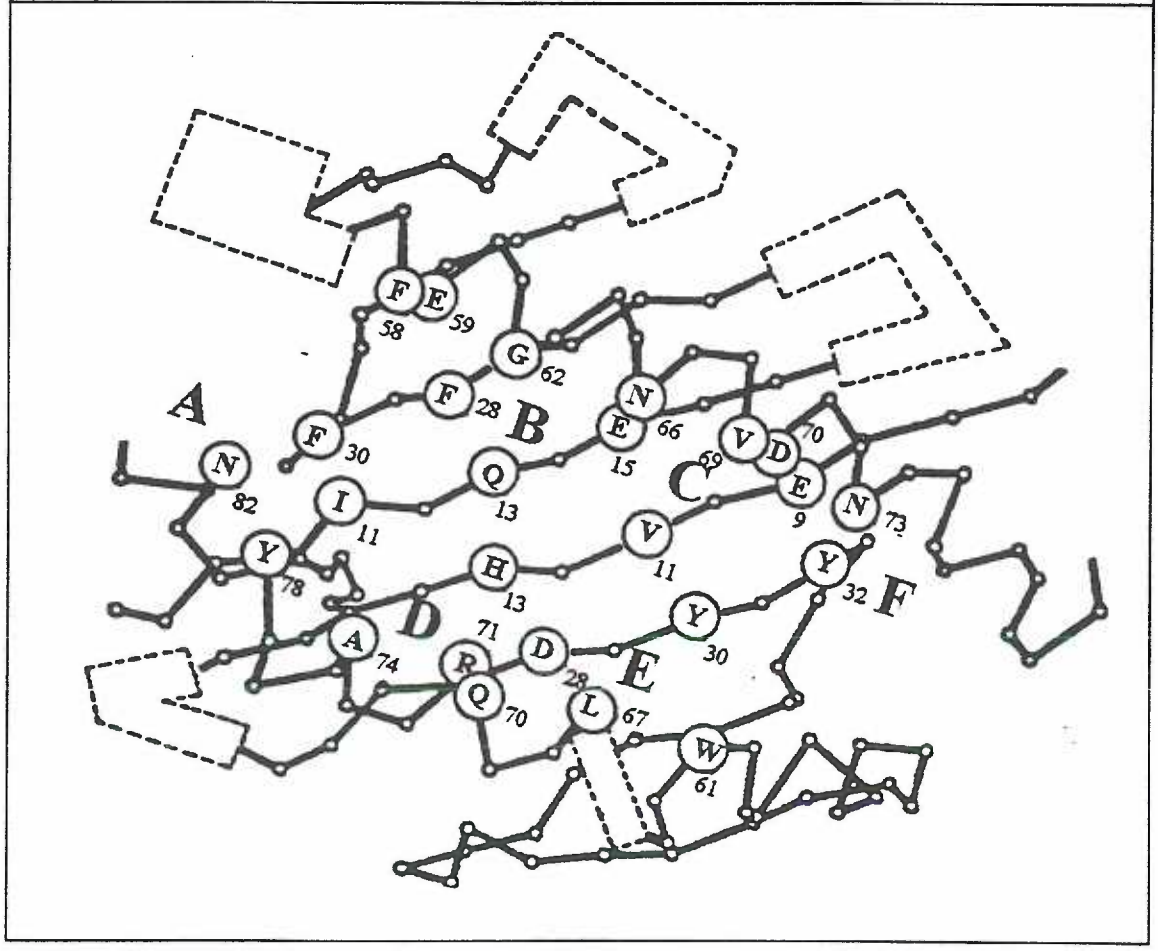
molecules (58). The crystal structure for B*2705 showed peptides aligned in the binding groove and the presence of a conserved R side chain which corresponded to the P2 position in the peptides (52). Four polymorphic residues are critical for constructing this pocket in the B*27 family. Three of these residues, 9H, 24T, 45E are positioned to form four hydrogen bonds with the arginine guanidinium group which would add substantially to the binding energy. The fourth residue, 67C is small enough to allow the arginine into the pocket and hydrophobic enough to help stabilize the aliphatic region of the R side chain. Model building experiments imply that changing only one residue, such as 67C to F or Y as is found in 8 other class I molecules, would fill the pocket leaving room for only a short hydrophobic side chain (52). Therefore, a B pocket capable of accepting an arginine in the P2 position of the peptide is a trait specific to the B27 family. The other polymorphic residues that separate the allelic forms in this family might influence the physiochemical properties of the other pockets in a similar manner leading to a unique series of binding pockets for each allele. However, the pockets of an allele will not change and that demands that the peptides binding to a particular allele all have the some common traits that allow them to interact with the critical pockets. The traits that define the specific peptide binding for a particular MHC molecule are called an allele specific binding motif.

Crystal Structure of Class II

A comparison of the framework regions between class I and class II suggested that the antigen binding cleft of the class II molecule would be similar to the one created by $\alpha 1$ and $\alpha 2$ domains of the class I heavy chain (59). The class II $\alpha 1$ domain donates a region like that of the class I $\alpha 1$ domain, and the class II $\beta 1$ domain participates like the class I $\alpha 2$ domain. The actual binding cleft would have an

Figure 10 Predicted Structure of the DRB1*0404 Antigen Binding Groove

Structure of class I based on the crystal structure of the class I molecule. The upper helix and the two strands of the β pleated sheet on the left of the antigen binding cleft are from the non-polymorphic α chain. The lower helix and the three strands of β pleated sheet on the right of the binding cleft are from the β chain. The large letters represent the predicted locations of the binding pockets. Pockets A and B will be determined mainly by the α chain. Pockets C and F will have contributions from both chains. Only pockets D and E will be determined almost exclusively by the β chain and should be critical for determining allele specific motifs for DR molecules. The RA associated sequence, QKRAA or QRRAA, at position 70-74 of the β chain, is located on the lower helix. with residues 70 and 71 point into the groove between the D and E pockets. Changing these residues could alter the properties of both pockets.



antiparallel β pleated sheet floor and be surrounded by two α helical regions, one from the α chain and one from the β chain. Recently, the crystal structure for HLA-DRB1*0101 was solved confirming these predictions (60). This is fortunate because the crystal structures for the RA associated alleles DRB1*0401 and DRB1*0404 are not yet available. Because the antigen binding cleft of the class II molecule is very similar to the class I, the prediction is that it will bind peptides in the same manner. However, there are some fundamental differences. The ends of the class II binding groove are different (60) which changes the configuration of the A and F pockets. These pockets do not have the absolute requirement of interacting with the N and C termini. For this reason, the processed peptides can be longer that 9 residues and can extend beyond the edges of the binding cleft. Despite these differences, the class II antigen binding groove can still be divided into 6 pockets whose physiochemical properties are dependent on the hypervariable residues as shown in fig 10. Therefore, an allele specific binding motif will still exist for each class II allele.

Since the DR molecule has a relatively nonpolymorphic α chain which contributes to approximately half of the binding groove, the allele specific binding motifs for all DR molecules would be expected to share some similarities. Most of pocket A is derived from the α chain, and pocket A has been shown to be very important in peptide binding to class II molecules. This would suggest that most DR molecules would accept the same side chain in pocket A just as all members of the B*27 family are expected to accept an arginine at B pocket. Residue 86 of the β chain also contributes to the A pocket. This position has been shown to be important for peptide binding to the class II molecule and in T cell reactivity (61). DRB1*0401 differs form *0404 and *0402 at this position (62) which is the reason this thesis focuses on only *0404 and *0402. However, most of the diversity in DR binding

motifs would occur through interactions with pockets more completely defined by the β chain.

As discussed above and summarized in table 2, amino acids 70-74 of the DR β chain predispose an individual to develop RA. This sequence is predicted to be located in the antigen binding cleft (59). Since the crystal structure of DRB1*0101 is known; and it also contains the critical sequence QRRAA, the predicted conformation has been strengthened (60). This critical sequence resides on the α helix that the β chain donates to the antigen binding cleft. Amino acids 70 and 71 point into the binding cleft, and their location is directly between the D and E pockets. From this vantage point, these residues may contribute to the physiochemical properties of one or both pockets. As mentioned earlier, another member of the DRB1*04 family defined as DRB1*0402 is very similar to DRB1*0401 and *0404. It differs from *0404 by only three amino acids (62). Position 67 has a conservative variation; *0402 has an I while *0404 has an L. The major differences reside at positions 70 and 71; *0402 has the sequence D, E while *0404 has the sequence Q, R. These residues can be seen on the helix donated by the β chain to the antigen binding cleft diagrammed in fig 10. The sequence variation results in a switch from a strongly positive charge to a net negative charge in the region between the D and E pockets. With only this change, DRB1*0402 lacks the critical sequence and is not associated with RA. This observation has led to the hypothesis that peptide presentation plays a role in the etiology of RA, and it is paramount to the hypothesis and experiments described in this thesis.

Figure 11 Amino Acid Sequences of DRB1*0401, *0402, *0404

The β1 domain forms the polymorphic potion of the DR antigen binding cleft. These three molecules, DRB1*0401, *0402, *0404, are homologous except for positions 67, 70, 71, and 86. *0402 and 80404 differ only at residues that contribute to the D and E pockets while *0401 also differs at 86 which influences the A pocket.

Allele	10	20	30
DRB1*0401	RFLEQ	VKHECHFFNG	TERVRFLDRY
DRB1*0402	-----	----------	----------
DRB1*0404	-----	----------	----------

	40	50	60
DRB1*0401	FYHQEEYVRF	DSDVGEYRAV	TELGRPDAEY
DRB1*0402	----------	----------	----------
DRB1*0404	----------	----------	----------

	70	80	90
DRB1*0401	WNSQKDLLEQ	KRAAVDTYCR	HNYGVGESFT
DRB1*0402	------I--D	E---------	-----V----
DRB1*0404	----------	R---------	-----V----

Allele Specific Motifs for Class I

Next to a crystal structure, one of the best ways to analyze the allele specific binding motif for a specific MHC gene product is to compare the sequences of naturally processed peptides. The class I molecule/peptide complex can be purified, and the entire array of naturally processed peptides can be acid eluted from the class I molecule. Because the class I molecule binds the amino and carboxy termini in the A and F pocket, almost all of the peptides from class I molecules are 9 residues long. The critical residues for specific peptide binding will have conserved positions relative to the amino termini, thus entire pool can sequenced. Residues making up the binding motif will be dominant at a particular position number. Motifs have been generated this way for the HLA-A2.1 (63), HLA-B53, HLA-B35 (64), and HLA-B8 (65) as well as the murine class I molecules H-2K^{d}, H-2D^{d}, and H-2K^{b} (63). To illustrate the differences in allele specific motifs, the motif from HLA-B35 was P in P2 and Y in P9 (64); for B8, it was K in P2, K in P5 and L in P9 (65). Others have used HPLC to fractionate the peptide pool then sequenced individual peptides by Edman degradation (66-68) or tandem mass spectroscopy (36, 69-71). When the individual peptides were aligned, the amino acids that occur repetitively in the same location across several of the peptides defined the allele specific motif. Table 3 contains examples of the peptides found the HLA-B*2705, and HLA-B7. HLA-B*2705 has already been shown to have an arginine dominant in the P2 position which was also demonstrated by this technique (67). A comparison of the peptides shows a strong propensity for both P1 and P9 to be R or K (67). The HLA-B7 molecule showed an almost absolute requirement for P2 to be proline, and a propensity for selected amino acids at the following positions: P1, A; P3, R; P9, L or I (70).

Table 3 Peptides Derived from HLA-B27 and HLA-B7

All the peptides eluted from HLA-B27 (67) have an R in position 2. There is also a high percentage of peptides with K or R at positions 1 and 9. This leads to the motif with an absolute requirement for R at P2, and a lesser requirement for P1 and P9 to be positively charged (K or R). HLA-B7 (72) had an almost absolute requirement for P2 to be P, and a lesser need for A at P1, R at P3 and L/I at P9. The B7 sequences were determined by mass spectrometry. The X represents I or L which have the same mass and can not be distinguished.

Peptides Eluted from HLA-B*2705

P1	P2	P3	P4	P5	P6	P7	P8	P9
R	R	Y	Q	K	S	T	E	L
R	R	I	K	E	I	V	K	K
R	R	V	K	E	V	V	K	K
G	R	I	D	K	P	I	L	K
R	R	S	K	E	I	T	V	R
R	R	W	L	P	A	G	D	A
A	R	L	F	G	I	R	A	K

Motif Derived for HLA-B*2705

K/R	R	-	-	-	-	-	-	K/R

Peptides Eluted from HLA-B7

P1	P2	P3	P4	P5	P6	P7	P8	P9
A	P	R	T	V	A	L	T	A
A	P	R	X	P	X	T	G	X
A	P	R	A	S	R	P	S	X
M	P	R	G	V	V	V	T	X
R	P	K	S	N	I	V	L	L
S	P	R	Y	I	F	T	M	L
L	V	M	A	R	P	T	V	L

Motif Derived for HLA-B7

A	P	R	-	-	-	-	-	L/I

Allele Specific Motifs for Class II

Two common techniques have been applied in an attempt to determine the binding motif of class II molecules. The first involves binding assays with a series of peptides containing selected substitutions. (73-79). The binding affinity of the original sequence is determined. One residue of the peptide is then replaced with another amino and the affinity of the altered peptide is measured against the original peptide. This process is continued until each residue in the peptide has been changed. This technique yields the relative importance of each residue to the binding of the peptide in the binding groove. The motifs for the murine class II molecules, I-A^d (73), I-E^d (74), and I-E^k (77) as well as the human class II molecules DRB1*0101 (75, 76), DRB1*0301 (78), DR2, DR5 and DR7 (75). The implications of these studies and a comparison of their results will be presented in the discussion section.

The second technique involves the purification of the MHC/peptide complex. Then the peptides are acid eluted from the MHC molecule. Because the peptides

Table 4 A Nested Cluster Derived from DRB1*0101

An example of a nested cluster from DRB1*0101 (80). The same core region of 12 amino acids is present in each peptide. There is variation in the length of both the amino- and the carboxy- termini.

V	G	S	D	W	R	F	L	R	G	Y	H	Q	Y	A	Y	D	G
V	G	S	D	W	R	F	L	R	G	Y	H	Q	Y	A			
V	G	S	D	W	R	F	L	R	G	Y	H	Q	Y				
	G	S	D	W	R	F	L	R	G	Y	H	Q	Y	A			
		S	D	W	R	F	L	R	G	Y	H	Q	Y	A			

binding to the class II molecule are not all the same length as in class I, the pool of peptides must be fractionated into individual peptides. The pure peptide is then sequenced by Edman degradation (80-82) or tandem mass spectrometry (71, 83). The peptides range from 10-34 residues long, and often the same core region is found with variations in length at the N and C termini as shown in table 4. This has been termed a nested cluster. A complete presentation of these studies will be done in the discussion section.

Antigen Presentation

Once the antigen is processed and the MHC class II/peptide complex is brought to the cell surface, it is ready for recognition by a T cell. Class II/peptide complexes are recognized by $CD4^+$ T cells. The peptide/MHC complex is specifically recognized by the T cell receptor (TCR) due to a conserved domain for $CD4^+$ binding on the β chain of the class II molecule. The TCR on $CD4^+$ cells is a heterodimer made up of an α and a β chain. Both chains are created through gene rearrangement. In the α chain, a variable (V) region combines with a joining (J) and constant (C) region; in the β chain, a V region combines with a diversity (D), a J and a C region. The diversity of the TCR is increased by the random addition of nucleic acids at the splice junctions, N-region addition, during the recombination process. The α and β chains fold in a manner that allows the VJ and VDJ regions to form a binding site for the MHC/peptide complex. When the T cell finds an APC presenting an antigen matching its TCR and expressing the proper adhesion and costimulatory molecules, the T cell becomes activated. The activated T cell clonally proliferates and differentiates into an effector cell. Any antigen would be expected to stimulate multiple T cells each with a unique affinity for the MHC/peptide complex resulting in an oligoclonal response.

Role of T cells in RA

The HLA-DR association is one of the reasons RA is proposed to be a T cell mediated disease. As mentioned earlier, HLA-DR expression is induced or increased on many cell types in the inflamed joint, and $CD4^+$ T cells are recruited into the diseased tissue. Manipulations of the T cell populations in RA patients have added credence to this theory. T cell numbers have been reduced by thoracic duct drainage (84, 85) and radiation therapy (86, 87) with both showing a decrease in the arthritis related symptoms. The use of anti-CD4 antibodies to specifically destroy $CD4^+$ T cells has had therapeutic success (88). If there was an antigen responsible for stimulating T cells and leading to the inflammatory response, the cells expressing the high affinity TCR recognizing this antigen would be expected to become dominant. The exact TCR would depend on the genetic background to the individual, in particular, the entire set of MHC class I and class II molecules and the self peptides responsible for T cell education in the thymus. Studies to identify a common element in the TCR of infiltrating T cells have been able to show a preponderance of particular β chain V regions within individual patients, but no consensus has been reached across multiple patients (89-92).

Rationale

The peptide selection model for the etiology of RA proposes that a T cell in a genetically susceptible individual who has been exposed to an antigenic challenge becomes activated then precipitates or maintains the disease. The activation of a T cell and subsequent effector functions requires the interaction of the antigen specific T cell receptor, a functional MHC molecule on the surface of the cell, and a processed peptide bound in the binding cleft of that MHC molecule. RA is associated with HLA-

DR alleles that function to present peptide antigens to $CD4^+$ T cells which are also implicated in the disease. The only variable remaining to validate this model is an antigen that is presented as a processed peptide. One method for learning about the antigen is to start with the relationship between peptides and the MHC molecule. In this case, natural mutations have provided a perfect model. The DRB1*0404 and DRB1*0402 molecule are identical except at the RA associated sequence and show completely different disease associations. Since the shared epitope is positioned to influence the physiochemical nature of the D and E pockets of the antigen binding cleft, it may play a key role in selecting which peptides bind and are subsequently presented. A patient with an RA associated allele might present a peptide that could stimulate the development of RA, but an individual with a non-RA associated allele might present a different peptide that would not precipitate the disease. Understanding more about the MHC/peptide interactions, specifically in MHC molecules associated with RA, might provide a better understanding about the nature of the antigen and etiology of the disease. This thesis uses sequences of naturally processed peptides to confirm that MHC molecules, in particular the DRB1*0404 and DRB1*0402, bind peptides derived from endogenous sources, that these peptides depend on the genetic background of the APC, and that the length of peptides binding to class II molecules is variable. Further, by a comparison of the predominant peptides and their donor proteins, the RA associated allele, DRB1*0404, selects different peptides than the non RA associated allele, DRB1*0402, possibly explaining the different disease association.

Materials and Methods

Cell Lines

Cell lines WT51(DRB1*0401), YAR (DRB1*0402), PE117 (DRB1*0404) and MZ070782 (DRB1*0101) were all obtained from E. Petersdorf. These cells are Epstein Barr Virus (EBV) transformed human B cell lines. They are listed in the ASHI repository and have been completely HLA typed. Cell lines HS-EBV (DRB1*0401/0401), JK-EBV (DRB1*0401/0404), RW-EBV (DRB1*0402/?) BL-EBV (DRB1*0404/?) are also EBV transformed human B cell lines. The transformation for these lines was done using a standard protocol (93). Basically, the mononuclear cells were separated from the red blood cells by ficoll hypaque density gradient. Live EBV was grown and prepared from the B95-8 marmoset cell line supernatant. The mononuclear cells and EBV containing cells were incubated together with cyclosporin A until a line began to grow. Human B cell lines media requirements were optimized by growing under several conditions. Minimum cell requirements were RPMI 1640 with 2 mM L-Gln, 4 % Fetal Calf Serum (FCS), and 10-20 mM HEPES. Penicillin, streptomycin and amphotericin B was also added. The cells were passaged continually in roller bottles containing 500 mls of media. The cells were grown to a maximum density that varied from 0.8-1.5 x 10^6 cells/ml depending on the cell line. When maximum density had been reached, the cells were harvested by centrifugation at 300 x g for 10 minutes. The cell pellets were resuspended at 1-2 x 10^8 cells/ml in lysis buffer consisting of PBS, 1% NP-40, 1 μM PMSF, 0.26 mg/ml 1,10-phenanthroline, 50 μg/ml pepstatin A, 2 mg/ml EDTA (94). Lysates were immediately frozen at -20 C for storage.

The monoclonal antibody clones L243 (anti DR α/β, murine IgG_{2a}) and OKT3 (anti-CD3, murine IgG_{2a}) were from American Type Cell Culture. Cells were grown in RPMI 1640 with 10 % FCS, 2 mM L-Gln, penicillin, streptomycin, amphotericin B.

Flow Cytometry

All EBV transformed B cell lines were periodically checked for HLA-DR expression by flow cytometry on a Beckenson and Dickenson FACSCAN. Approximately 1 x 10^6 cells were pelleted at 300 x g for 8 minutes and the supernatant poured off. The cell pellet was resuspended in 50 µl media with 1-10 µg L243 and incubated on ice for 30 minutes. Cell mixture was then diluted with 4 mls PBS, and the cells repelleted. Cells were resuspended in 50 µl PBS with 1 µg goat anti-mouse IgG coupled to FITC for a 30 minute incubation on ice. Cell suspension was again diluted with 4 mls PBS and pelleted. The supernatant was decanted, and the cell pellet resuspended in 500 µl PBS for flow cytometry. Acquisition gate was set using forward and side scatter, and 5000 events were collected. Analysis was done with no gates set (Data not shown).

HLA Typing

The DR type subtype of the HS-EBV, JK-EBV, RW-EBV, and BL-EBV cell lines created by EBV transformation was done by allele specific oligonucleotide hybridization (ASO) as previously described (95). DNA was amplified by polymerase chain reaction (PCR) with a 5' primer that recognized all DR4 alleles and a 3' primer that recognized all DRB1 sequences. The amplified DNA was blotted onto nylon filters and hybridized with ^{32}P labeled oligonucleotides specific for a particular DR4

allelic variant under stringent conditions. The second DRB1*04 allele was recently designated DRB1*0408. The oligo used to identify DRB1*0404 is not able to not separate these two. This work was done by others in the M. Davey laboratory (Data not shown). The HLA type of the HS-EBV line and the JK-EBV lines were confirmed by E Petersdorf at Fred Hutchinson Cancer Institute. Dr. Petersdorf uses a larger set of primers, and she was able to determine that JK-EBV expresses DRB1*0404 and not DRB1*0408. The HLA types on the ASHI cell lines was taken from published reports.

Antibody Purification and Column Production

Antibody supernatant was collected by centrifugation at 300 x g for 10 minutes and filtered though 0.22 μm filter. 1 M NaOH was added to bring the pH to approximately 8.0. Antibodies were purified on a protein A column by the following protocol. The protein A column was equilibrated with PBS. After the supernatant was loaded, the column was washed with at least 10 column volumes of PBS. Antibody was eluted in 0.5 M NaCl, 0.58 % glacial acetic acid. Fractions containing antibody were determined by UV absorbance at 280 nm. The fractions with antibody were pooled then dialyzed and concentrated in a Pro DiCon vacuum dialysis chamber against coupling buffer, 0.15 M NaCl, 0.1 M $NaHCO_3$ at pH 8.4. Antibodies were stored at -20° C until used.

Three columns were made using CNBr activated sepharose beads (Pierce): an anti-DR column with L243, a nonspecific antibody column with OKT3 and a general non-specific column with Tris. Activated beads were washed with 0.1 M HCl followed by coupling buffer. Six mls of beads were mixed with approximately 30 mgs of the concentrated antibodies and incubated at 4° C. The coupling reaction was monitored by withdrawing a small amount of fluid and checking UV absorbance at 280 nm. When

antibody concentration had decreased 90 % of the original solution, the reaction was terminated and any unreacted sites blocked by the addition of Tris. The Tris column was made by simply adding the Tris to the beads after the acid wash and then re-equilibrating in coupling buffer. Columns were washed extensively in PBS and stored at 4° C in PBS with azide.

Immunoaffinity Purification

Frozen lysate from approximately 1 x 10^{10} cells was placed in a 37° C water bath until 50 % thawed. The lysate was transferred to a small beaker and mixed on a magnetic stir plate for 30 minutes at 4° C. Large debris was removed by centrifugation at 2000 x g for 10 minutes, and lysate was further clarified by ultracentrifugation at 150,000 x g for 30 minutes (63). DR/peptide complexes were purified as described with modifications (94). The affinity columns were configured in series with the general non-specific Tris column first, the antibody non-specific OKT3 column second, and the anti-DR column third. After the series was equilibrated with PBS, 1% NP-40, the clarified supernatant was loaded. The series was then washed with 10 volumes PBS, 1% NP-40 followed by 5 volumes PBS 0.5 % NP-40, 0.1 % SDS. At this point the series was broken and the non-specific columns were eluted with elution buffer, 0.5 M NaCl, 0.15 M diethylamine, pH 10.2 then regenerated with PBS, azide. The anti-DR column was washed with at least 5 volumes PBS, 1% n-octyl glucopyranoside. Elution was accomplished with elution buffer, 1% n-octyl glucopyranoside, and 2 ml fractions were collected. Fractions containing protein were detected by BCA assay (Pierce) using bovine serum albumin as a standard and performed in a microtiter plate according to manufacturer's directions. Average protein yield was 1.9 mg from 1 x

10^{10} cells. Purity of the DR/peptide complex was checked by SDS-PAGE (Data not shown).

Acid Elution

All fractions containing DR/peptide complexes were pooled for approximately 8 mls total and concentrated in a Centricon 10 (Amicon) to a final volume of 200-300 µl by centrifugation according to manufacturer's directions. The large volume filtered through a single filter helps remove potential contaminants from the filter and limits sample losses that might occur if multiple filters were used. The filtrate is discarded and a clean collection tube is put on the filter. Glacial acetic acid was added to the concentrated DR/peptide mixture remaining in the Centricon 10 to give a final concentration of 2.5 M (96) and mixed thoroughly with a pipette. The acid eluted peptides are separated from the α and β chains of the DR molecule by filtration through the original Centricon 10 membrane (71).

HPLC Fractionation

All the peptide fractionation was done by reversed phase on an Helwett and Packard 1090 HPLC. Initially, a 4.6 x 250 mm C18 column as reported in early peptide elution studies was used (63, 66). This column had the resolution capability but not the sensitivity needed for detecting the small quantities of each peptide in the complex pools. Since the sensitivity of a column is inversely proportional to the square of the diameter, columns of 2.1 and 1.0 mm x 250 mm were eventually used. All columns were from Vydac and used the same packing material which allowed gradient conditions other than dead volume and flow rate to be directly transferred from one

column to the next. The HP 1090 HPLC needed modifications to reduce the dead time that grew as the flow rate decreased with the smaller columns. Dead time was reduced by altering the length and bore size of the internal tubing and rerouting the flow directly from the flow cell to a fraction collector.

The peptide pool from 0.5 x 10^{10} cells was used for each separation. Buffer A consisted of water with 0.07% TFA; and buffer B consisted of 80:20, ACN:water, with 0.06% TFA. The percentage of TFA added to B was optimized to give the same UV absorbance at 220 nm as buffer A. The gradient in buffer A was: 1-45 min, 2-40% B; 45-60 min, 40-75% B; 60-65 min, 75-98% B. Flow rate for the 4.6 mm column was 0.5 ml/min, for the 2.1 mm 200 µl/min, and for the 1.0 mm 70 µl/min. All the separations were done at room temperature. Fractions were collected in 1 minute intervals.

The complexity of the HPLC chromatograms and capillary electrophoresis (CE) analysis done on the fractions indicated that these fractions contained multiple peptide species. Attempts were made to devise a new gradient by optimizing the variables of temperature, rate of change, and the ion pairing agent using peptide standards. The optimum temperature was determined by multiple analysis using the same gradient while varying the temperature and plotting the retention time and peak width versus temperature. Rate of change was determined by multiple analysis using progressively slower gradients and plotting the peak width versus rate of gradient change. Different ion pairing agents, heptafluorobutric acid and sodium acetate, were tried on multiple gradients, but the results proved too variable for use with actual samples. Eventually, a second gradient was devised using the same solutions A and B. In buffer A the gradient was: 1-87.5 min, 10-55% B; 87.5-97.5 min, 55-90% B at temperature of 45° C. The second gradients were done on the 1.0 mm x 250 mm C18 column. Identical fractions from multiple primary separations were pooled. The organic solvents were

removed and the pool was concentrated on a Savant speed vacuum. These concentrated pools were separated again using the second gradient with fraction collecting governed by a peak detector use UV at 220 nm.

Capillary Electrophoresis Separation

The purity of the fractions collected from the HPLC was assessed using free solution CE on an Applied Biosystems Inc. 270 A Capillary Electrophoresis system. The organic solvents were removed and the individual fractions concentrated to 10-20 μl in a Savant Speed vacuum. A 50 μm x 77 cm capillary was washed with 0.1 M NaOH for 3 minutes by vacuum then equilibrated with 20 mM sodium citrate, pH 2.5 (Applied Biosystems Inc.) also for 3 minutes by vacuum. The fractions, now with a solvent phase of water as a result of the concentrating, were vacuum loaded for 15 seconds giving a loading volume of approximately 180 nl. The low ion concentration in the loaded sample caused the sample to be concentrated at the sample front in a manner similar to a stacking gel in an SDS-Page system. The separation was done at 30 kV for 12 minutes at 30 C and peaks were monitored by UV at 200 nm. Between each separation, the capillary was washed with 0.1 M NaOH and then equilibrated with citrate buffer by vacuum for 3 minutes.

Peptide Sequence Determination

Fractions that appear to contain a single peak on both HPLC and CE were considered pure. Often fractions would contain a dominant peak and several minor peaks. Fractions containing a pure peptide species and those containing a clearly dominant peptide were sequenced by Edman degradation on a modified Applied

Biosystems Inc. 477A protein sequencer according to manufacturer's protocols. A modification was made to the sample injection of the 120, the phenylthiohydantoin (PTH) analyzer used to detect the cleaved amino acids, to increase the loaded sample by doubling the injection loop size. All sequences derived were read manually. Approximately 1 μl was retained for confirmation of the end points and further analysis by mass spectrometry. Mass spectrometry sequencing was done by J. Yates and A. McCormack at the University of Washington as previously described (83, 97).

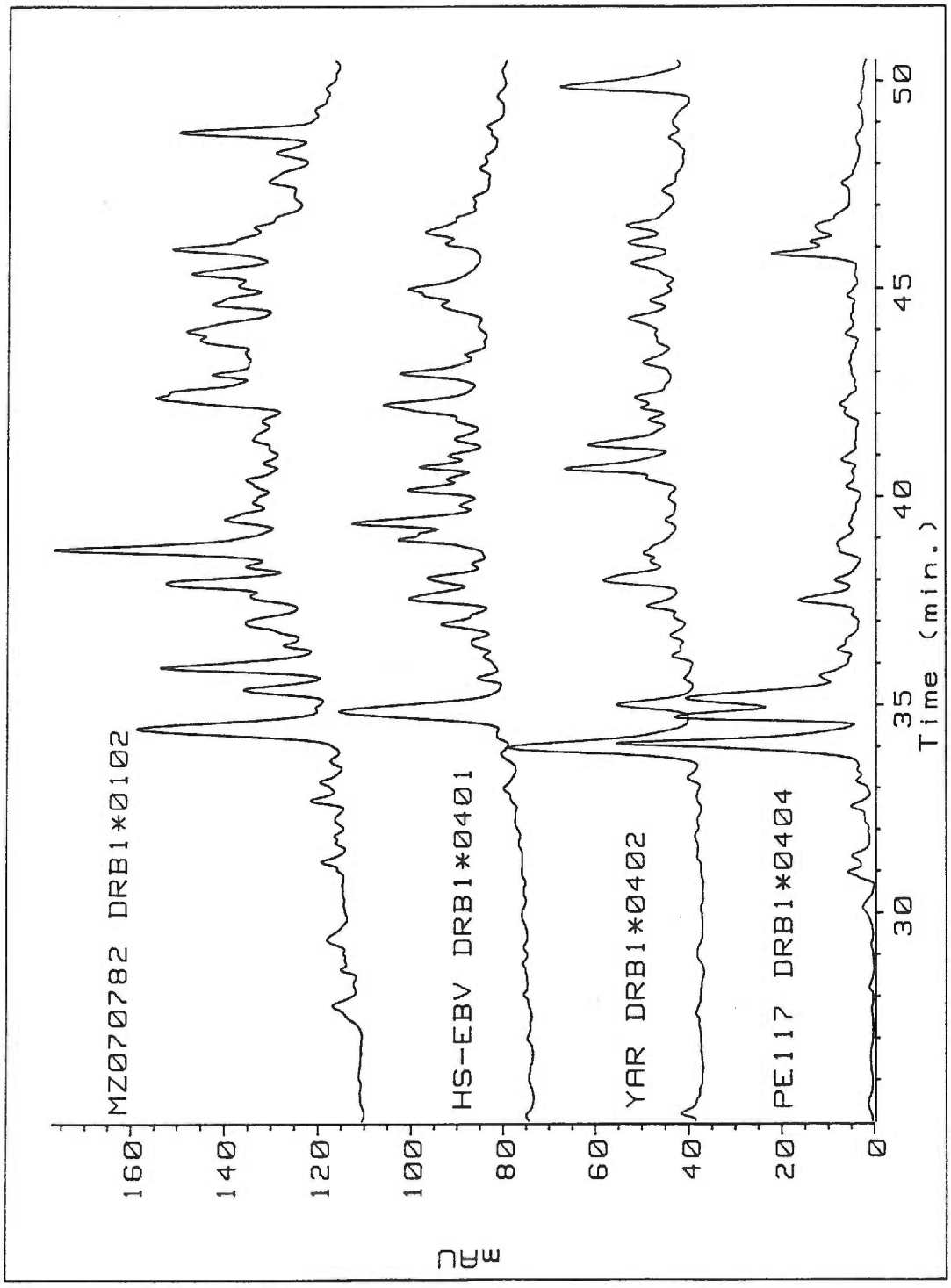
Results

Comparison of Peptides Derived from Different DR Alleles

Four different homozygous EBV transformed B cell lines were grown under the same culture conditions. The MHC/peptide complexes were immunoaffinity purified and the peptides were acid eluted from the MHC. The peptide pool from approximately 0.8 mg of DR molecules was fractionated by reversed phase HPLC. A comparison of the chromatograms is presented in fig 12. The cell lines MZ070782 (DRB1*0101), HS-EBV (DRB1*0401), and PE117 (DRB1*0404) all express RA associated alleles, while YAR (DRB1*0402) expresses a closely related but non-RA associated allele. Reversed phase HPLC separates on the basis of hydrophobicity with the more hydrophobic molecules requiring increased organic solvent for elution and, therefore, elute later in the separation. Molecules that elute at the same time have a similar hydrophobicity. The peptides from all the DR molecules eluted over a range of 25-40 % of the organic solvent used. The wide elution conditions demonstrates that the DR molecules can bind peptides with a broad range of hydrophobicity. Interestingly, the range of peptide hydrophobicity varies little from one molecule to another. It is important to note that all the chromatograms are unique. Class II molecules function to present exogenous antigens. Since all the cell lines were grown in the same media, they all were exposed to the same antigen pool. Despite this, each cell line gave a unique profile suggesting the presence of cell specific or MHC specific peptides.

Figure 12 HPLC Chromatogram of Peptides Derived from Different DR Alleles

DRB1*0102, *0401, and *0404 are all RA associated alleles. DRB1*0402 is a not-RA associated allele. All the alleles bind peptides that have similar hydrophobicity, but each allele has a unique repertoire of peptides.

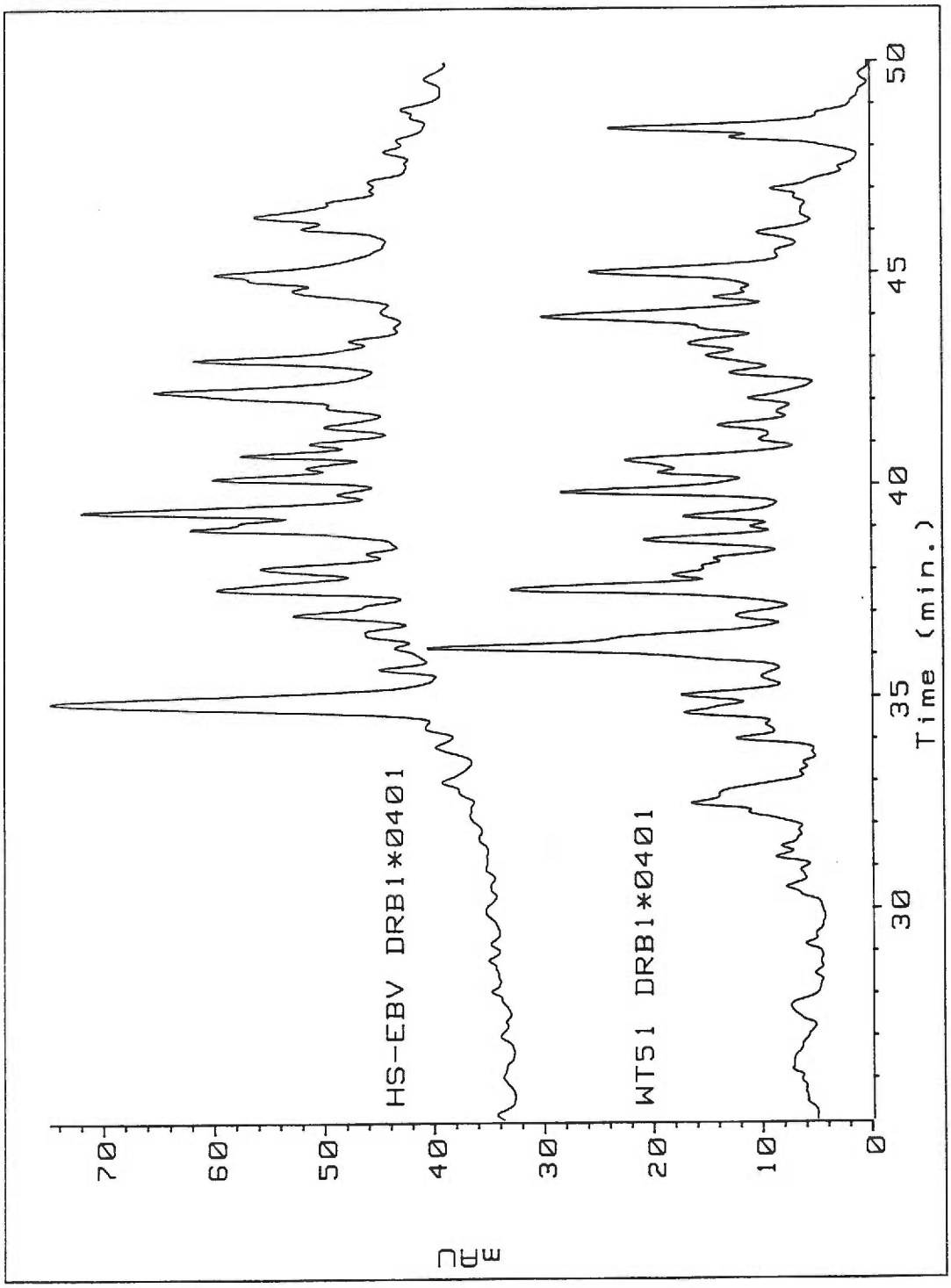

MZ070782 DRB1*0102
HS-EBV DRB1*0401
YAR DRB1*0402
PE117 DRB1*0404
mAU
160
140
120
100
80
60
40
20
0
30
35
40
45
50
Time (min.)

Comparison of Peptides Derived from a Single DR Allele

In an effort to determine the contribution of cellular proteins other than MHC in the chromatogram variations seen in fig 12, the peptide pool derived from identical class II molecules harvested from two different cell lines was analyzed. Fig 13 shows a comparison of the profile of peptides derived from HS-EBV (DRB1*0401) and WT51 (DRB1*0401). The elution characteristics of peptide pool was similar to those in fig 12. Even though these two cell lines express the same HLA-DR molecule, they yield very different peptide profiles. This demonstrates that the uniqueness of the peptide repertoire is influenced by the genetic background of the cell line.

Figure 13 HPLC Chromatogram of Peptides Derived for DRB1*0401

Peptides were eluted from two different cell lines that express DRB1*0401, HS-EBV and WT51. As demonstrated for different alleles, the peptides had similar range of hydrophobic properties. However, each cell line had a unique repertoire of peptides bound to it's DRB1*0401 protein which indicates that the peptides presented are influenced by the genetic complement of the cell.

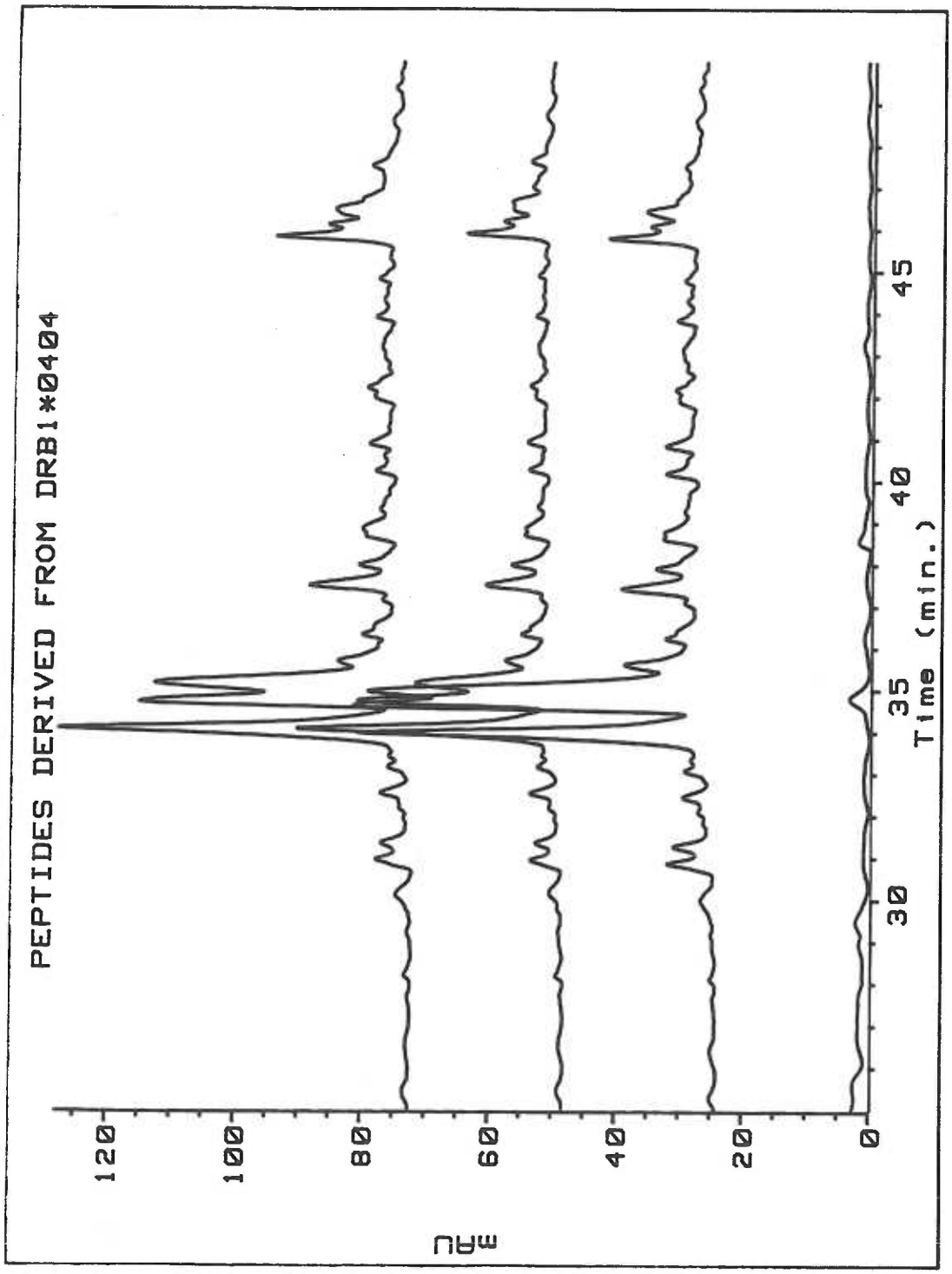

mRU
70
60
50
40
30
20
10
0
30
35
40
45
50
Time (min.)
HS-EBV DRB1*0401
WT51 DRB1*0401

Reproducibility of HPLC Separations

The PE117 (DRB1*0404) and the YAR (DRB1*0402) cell lines were grown over a six month period. Fig 14 shows several chromatograms of peptides derived from the PE117 cell line at various points over this time period. The chromatograms stayed constant, suggesting that the most abundant naturally processed peptides presented by these cells do not change. If the dominant peptides were from exogenous proteins and the culture conditions were constant, then the chromatograms would be expected to be identical. However, human cell lines occasionally change growth characteristics after many passages. This was observed with the PE117 cell line that began to grow much slower and was not able to grow to as high a cell density suggesting an alteration in protein expression. As can be seen in the figure, this did not change the chromatograms in any detectable manner.

Figure 14 HPLC Chromatogram of Peptides Derived from DRB1*0404

To obtain enough starting material for these separations, the cells had to be grown over a long period of time. Multiple isolations of DRB1*0404 from the cell line PE117 showed identical chromatographic profiles. The bottom tracing is a negative control separation

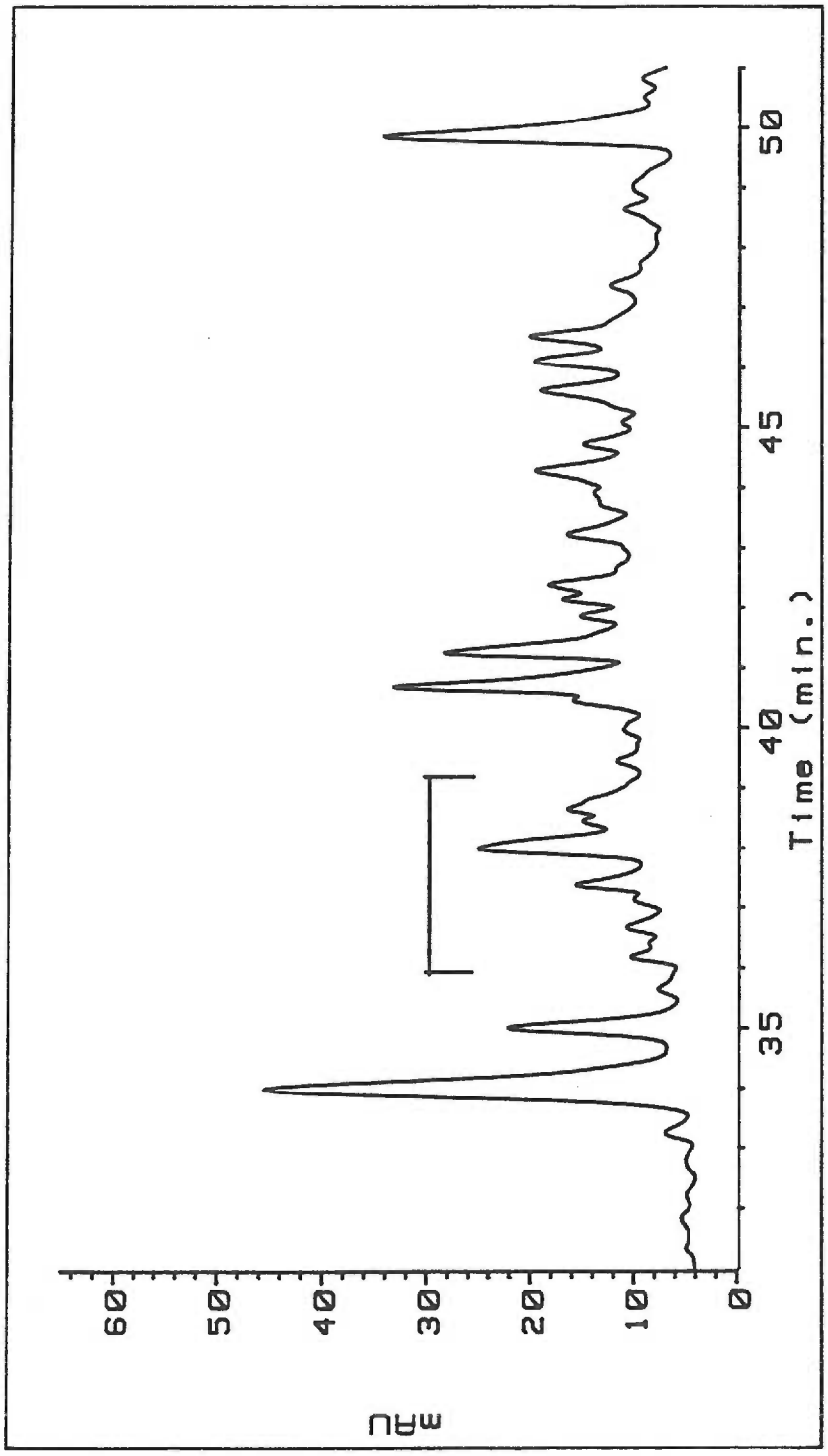

PEPTIDES DERIVED FROM DRB1*0404
mAU
120
100
80
60
40
20
0
30
35
40
45
Time (min.)

Complexity of Peptide Mixtures

Peptide fractions from the primary separations still contained a complex mixture of peptides, and the amount of any one species was too small to attempt sequencing by Edman degradation. However, since the isolations were reproducible, multiple separations could be pooled and fractionated again with different conditions. The peptides derived from DRB1*0404 were done first. Individual fractions from 7 isolations using 1 x 10^{10} PE117 cells were pooled. Since the initial separations had been collected by time, often a peak was split between two fractions. When these two fractions were separated again, a peak was found that had identical elution characteristics in both chromatograms. To avoid this problem with the peptides derived from DRB1*0402, the fractions from the primary separation were grouped according to natural breaks in the chromatography. Figure 15 shows one of these groupings. The entire group was pooled across 7 different isolations of 1 x 10^{10} cells, and this pool was fractionated a second time as shown in figure 16. The combination of adding multiple primary separations and a different gradient dramatically increased the number of peaks visible. The initial group appeared to contained perhaps 6-10 peptides often in overlapping peaks. The secondary separation demonstrates at least 14 major peaks and numerous minor peaks. The fractions from this separation were collected with a peak detector to minimize peaks being split between fractions.

Figure 15 HPLC Chromatogram of Peptides Eluted from DRB1*0402

Fractions from this separation of *0402 were collected at one minute intervals. These fractions were then pooled based on natural breaks in the chromatogram. One pool is identified from 36 to 39 minutes. The corresponding fraction from 14 different primary separations were added to this pool prior to the secondary separation.


mAU
60
50
40
30
20
10
0
35
40
45
50
Time (min.)

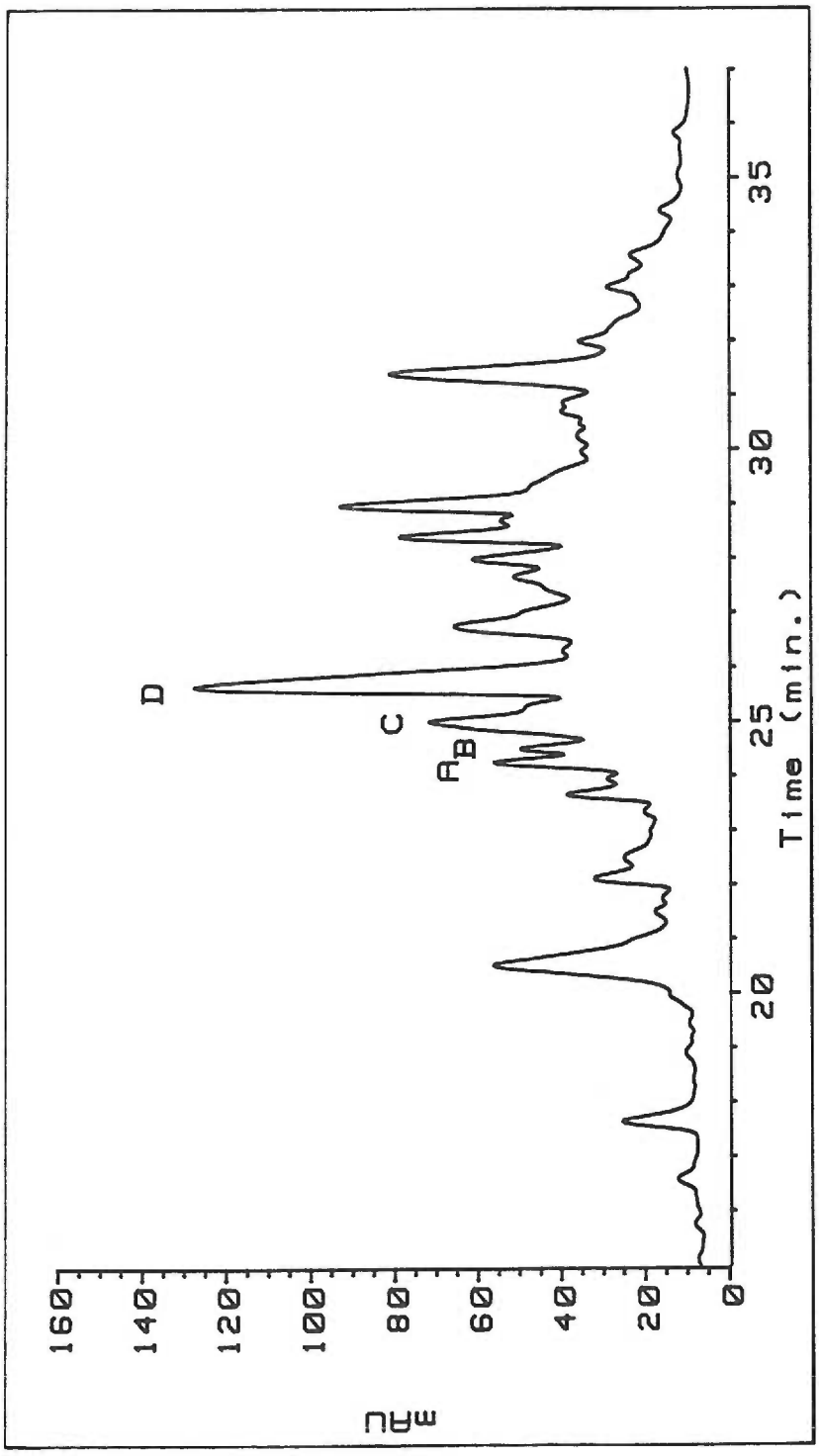

mAU
160
140
120
100
80
60
40
20
0
20
25
30
35
Time (min.)
D
C
A
B

Figure 16 Chromatogram of a Secondary Separation of DRB1*0402 Peptides

The pool identified in figure 15 was separated a second time using different gradient conditions. The number of identifiable peaks increased significantly with the second gradient. Fractions from this separation were collected with a peak activated fraction collector. Fractions containing the peaks marked A through D are representative of factions analyzed by CE.

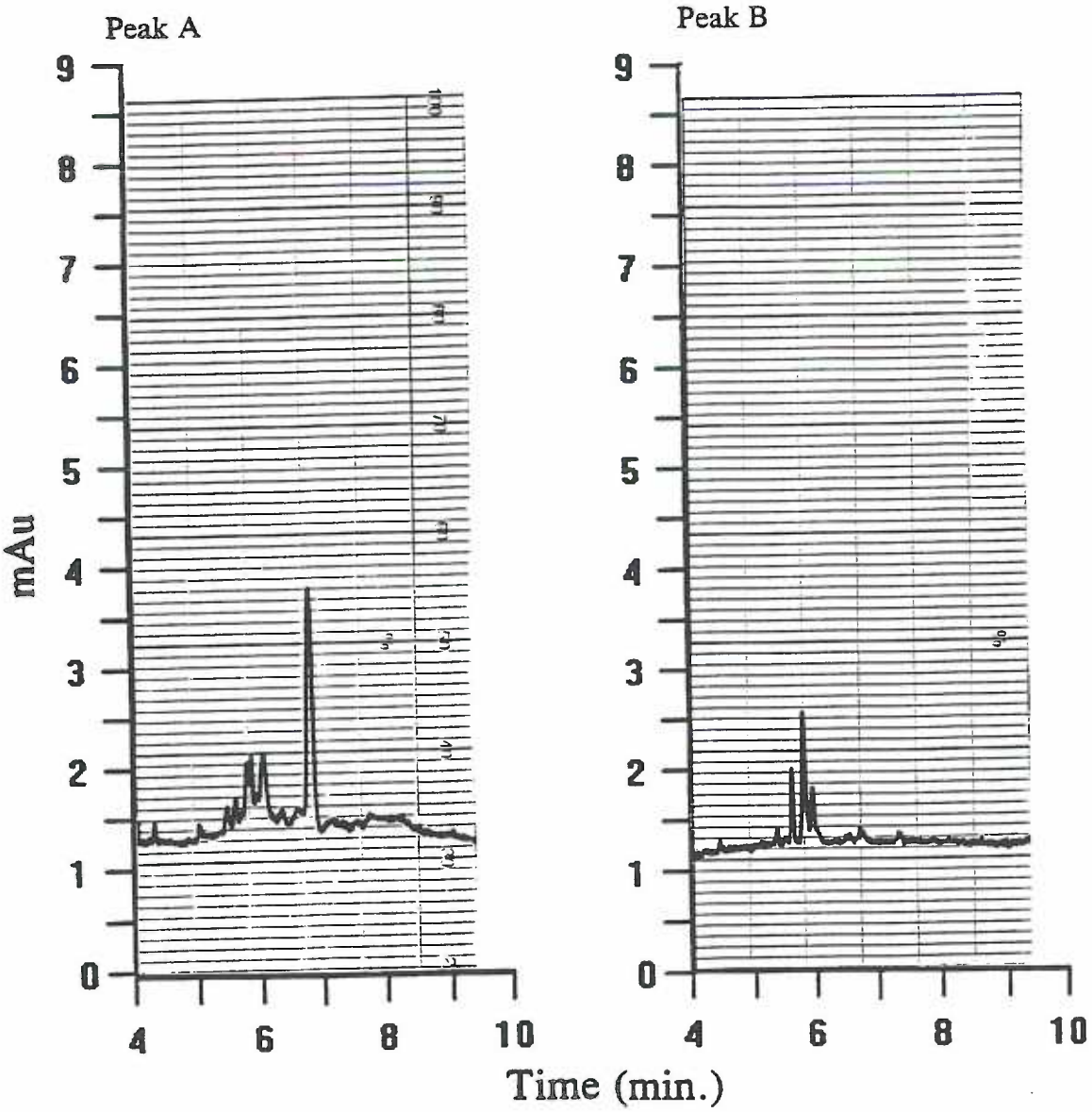
Purity of Fractions by CE

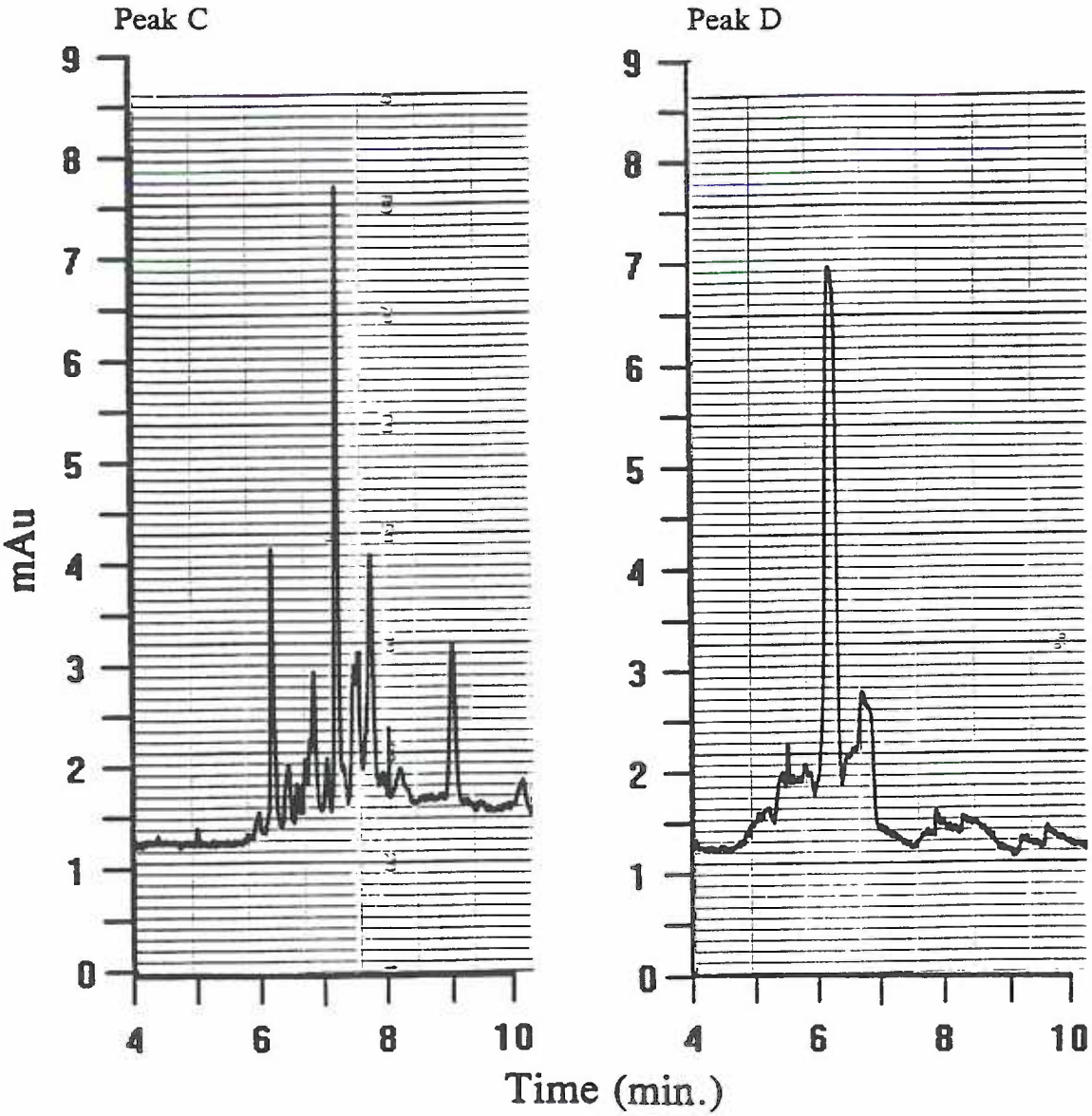
The individual fractions from the secondary separation were then concentrated and checked for purity. Free solution capillary electrophoresis separates on the basis of a charge to mass ratio. When a buffer system of citrate pH 2.5 is used, all peptides will have a net positive charge and migrate to the anode. Since this separation is independent of the hydrophobicity, it is an ideal system to check the purity of the reversed phase HPLC fractions. Figure 16 has four peaks labeled A, B, C, and D. The CE chromatograms for these peaks are shown in figure 17. CE analysis shows that all these fractions contain multiple peptides. For instance peak A is a balanced peak on the HPLC suggesting it is relatively pure. By CE that fraction contains one major species and 8-9 contaminating peptides. The fraction containing peak B has 2 other contaminating peaks by CE that make up a significant portion of the total protein. Peak C is a broad peak with shoulders on both sides in the HPLC chromatogram indicating a least three peptides in the fraction. CE demonstrates that this fraction is extremely complex containing numerous peptides. The peak marked D is ideal. On HPLC it is a large balanced peak. The CE shows one broad dominant peak with a few small contaminants indicating an almost pure peptide.

Purity of the peptides is critical for sequencing by Edman degradation. Edman degradation is a process that cleaves one amino acid at a time from the amino terminus of a protein. The cleaved product is then detected and identified, usually by reversed phase HPLC analysis. If there are multiple amino termini present, there will be more than one amino acid detected by the HPLC. Unless the amount of the two amino acids differs significantly, it is impossible to determine which residue belongs to a particular parent peptide. This is where the analysis by CE is critical. In the fraction that contains peak B, the CE dominant peak may represent only half the total protein. It would be difficult to determine which amino acid should be assigned to the dominant peptide, and the assignments to the two contaminating peptides would be impossible. The fraction containing peak C is extremely complex. There is no hope of making accurate assignments on the peptides in this fraction. However, fraction D is ideal. One clearly dominant peptide removes the problem of which peptide should be assigned the residue. The fraction with peak A has one dominant CE peak and several smaller peaks. The assignments can usually be made to the most prevalent species, but the complexity of the sequencing chromatograms demands that they be manually interpreted.

Figure 17 CE Chromatograms of Fractions from DRB1*0402

The fractions identified in figure 16 as A through D were assessed for purity using CE. Peak A has one dominant peptide and several contaminants. Because the peptide represented by the dominant peak is present in a much larger quantity, the fraction can be used for sequencing. The fraction containing peak B does not have a dominant peak. The fraction with peak C contains many peptides. Both B and C can not be sequenced by Edman degradation. The fraction containing peak D is ideal. It contains one large, dominant peptide peak


Peak A
Peak B
mAu
9
8
7
6
5
4
3
2
1
0
4
6
8
10
Time (min.)


Peak C
Peak D
mAu
9
8
7
6
5
4
3
2
1
0
4
6
8
10
Time (min.)

Peptides Derived from DRB1*0404

The peptides derived from DRB1*0404 are presented in table 5. They showed two unexpected findings. The first that the same core region was present multiple times with variations at the amino and carboxy termini. This combination of similar peptides has been called a nested cluster. The second was the origin of the nested clusters. Both clusters were derived from the heavy chain of class I molecules. The third peptide matches with a sequence from an Ig kappa light chain.

Table 5 Peptides Eluted from DRB1*0404

Ten Peptides derived from DRB1*0404. Two closely related nested clusters were found. Sequence matches were found for all sequences. The two clusters originate from the HLA class I molecules known to be expressed by this cell line. The third matches with a known Ig kappa light chain sequence, but it is unknown if this light chain is expressed by the cell line.

HLA-B40

	S	H	S	M	R	Y	F	H	T	A	M	S	R	P		
	S	H	S	M	R	Y	F	H	T	A	M	S	R	P	G	
G	S	H	S	M	R	Y	F	H	T	A	M	S	R	P	G	
G	S	H	S	M	R	Y	F	H	T	A	M	S	R	P	G	R	G

HLA-C10

	S	H	S	M	R	Y	F	Y	T	A	V	S	R	P		
	S	H	S	M	R	Y	F	Y	T	A	V	S	R	P	G	
	S	H	S	M	R	Y	F	Y	T	A	V	S	R	P	G	R	G
G	S	H	S	M	R	Y	F	Y	T	A	V	S	R	P	G		
G	S	H	S	M	R	Y	F	Y	T	A	V	S	R	P	G	R	

Ig Kappa Chain

P	G	K	A	P	K	R	L	I	Y	A	A

Peptides Derived from DRB1*0402

The peptides sequenced from DRB1*0402 are presented in table 6. As with DRB1*0404 the most prominent feature of the peptides is a nested cluster. This cluster matches sequences HLA-B38 which is expressed in this cell line. The other three peptides are interesting. The next two have 8 of 14 residues in common in a pattern possibly suggestive of Ig gene rearrangements. However, no sequence matches have been found for these peptides.

Table 6 Peptides Eluted from DRB1*0402

The nested cluster matches a sequence from HLA-B38 which this cell line is known to express. The other sequences did not have complete matches, although it should be noted that there is greater than ninety percent homology in the first two peptides with proteins from the glycolytic pathway.

HLA-B38

			G	P	D	G	R	L	L	R	G	H	N	Q	F	A	Y	D	G	K		
G	P	D	G	P	D	G	R	L	L	R	G	H	N	Q	F	A	Y	D	G	K		
G	P	D	G	P	D	G	R	L	L	R	G	H	N	Q	F	A	Y	D				
G	P	D	G	P	D	G	R	L	L	R	G	H	N	Q	F	A	Y	D	G	K	D	Y
V G	P	D	G	P	D	G	R	L	L	R	G	H	N	Q	F	A	Y	D				

Unknown Donor Proteins

	F	G	R	I	G	R	L	V	T	R	A	A	F	N	S	
G	P	G	I	I	G	R	L	V	T	Y	A	?	F	N		
	S	N	Q	E	I	K	F	L	D	S	H	A	S	H	Y	G

Comparison of Nested Clusters

A comparison of the nested clusters from DRB1*0404 and DRB1*0402 suggest a fundamental difference in peptide binding. The sequences corresponding to the class I alleles donating nested clusters are compared in table 7. The predominant nested clusters found in DRB1*0404 are derived from the same region of the HLA-B and C molecules, stretching from residues 1-18 but with a core region of 2-14. The cluster donated by HLA-C10 is a dominant with 6 peptides found compared to the 4 derived from B40. As can be seen in table 7, the sequences are very homologous. Residue 9 is a His in the B40 peptide and a Y in C10 while residue 12 is a Met in the B40 and a V in the C10. The sequence from the same region of the B38 molecule expressed by the line used for the *0402 analysis closely resembles the peptides found in *0404. It differs from *0404 dominant cluster by only at residue 11, an A to S switch. Because the B38 molecule is the donor of the dominant nested cluster found in *0402, it does arrive in the CPL and is processed. However, the 1-18 peptide was not found in the *0402 molecule. The dominant nested cluster found in *0402 is derived from 103-123 of the B38 molecule with a core region of 107-119. The homology between this region and the identical region in the B40 molecule from the *0404 cell line is nearly complete. The only difference is highly conservative with 116 of the B38 having a F, and the B40 having a Y. Despite the similarity between these two peptides, the 103-123 region is not found in the *0404 allele. These two antigens that have the same tertiary structure and nearly homologous primary structures in the epitope regions donate very different peptides to the *0404 and *0402 class II molecules indicating that these class II molecules select very different peptides for presentation.

Table 7 Comparison of the Class I Donor Protein Amino Acid Sequences

This table compares the entire length of the dominant nested clusters from DRB1*0402 and *0404. The first two lines show the sequences from B40 and C10 on the PE117 cell line (*0404). The third line shows the B38 sequences expressed by the YAR line (*0402). Sequence 1-18 represents the dominant peptides from *0404, and 103-123 is the one from *0402. Since these all of these peptides are derived from the framework regions of the class I molecule, there is a great deal of homology between the different alleles. Despite the homology, only 1-18 is found in *0404, and only 103-123 is found on *0402.

Class I allele	1 10 18	103 113 123
B40 (DRB1*0404)	GSHSMRYFHT AMSRPGRG	VGPDGRLLRGH NQYAYDGKDY
C10 (DRB1*0404)	--------Y- -V------	
B38 (DRB1*0402)	--------Y- SV------	---------- --F-------

Discussion

Class I in Class II

One of the most striking findings of this thesis and the recent work of others in human model systems is that large nested clusters derived from class I heavy chains make up the most dominant peptides found in class II molecules (80, 82). Peptides derived from murine class II molecules have not shown this propensity for nested clusters or for the class I heavy chain (71, 81, 98). Peptides eluted from murine class II molecules have rarely had more than two variations of the same core region of any peptide. Sequences matching β2 microglobulin have been found, but no peptides from the class I heavy chain have been identified. However, there are many similarities. Both murine and human class II molecules seem to be dominated by peptides donated from a small number of parent proteins. In mice 50-75 percent of all the class II sites are reportedly filled with peptides derived from only three parent proteins (71). In this thesis 9 of 10 peptides eluted from DRB1*0404 were from 2 donor proteins, HLA-B40 and HLA-C10, and 5 of 8 peptides derived from DRB1*0402 were from one parent protein, HLA-B38. Others have observed the same preference for donating proteins with 16 of 20 peptides derived from DRB1*0101 originating from 2 parent proteins (80). Despite the dominance of a few peptides, both murine and human class II molecules have a large number of unique peptides. Between 650 and 2000 unique peptide masses have been reported from murine class II molecules (71), and five DR alleles averaged over 200 unique, detectable masses (82).

Perhaps the propensity for nested clusters in human class II molecules that is not found in murine class II molecules represents a difference in the proteolytic machinery used for processing. Unfortunately, the actual machinery for antigen processing has not been described in mice or in humans, so this is just speculation. Another explanation is the source of the MHC/peptide complexes. The human

experiments have focused on EBV transformed B cells, while the murine work has been done with B cell lymphomas, spleen preparations, and thymus lysates (71, 98). EBV transformation leads to the expression of an activated phenotype that may not be present in lymphomas or resting cells from the spleen or thymus. Perhaps the activation alters the route into the CPL, the diversity of antigen processing, or simply the expression levels of the class I donor proteins.

The finding of class I peptides in human class II molecules has been found by more than one laboratory and with multiple cell lines. Of six HLA-DR molecules previously reported and the two more included in this thesis, only 1, DRB1*0801, failed to have the most common peptides derived from MHC related molecules (80, 82). One theory to explain the presence of class I peptides in class II molecules is that the class I peptides could simply be the most common fragment generated in the degradative pathway of cell surface proteins. It may be a disadvantage to the host to have no peptides available to occupy the class II molecule, and these common fragments could fill that need. If this were true, nested clusters of the β2 microglobulin would be expected to be seen as well. Even though the β2 microglobulin protein is smaller, it would be present with the heavy chain and would be in an equal molar ratio. Since β2 microglobulin has not been found in any human class II molecules, the class I heavy chain fragments probably provide a more important role.

A second theory is that class I derived peptides in class II molecules control the alloreactive process. The high precursor frequency of T cells that respond in an allogeneic assay may in part be caused by the high density of class I peptides being presented by the class II molecules. However, it is difficult to define a reason for such a system to evolve. Alloreactive responses have very little survival benefit and even inhibit survival in the era of transplantation. In a world requiring alloreactivity, the optimum system for identifying self would be to have class I hypervariable regions

presented in class II. Using the most variable regions of highly polymorphic class I molecules for presentation in the highly polymorphic class II molecules would certainly provide the most unique identification of self. Instead, the class I peptides found in *0404 and *0402 are from rather constant framework regions of the class I molecule.

Though the presence of class I peptides in class II molecules probably did not evolve to control the alloreaction, it may play a critical role Although there is no evidence of alloreactivity against class I peptides in class II, there is evidence of T cell reactivity against class II peptides in presented in class II (99, 100). Since peptides derived from class I are clearly the most dominant peptide presented, any T cell with a TCR that has some reactivity may be able to reach the activation threshold. This would allow a large number of low affinity TCR to be stimulated. Murine models demonstrate this in practice where it has been estimated that up to 10% of circulating T cells will respond to an alloreactive challenge (101).

A third theory is that the class I heavy chain fragments play an important immunological role. The class II molecule with a class I peptide may be important for tolerance of T cells. T cells are educated through positive and negative selection in the thymus by an interaction of TCR with MHC molecules (101). Positive selection is thought to occur in the thymic cortex when T cells are at a $CD4^+/CD8^+$ stage. These cells express low levels of TCR which interact with MHC/peptide complexes displayed by the cortical epithelial cells (101). If class I peptides are the dominant species presented by the cortical epithelial cells, then these complexes would be expected to play a dominant role in the selection process of T cells. After positive selection the T cells increase the expression of TCR, become singly $CD4^+$ or $CD8^+$, and move to the thymic medulla (101). In the medulla these cells undergo negative selection by interacting through their TCR with the MHC expressing cells (101). Again, if the dominant peptides found in the class II molecules were class I derived, then these

peptides would have a profound influence on the selection of the $CD4^+$ T cells. A comparison of peptides eluted from murine spleen, thymic cortical epithelium and thymic medullary epithelium found that the dominant peptides eluted from the class II molecules derived from all three tissues were the same (98). Another demonstration of the predominance of a single self peptide and its possible role in selection was done with an antibody that recognizes a self peptide bound to the murine class II molecule, $I\text{-}A^b$ (102). This antibody recognizes approximately 12% of the class II molecules on an APC. It also binds to cells in the thymic medulla, the location of negative selection (103).

Perhaps the immunological role of these predominant peptides is to reduce the complexity of self (102). Estimates suggest that human beings have the potential to make 100,000 unique proteins spread over various tissue distributions. An incredible amount of energy would be expended if the thymus had to produce or capture all the potential epitopes and present them in high enough levels to trigger positive and negative selection. It would be much simpler if only a few peptides dominated the repertoire displayed on a normal APC leading to the exclusion of other epitopes at the levels needed to activate T cells. Then thymic education would only have to accommodate this reduced number of epitopes. Those cells that escaped thymic education would be faced with a sea of APC also expressing extremely high levels of the same MHC/peptide complex, a paradigm known to result in tolerance. Any cell that had altered its repertoire of peptides, reducing the few normally dominant species and increasing the normally nondominant or foreign peptides, would lead to T cell stimulation. By presenting only a few peptides throughout the body, most of the cellular proteins would then reside in immunologically privileged sites until the system was perturbed.

Dominant peptides from any protein donor could be used for thymic education and peripheral tolerance. However, class I molecules are an excellent choice. They are expressed on all nucleated cells, so they will be available should class II become induced on cells other than traditional APC. Class I molecules are a surface membrane protein that could be directed to the CPL using the normal pathways for receptor mediated antigen internalization. Most importantly, class I molecules tend to be coordinately regulated with class II molecules. This would be critical for controlling self responses when an APC is activated and class II is upregulated. If there is not enough foreign antigen to fill the increased number of binding grooves, the cell might resort to other self proteins leading to an increase in self peptide presentation. This might allow the development of pathological T cell response. However, since the class I is upregulated at the same time, it would be available to continue controlling the self response.

Class II Allele Specific Motifs

Since the demonstration of class I allele specific motifs, it has been predicted that class II molecules should exhibit the same phenomenon of a motif to fit a particular allele. This has been difficult to demonstrate. First, most of the work published on class II binding has been done on HLA-DR molecules. As mentioned previously, the DR α chain in relatively non-polymorphic. Since the α chain contributes half of the antigen binding cleft, it would be expected that many peptides would be able to bind to more than one allele. Second, the theory of broad specificity, the interaction of both the α and β chains with the peptide backbone rather than specific side chains, should make it possible for peptides to bind on multiple class II molecules. This has been demonstrated clearly with peptides derived from hemagglutinin and tetanus toxoid that

are considered universal binding peptides. It has also shown with binding assays. In one series, 133 peptides from various proteins were assayed for binding on 4 different DR alleles with 20 % of peptides binding to 2 class II molecules and 25 % binding to 3 or 4 (104). Even if the α chain were not the same in each DR molecule and broad specificity did not contribute to peptide binding, overlap of peptide binding should be expected. Since the ends of the class II binding cleft are open and most of the peptides used in binding assays are longer than the minimum length to fit in the groove, 9-10 residues, peptides may be using different registers for binding on the various class II molecules.

Despite the complexities of working with HLA-DR, several approaches have been tried for finding an allele specific motif. One frequently tried method uses binding assays. Binding assays to determine an allele specific motif can not be done with the specificity of a T cell as a detection system. Since a T cell detects the three dimensional conformation of the MHC/peptide complex, there is no way to determine if a change seen in the T cell response is due to variations in the binding of the peptide to MHC or to alterations of the TCR binding to MHC/peptide complex. Therefore, a binding assay must utilize a detection system that is dependent solely on the peptide interaction with the MHC. Assays have been designed using cells expressing a particular class II molecule or using immunoaffinity purified class II molecules. To determine which residues are critical to the motif, amino acids are altered one at a time then the binding is compared to the parent molecule. All of these systems require some type of modification for bound peptide to be followed. The detection systems have varied from direct binding detected with fluorescence to competition assays using a radiolabeled competitor peptide. Common alterations are iodination of a tyrosine residue either one inherent to the peptide or one added to either termini, or it could be biotinylation or fluorescent labeling of a primary amine, either the N termini or an

internal lysine. These changes to the peptide may alter its binding properties, and they may or may not exert a constant influence across all the assays due to the new chemical moiety.

The second problem is that the complete analysis is usually done with only one or two peptides. If the interactions of one residue with a particular pocket influences the nature of interactions of the next residue with its pocket, then an analysis done in only one context may not be valid. A third problem is that most MHC molecules come to the surface only when they have a peptide bound and have achieved a stable conformation. Any peptide being assayed will have to displace the endogenously bound peptides. In essence, any binding assay using MHC molecules synthesized by non-mutated mammalian cells will be a complex competition assay. As the data presented in this thesis demonstrates, the array of competitor peptides can be very diverse. This pool of potential competitors will remain quite stable as long as the same cell line is used and the growth conditions are conserved. However, a unique set of competitors may be bound to the same MHC allele derived from a cell line with a different genetic background and, therefore, different donating proteins. It remains to be determine if the binding affinities of peptides will differ when repertoire of endogenously bound, competitor peptides is changed.

Despite these problems allele specific motifs derived through binding assays have been reported for human class II molecules DRB1*0101, 0301, 0401, DR2, DR5, and DR7 (75, 76, 78, 79) As an example, a motif was generated for DRB1*0401 using hemagglutinin 307-319 (HA) as the parent peptide (79). Each position of HA was exchanged individually with 4 to 8 different amino acids for a total of 81 unique derivatives. The derivatives were all assayed against an iodinated competitor, HA with a tyrosine added to the N termini. A dose providing 50% inhibition of the labeled HA binding was used as a comparison value between peptides. Because the position of the

first critical residue varies with length of the N termini, it will be described as the i position with the other residues defined in relation to the i position such as i+5. For DRB1*0401, the a tyrosine or phenylalanine in the i position gave the strongest binding while 6 of 9 variations from Y or F reducing the binding 100 fold. At the i+5 position, 5 of 8 derivatives reduced the binding from the original threonine. Other positions had limitations on which residues could be present but not clear requirements. The residues at i+3 and i+6 could not be positively charged, and i+8 could not have any charge.

The second method used to determine the allele specific motif for class II molecules involves methods identical to those presented in this thesis. MHC class II/peptide complexes are immunoaffinity purified, the peptides are acid eluted, and the peptide pool is fractionated. Individual peptides are sequenced by Edman degradation or by tandem mass spectrometry. Unlike the class I molecule, the ends of the antigen binding cleft are open in a class II molecule. This complicates the analysis. First, the open end allows the length of the N termini to vary, so the critical residues will occur at a variable number of residues from the amino termini. This means that sequencing the peptide pool in its entirety as was done for pools eluted from class I molecules will not provide any useful information. Second, the actual binding region of the class II molecule is no longer than the class I system, yet the peptides sequenced from class II molecules can be greater than 3 times longer than those found in class I molecules. The average peptide sequenced from *0404 was 15 residues long, and from *0402 it was 16 residues long. If only the core regions contained in all the peptides contributing to a nested cluster are considered, then the average length of each peptide from both molecules is 13. If the binding cleft accommodates a linear portion of the peptide 9 residues in length, then for a 13 residue peptide there are 5 different orientations or registers. For example, the fragment from 1 to 9 may be in the cleft, or perhaps the region of 3 to 11 is. Without the A pocket or F pocket fixing an end point as in class I,

there is no way from the sequence data to determine which 9 residue stretch actually interacts with the binding groove.

The problem of variable registers could be overcome by sequencing more peptides. With a larger database of peptides to analyze, the consistent pattern of the allele specific motif should still become apparent. Unfortunately, nested clusters have been found to dominate the peptide pool. A nested cluster complicates the analysis because all the peptides have the same core region. The variations at the ends of the cluster may help eliminate some possible registers. As an example, the longest peptide sequenced from *0402 contained 20 residues which has 12 different sections 9 amino acids long. However, other peptides in the cluster are shorter and indicate that the core region is only 13 residues. While this reduces the possible registers from 12 to 5, it is still too large for a simple comparison. In the class I system where nearly all the peptides reported are unique, a motif has been determined from 8 to 10 peptides. Sequencing 10 peptides from *0404 and 8 from *0402 yielded only 3 unique binding regions for each molecule. The presence of nested clusters effectively reduces the number of unique peptides and compounds the problem of determining the proper register.

Despite the complications, sequenced peptides have lead to proposed allele specific motifs for DRB1*0101 (80), DRB1*0301, DRB1*0401, DRB1*0701, DRB1*0801, and DRB1*1501 (82). For example, 39 peptides were sequenced from DRB1*0401, but because of the nested clusters, there was only 8 unique core regions used for analysis (82). The motif generated was the position of I being, F, L or V; followed by i+5 being N, Q, S, or T. Despite the variability of this motif, only 7 of 8 core regions completely matched. However, this motif fits well with one proposed from binding assays using HA derivatives. The HA assays also had limitations on residues at i+3, i+6, and i+8, but 7 of 8 sequenced peptides conform to these

limitations. However, at the i position the HA assays suggested tyrosine was very important, but the elution study did not. At the i+5 position, the HA assays proposed a small hydrogen bond donator (S or T) or a hydrophobic residue (V, L, I, or M) but the eluted peptides suggested a wider range of hydrogen bond donors, (Q, N, S, or T), and no hydrophobic residues.

Generating allele specific motifs for the class II molecules is a technically demanding process. No motif presented for a class II molecule has had the certainty of the class I motifs. The conformation and the formation of the pockets in the class I and class II binding clefts are extremely similar suggesting that a clear motif does exists for the class II molecules. However, the open ends with the variability in peptide length coupled with the presence of nested clusters that reduce the apparent number of unique core regions complicate the analysis of the class II motif.

Different Peptides Derived from Homologous Proteins

There is another way to compare the antigen binding properties of two alleles. The class I system has shown some polymorphism in genes necessary for antigen processing, such as the TAP genes, but the variations are limited. Using the class I system as a model suggests that any polymorphism present in the class II processing pathway would be equally limited. Therefore, the same antigen processed by two different APC that express the same class II molecule would be expected to generate the same peptide. Any difference in the peptide selected for presentation would be dependent on the peptide interactions with the class II molecule and not the processing machinery. Since antigen processing can be influenced by the mechanism for delivering the antigens into the appropriate cell compartment and by the conformation of the antigen, comparisons of this nature would require the exact same antigen or antigens

that had the same collection mechanism, the same three dimensional conformation, as well as a homologous peptide sequence. The peptides sequenced from class II molecules show a preponderance of nested clusters derived from class I molecules. The class I molecules meet the criteria for similarity needed to compare the peptides selected by one class II allele to those chosen by another. The class I molecules have extremely similar three dimensional structures. This suggests that any mechanism for bringing the class I molecule into the CPL for antigen processing would be the same for all class I molecules. It also suggests any processing machinery would have to accommodate the same secondary and tertiary structure. The class I molecules are very polymorphic, but the polymorphism falls in hypervariable regions, and peptides generated outside these regions would have very homologous sequences. The comparison of *0402 and *0404 derived peptides from class I molecules performed here meets the criteria and the model.

This method of comparing the peptides selected from class I molecules by DRB1*0404 and DRB1*0402 is very revealing, refer to table 7. The predominant nested clusters found in DRB1*0404 are derived from the HLA-B and C class I molecules, stretching from residues 1-18 but with a core region of 2-14. These two sequence are 86% homologous. The sequence from the same region of the B38 molecule expressed by the line used for the *0402 analysis is 93% homologous to the HLA-C peptide. However, the 1-18 peptide was not found not found in the *0402 molecule. The dominant nested cluster found in *0402 is derived from 103-123 of the B38 molecule with a core region of 107-119. The homology between this region and the identical region in the B40 molecule from the *0404 cell line is 95%. The only difference is highly conservative with 116 of the B38 having a F, and the B40 having a Y. Despite the similarity between these two peptides the 103-123 region is not found in the *0404 allele. HLA class I molecules that have the same tertiary structure and

nearly homologous primary protein sequences in the epitope regions, yet donate very different peptides to the *0404 and *0402 class II molecules indicating that these two molecules select very different peptides for presentation.

Implications for Rheumatoid Arthritis

The DRB1*0404 and DRB1*0402 molecules have a very closely related sequence, but demonstrate very different disease susceptibilities and select different peptides for presentation. A similar association exists in an animal model. The B6.C-H-2^{bm12} (bm-12) mouse differs from the parental strain, C57BL/6, by three amino acids in the I-Aβ chain, which is the murine equivalent of HLA-DRB. The three residues correspond to 67, 71, and 72 of the β chain which places them at the same location in the antigen binding cleft as the differences between *0404 and *0402 (105). Like the *0404/*0402 variations, position 67 is a relatively conservative I to F change, but 71 and 72 are less conservative, R to Q and T to K, respectively. If these two strains are immunized with the acetylcholine receptor from *Torpedo californica* (AChR), the parental strain develops experimental autoimmune myasthenia gravis, but the bm12 strain does not (106, 107). Both strains recognize AChR, demonstrating both $CD4^+$ T cell responses and anti-AChR antibody production (107). However, the parental strain recognizes an epitope contained in the peptides 150-169 (107) and 141-162 (106) which the bm12 strain does not. Even when immunized with this peptide, the bm12 mouse lacks a response. Further analysis has shown that these two strains respond differently to many antigens such as insulin, hemoglobin, and H-Y antigen (108). Even though this analysis uses the T cell response, it indicates that the three residue difference in the class II binding cleft leads to different epitopes being selected and recognized.

The bm12 model and the data presented of *0404 and *0402 peptide selection offer a simplistic explanation for the different disease susceptibilities seen in RA. If an individual expresses *0404 and is exposed to some antigen, the antigen will be processed and presented in the class II binding cleft to T cells. This will lead to an oligoclonal T cell proliferation with T cells of various effector functions. This thesis and the work of others has shown that the class II molecules also present self antigens. If a self antigen is similar enough to the foreign antigen, some of the T cells may cross react with it. A person who expressed *0402 would select a different epitope on both the foreign antigen as well as the self antigen. It would be unlikely that the two different antigens would lead to development of the same cross reactive T cell population. Since T cell activation is dependent on the amount of peptide/MHC complex on the surface of the APC, an individual that expressed two DRB1*0404 alleles and no other DRB1 genes to compete for epitopes would be more likely to reach the threshold of T cell activation. Homozygosity may lead to more MHC/peptide complexes capable of stimulating the cross reactive T cells at tissue sites that have a lower concentration of the self antigen. Hence, a homozygous patient would be expected to have greater activation and destruction at the common sites, diarticular joints, and an increased chance of activation at extraarticular sites such as RA nodules, heart, and lung.

Summary and Conclusions

The chromatogram from the reversed phase separations of peptides demonstrated that DRB1*0102, DRB1*0401, DRB1*0402 and DRB1*0404 all bound peptides with similar hydrophobic properties. However, each HLA-DR allele bound peptides that gave a unique chromatogram. A comparison of profile generated from the separation of peptides eluted from DRB1*0401 molecules harvested from different cell lines demonstrated that unique chromatography were dependent on the genetic background of the cell.

Further analysis of the peptides eluted from DRB1*0404 and DRB1*0402 demonstrated that these molecules contained a diverse repertoire of peptides. The array of peptides eluted from each of the class II molecules was dominated by a few unique core regions making a nested cluster. The nested clusters eluted from both HLA-DR alleles were derived from the class I proteins expressed by the respective cell line. The over representation of these peptides in the eluted pool suggest they have some important role.

A comparison the nested clusters demonstrated that they were from regions of the class I molecule that contained relatively conserved sequences. Sequences similar to the nested clusters eluted from DRB1*0404 were found in the class I molecules expressed by the cell line used for DRB1*0402 production. However, these sequences were not found bound to DRB1*0402 molecule. Conversely, the sequences that made up the nested cluster found in DRB1*0402 were similar to regions in the class I molecules expressed by the cell line used for DRB1*0404 production. Theses sequences were not found in DRB1*0404. The class I molecules are very similar in intracellular domains and three dimensional structure, characteristics that might influence antigen processing. Since the self proteins are so homologous, this strongly implies that these two different HLA-DR alleles bind different processed peptides.

Rheumatoid arthritis is associated with HLA-DR alleles that contain the sequence QRRAA or QKRAA at the 70-74 position in the β chain. This sequence is located between the D and E pockets in the antigen binding cleft of the HLA-DR molecule. The antigen binding cleft is created by the relatively non-polymophic α chain and the β chain. With the critical sequence for RA and the α chain conserved across all RA associated DR alleles, the peptide binding characteristics of these molecules should be quite similar, and they may be able to present the same peptides. A closely related allele that contains a different sequence at 70-74, DERAA, is not associated with the disease. If the model of RA being caused or perpetuated by an antigen presented in an associated allele is correct, then a non-RA associated allele would not be expected to present the same peptide. This thesis has demonstrated that an RA associated allele and the closely related non-RA associated allele naturally select different peptides from nearly identical parent proteins which strongly supports the proposed model that RA is a disease initiated or maintained by an antigenic challenge.

References

1. **Arnett, F. C., S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fires, N. S. Cooper, L. A. Healey, S. R. Kaplan, M. H. Liang, H. S. Luthra, T. A. J. Medsger, D. M. Mitchell, D. H. Neustadt, R. S. Pinals, J. G. Schaller, J. T. Sharp, R. L. Wilder, and G. G. Hunder**. 1988. The american rheumatism association 1987 revised criteria for the classificaion of rheumatoid arthritis. *Arth. Rheum. 31(3):315.*

2. **Schumacher, H. R., Jr., J. H. Klippel, and D. R. Robinson, eds**. 1988. *Primer on the Rheumatic Diseases, 9th ed.* Arthritis Foundation, Atlanta, GA.

3. **Harris, E. D. J.** 1993. Clinical Features of Rheumatoid Arthritis. In *Textbook of Rheumatology, 4th ed., vol. 1.* W. N. Kelley, E. D. J. Harris, S. Ruddy, and C. B. Sledge, eds. WB Saunders Company, Philadephia, p. 874.

4. **Heath, C. W. J., and P. R. Fortin**. 1992. Epidemiologic studies of rheumatoid arthritis: future directions. *J. Rheumatol. 19(Suppl. 32):74.*

5. **Hochberg, M. C., and T. D. Spector**. 1990. Epidemiology of rheumatoid arthritis: update. *Epidemiologic Reviews 12:247.*

6. **Deighton, C. M., J. Wentzel, G. Cavanagh, D. F. Roberts, and D. J. Walker**. 1992. Contribution of inherited factors to rheumatoid arthritis. *Ann. Rheum. D. 51:182.*

7. **Deighton, C. M., and D. J. Walker**. 1991. The familial nature of rheumatoid arthritis. *Ann. Rheum. D. 50:62.*

8. **Bodmer, J. G., S. G. E. Marsh, E. D. Albert, W. F. Bodmer, B. Dupont, H. A. Erlich, B. Mach, W. R. Mayr, P. Parham, T. Sasazuki, G. M. T. Schreuder, J. L. Strominger, A. Svejgaard, and P. I. Terasaki**. 1992. Nomenclature for factors of the HLA system, 1991. *Human Immun. 34:4.*

9. **Gregersen, P. K., J. Silver, and R. J. Winchester**. 1987. The shared epitope hypothesis. *Arth. Rheum. 30(11):1205.*

10. **Wordsworth, B. P., J. S. S. Lanchbury, L. I. Sakkas, K. I. Welsh, G. S. Panayi, and J. I. Bell**. 1989. HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc. Natl. Acad. Sci. USA 86:10049.*

11. **Nepom, G. T., P. Byers, C. Seyfried, L. A. Healey, K. R. Wilske, D. Stage, and B. S. Nepom**. 1989. HLA genes associated with rheumatoid arthritis. *Arth. Rheum. 32(1):15.*

12. **Hiraiwa, A., K. Yamanaka, W. W. Kwok, E. M. Mickelson, S. Masewicz, J. A. Hansen, S. F. Radka, and G. T. Nepom**. 1990. Structural requrements for the recognition of the HLA-Dw14 class II epitope: a key HLA determinant associated with rheumatoid arthritis. *Proc. Natl. Acad. Sci. USA 87:8051.*

13. **Woodrow, J. C., F. E. Nichol, and G. Zaphiropoulos**. 1981. DR antigens and rheumatoid arthritis: a study of two populations. *Br. Med. J. 283:1287.*

14. **Willkens, R. F., G. T. Nepom, C. R. Marks, J. W. Nettles, and B. S. Nepom**. 1991. Association of HLA-Dw16 with rheumatoid arthritis in Yakima Indians. *Arth. Rheum. 34(1):43.*

15. **Watanabe, Y., K. Tokunaga, K. Matsuki, F. Takeuchi, K. Matsuta, H. Maeda, K. Omoto, and T. Juji**. 1989. Putative amino acid sequence of HLA-DRB chain contributing to rheumatoid arthritis susceptibility. *J. Exp. Med. 169:2263.*

16. **Van Zeben, D., J. M. W. Hazes, A. H. Zwinderman, A. Cats, G. M. T. Schreuder, J. D'Amaro, and F. C. Breedveld**. 1991. Association of HLA-DR4 with a more progressive disease course in patients with rheumatoid arthritis. *Arth. Rheum. 34(7):822.*

17. **McCusker, C. T., B. Reid, D. Green, D. D. Gladman, W. W. Buchanan, and D. P. Singal**. 1991. HLA-D region antigens in patients with rheumatoid arthritis. *Arth. Rheum. 34(2):192.*

18. **Zoschke, D., and M. Segall**. 1986. Dw subtypes of DR4 in rheumatoid arthritis: evidence for a preferential association with Dw4. *Human Immun. 15:118.*

19. **Weyand, C. M., C. Xie, and J. J. Goronzy**. 1992. Homozygosity for the HLA-DRB1 allele selects for extraarticular manifestations in rheumatoid arthritis. *Journal of Clinical Investigation 89:2033.*

20. **Weyand, C. M., K. C. Hicok, D. L. Conn, and J. J. Goronzy**. 1992. The influenced of HLA-DRB1 genes on disease severity in rheumatoid arthritis. *Annals of Internal Medicine 117(10):801.*

21. **Hansen, T. H., B. M. Carreno, and D. H. Sach**. 1993. The major histocompatibility complex. In *Fundamental Immunology, 3rd ed.* W. E. Paul, ed. Raven Press, New York, p. 577.

22. **Who Nomenclature Committee**. 1992. Nomenclature for factors of the HLA system, 1991. *Immunogenet. 36:135.*

23. **Howard, M. C., A. Miyajima, and R. Coffman**. 1993. T-cell-derived cytokines and their receptors. In *Fundamental Immunology, 3rd ed.* W. E. Paul, ed. Raven Press, Ltd., New York, p. 763.

24. **Cush, J. J., and P. E. Lipsky**. 1988. Phenotypic analysis of synovial tissue and peripheral blood lymphocytes isolated from patients with rheumatoid arthritis. *Arth. Rheum. 31(10):1230.*

25. **van Dinther-Janssen, A. C. H. M., S. T. Pals, R. Scheper, F. Breedveld, and C. J. L. M. Meijer**. 1990. Dendritic cells and high endothelial venules in the rheumatoid synovial membrane. *J. Rheumatol. 17:11.*

26. **Miyasaka, N., K. Sato, M. Goto, M. Sasano, M. Natsuyama, K. Inoue, and K. Nishioka**. 1988. Augmented interleukin-1 production and HLA-DR

expression in the synovium of rheumatoid arthritis patients. Possible involvement in joint destruction. *Arth. Rheum. 31(4):480.*

27. **Nakao, H., K. Eguchi, A. Kawakami, K. Migita, T. Otsubo, U. Yukitaka, C. Shimomura, H. Tezuka, M. Matsunaga, K. Maeda, and S. Nagataki**. 1990. Phenotypic characterization of lymphocytes infiltrating synovial tissue from patients with rheumatoid arthritis: analysis of lymphoctyes isolated from minced synovial tissue by dual immunofluorescent staining. *J. Rheumatol. 17:142.*

28. **Haynes, B. F., B. J. Grover, L. P. Wichard, L. P. Hale, J. A. Nunley, D. E. McCollum, and K. H. Singer**. 1988. Synovial microenvironment-T cell interactions. *Arth. Rheum. 31(8):947.*

29. **Firestein, G. S., M. M. Paine, and B. H. Littman**. 1991. Gene expression (collagenase, tissue inhibitor of metalloporteinases, complement, and HLA-DR) in rheumatoid arthritis and osteoarthritis synovium quantitative analyiss and effect of intraarticular corticosteroids. *Arth. Rheum. 34(9):1094.*

30. **Germain, R. N.** 1994. MHC-dependent antigen processing and peptide presentation: providen ligands for T lymphocyte activation. *Cell 76:287.*

31. **Germain, R. N.** 1993. Antigen processing and presentation. In *Fundamental immunology, 3rd ed.* W. E. Paul, ed. Raven Press, Ltd, New York, p. 629.

32. **Brown, M. G., J. Driscoll, and J. J. Monaco**. 1991. Structural and serological similarity of MHC-linked LMP and proteasome (multicatalytic proteinase) complexes. *Nature 353:355.*

33. **Glynne, R., S. H. Powis, S. Beck, A. Kelty, L. Kerr, and J. Trowsdale**. 1991. A proteasome-related gene between the two ABC transporter loci in the class II region of the human MHC. *Nature 353:357.*

34. **Spies, T., M. Bresnahan, S. Bahram, D. Arnald, G. Blanck, E. Mellins, D. Pious, and R. DeMars**. 1990. A gene in the human major histocompatibility complex class II region controlling the class I antigen presentation pathway. *Nature 348:744.*

35. **Trowsdale, J., I. Hanson, I. Mockridge, S. Beck, A. Townsend, and A. Kelty**. 1990. Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters. *Nature 348:741.*

36. **Henderson, R. A., H. Michel, K. Sakaguchi, J. Shabanowitz, E. Appella, D. F. Hunt, and V. H. Engelhard**. 1992. HLA-A2.1-associated peptides from a mutant cell line: a second pathway of antigen presentation. *Science 255:1264.*

37. **Tulp, A., D. Verwoerd, B. Dobberstein, H. L. Ploegh, and J. Pieters**. 1994. Isolation and characterization of the intracellular MHC class II compartment. *Nature 369:120.*

38. **Newcomb, J. R., and P. Cresswell**. 1993. Characterization of endogenous peptides bound to purified HLA-DR molecules and their absence from invariant chain-associated Âß dimers. *J Immunol. 150(2):499.*

39. **Bakke, O., and B. Dobberstein**. 1990. MHC class II-associated invariant chain contains a sorting signal for endosomal compartments. *Cell 63:707.*

40. **Schmid, S. L., and M. R. Jackson**. 1994. Making class II presentable. *Nature 369:103.*

41. **Germain, R. N., and D. H. Margulies**. 1993. The biochemistry and cell biology of antigen processing and presentation. *Annu. Rev. Immunol. 11:403.*

42. **Lanzavecchia**. 1985. Antigen-specific interaction between T and B cells. *Nature 314:537.*

43. **Unanue, E. R.** 1993. Macrophages, antigen-presenting cells, and the phenomena of antigen handling and presention. In *Fundamental Immunology, 3rd ed.* W. E. Paul, ed. Raven Press, Ltd., New York, p. 111.

44. **Chu, C. T., and S. V. Pizzo**. 1993. Receptor-mediated antigen delivery in to macrophages complexing antigen to $\hat{A}_2$-macroglobulin enhances presentation to T cells. *J Immunol. 150(1):48.*

45. **Morris, P., J. Shaman, M. Attaya, M. Amaya, S. Goodman, C. Bergman, J. J. Monaco, and E. Mellins**. 1994. An essential role for HLA-DM in antigen presentation by class II major histocompatibiltiy molecules. *Nature 368:551.*

46. **Fling, S. P., B. Arp, and D. Pious**. 1994. HLA-DMA and DMB genes are both required for MHC class II/peptide complex formation in antigen-presenting cells. *Nature 368:554.*

47. **Lanzavecchia, A.** 1990. Receptor-mediated antigen uptake and its effect on antigen presentation to class II-restricted T lymphocytes. *Annu. Rev. Immunol. 8:773.*

48. **Gammon, G., J. Klotz, D. Ando, and E. E. Sercarz**. 1990. The T cell repertoire to a multideterminant antigen. Clonal heterogeneity of the T cell response, variation between syneneic individuals, and in vitro selection of T cell specificities. *J Immunol. 144(5):1571.*

49. **Rouas, N., S. Christophe, F. Housseau, D. Bellet, J.-G. Guillet, and J.-M. Bidart**. 1993. Influence of protein-quaternary structure on antigen processing. *J Immunol. 150(3):782.*

50. **Tampé, R., and H. M. McConnell**. 1991. Kinetics of antigenic peptide binding to the class II major histocompatibility molecule I-A^d. *Proc. Natl. Acad. Sci. USA 88:4661.*

51. **Stern, L. J., and D. C. Wiley**. 1992. The human class II MHC protein HLA-DR1 assembles as empty αß heterodimers in the absence of antigenic peptides. *Cell 68:465.*

52. **Madden, D. R., H. C. Gorga, J. L. Strominger, and D. C. Wiley**. 1992. The three-dimensional strucute of HLA-B27 at 2.1 Å resolution suggest a general mechanism for tight peptide binding to MHC. *Cell 70:1035.*

53. **Bjorkman, P. J., M. A. Saper, B. Samraoui, W. S. Bennett, J. L. Strominger, and D. C. Wiley**. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature 329:506.*

54. **Garrett, T. P. J., M. A. Saper, P. J. Bjorkman, J. L. Strominger, and D. C. Wiley**. 1989. Specificty pockets for the side chains of peptide antigens in HLA-Aw68. *Nature 342:692.*

55. **Matsumura, M., D. H. Fremont, P. A. Peterson, and I. A. Wilson**. 1992. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science 257:927.*

56. **Zhang, W., A. C. M. Young, M. Imarai, and S. G. Nathenson**. 1992. Crystal structure of the major histocompatibility complex class I H-2K^{b} molecule containing a single viral peptide: implications for peptide binding and T-cell receptor recognition. *Proc. Natl. Acad. Sci. USA 89:8403.*

57. **Madden, D. R., J. C. Gorga, J. L. Strominger, and D. C. Wiley**. 1991. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature 353:321.*

58. **Buxton, S. E., R. J. Benjamin, C. Clayberger, P. Parham, and A. M. Krensky**. 1992. Anchoring pockets in human histocompatibility comples leukocyte antigen (HLA) class I molecules: analysis of the conserved B ("45") pocket of HLA-B27. *J. Exp. Med. 175:809.*

59. **Brown, J. H., T. Jardetzky, M. A. Saper, B. Samraoui, P. J. Bjorkman, and D. C. Wiley**. 1988. A hypothetical model of the foreign antigen binding site of class II histocompatibility molecules. *Nature 332:845.*

60. **Brown, J. H., T. S. Jardetzky, J. C. Gorga, L. J. Stern, R. G. Urban, J. L. Strominger, and D. C. Wiley**. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature 364:33.*

61. **Zeliszewski, D., J. J. Golvano, P. Gaudebout, I. Dorval, C. Freidel, L. Gebuhrer, H. Betuel, Borras-Cuesta, and G. Sterkers**. 1993. Implication of HLA-DR residues at positions 67, 71, and 86 in interaction between HLA-DR11 and peptide HA306-320. *J Immunol. 151(11):6237.*

62. **Petersdorf, E. W., A. G. Smith, E. M. Mickelson, P. J. Martin, and J. A. Hansen**. 1991. Ten HLA-DR4 alleles defined by sequence polymorphisms within the DRB1 first domain. *Immunogenet. 33:267.*

63. **Falk, K., O. Rötzschke, S. Stevanovic, J. Günther, and H.-G. Rammensee**. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature 351:290.*

64. **Hill, A. V. S., J. Elvin, A. C. Willis, M. Aidoo, C. E. M. Allsopp, F. M. Gotch, X. M. Gao, M. Takiguchi, B. M. Greenwood, A. R. M. Townsend, A. J. McMichael, and H. C. Whittle**. 1992. Molecular analysis of the association of HLA-B53 and resistance to severe malaria. *Nature 360:434.*

65. **Sutton, J., S. Rowland-Jones, W. Rosenberg, D. Nixon, F. Gotch, X. M. Gao, N. Murray, A. Spoonas, P. Driscoll, M. Smith, A. Willis, and A. McMichael**. 1993. A sequence pattern for peptides presented to cytotoxic T lymphocytes by HLA B8 revealed by analysis of epitopes and eluted peptides. *Eur. J. Immun. 23:447.*

66. **Van Bleek, G. M., and S. G. Nathenson**. 1990. Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule. *Nature 348:213.*

67. **Jardetzky, T. S., W. S. Lane, R. A. Robinson, D. R. Madden, and D. C. Wiley**. 1991. Identification of self peptides bound to purified HLA-B27. *Nature 353:326.*

68. **DiBrino, M., K. C. Parker, J. Siloach, M. Knierman, J. Lukszo, R. V. Turner, W. E. Biddison, and J. E. Coligan**. 1993. Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proc. Natl. Acad. Sci. USA 90:1508.*

69. **Hunt, D. F., R. A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella, and V. H. Engelhard**. 1992. Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science 255:1261.*

70. **Rosloniec, E. F., K. S. Beard, and J. H. Freed**. 1993. Functional analysis of the antigen binding region on an MHC class II molecule. *Mol. Immunol. 30(5):491.*

71. **Hunt, D. F., H. Michel, T. A. Dickinson, J. Shabanowitz, A. L. Cox, K. Sakaguchi, E. Appella, H. M. Grey, and A. Sette**. 1992. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-A^d. *Science 256:1817.*

72. **Huczko, E. L., W. M. Bodnar, D. Benjamin, K. Sakaguchi, N. A. Zhu, J. Shabanowitz, R. A. Henderson, E. Appella, D. F. Hunt, and V. H. Engelhard**. 1993. Characteristics of endogenous peptides eluted from the class I MHC molecule HLA-B7 determined by mass spectrometry and computer modeling. *J Immunol. 151(5):2572.*

73. **Sette, A., S. Buus, S. Colon, J. A. Smith, C. Miles, and H. M. Grey**. 1987. Structural characteristics of an antigen required for its interaction with Ia and recognition by T cells. *Nature 328:395.*

74. **Sette, A., L. Adorini, E. Appella, S. M. Colón, C. Miles, S. Tanaka, C. Ehrhardt, G. Doria, Z. A. Nagy, S. Buus, and H. M. Grey**. 1989. Structural requirements for the interaction between peptide antigens and I-E^d molecules. *J Immunol. 143(10):3289.*

75. **O'Sullivan, D., T. Arrhenius, J. Sidney, M. Del Guercio, M. Albertson, M. Wall, C. Oseroff, S. Southwood, S. M. Colón, F. C. A. Gaeta, and A. Sette**. 1991. On the interaction of promiscuous antigenic peptides with different DR alleles. *J Immunol. 147(8):2663.*

76. **Hill, C. M., J. D. Hayball, A. A. Allison, and J. B. Rothbard**. 1991. Conformational and structural characteristics of peptides binding to HLA-DR molecules. *J Immunol. 147(1):189.*

77. **Leighton, J., A. Sette, J. Sidney, E. Appella, C. Ehrhardt, S. Fuch, and L. Adorini**. 1991. Comparison of structural requirements for interaction of the same peptide with I-E^k and I-E^d molecules in the activation of MHC class II-restricted T cells. *J Immunol. 147(1):198.*

78. **Sidney, J., C. Oseroff, S. Southwind, M. Wall, G. Ishioda, F. Koning, and A. Sette**. 1992. DRB1*0301 molecules recoginize a structural motif distinct from the one recognized by most $DR\beta_1$ alleles. *J Immunol. 149(8):2634.*

79. **Sette, A., J. Sidney, C. Oseroff, M.-F. del Guercio, S. Southwood, T. Arrhenius, M. F. Powell, S. M. Colon, F. C. A. Gaeta, and H. M. Grey**. 1993. HLA DR4w4-binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide-DR interactions. *J Immunol. 151(6):3163.*

80. **Chicz, R. M., R. G. Urban, W. S. Lane, H. C. Gorga, L. J. Stern, D. A. A. Vignall, and J. L. Strominger**. 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature 358:764.*

81. **Rudensky, A. Y., P. Preston-Hurlburt, S.-C. Hong, A. Barlow, and C. A. J. Janeway**. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature 353:622.*

82. **Chicz, R. M., R. G. Urban, J. C. Gorga, D. A. A. Vignali, W. S. Lane, and J. L. Strominger**. 1993. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J. Exp. Med. 178:27.*

83. **Yates, J. R., A. L. McCormack, J. B. Hayden, and M. P. Davey**. In Press. Sequencing peptides derived from the class II major histocompatibility complex by tandem mass spectrometry. In *Cell biology: A laboratory handbook*. J. E. Celis, ed. Academic Press, San Diego, CA.

84. **Paulus, H. E., H. I. Machleder, S. Levine, D. T. Y. Yu, and N. S. DacDonald**. 1977. Lymphocyte involvement in rheumatoid arthritis: studies during thoracic duct drainage. *Arth. Rheum. 20(6):1249.*

85. **Ueo, T., Y. Tominaga, H. Ogawa, and T. Sakurami**. 1979. The effect of thoracic duct drainage on lymphocyte dynamics and clinical symptoms of patients with rheumatoid arthritis. *Arth. Rheum. 22(12):1405.*

86. **Kotzin, B. L., S. Strober, E. G. Engleman, A. Calin, R. T. Hoppem, G. S. Kansas, C. P. Terrell, and H. S. Kaplin**. 1981. Treatment of intractable rheumatoid arthritis with total lymphoid irradiation. *N. Eng. J. Med. 305(17):969.*

87. **Trentham, D. E., J. A. Belli, R. J. Anderson, J. A. Buckley, E. J. Goetzl, J. R. David, and K. F. Austen**. 1981. Clinical and immunologic effects of fractionated total lymphoid irradiation in refractory rheumatoid arthritis. *N. Eng. J. Med. 305(17):976.*

88. **Horneff, G., G. R. Burmester, F. Emmrich, and J. R. Kalden**. 1991. Treatment of rheumatoid arthritis with an anti-CD4 monoclonal antibody. *Arth. Rheum. 34(2):129.*

89. **Paliard, X., S. G. West, J. A. Lafferty, J. R. Clements, J. W. Kappler, P. Marrack, and B. L. Kotzin**. 1991. Evidence for the effects of a superantigen in rheumatoid arthritis. *Science 253:326.*

90. **Bucht, A., J. R. Oksenberg, S. Lindblad, A. Grönberg, L. Steinman, and L. Klareskog**. 1992. Characterization of T-cell receptor Âß repertoire in synovial tissue from different temporal phases of rheumatoid arthritis. *Sc. J. Immun. 35:159.*

91. **Miltenburg, A. M. M., J. M. Van Laar, M. R. Daha, R. R. P. De Vries, P. J. Van Den Elsen, and F. C. Breedveld**. 1990. Dominant T-cell receptor ß-chain gene rearrangements indicate clonal expansion in the rheumatoid joint. *Sc. J. Immun. 31:121.*

92. **Davey, M. P., and D. D. Munkirs**. 1993. Patterns of T-cell receptor variable ß gene expression by synovial fluid and peripheral blood T-cells in rheumatoid arthritis. *Clinical Imunology and Immunopathology 68(1):79.*

93. **Coligan, J. E., A. M. Kruisbeck, D. H. Margulies, E. M. Shevach, and W. Strober, eds**. 1992. *Current protocols in immunology*. Greene Publishing and Wiley-Interscience, New York.

94. **Buus, S., A. Sette, S. M. Colon, D. M. Jenis, and H. Grey**. 1986. Isolation and characterization of antigen-Ia complexes involved in T cell recognition. *Cell 47:1071.*

95. **Gao, X., M. Fernandex-Vina, W. Shumway, and P. Stastny**. 1990. DNA typing for class II HLA antigens with allele-specific or group-specific amplification. I. Typing for subsets of HLA-DR4. *Human Immun. 27:40.*

96. **Demotz, S., H. M. Grey, E. Appella, and A. Sette**. 1989. Characterization of a naturally processed MHC class II restricted T-cell determinant of hen egg lysozyme. *Nature 342:682.*

97. **Griffin, P. R., J. A. Coffman, L. E. Hood, and I. I. I. Yates JR**. 1991. Structural analysis of proteins by HPLC-MS and HPLC-MS/MS using electrospray ionization on a triple quadrupole mass stectrometer. *Int. J. Mass Spectrom. Ion Processes 111:131.*

98. **Marrack, P., L. Ignatowicz, J. W. Kappler, J. Boymel, and J. H. Freed**. 1993. Comparison of peptides bound to spleen and thymus class II. *J. Exp. Med. 178:2173.*

99. **LiumZ, Y.-D. Sun, Y.-P. Xi, P. Harris, and N. Suciu-Fuca**. 1992. T cell recognition of self-human histocompatibility leukocyte antigens (HLA)-DR peptides in context of syngeneic HLA-DR molecules. *J. Exp. Med. 175:1663.*

100. **De Koster, H. S., D. C. Anderson, and A. Termijtelen**. 1989. T cells sensitized to synthetic HLA-DR3 peptide give evidence of continuous presentation of denatured HLA-DR3 molecules by HLA-DP. *J. Exp. Med. 169:1191.*

101. **Sprent, J.** 1993. T lymphocytes and the thymus. In *Fundamental Immunology, 3rd ed.* W. E. Paul, ed. Raven Press, Ltd., New York, p. 75.

102. **Rudensky, Y. A., S. Rath, P. Preston-Hurlburt, D. B. Murphy, and C. A. J. Janeway**. 1991. On the complexity of self. *Nature 353:660.*

103. **Murphy, D. B., D. Lo, S. Rath, R. L. Brinster, R. A. Flavell, A. Slanetz, and C. A. J. Janeway**. 1989. A novel MHC class epitope expressed in thymic medulla but not cortex. *Nature 338:765.*

104. **O'Sullivan, D., J. Sidney, E. Appella, L. Walker, L. Phillips, S. Colón, C. Miles, R. W. Chesnut, and A. Sette**. 1990. Characterization of the specificity of peptide binding to four DR haplotypes. *J Immunol. 145(6):1799.*

105. **McIntyre, K. R., and J. G. Seidman**. 1984. Nucleotide sequence of mutant *I-Aß*$^{bm-12}$ gene is evidence for genetic exchange between mouse immune response genes. *Nature 308:551.*

106. **Infante, A. J., P. A. Thompson, K. A. Krolick, and K. A. Wall**. 1991. Determinant selection in murine experimental autoimmune myasthenia gravis effect of the bm12 mutation of T cell recognition of acetylcholine receptor epitopes. *J Immunol. 146(9):2977.*

107. **Bellone, M., N. Ostlie, S. Lei, X.-D. Wu, and B. M. Conti-Tronconi**. 1991. The I-A^{bm12} mutation, which confers resistance to experimental myasthenia gravis, drastically affects the epitope repertoire of murine $CD4^+$ cells sensitized to nicotinic acetylcholine receptor. *J Immunol. 147(5):1484.*

108. **Hansen, T. H., and H. Y. Tse**. 1987. Insights into immune-response gene function using an Ia mutant mouse strain. *CRC Crit. R. Immun. 7(3):169.*

Appendix

Properties of Class I and Class II MHC Molecules

Feature	Class I	Class II
Location of synthesis	Ribosomes on endoplasmic reticulum	Ribosomes on endoplasmic reticulum
Formation of dimer	heavy chaing and β2 microglobulin	α and β from HLA locus
Proteins that guide folding	Calnexin	Invariant Chain
Chaperone protein	not applicable (uses default pathway)	Invariant Chain
Source of antigen	Intracellular	Endocytosed extracellular, secreted, or membrane bound proteins
Site of proteolysis	Cytoplasm	Endosome, Compartment for Peptide Loading
Peptide transport	Transporter for Antigen Procesing (ATP dependent)	not applicable
Site of peptide loading	Endoplasmic Reticulum	Compartment for Peptide Loading
Shape of antigen binding pocket	Closed by A and F pockets	Open allowing long peptides
Lengths of processed peptides	9 amino acids (8-13 range)	12-18 amino acids (12-34 range)
Present to T cell	$CD8^+$	$CD4^+$

Amino Acid Codes

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan