Specificity and Promiscuity among Naturally Processed Peptides Bound to HLA-DR Alleles

By Roman M. Chicz, Robert G. Urban, Joan C. Gorga, Dario A. A. Vignali, William S. Lane,* and Jack L. Strominger

*From the Department of Biochemistry and Molecular Biology and the *Harvard Microchemistry Facility, Harvard University, Cambridge, Massachusetts 02138*

Summary

Naturally processed peptides were acid extracted from immunoaffinity-purified HLA-DR2, DR3, DR4, DR7, and DR8. Using the complementary techniques of mass spectrometry and Edman microsequencing, >200 unique peptide masses were identified from each allele, ranging from 1,200 to 4,000 daltons (10–34 residues in length), and a total of 201 peptide sequences were obtained. These peptides were derived from 66 different source proteins and represented sets nested at both the amino- and carboxy-terminal ends with an average length of 15–18 amino acids. Strikingly, most of the peptides (>85%) were derived from endogenous proteins that intersect the endocytic/class II pathway, even though class II molecules are thought to function mainly in the presentation of exogenous foreign peptide antigens. The predominant endogenous peptides were derived from major histocompatibility complex–related molecules. A few peptides derived from exogenous bovine serum proteins were also bound to every allele. Four prominent promiscuous self-peptide sets (capable of binding to multiple HLA-DR alleles) as well as 84 allele-specific peptide sets were identified. Binding experiments confirmed that the promiscuous peptides have high affinity for the binding groove of all HLA-DR alleles examined. A potential physiologic role for these endogenous self-peptides as immunomodulators of the cellular immune response is discussed.

MHC class I and II molecules are membrane-bound glycoproteins that present processed antigen to T cells and initiate an immune response (1). Crystallographic analysis of several class I molecules identified a groove composed of two α helices supported by an eight-strand β-pleated sheet containing electron-dense material that represents bound antigenic peptide (2–4). Several groups have characterized the complex mixtures of acid-extracted class I–bound peptides by HPLC fractionation and sequencing (5–9). The majority of these peptides were 8–11 amino acids long and possessed a binding motif characteristic of peptides that bind to a given class I allele.

The characterization of naturally processed peptides bound to class II molecules provides an approach towards understanding both antigen processing and peptide binding events in vivo. The stability of class II molecules requires peptide binding (10, 11); however, the precise class II molecule–peptide contacts that provide this energy are not yet well defined. Identification of naturally processed peptides extracted and sequenced from class II molecules revealed that the bound peptides were longer (13–25 residues) than those bound to class I (12–15) and nested at the amino- and/or carboxy-terminal ends, suggesting that the peptide binding groove on class II molecules is open at both ends (13, 15). Although only a limited number of source proteins were reported, peptides derived from both endogenous proteins and exogenous serum proteins were identified. The association constants (measured by competitive inhibition) for several of these peptides were in the nanomolar range, confirming the high affinity of these peptides for class II molecules.

To further our understanding of class II antigen processing and peptide binding, several different allelic forms of HLA-DR molecules were purified and the bound peptides were identified using the complementary techniques of mass spectrometry and Edman microsequencing. Each allele has a complex, yet readily distinguishable and reproducible peptide profile. The range of peptide masses is comparable among the alleles examined. To establish a representative collection of bound peptides and to identify their respective source proteins, individual fractions corresponding to both major and minor chromatographic peaks from the extracted peptide pool separations were chosen for sequence analysis. Various features relating to the allelic specificity and promiscuity of naturally processed peptides bound to different DR alleles will be described.

Materials and Methods

Reagents and Materials. HLA-DR molecules were purified from homozygous EBV-transformed human B lymphocyte lines:

J. Exp. Med. © The Rockefeller University Press • 0022-1007/93/07/0027/21 $2.00
Volume 178 July 1993 27–47

DR1 (DRB1*0101) from LG-2 and HOM-2 cells, DR2b/DR2a (DRB1*1501/DRB5*0101) from MST cells, DR3/DRw52 (DRB1*0301/DRB3*0201) from WT20 cells, DR4/DRw53 (DRB1*0401/DRB4*0101) from Priess cells, DR7/DRw53 (DRB1*0701/DRB4*0101) from Mann cells and DR8 (DRB1*0801) from 23.1 cells. DR1 was also purified from the MHC antigen-loss mutant cell lines 721.45 (kindly provided by Robert DeMars, University of Wisconsin, Madison, WI) and 721.221 (CRL 1855; American Type Culture Collection, Rockville, MD). The cells were grown in RPMI 1640 supplemented with 10% FCS, 2 mM glutamine, 50 U/ml penicillin G, and 50 μg/ml streptomycin in spinner flasks or roller bottles and stored as pellets at −80°C.

The anti-HLA-DR hybridoma LB3.1 (IgG2b) was produced in this laboratory (16). The anti-HLA-DRw53 mAb 109d6 (17) was kindly provided by R. Winchester (Columbia University, New York, NY).

Peptides representing residues 97–120 of the invariant chain (Ii) (LPKPPKPVSKMRMATPLLMQALPM), residues 182–198 of HLA-DR α chain (APSPLPETTENVVCALG), residues 188–202 of the Ig κ chain C region (KHKVYACEVTHQGLS), residues 103–117 of the HLA-A2-like sequence (VGSDWRFLRGYHQYA), and residues 1273–1291 of bovine apolipoprotein B-100 (IPDNLFLKSDGRIKYTLNK) were synthesized on a peptide synthesizer (430; Applied Biosystems, Inc., Foster City, CA) using solid-phase Fmoc/HBTU chemistry. Biotinylations were done on resin-bound free amino-terminal peptides. Briefly, a fivefold molar excess of *N*-hydroxysuccinimidyl-crosslinked biotin (Pierce Chemical Co., Rockford, IL) was added to a 5-mg/ml suspension of resin-bound peptide in imidazole-buffered DMSO and rocked at room temperature for 2 h. The slurry was then washed four times with 25 ml of DMSO followed by four washes with 50 ml HPLC-grade water, air dried, and stored at 4°C until deprotection/cleavage. All peptides were purified by HPLC (HPXL solvent delivery system with UV-M detector; Rainin Instrument Co., Inc., Woburn, MA) on a preparative C_{18} reversed-phase chromatography (RPC)[1] column (25 × 2.2 cm, 300 Å, 10–15 μm; Vydac, Hesperia, CA) in 0.06% TFA/water with a 20–80% 50-min linear 0.055% TFA/acetonitrile gradient. The integrity of the purified peptides was confirmed by amino acid analysis (420A/130A derivatizer/HPLC after hydrolysis with 6 N HCl for 24 h in vacuo; Applied Biosystems, Inc.) and mass spectrometry on a triple quadrupole mass spectrometer equipped with an electrospray ion source (ESI-MS) (TSQ 700; Finnigan MAT, San Jose, CA).

Protein Purification. Immunoaffinity purification of HLA-DR alleles was performed as reported (18) with minor modifications. The detergent-soluble lysates from Priess and Mann cells were loaded onto an immunoaffinity column prepared with the mAb 109d6 to remove the DRw53 (DRB4*0101) allele–linked molecules before passage through the LB3.1 immunoaffinity column. DRw52 (DRB3*0201) was not removed from the WT-20 lysates and is a minor contaminant (<10%) of the DR3 preparations. Similarly, the DR2 preparations contain both the major and minor isotypes: DR2a (DRB5*0101) and DR2b (DRB1*1501).

Water-soluble HLA-DR was produced by limited papain digestion, as prepared for the crystallization of HLA-DR molecules, then purified using gel filtration, concentrated by vacuum dialysis to ~10 mg/ml, and stored at 4°C (19). SDS-PAGE was performed by the method of Laemmli (20) with the single modification that samples were not boiled before analysis. Protein concentrations were determined by bicinchoninic acid assay (Pierce Chemical Co.) and/or quantitative amino acid analysis (Applied Biosystems, Inc.). Immediately before peptide extraction protein samples were further purified by HPLC (6000A solvent delivery system; Waters, Milford, MA) on a size-exclusion column (300 × 7.5 mm, 250 Å, 10 μm, TSK-3000SW; Toyo Soda, Philadelphia, PA) in 25 mM 2-[*N*-morpholino] ethanesulfonic acid, 150 mM NaCl, pH 6.5, at a flow rate of 1 ml/min and spin-concentrated using Centricon 10 ultrafiltration devices (Amicon, Danvers, MA) to remove any residual small molecular weight contaminants.

Peptide Extraction and Separation. All protein samples (1 mg) were spin-concentrated to a final volume between 50 and 100 μl using Centricon 10 ultrafiltration devices. Bound peptides were eluted from HLA-DR by addition of acetic acid (10%, 1 ml) and incubation at 70°C for 15 min. Isolation of peptide pools from HLA-DR was accomplished by ultrafiltration. The acid-denatured protein was retained in the Centricon 10 microconcentrators, while extracted peptides were collected from the flow-through. The peptide pools were vacuum concentrated to 50 μl in a Speed-Vac (Savant Instrs., Inc., Farmingdale, NY) before HPLC separation.

Peptides were separated by HPLC (Rainin Instrument Co. Inc.) on a microbore C_{18} RPC column (250 × 2.1 mm, 300 Å, 5 μm; Vydac) as previously described (15). Chromatographic analysis was monitored at multiple UV wavelengths simultaneously, permitting spectrophotometric evaluation of purified species before mass spectral and sequence analyses. Fractions were collected and stored at −20°C until subsequent analyses. A sample of 25 mM MES was processed in parallel from the extraction procedure through the HPLC analysis for each allele to control for possible contamination from reagents or materials.

Mass Spectrometry and Edman Sequence Analyses. RPC fractions were chosen for mass spectrometry analyses based on UV absorption profiles. Briefly, optimum fractions from the RPC separation were chosen based on differential UV absorbance at 210, 254, 277, and 292 nm, peak symmetry, and resolution (21). Selected fractions were further analyzed by matrix-assisted laser-desorption mass spectrometry (MALD-MS) to determine the individual mass values for the predominant peptides. 4% (4 μl) of the collected fractions was mixed with 1 μl of matrix (α-cyano-4-hydroxycinnamic acid, 10 mg/ml in 50% acetonitrile; Sigma Chemical Co., St. Louis, MO), applied to a gold plated sample disc, and dried. Mass determinations were made using a LASERMAT mass spectrometer (Finnigan MAT). Selected fractions were analyzed by sample infusion using a TSQ 700 triple quadrupole mass spectrometer.

Amino-terminal sequence analysis was accomplished by automated Edman microsequencing using a pulsed-liquid protein sequencer (477A; Applied Biosystems, Inc.). The resultant phenylthiohydantoin amino acid derivatives were identified manually using an on-line HPLC (120A; Applied Biosystems, Inc.). Initial sequence yields measured in picomoles were used to determine the total yield of peptides derived from single source proteins. Combining the mass spectral data with the sequencing analyses enables crucial verification of both the amino- and carboxy-terminal amino acids of peptides within a single sample. For example, given a low picomolar mixture of two theoretical peptides, SMILE and SMIL, the presence of the carboxy-terminally truncated peptide SMIL would not be apparent by conventional Edman microsequencing alone. The observed sequence data would be SMILE and any reduction in yield of the carboxy-terminal residue could be attributed to normal losses during procession of the chemistry. All identified peptides were aligned to regions of proteins stored in the database using the BLAST network at the National Center for Biotechnology Information (22).

Peptide Binding. The assay was performed as described previously (23, 24). Briefly, 5 × 10^6 cells were resuspended in 3%

[1] *Abbreviation used in this paper:* RPC, reversed-phase chromatography.

paraformaldehyde and incubated at room temperature for 10 min followed by a wash with 50 ml PBS. Residual paraformaldehyde was quenched by addition of 5 mM glycylglycine (incubated at room temperature for 5 min) and washed with another 50 ml PBS. 2×10^5 cells were next resuspended in 100 μl of complete medium, added to 100 μl PBS containing biotinylated peptide, and incubated at 37°C for 10 h. Cells were then washed and stained with FITC-streptavidin (4.22 μg/ml PBS, 0.1% BSA) at 4°C for 30 min. Each incubation was followed by two washes with 250 μl PBS, 0.1% BSA at 4°C. Stained cells were subjected to flow cytometry on a FACScan® analyzer (Becton Dickinson & Co., Mountain View, CA). To measure the relative amount of fluoresceinated streptavidin bound, the mean fluorescence of 10,000 stained cells was determined.

Comparison of different homozygous B lymphoblastoid cells for their relative peptide binding capacity was done by normalizing the class II expression levels using biotinylated LB3.1 mAb. Specificity was demonstrated by blocking using the L243 anti-HLA-DR mAb and/or by competition with nonbiotinylated peptide.

Results

Peptide Isolation from Immunoaffinity-purified HLA-DR Alleles

The bound peptide pools were released from 1 mg of DR2, DR3, DR4, DR7, and DR8 (~16 nmol) by acid elution with the total extracted peptide yields (~11–13 nmol, 70–80% recovery) for each DR allele similar to those previously described for DR1 (15). Moreover, purified DR molecules were stable α/β complexes (>95% in each case), as measured by SDS-PAGE, indicating that the starting material was fully loaded with peptide (10, 25, 26); no separate α and β chains were detectable in any preparation. Only single free α and β chains were observed after acid elution of peptides, suggesting the complete extraction of peptides (data not shown).

Separations of the respective peptide pools were achieved using RPC and the peptide profiles from multiple extraction/separations for each of the five alleles were highly reproducible. These profiles exhibit the complexity of the bound peptide pools and the differences highlight the selectivity of individual peptides for specific alleles (Fig. 1). The bound peptides from papain-digested and detergent-solubilized DR had identical HPLC profiles. To confirm that papain solubilization of HLA-DR did not generate or alter the bound peptide repertoire, the predominant peaks from both the detergent-soluble and papain-digested DR1 profiles were analyzed and found to contain peptides identical in mass and sequence (15).

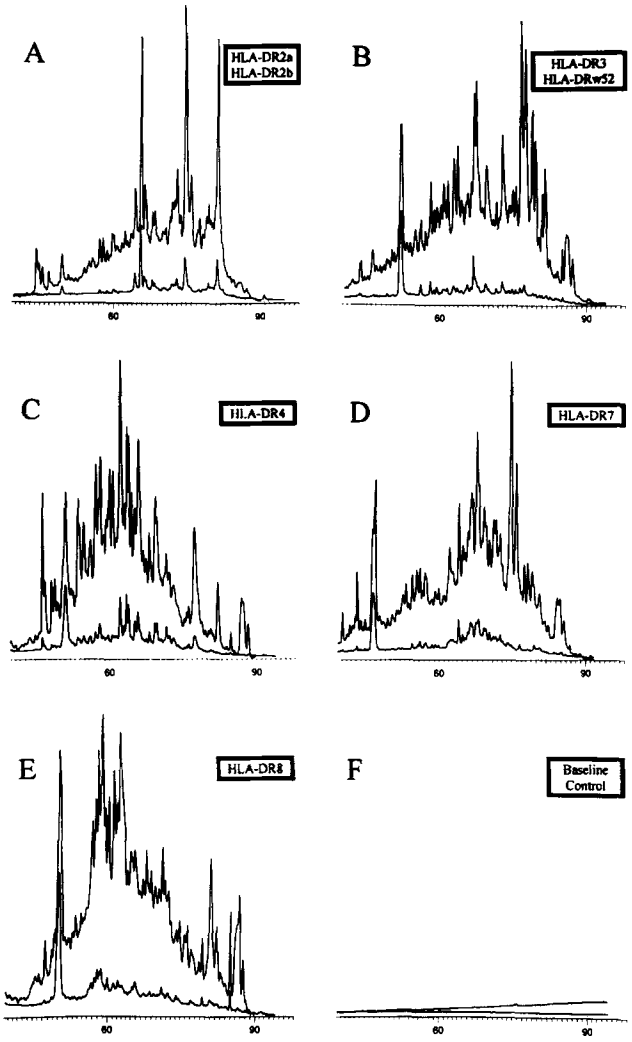

Figure 1. Reversed-phase HPLC separation profiles of peptide pools extracted from papain-digested HLA-DR. Each chromatogram represents the peptide repertoire as detected by UV absorbance for both 210 (sensitivity) and 277 nm (tryptophan/tyrosine content) at a full scale absorbance of 200 mAU. The profile patterns differ substantially, thus representing the various peptide fingerprints associated with each DR allele. (*A*) Peptides bound to DR2, including two prominent peaks at 75 and 81 min that were identified by sequence and mass analyses as amino-terminal fragments (papain overdigestion products) of the DR2a and DR2b β chain, respectively; (*B*) peptides bound to DR3; (*C*) peptides bound to DR4; (*D*) peptides bound to DR7; (*E*) peptides bound to DR8; (*F*) mock extraction buffers. The peak in *B–E* at ~50 min with a high 277-nm absorbance was identified as an organic contaminant from the Centricon ultrafiltration devices.

Size Distribution of Allele-specific Peptide Pools

Aliquots from the individual RPC separation fractions (20–200 fmol) were analyzed by MALD-MS and 1,101 unique masses were identified from the peptide pools extracted from the five DR alleles (an average of 220/allele). Many individual fractions contained multiple species of varying mass, indicating coelution of peptides in single chromatographic peaks, further corroborating the complexity of the separated mixtures. The molecular mass distribution and estimated peptide length (based on an amino acid average mass value of 118 daltons) varied from 1,239 to 4,091 daltons (10–34 residues in length) with the mode lying between 1,700 and 2,100 daltons (15–18 residues) for each allele (Fig. 2).

Identification and Sequence Characterization of Naturally Processed Bound Peptides

Complete sequences for 201 peptides isolated from five DR alleles were obtained by automated Edman microsequencing

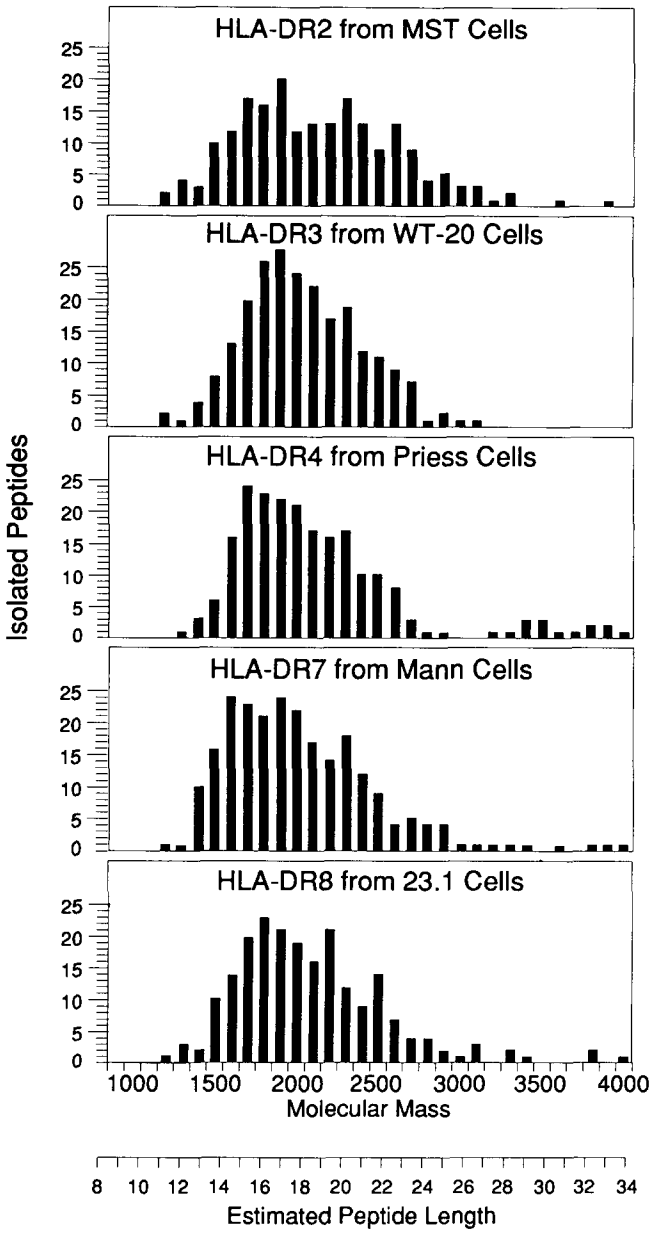

Figure 2. Size distribution of peptides bound to different HLA-DR alleles. A minimum aliquot (see text) was removed from each RPC fraction containing peptidic material and analyzed by MALD-MS. The molecular masses are plotted against the number of isolated peptides for each allele. Peptide length was estimated by dividing the experimental mass by an average amino acid mass of 118 daltons. The mass range varied between 1,200 and 4,100 daltons with the mode lying between 1,700 and 2,100 daltons of each allele.

in combination with mass spectrometry. Source proteins of the sequenced peptides were identified by homology search of the current database (22, 27). Over 29 peptides were characterized from each DR allele. Partial sequences for an additional 94–110 peptides from unknown source proteins were also obtained.

The predominant source proteins of the naturally processed peptides sequenced from the class II molecules of human B cells were MHC-related molecules. Self-peptides derived from MHC class I and/or II molecules were isolated from every class II allele, including peptides derived from the three class I isotypes (HLA-A, HLA-B, and HLA-C) as well as from the three class II isotypes (HLA-DR, HLA-DQ, and HLA-DP). From 201 sequenced peptides, 78 were derived from class I molecules, class II molecules, or the invariant chain (Ii). In addition, the relative abundance of these MHC-related self-peptides, as measured by initial yields from sequencing analyses, was the highest observed.

Peptides Identified in a Single Allele. Peptides derived from 44 integral membrane proteins (including two EBV-encoded proteins), four proteinase inhibitors, three cathepsins, six cytosolic proteins, and six bovine serum proteins representing 84 different core sequences were only detected in one allele. Peptides homologous with tubulin, an extracellular matrix–related protein not normally expressed in human lymphocytes but described as appearing on the cell surface of EBV transformed B cells (28), were also identified. Self-proteins thus contributed the large majority of bound peptides and were derived from both membrane proteins (which could have reached the endocytic pathway by endocytosis) and from cytosolic proteins (which must have intersected the endocytic/class II pathway in some other manner). Sets of peptides with common core sequences and nested at either the amino- and/or carboxy-terminal ends were observed for 38 of these peptides. Detailed descriptions of the identified peptides are listed by allele in Tables 1–5.

Additional peptides were derived from a limited set of secreted but not surface-expressed proteins, including endocytic processing enzymes and proteinase inhibitors that are present in endocytic or lysosomal vesicles. Five peptides were identified from the human cathepsins E and S and from a sequence homologous to rat cathepsin C (human cathepsin C has not been described). Peptides derived from two cytosol resident proteins implicated in intracellular transport and trafficking were also identified. A single peptide from the IFN-induced guanylate-binding protein (GBP-2), distinct from the Rab family of canonical regulatory heterotrimeric GTP-binding proteins (29), was identified, as was a nested set of peptides derived from Hsp 70. Guanylate-binding proteins can direct vesicular traffic by stimulating endosomal fusion, transduce hormonal and sensory signals across the plasma membrane, and act as initiation/elongation factors during protein synthesis (30, 31). The Hsp 70 source protein may be prp72/74 (a member of the Hsp 70 family with no sequence available from any databank), which has been implicated in the intracellular assembly of processed antigen–class II complexes (32). The other four peptides from cytosolic proteins were from cytochrome-b_5 reductase, EBV capsid protein, c-*myc*, and K-*ras*.

Peptides from exogenous bovine serum proteins were also isolated from each allele. 25 peptides representing 11 different core sequences that were homologous to human apolipoprotein B-100 were identified. The core sequences varied among alleles indicative of allelic specificity. Although the sequence for bovine apolipoprotein B-100 has not been described, this protein is the principal apolipoprotein in chylomicrons, very low density lipoprotein and low density lipoprotein in humans

and mice, all abundant in serum-supplemented culture media. Peptides derived from other proteins in bovine serum include fetuin (15), factor VIII, hemoglobin, complement C9, von Willebrand factor, and transferrin. With the exception of apolipoprotein B-100, the relative abundance of the serum-derived peptides was significantly lower than that observed for MHC-related self-peptides.

Sequence analyses of the peptides extracted from these five DR alleles confirm that class II antigen processing produces heterogeneous-sized peptides nested at both the amino- and carboxy-terminal ends. The characterization of potential enzymatic cleavage points was attempted by alignment of the flanking regions surrounding the identified peptide sequences. No common pattern was recognized, probably due to the exoproteolytic removal of amino acids from each end, and the shifting nature of the binding core sequences. However, a marked preference for heterogeneity, both in number of amino acids and total number of peptides, was observed at the carboxy-terminal end as compared with the amino-terminal end.

Promiscuous Self-peptides. The ability of individual peptides to bind to multiple DR alleles is defined as promiscuity (also referred to in the literature as degeneracy) and has been previously described for antigenic peptides presented to DR-restricted T cells (23, 33–35) as well as in in vitro binding experiments (36–40). 37 peptides derived from four promiscuous self-peptide families were identified (Table 6). These peptides were derived from Ii, an HLA-A2-like sequence (described below), HLA-DR α chain, and Ig κ chain. Each was represented by a nested set of peptides similar to those de-

Table 1. *Naturally Processed Peptides Bound to HLA-DR2*

Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	Observed	Yield*
					min	*m/z*	*m/z*	*pmol*
HLA-DQ α chain	Membrane	97–119	NIVIKRSNSTAATNEVPEVTVFS	23	52	2,477	2,478	–‡
		97–112	NIVIKRSNSTAATNEV	16	49	1,717	1,717	–‡
HLA-DQ β chain	Membrane	42–59	SDVGVYRAVTPQGRPDAE	18	49	1,917	1,920	14.0
		43–59	DVGVYRAVTPQGRPDAE	17	49	1,830	1,833	–§
		43–57	DVGVYRAVTPQGRPD	15	49	1,630	1,633	–§
HLA-DR2b β chain	Membrane	94–111	RVQPKVTVYPSKTQPLQH	18	48	2,107	2,114	12.5
		94–108	RVQPKVTVYPSKTQP	15	48	1,728	1,731	–§
FnR α chain	Membrane	586–616	LSPIHIALNFSLDPQAPVDSHGLR PALHYQ	30	71	3,308	3,313	11.1
K^+ channel protein	Membrane	173–190	DGILYYYQSGGRLRRPVN	18	71	2,127	2,133	–‡
		173-189	DGILYYYQSGGRLRRPV	17	71	2,013	2,018	–‡
Mannose binding protein	Membrane	174–193	IQNLIKEEAFLGITDEKTEG	20	80	2,248	2,248	–‡
MET	Membrane	59–81	EHHIFLGATNYIYVLNEEDLQKV	23	75	2,746	2,747	–‡
GBP-2	Cytosolic	434–450	QELKNKYYQVPRKGIQA	17	81	2,063	2,074	–‡
Apolipoprotein B-100‖	Exogenous	1200–1220	FPKSLHTYANILLDRRVPQTD	21	71	2,485	2,491	22.6
		1200–1218	FPKSLHTYANILLDRRVPQ	19	71	2,269	2,277	–§
Factor VIII‖	Exogenous	1775–1790	LWDYGMSSSPHVLRNR	16	53	1,918	1,922	–‡

Amino acid sequences (single-letter code) and mass determinations are shown for peptides isolated from HLA-DR2. All sequences were determined by automated Edman degradation using a protein sequencer (ABI 477A). RT, retention time from RPC separation; FnR, fibronectin receptor; MET, MET protooncogene member of the tyrosine kinase family of growth factor receptors; GBP-2, IFN-induced guanylate-binding protein 2. Protein category refers to the type of source protein; membrane refers to endogenous membrane associated/secretory proteins; cytosolic refers to endogenous resident cytosolic proteins; and exogenous refers to bovine serum proteins. $[M + H]^+$ refers to the calculated mass of the peptide in daltons, while the observed m/z refers to the mass in daltons determined by either a Finnigan TSQ 700 or LASERMAT mass spectrometer.

* Yield refers to initial sequence levels of individual peptides. Total yield of peptides from a single source protein is summed.

‡ A yield <8 pmol.

§ Yields for coeluting peptides with identical amino termini could not be distinguished.

‖ Proposed bovine serum sources with sequence homology to known human proteins.

Table 2. *Naturally Processed Peptides Bound to HLA-DR3*

Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	Obs.	Yield*
					min	*m/z*	*m/z*	*pmol*
HLA-A30¶	Membrane	28–?	VDDTQFVRFDSDAASQ . . .	ND	58	ND	ND	–‡
HLA-DR α chain	Membrane	111–129	PPEVTVLTNSPVELREPNV	19	58	2,090	2,093	–‡
		111–128	PPEVTVLTNSPVELREPN	18	58	1,991	1,990	–‡
Invariant chain	Membrane	131–149	ATKYGNMTEDHVMHLLQNA	19	73	2,173	2,179	41.5
Acetylcholine receptor	Membrane	289–304	VFLLLLADKVPETSLS	16	69	1,745	1,750	18.5
Glucose transporter	Membrane	459–474	TFDEIASGFRQGGASQ	16	58	1,671	1,673	–‡
Na^+ channel protein	Membrane	384–397	YGYTSYDTFSWAFL	14	43	1,721	1,721	–‡
CD45	Membrane	1071–1084	GQVKKNNHQEDKIE	14	43	1,667	1,667	–‡
ICAM-2	Membrane	64–76	LNKILLDEQAQWK	13	54	1,599	1,602	12.2
IFN-γ receptor	Membrane	128–148	GPPKLDIRKEEKQIMIDIFHP	21	82	2,505	2,510	25.4
		128–147	GPPKLDIRKEEKQIMIDIFH	20	82	2,408	2,412	–§
EBV gp220	Membrane	592–606	TGHGARTSTEPTTDY	15	43	1,593	1,593	–‡
EBV tegument p140	Membrane	1395–1407	KELKRQYEKKLRQ	13	54	1,747	1,750	–‡
IP-30	Membrane	38–59	SPLQALDFFGNGPPVNYKTGNL	22	82	2,350	2,353	–‡
		38–57	SPLQALDFFGNGPPVNYKTG	20	82	2,122	2,124	–‡
Cyt-b5	Cytosolic	155–172	GKFAIRPDKKSNPIIRTV	18	54	2,040	2,043	–‡
Apolipoprotein B-100‖	Exogenous	1276–1295	NLFLKSDGRIKYTLNKNSLK	20	67	2,353	2,360	43.1
		1273–1292	IPDNLFLKSDGRIKYTLNKN	20	69	2,350	2,355	27.5
		1273–1291	IPDNLFLKSDGRIKYTLNK	19	69	2,236	2,245	–§
		1273–1290	IPDNLFLKSDGRIKYTLN	18	69	2,107	2,097	–§
		1273–1289	IPDNLFLKSDGRIKYTL	17	69	1,993	2,001	–§
		1276–1291	NLFLKSDGRIKYTLNK	16	64	1,910	1,911	43.8
		1276–1290	NLFLKSDGRIKYTLN	15	64	1,782	1,786	–§
		1207–1224	YANILLDRRVPQTDMTF	17	67	2,053	2,059	–‡
		1794–1810	VTTLNSDLKYNALDLTN	17	73	1,895	1,896	–‡
								114.4

Sequence and mass determinations for peptides bound to HLA-DR3 are shown. Experimental conditions, abbreviations, and notes are as listed in Table 1 with the following additions: IP-30, IFN-γ-induced protein; Cyt-b5, NADH-cytochrome b_5 reductase.
¶ Partial sequence not verified by mass spectrometry.

scribed earlier for the allele-specific peptides. Mass analysis of fractions collected adjacent to those sequenced confirmed the presence of additional peptides corresponding precisely in mass to each of the promiscuous peptides with extensions and/or truncations at either terminal end.

The second most abundant set of characterized promiscuous self-peptides was derived from an HLA-A2-like sequence with residues 105–117 as the core. However, only two of the five cell lines studied express HLA-A2 (Priess A2/DR4 and 23.1 A2/DR8). Nevertheless, HLA-A2-like peptides were isolated from DR2 (MST cells, A3), DR3 (WT-20 cells, A30), and DR7 (Mann cells, A29), none of which expresses HLA-A2 (reconfirmed by staining with specific mAb; data not shown). These same peptides have also been identified from peptide pools extracted from DR1 expressed in the HLA-A2-positive cell line, LG-2 (15). This sequence, in which tryptophan 107 is a unique residue and facilitates spectrophotometric detection, has been found only in HLA-A2 and HLA-A69 (which appears to have arisen by an exon exchange involving the third exon of an HLA-A2 gene encoding its α2 domain) (41). DR1 expressed in LG-2 (HLA-A2 homozygous), HOM-2 (HLA-A3 homozygous), and the mutant cell lines 721.45 (HLA-A2-positive hemizygous deletion mutant) and 721.221 (a class I negative variant of 721.45, reference 42) was purified and the bound peptide pools characterized. The HLA-A2-like peptides were found in each of the cell lines, but in varying amounts (Fig. 3). Identification of these peptides was accomplished by sequence and mass identity of the HLA-A2-like

Table 3. *Naturally Processed Peptides Bound to HLA-DR4*

Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	Obs.	Yield
					min	*m/z*	*m/z*	*pmol*
HLA-A2	Membrane	28–50	VDDTQFVRFDSDAASQRMEPRAP	23	61	2,639	2,641	–‡
		28–48	VDDTQFVRFDSDAASQRMEPR	21	59	2,471	2,473	–§
		28–47	VDDTQFVRFDSDAASQRMEP	20	62	2,315	2,319	13.5
		28–46	VDDTQFVRFDSDAASQRME	19	57	2,217	2,219	28.9
		30–48	DTQFVRFDSDAASQRMEPR	19	58	2,256	2,263	51.7
		31–49	TQFVRFDSDAASQRMEPRA	19	59	2,212	2,212	12.5
		28–44	VDDTQFVRFDSDAASQR	17	58	1,957	1,963	–§
		31–47	TQFVRFDSDAASQRMEP	17	59	1,985	1,987	–§
		31–45	TQFVRFDSDAASQRM	15	57	1,759	1,761	–§
		31–42	TQFVRFDSDAAS	12	57	1,343	1,343	–§
								106.6
HLA-Cw9	Membrane	28–50	VDDTQFVRFDSDAASPRGEPRAP	23	59	2,534	2,537	–§
		31–52	TQFVRFDSDAASPRGEPRAPWV	22	57	2,490	2,491	–§
		28–48	VDDTQFVRFDSDAASPRGEPR	21	57	2,366	2,368	–§
		28–47	VDDTQFVRFDSDAASPRGEP	20	59	2,209	2,211	–§
		28–46	VDDTQFVRFDSDAASPRGE	19	59	2,112	2,114	–§
		28–45	VDDTQFVRFDSDAASPRG	18	59	1,983	1,987	–§
		31–48	TQFVRFDSDAASPRGEPR	18	55	2,036	2,041	11.0
		28–44	VDDTQFVRFDSDAASPR	17	58	1,926	1,932	–§
		30–46	DTQFVRFDSDAASPRGE	17	55	1,898	1,902	–§
		31–44	TQFVRFDSDAASPR	14	55	1,597	1,604	–§
		130–150	LRSWTAADTAAQITQRKWEAA	21	59	2,375	2,376	–‡
		129–147	DLRSWTAADTAAQITQRKW	19	61	2,218	2,220	147.2
		130–147	LRSWTAADTAAQITQRKW	18	61	2,103	2,105	–§
		129–145	DLRSWTAADTAAQITQR	17	62	1,904	1,909	–§
		129–144	DLRSWTAADTAAQITQ	16	62	1,748	1,752	–§
		129–143	DLRSWTAADTAAQIT	15	62	1,620	1,622	16.8
								175.0
HLA-Bw62	Membrane	129–150	DLSSWTAADTAAQITQRKWEAA	22	69	2,421	2,423	–‡
		129–148	DLSSWTAADTAAQITQRKWE	20	70	2,278	2,285	13.6
		129–146	DLSSWTAADTAAQITQRK	18	69	1,963	1,966	–§
		129–145	DLSSWTAADTAAQITQR	17	63	1,835	1,838	84.8
								98.4
VLA-4	Membrane	229–248	GSLFVYNITTNKYKAFLDKQ	20	69	2,351	2,353	–§
		229–244	GSLFVYNITTNKYKAF	16	69	1,866	1,868	20.1
HLA-DQ3.2 β chain	Membrane	24–38	SPEDFVYQFKGMCYF	15	83	1,861	1,862	–‡
PAI-1	Membrane	261–281	AAPYEKEVPLSALTNILSAQL	21	69	2,228	2,229	–‡
		261–278	AAPYEKEVPLSALTNILS	18	69	1,916	1,917	–‡
Cathepsin C¶	Membrane	151–167	YDHNFVKAINADQKSWT	17	74	2,037	2,039	–‡
		151–166	YDHNFVKAINADQKSW	16	74	1,936	1,938	–‡
Ig Heavy chain**	Membrane	121–?	GVYFYLQWGRSTLVSVS . . .	ND	70	ND	ND	–‡
Bovine hemoglobin	Exogenous	26–41	AEALERMFLSFPTTKT	16	83	1,842	1,836	–‡

Sequence and mass determinations for peptides bound to HLA-DR4 are shown. Experimental conditions, abbreviations, and notes are as listed in Table 1 with the following additions. VLA-4, a cell surface heterodimer in the integrin superfamily of adhesion receptors; PAI-1, plasminogen activator inhibitor 1.

¶ Proposed human protein source with sequence homology to the reported rat cathepsin C sequence.

** Partial sequence not verified by mass spectrometry.

Table 4. *Naturally Processed Peptides Bound to HLA-DR7*

Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	*Obs.*	*Yield**
					min	*m/z*	*m/z*	*pmol*
HLA-A29	Membrane	234–253	RPAGDGTFQKWASVVVPSGQ	20	68	2,087	2,092	–§
		234–249	RPAGDGTFQKWASVVV	16	65	1,718	1,718	17.8
		237–258	GDGTFQKWASVVVPSGQEQRYT	22	68	2,441	2,440	–§
		237–254	GDGTFQKWASVVVPSGQE	18	68	1,892	1,892	–§
		239–252	GTFQKWASVVVPSG	14	68	1,463	1,465	–§
		239–253	GTFQKWASVVVPSGQ	15	68	1,720	1,721	–§
		239–261	GTFQKWASVVVPSGQEQRYTCHV	23	68	2,606	2,606	116.5
								134.3
HLA-B44	Membrane	83–99	RETQ I SKTNTQTYRENL	17	35	2,082	2,086	–‡
		83–98	RETQ I SKTNTQTYREN	16	35	1,969	1,971	–‡
		83–97	RETQ I SKTNTQTYRE	15	35	1,855	1,857	–‡
HLA-DR α chain	Membrane	101–126	RSNYTP I TNPPEVTVLTNSPVELREP	26	35	2,924	2,927	16.6
		58–78	GALAN I AVDKANLE I MTKRSN	21	68	2,229	2,221	–‡
HLA-DQ α chain¶	Membrane	179–?	SLQSP I TVEWRAQSESAQSKMLSG I GGFVL	ND	35	ND	ND	–‡
4F2	Membrane	318–338	VTQYLNATGNRWCSWSLSQAR	21	74	2,442	2,445	–‡
		318–334	VTQYLNATGNRWCSWSL	17	74	1,999	2,002	–‡
LIF receptor	Membrane	854–866	TS I LCYRKREW I K	13	35	1,696	1,701	–‡
Thromboxane-A synthase	Membrane	406–420	PAFRFTREAAQDCEV	15	74	1,740	1,743	–‡
K^+ channel protein	Membrane	492–516	GDMYPKTWSGMLVGALCALAGVLT I	25	74	2,567	2,567	–‡
Hsp 70	Cytosolic	38–54	TPSYVAFTDTERL I GDA	17	71	1,856	1,857	–‡
		38–52	TPSYVAFTDTERL I G	15	71	1,670	1,672	–‡
EBV MCP	Cytosolic	1,264–1,282	VPGLYSPCRAFFNKEELL	18	56	2,082	2,081	19.3
		1,264–1,277	VPGLYSPCRAFFNK	14	56	1,598	1,599	–§
Apolipoprotein B-100‖	Exogenous	1,586–1,608	KVDLTFSKQHALLCSDYQADYES	23	56	2,661	2,662	18.5
		1,586–1,600	KVDLTFSKQHALLCS	15	56	1,689	1,688	–§
		1,942–1,954	FSHDYRGSTSHRL	13	43	1,563	1,567	16.6
		2,077–2,089	LPKYFEKKRNT I I	13	63	1,650	1,654	39.2
								74.3
Complement C9‖	Exogenous	465–483	APVL I SQKLSP I YNLVPVK	19	63	2,080	2,084	–‡

Sequence and mass determinations for peptides bound to HLA-DR7 are shown. Experimental conditions, abbreviations, and notes are as listed in Table 1 with the following additions. 4F2, human cell surface antigen involved in normal and neoplastic cell growth; LIF receptor, leukemia inhibitory factor receptor; Hsp 70, heat-shock protein 70; EBV MCP, Epstein-Barr virus major capsid protein.
¶ Partial sequence not verified by mass spectrometry.

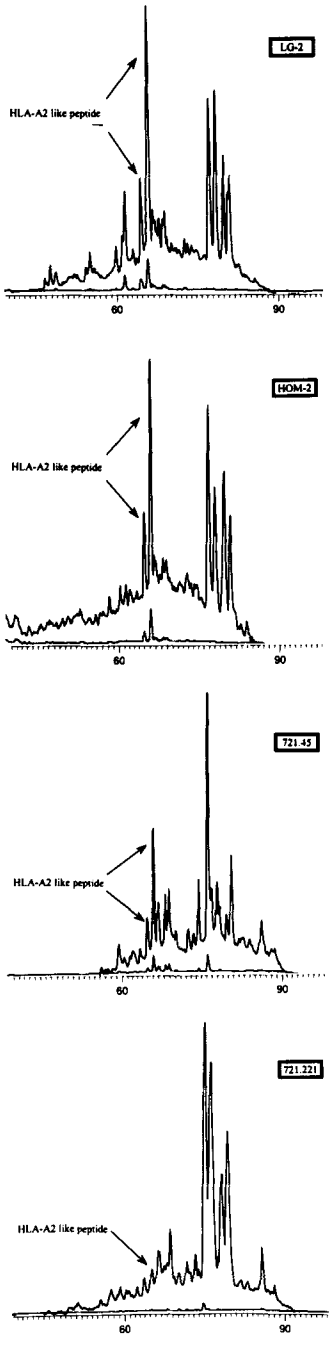

Figure 3. Reversed-phase HPLC separation profiles of peptides bound to HLA-DR1 purified from both HLA-A2-positive and -negative cell lines. Chromatographic peaks containing the HLA-A2-like peptides are easily discerned by their prominent 277-nm UV absorbance due to tryptophan. HLA-A alleles of the homozygous cell lines are: LG-2 (HLA-A2), HOM-2 (HLA-A3), 721.45 (HLA-A2 hemizygous), and 721.221 (class I–negative deletion mutant).

labeled peaks (data not shown). Experiments to characterize a putative locus encoding the HLA-A2-like sequence in HLA-A2-negative cells are underway.

The peptides derived from Ii varied in length from 15 to 25 residues with a common core sequence at position 107–120, and were found in three of the five alleles studied here. They were prominent peptides in each of the alleles in which they were found. An interesting physical characteristic of the Ii peptides is their lack of aromatic amino acids, especially since this feature is prominent in several putative DR-binding motifs (39, 43–46).

The two remaining sets of promiscuous peptides both derived from nonpolymorphic regions of their respective source proteins. A set of MHC-related peptides was derived from the HLA-DR α chain (positions 182–194), common to all DR alleles. Three DR α peptides varying in length from truncations at the carboxy terminus have been sequenced from four of the five DR isotypes. These peptides are similar in composition to those derived from Ii in that both groups have a relatively high proline content and no aromatic residues, but neither of these two properties are shared by the other two promiscuous self-peptides. The only peptides isolated from multiple DR isotypes not associated with the MHC were derived from the Ig κ chain (present in varying quantities on the surface of most of the cell lines from which the DR molecules were purified; data not shown).

Peptide Binding to Surface-expressed HLA-DR

Synthetic peptide analogues representing promiscuous and allele-specific peptides were used in surface-binding assays. Before biotinylation each synthetic peptide was observed to have the same RPC retention time as the naturally processed extracted self-peptide (data not shown). Each peptide contained a single biotin attached via a 22-Å spacer to the amino terminus. The cell lines used for the binding assay (23, 24) were those from which the respective DR isotypes were purified. As a negative control, the class II deletion mutant T2 was used to monitor nonspecific binding of peptide to the cell surface.

Direct binding of biotinylated peptide to class II MHC molecules on the surface of fixed cells was detected by FITC-streptavidin using flow cytometry. Peptide concentrations between 6.25 and 100 μM were used to measure relative binding (Table 7). Specific peptide binding could be blocked by mAb L243 before addition of peptide (Fig. 4). The biotinylation of the HLA-A2-like peptide 103–117 interfered with the normal mode of binding for this peptide; however, the non-biotinylated synthetic peptide did competitively inhibit the specific binding of the other peptides. Similarly, competitive inhibition using an excess of nonbiotinylated peptide was observed for each peptide (data not shown). Thus, it appears that the promiscuous peptides were capable of binding to multiple DR alleles. In contrast, strict allelic specificity was demonstrated by the apolipoprotein B-100 peptide 1273–1291; it bound only to DR3. No specific binding of any of the peptides to the class II–negative cell line, T2, was observed.

Discussion

Four properties illustrate the functional dichotomy between class I and II proteins with respect to peptide binding: (*a*) although both class I and II proteins assemble in the endoplasmic reticulum, only class I binds peptides before reaching the golgi; (*b*) cellular trafficking ensures that class II molecules travel a pathway to the cell surface separate from that of class I and encounter different types of source proteins in the endocytic structures where peptides are loaded; (*c*) structural differences in the binding groove/pocket geometry define the sizes of peptides from a large random pool that can bind

Table 5. *Naturally Processed Peptides Bound to HLA-DR8*

Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	Obs.	Yield
					min	*m/z*	*m/z*	*pmol*
HLA-DR α chain	Membrane	158–180	SETVFLPREDHLFRKFHYLPFLP	23	63	2,889	2,889	–‡
HLA-DP β chain	Membrane	80–92	RHNYELDEAVTLQ	13	81	1,588	1,591	–‡
LAM Blast-1	Membrane	88–108	DPQSGALY I SKVQKEDNSTY I ‡‡	21	57	2,544	2,549	32.6
		92–108	GALY I SKVQKEDNSTY I ‡‡	17	55	2,116	2,118	29.0
		129–146	DPVPKPV I K I EK I EDMDD	18	61	2,081	2,086	–§
		129–143	DPVPKPV I K I EK I ED	15	61	1,720	1,725	42.6
								104.2
Ig κ chain	Membrane	63–80	FTFT I SRLEPEDFAVYYC	18	61	2,202	2,204	–‡
		63–77	FTFT I SRLEPEDFAV	15	61	1,772	1,777	–‡
LAR	Membrane	1,302–1,316	DPVEMRRLNYQTPG	14	81	1,676	1,680	–‡
LIF receptor	Membrane	709–726	YQLLRSM I GY I EELAP I V	18	70	2,109	2,112	–‡
IFN-α receptor	Membrane	271–287	GNHLYKWKQ I PDCENVK	17	70	2,073	2,075	–‡
IL-8 receptor	Membrane	169–188	LPFFLFRQAYHPNNSSPVCY	20	63	2,401	2,403	–‡
Ca^{2+} release channel	Membrane	2,614–2,623	RPSMLQHLLR	10	73	1,251	1,255	–‡
CD35	Membrane	359–380	DDFMGQLLNGRVLFPVNLQLGA	22	78	2,418	2,421	–‡
CD75	Membrane	106–122	I PRLQK I WKNYLSMNKY	17	70	2,196	2,202	–‡
Calcitonin receptor§§	Membrane	38–53	EPFLY I LGKSRVLEAQ	16	83	1,863	1,848	–‡
TIMP-1	Membrane	101–118	NRSEEFL I AGKLQDGLLH	18	70	2,040	2,043	–‡
		102–117	RSEEFL I AGKLQDGLL	16	75	1,789	1,800	–‡
		103–117	SEEFL I AGKLQDGLL	15	77	1,633	1,646	–‡
		101–112	NRSEEFL I AGKL	12	70	1,377	1,382	–‡
TIMP-2	Membrane	187–214	QAKFFAC I KRSDGSCAWYRGAAPPKQEF	28	67	3,162	3,165	–‡
		187–205	QAKFFAC I KRSDGSCAWYR	19	67	2,235	2,234	
PAI-1	Membrane	378–396	DRPFLFVVRHNPTGTVLFM	19	63	2,247	2,247	–‡
		133–148	MPHFFRLFRSTVKQVD	16	75	2,008	2,116	–‡

Cathepsin E	Membrane	89–112	QNFTV I FDTGSSNLWVPSVYCTSP	24	63	2,663	2,664	–‡
		89–104	QNFTV I FDTGSSNLWV	15	63	1,828	1,830	–‡
Cathepsin S	Membrane	189–205	TAFQY I I DNKG I DSDAS	17	68	1,858	1,857	–‡
Cystatin SN	Membrane	41–58	DEYYRRLLRVLRAREQ I V	18	68	2,349	2,348	–‡
Tubulin α-1 chain	Membrane	207–223	EA I YD I CRRNLD I ERPT	17	68	2,077	2,078	–‡
		207–219	EA I YD I CRRNLD I	13	68	1,594	1,595	–‡
Myosin β heavy chain	Membrane	1,027–1,047	HELEK I KKQVEQEKCE I QAAL	21	63	2,494	2,494	–‡
α-enolase‖‖	Membrane	23–?	AEVYHDVAASEFF . . .	nd	58	ND	ND	–‡
c-myc	Cytosolic	371–385	KRSFFALRDQ I PDL	14	73	1,706	1,709	–‡
K-*ras*	Cytosolic	164–180	RQYRLKK I SKEEKTPGC	17	58	2,064	2,066	–‡
Apolipoprotein B-100‖	Exogenous	1,724–1,743	KN I FHFKVNQEGLKLSNDMM	20	66	2,394	2,399	16.1
		1,724–1,739	KN I FHFKVNQEGLKLS	16	61	1,902	1,904	19.8
		1,780–1,799	YKQTVSLD I QPYSLVTTLNS	20	58	2,272	2,274	12.3
		2,646–2,662	STPEFT I LNTLH I PSFT	17	85	1,918	1,929	–‡
		2,647–2,664	TPEFT I LNTLH I PSFT I D	18	85	2,059	2,073	–‡
		2,647–2,662	TPEFT I LNTLH I PSFT	16	85	1,831	1,842	18.4
		2,885–2,900	SNTKYFHKLN I PQLDF	16	73	1,965	1,970	–‡
		2,072–2,088	LPFFKFLPKYFEKKRNT	17	80	2,204	2,207	–‡
		2,072–2,086	LPFFKFLPKYFEKKR	15	81	1,988	1,992	10.1
		4,022–4,036	WNFYYSPQSSPDKKL	15	63	1,860	1,863	–‡
								76.7
Bovine transferrin	Exogenous	261–281	DV I WELLNHAQEHFGKDKSKE	21	81	2,524	2,525	–‡
		261–275	DV I WELLNHAQEHFG	15	83	1,808	1,818	16.1
		261–273	DV I WELLNHAQEH	13	78	1,604	1,609	20.5
								36.6
von Willebrand factor‖	Exogenous	617–636	I ALLLMASQEPQRMSRNFVR	20	63	2,360	2,360	–‡
		617–630	I ALLLMASQEPQRM	14	63	1,601	1,601	–‡

Sequence and mass determinations for peptides bound to HLA-DR8 are shown. Experimental conditions, abbreviations, and notes are as listed in Table 1 with the following additions: LAM, *lymphocyte activation marker*; LAR, *leukocyte antigen-related protein*; TIMP, tissue inhibitor of metalloproteinases; *c-myc*, protooncogene member of DNA binding proteins; K-*ras*, protooncogene member of the GTP binding proteins.

‡‡ The combination of sequence analysis and mass spectrometry was able to confirm the posttranslational addition of *N*-acetylglucosamine on asparagine 104 of these peptides. Presumably, the rest of the carbohydrate attached at this site was hydrolyzed either during antigen processing or peptide extraction/separation.

§§ Proposed human source with sequence homology to known porcine calcitonin receptor.

‖‖ Complete sequence of α-enolase is unknown therefore this peptide sequence cannot be verified by mass spectrometry.

Table 6. *Promiscuous Self-Peptides Bound to HLA-DR Alleles*

HLA-DR Allele(s)	Source protein	Protein category	Residues	Sequence	Length	RT	$[M + H]^+$	Obs.	Yield*
						min	*m/z*	*m/z*	*pmol*‖‖
DR7	HLA-A2 like	Membrane	105–124	SDWRFLRGYHQYAYDGKDY I	20	68	2,554	2,556	39.2
DR1¶, DR2, DR7			103–120	VGSDWRFLRGYHQYAYDG	18	67	2,190	2,190	28.3
DR2			103–119	VGSDWRFLRGYHQYAYD	17	67	2,133	2,132	–§
DR2			104–119	GSDWRFLRGYHQYAYD	16	66	2,034	2,040	–§
DR1, DR2, DR3, DR7			103–117	VGSDWRFLRGYHQYA	15	66	1,855	1,854	52.6
DR1, DR2			103–116	VGSDWRFLRGYHQY	14	65	1,784	1,784	–§
DR1, DR2, DR7			104–117	GSDWRFLRGYHQYA	14	63	1,755	1,755	25.6
DR7			104–116	GSDWRFLRGYHQY	13	63	1,685	1,688	–§
DR1, DR2, DR7			105–117	SDWRFLRGYHQYA	13	62	1,698	1,699	15.3
									161.0
DR1	Invariant chain	Membrane	97–121	LPKPPKPVSKMRMATPLLMQALPMG	25	81	2,734	2,734	nr
DR1, DR2			97–120	LPKPPKPVSKMRMATPLLMQALPM	24	80	2,676	2,676	11.3
DR1			98–121	PKPPKPVSKMRMATPLLMQALPMG	24	79	2,620	2,620	nr
DR1, DR3			97–119	LPKPPKPVSKMRMATPLLMQALP	23	78	2,545	2,544	36.5
DR1, DR2			98–120	PKPPKPVSKMRMATPLLMQALPM	23	80	2,563	2,562	13.2
DR1, DR2			99–120	KPPKPVSKMRMATPLLMQALPM	22	80	2,466	2,466	–§
DR1, DR2, DR3, DR7			98–119	PKPPKPVSKMRMATPLLMQALP	22	76	2,432	2,432	50.7
DR1, DR2, DR3, DR7			99–119	KPPKPVSKMRMATPLLMQALP	21	78	2,335	2,334	–§
DR1, DR2			100–119	PPKPVSKMRMATPLLMQALP	20	80	2,207	2,207	–§

DR2			106–124	KMRMATPLLMQALPMGALP	19	81	2,071	2,074	–‡
DR1, DR2			106–120	KMRMATPLLMQALPM	15	80	1,732	1,732	18.9
DR1			106–119	KMRMATPLLMQALP	14	81	1,601	1,600	nr
									130.6
DR7	HLA-DR α chain	Membrane	182–200	APSPLPETTENVVCALGLTV	20	43	1,912	1,918	18.4
DR2, DR4, DR8			182–198	APSPLPETTENVVCALG	17	45	1,698	1,704	13.4
DR2			182–194	APSPLPETTENVV	13	44	1,354	1,362	80.0
									111.8
DR4	Ig κ chain	Membrane	188–208	KHKVYACEVTHQGLSSPVTKS	21	47	2,300	2,304	44.0
DR4			188–207	KHKVYACEVTHQGLSSPVTK	20	50	2,212	2,213	32.1
DR4			189–206	HKVYACEVTHQGLSSPVT	18	45	1,955	1,952	23.4
DR4			188–204	KHKVYACEVTHQGLSSP	17	47	1,883	1,883	–§
DR4			187–203	EKHKVYACEVTHQGLSS	17	47	1,915	1,922	–§
DR4			188–203	KHKVYACEVTHQGLSS	16	57	1,787	1,787	14.8
DR4			189–204	HKVYACEVTHQGLSSP	16	50	1,755	1,768	–§
DR4			187–202	EKHKVYACEVTHQGLS	16	45	1,828	1,823	–§
DR4			188–202	KHKVYACEVTHQGLS	15	54	1,700	1,708	58.1
DR4			189–203	HKVYACEVTHQGLSS	15	47	1,658	1,667	–§
DR4			187–200	EKHKVYACEVTHQG	14	54	1,629	1,633	–§
DR7			188–201	KHKVYACEVTHQGL	14	53	1,613	1,616	–§
DR7			188–200	KHKVYACEVTHQG	13	53	1,499	1,501	10.9
									183.3

Sequence and mass determinations for naturally processed promiscuous peptides bound to HLA-DR alleles are shown. Experimental conditions, abbreviations, and notes are as listed in Table 1 with the following additions. nr, yields from DR1 are not reported.

¶ DR1 refers to previously published work (15).

‖ For peptides observed binding to multiple alleles, the yields are reported for HLA-DR2-isolated peptides (similar yields were observed for the same peptides isolated from multiple DR alleles).

Table 7. *Surface Staining of EBV-transformed Lymphoblastoid Cell Lines*

Cell lines	Ii (97–120)	Ig κ (188–202)	HLA-DR α (182–198)	Apo B (1273–1291)
LG-2 (DR1)	123	191	345	50
MST (DR2)	288	178	148	18
WT-20 (DR3)	170	174	176	450
Priess (DR4)	215	189	189	57
Mann (DR7)	288	171	nd	50
23.1 (DR8)	213	221	172	19
T2 (no DR)	35	42	39	40

Peptide binding to cell surface expressed class II MHC molecules. The cell lines used to produce the HLA-DR molecules in this study were incubated with biotinylated synthetic peptide and specific binding was measured by flow cytometry. The data are presented as the relative mean log fluorescence. Promiscuous peptides were determined to bind to all six surface expressed DR alleles, while the apolipoprotein B-100 peptide 1273–1291 was found only to bind DR3-positive cells.

class II (10–34 residues) rather than class I molecules (8–11 residues); and (*d*) since the size distributions of bound peptides are different, the antigen processing machinery used to supply peptide for each is likely to be different. Peptides from antigen produced endogenously are primarily presented by class I, while peptides from exogenously administered antigens are presented by class II MHC molecules (47). However, class II MHC molecules also present peptides derived from proteins synthesized by the APC itself, although primarily derived from source proteins translated with signal peptides (membrane associated or secretory proteins) (48–55). Class II molecules can in addition bind endogenous peptides derived from cytosolic source proteins, but only after such proteins enter the class II processing compartment by either translocation or autophagy (56, 57). The relative proportion of class II–bound peptides derived from these sources has not been clear.

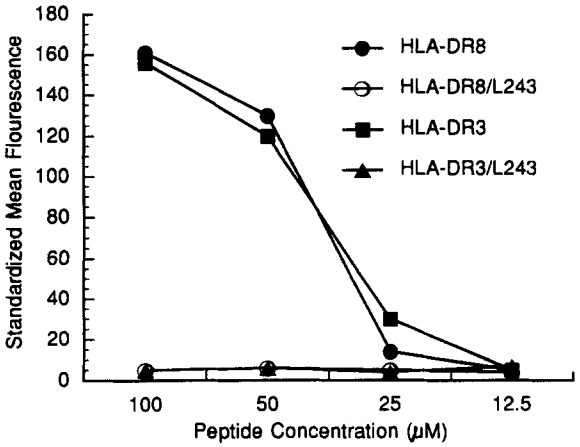

Figure 4. Ii peptide 97–120-specific binding to surface-expressed HLA-DR was determined by direct measurement of added biotinylated peptide using flow cytometry. APC were preincubated with either the anti-DR mAb L243 to sterically obstruct the peptide binding site or a mock buffer before the addition of biotinylated peptide. Ii peptide was blocked at all concentrations measured, representative of the L243 antibody–induced blocking of peptide binding to surface-expressed DR alleles.

The peptides bound to class II MHC molecules purified from APC grown in culture without exposure to "antigen" are described here. Peptides derived from endogenous proteins (both membrane associated/secretory and cytosolic proteins) and exogenous serum sources were identified. If the peptides presented by EBV-transformed B cells grown in vitro are truly representative, then >85% of the peptides bound to class II MHC molecules in vivo are derived from endogenous self-proteins. The frequency of the different types of peptides bound to the DR alleles examined is represented in Fig. 5. The predominant endogenous peptides were derived from MHC-related source proteins, supporting earlier observations on sequenced peptides bound to DR1 and H-2A^d (13, 15) and functional work describing the presentation of both class I– and class II–derived peptides to CD4$^+$ T cells (53, 58, 59). Interestingly, only a single MHC-related peptide was identified as bound to either DR11/DRw52 or H2-A^k, while none at all were reported among the peptides sequenced from H-2A^s or H-2E^b, even though these MHC molecules were purified from similar APC. However, only a limited number of peptides were analyzed in these studies (12, 14, 46, 60). Cell-specific differences in class II antigen processing or protein immunoaffinity purification could also have introduced peptide bias.

These peptide sequences may also further our understanding of class II antigen processing in vivo. Only 27 of the 201 sequenced peptides were <15 residues in length, even though in vitro experiments have shown minimal required peptide lengths of 9–10 residues (1, 34, 38, 45, 61), thus supporting the notion that class II molecules protect bound peptides from proteolytic degradation in vivo (15) as well as in vitro (62, 63). The identification of peptides containing a single *N*-acetylglucosamine (see Table 5) confirmed an earlier in vitro functional study, but more importantly further illustrates the role of lysosomal degradation in antigen processing (64, 65). Since

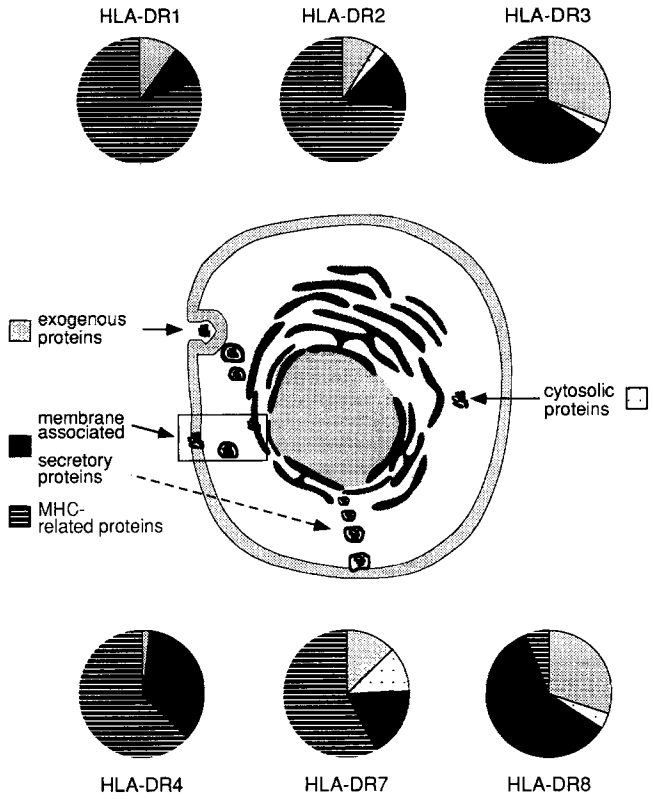

Figure 5. Relative frequency and source of peptides bound to six HLA-DR alleles. Source proteins were divided into three groups; exogenous serum proteins, endogenous membrane associated/secretory proteins, and endogenous cytosolic proteins, and the number of sequenced peptides (not the molar abundance) for each allele,was plotted. The DR1 peptides were as previously described (15). Although this representation is only based on the 223 peptides sequenced from six DR alleles (between 1.6 and 5% of the estimated total number of peptides bound to any single class II allele), it appears to show a consistent preference for certain types of source proteins. The predominance of MHC-related peptides is highlighted as a subset of the membrane-associated/secretory source protein group.

neither the immunoaffinity purification nor the mild acid extraction procedure altered the glycosylation of the class II α or β chains (data not shown), the complex carbohydrate originally attached to the peptide must have been enzymatically cleaved in a lysosomal carbohydrase-containing compartment. Further corroboration of lysosomal processing was demonstrated by the identification of cathepsin C–, E–, and S–derived peptides and the observation of 17 cysteine-containing peptides, presumably generated after lysosomal disulfide reduction (66). Finally, the marked preference for heterogeneity at the carboxy-terminal end is consistent with the concerted action of the endosomal proteases cathepsin D and B (67). However, because no fixed cleavage sites were determined, additional exo- and endoproteases such as aminopeptidase N should also continue to be considered as part of the processing machinery.

Allelic Specificity in Peptide Binding. Allelic specificity was suggested by distinct HPLC profiles of peptide isolated for each allele and by the identification for single DR alleles of 84 peptides with different core sequences. In addition to the promiscuous class I– and class II–derived peptides, 41 other peptides from six different class I molecules and 15 peptides from eight different class II molecules were processed and efficiently bound by the DR alleles. Although the predominant peptides were derived from MHC-related molecules, peptides in sufficient quantity for sequence analysis were identified from 52 source proteins not associated with the MHC. The majority of these were from integral membrane proteins expressed on the cell surface or serum-derived proteins. If non-MHC-related membrane-bound proteins were internalized and degraded, the peptide products would travel through the endocytic pathway and encounter DR molecules en route to the cell surface in endocytic vesicles (68–70). Similar source proteins were described for the peptides isolated from murine H-2A alleles (12, 13, 60). Of the few cytosolic peptides identified, a partial sequence overlap was found with the Hsp 70 self-peptide bound to H-2A (14), while a different Hsp 70 peptide sequence was reported from DR11/DRw52 (46). These peptides may be allele specific since no common sequences from these peptides were observed in more than one DR isotype.

Apolipoprotein B-100 from serum in the medium was the most abundant exogenous peptide source protein. This probably occurred as a result of the efficient internalization of apolipoprotein B-100 by the LDL receptor in clathrin-coated pits (71) and subsequent delivery of the protein to the endosomes. The binding specificity for apolipoprotein B-100 1273–1291 was determined to be limited to DR3, but additional synthetic peptides must be made to confirm the specificity of the remaining 10 apolipoprotein B-100 core sequences. This is especially significant in light of two partial sequence overlaps from apolipoprotein B-100 observed in DR2/DR3 and DR7/DR8, respectively. These sequences cannot be eliminated as candidates for promiscuity until their binding specificity is tested with synthetic analogues. Interestingly, an apolipoprotein B-100 peptide sequence not isolated from any of the five alleles reported here was identified in DR11/DRw52 (46), further suggestive of allele specificity. Additional sequence analyses and/or binding studies with synthetic peptides should determine whether all of the allele-specific observations represent truly specific binding to different DR alleles.

Promiscuity in Peptide Binding to Class II MHC Molecules. The identification of peptides bound to different DR alleles and confirmation of their binding capacity using synthetic peptides demonstrated the promiscuity of four of the self-peptides. The remarkable efficiency of both processing and binding of promiscuous self-peptides may suggest a physiologic role for these peptides in modulation of the immune response. Perhaps they ensure that only antigenic peptides with sufficient affinity are presented to $CD4^+$ T cells. Recognition of antigen has been demonstrated to be extremely sensitive, with 80–500 immunogenic class II/peptide complexes being sufficient for T cell stimulation (72, 73). Thus, a few immunologically active peptides could overshadow a plethora of peptides (estimated to be between 650 and 2,000;

references 13 and 15) bound to surface-expressed MHC molecules. Since overstimulation by antigen can induce peripheral tolerance (74–78), these self-peptides may serve not only as competitors to ensure that only antigenic peptides capable of long half-lives bind, but also to prevent overpresentation of a foreign peptide epitope in vivo. Generation of self-peptides as immunomodulators could be an efficient mechanism within the cell since the source proteins are available in sufficient concentration and the immune system in general would have developed tolerance to these sequences during maturation.

Perhaps the most interesting set of promiscuous self-peptides are those derived from Ii. Ii forms a complex with class II MHC molecules early during biosynthesis and has several roles in antigen processing/presentation. Ii functions as a chaperone, blocks premature peptide binding, and affects class II intracellular trafficking (79–86). Could Ii have evolved within its primary structure an immunomodulatory peptide sequence capable of high affinity binding to the class II binding groove? Is that sequence the site at which intact Ii binds to class II MHC molecules initially? Ii is a type II transmembrane glycoprotein with its amino terminus as the cytoplasmic tail. There are 34 residues between the ER lumenal side of its transmembrane domain (residue 72) and the beginning of the core sequence of the Ii-nested set of peptides bound to class II MHC molecules (residue 106). This length would be adequate for the Ii polypeptide to reach the cleft at the top of the class II MHC molecule, estimated to be at least 45 Å from the membrane (J. Brown, personal communication) as an α-helix (51 Å at 1.5 Å/amino acid), or more than adequate as an extended polypeptide (119 Å at 3.5 Å/amino acid for a β-pleated sheet). Alternatively, the Ii peptides may be formed efficiently by proteolysis of Ii in the endosomes and then bind to α/β dimers. In either case, since proteolytic cleavage efficiently dissociates Ii from class II in the antigen processing compartment, the final role of Ii may be to provide high affinity peptides to serve as competitors during the binding of antigenic peptides. At present, the additional possibility that these Ii peptides act as allosteric effectors and bind to the α/β dimers at some site other than the binding cleft is still imaginable. Regardless of how peptide binding occurs, both the high affinity (15, 87) and promiscuous binding of the Ii peptides to DR molecules are clearly evident. Furthermore, the processing machinery for the proteolysis of Ii does not differ significantly among cell lines, since the same nested set of peptides was isolated from multiple cell lines, as well as DR1 (15), and in antigen-processing mutants expressing DR3 (87, 88). Interestingly, homologous Ii peptides from two different H-2A alleles have also been identified (12, 13), suggesting a universal function for the relatively well-conserved high affinity peptides derived from this molecule and bound to class II proteins.

Peptides derived directly from class II MHC molecules by proteolysis are also among the naturally processed bound peptides. A set of peptides from the conserved DR α chain exhibited promiscuous binding like the Ii peptides. Similarly, a peptide set from the nonpolymorphic murine homologue, H-2E α chain, was identified as bound to H-2A molecules that were immunoaffinity purified from either cultured B cells (12, 13) or spleen cells, thus demonstrating that peptides from MHC-related proteins also bind to class II molecules in vivo (60).

The set of class I MHC-related promiscuous peptides is by far the most enigmatic. Although this peptide set is clearly related to HLA-A2 (or HLA-A69, which contains the same α2 domain as HLA-A2), it was also found in cell lines that were HLA-A2 negative. This set of peptides was also present in the class II molecules purified from the deletion mutant 721.221, which lacks expression of HLA-A, -B, and -C molecules, although some class I loci are still present (42). Thus, a putative locus encoding the HLA-A2-like protein could be located within the class I subregion of the MHC. Exon shuffling in vivo can generate novel HLA class I molecules (89), and one possibility for the source of the HLA-A2-like self-peptides could be a nonfunctional pseudogene encoded within this region of the MHC. The reduced yield of the A2-like peptide in 721.221 could be due to a deletion in a regulatory element or promoter region located in the class I MHC loci, since large gene deletions exist in this mutant cell line. The amount of identified HLA-A2-like peptides varied among the DR alleles, suggesting two possibilities: (*a*) the number of fractions sequenced was too small to permit statistical evaluation; or (*b*) the concentration-dependent equilibrium essential for peptide binding was perturbed by either poor processing of the HLA-A2-like molecule, resulting in less peptide generated or more efficient processing of other protein sources.

The only set of promiscuous binding peptides derived from a non-MHC-related protein was from the Ig κ chain. Relatively high concentrations of this protein would be expected to enter the endocytic pathway since its role is to capture intact antigen on the surface of APC and traffic it to endosomes. Presentation of endogenously processed Ig light chain by class II MHC molecules has been demonstrated in B lymphoma cells with T cell stimulation assays (49, 54) and by sequence analysis of bound peptides (60). The high affinity binding of the Ig κ synthetic analogue to six DR alleles (including DR1 from LG-2, which is λ positive) supports the binding promiscuity even though Ig κ peptides were only isolated from two alleles. The recent sequence analysis of DR11/DRw52-bound peptides also reported this same peptide (46), further supporting its promiscuity.

The ability of peptides to bind multiple DR alleles must be dependent on the composition and location of several key amino acids within the primary structure. Peptide binding to class I MHC molecules has been demonstrated to primarily involve hydrogen bond networks between the MHC molecule and peptide backbone atoms. If a similar mechanism can be extended to peptide binding in class II molecules, then side chain exclusion in the binding cleft region could account for peptide selectivity. Hence, steric hinderance could be avoided by promiscuous peptides possessing a minimum number of bulky side chain residues leading to high affinity

binding to a maximum number of DR alleles. A similar proposal based on synthetic peptide analogue studies has been made (1).

Motifs for Binding of Peptides to HLA-DR Alleles. Strict allele-specific binding motifs, as defined for class I MHC molecules, may not exist for class II–bound peptides. Identification of class I motifs was aided: (*a*) by the relatively small size of the bound peptides (nonamers); (*b*) the fact that the majority of these peptides have their amino and carboxy termini fixed within the A and F pockets of the class I MHC molecules; and (*c*) the uniform length of the peptides bound to the alleles initially studied. The problem of deducing motifs from sequence information is more difficult for class II molecules. Bound peptides vary in size (10–34 residues) and are represented by nested sets with varying numbers of flanking residues added at either the amino or carboxy termini, making interpretation of pooled sequencing data nearly impossible. Thus, the longer peptides do not appear to have an end anchored within the binding cleft of class II MHC molecules and may extend outside the cleft at either or both ends. The minimal-length peptide could contain a motif within its core region (9–10 residues), but the core and its anchor residues may be located at various positions within different peptides. Thus, the identification of motifs from sequence information alone is difficult and some additional data are required, such as binding studies varying the length and composition of the peptides and even more definitively (and with correspondingly greater difficulty) refined crystallographic information using single peptide complexes (90–92).

Prior attempts at proposing DR binding motifs using various experimental methods have generated conflicting conclusions (15, 24, 39, 43, 45, 93–95). For example, aside from the probable presence of a hydrophobic residue important in anchoring the amino-terminal region of the peptide, neither of the motifs recently proposed for DR3 (94, 95) fits all of the naturally processed peptides found in this molecule (Table 2), illustrating the difficulty of assigning motifs based on limited sequence information.

With these limitations, can any information relating to preferred amino acids be obtained from the sequences of the naturally processed peptides bound to the different DR alleles? The peptide is likely to be oriented with its amino terminus at the left hand end of the binding cleft, as it is usually represented, and its carboxy terminus at the right hand end (44), as in class I MHC molecules. The left hand end of the cleft is composed mainly of residues contributed by the nonpolymorphic DR α chain. A hydrophobic pocket may exist in this region, for example, in the region of DR α F26, F54, and F51 (equivalent to Y59 in the A pocket of class I molecules) and DR β G86 (equivalent to Y171) (96). The diallelic polymorphism of residue 86 (G or V in different alleles) in the β chain may affect this pocket, and could well contribute to the size of the pocket, which might influence whether a large hydrophobic amino acid (I, L, M, or V) or an aromatic residue (F, Y, or W) could be accommodated in this region (see Fig. 3 *C* in reference 44). Analysis of the minimum core sequences from Tables 1–6 reveals that there is a preference for hydrophobic residues near the amino-terminal end

Table 8. *Sequence Alignment of Allele-specific Preferred Amino Acids*

Allele	Sequence alignment	Conformity
		%
	ixxxxxxxxx↓	70
DR2a/DR2b	(i + 10)	
	i = *I, L, or V* ↓ = *H, K, or R*	(7/10 core sequences)
	ixx↓	81
DR3/DRw52	(i + 3)	
	i = *F, I, L, V, or Y* ↓ = *D, N, Q, or T*	(13/16 core sequences)
	ixxxxxxx↓	88
DR4	(i + 8)	
	i = *F, L, or V* ↓ = *N, Q, S, or T*	(7/8 core sequences)
	ixxxx↓	87
DR7	(i + 5)	
	i = *F, I, L, V, or Y* ↓ = *N, S, or T*	(13/15 core sequences)
	ixxx↓	73
DR8	(i + 4)	
	i = *F, I, L, V, or Y* ↓ = *H, K, or R*	(24/33 core sequences)

The minimum core sequences from the allele-specific bound peptides (Tables 1–5) were aligned to determine the existence of any prominent amino acids/positions. Conformity refers to the percentage of identified sequences in exact alignment. The spacing between key residues is listed under the alignment, with *i* referring to the index position for the listed hydrophobic residues, and ↓ referring to the preferred residues at the designated position.

for most of the bound peptides, supporting the proposal orientation of bound peptides. In fact, <1% of the bound peptides are without some type of hydrophobic residue in the amino-terminal third of any given minimum core sequence.

Allele-specific preferences for peptide residue content and location are not obvious. Six class II allele-specific binding motifs have been proposed by correlating sequence analyses from identified bound peptides with those of synthetic analogue studies (13, 15, 46, 60). However, in each instance there are peptides that bind to the particular class II molecules without an absolute compliance to the proposed preferences. To investigate if preferred residues exist in the peptides bound to individual alleles, sequence alignments were completed on the minimum core sequences for the peptides in Tables 1–5 (10, 16, 8, 15, and 33 sequences, respectively). The conformity varied from a 70% alignment in DR2 (7/10 peptides, perhaps due in part to the minor contamination of DR2a-bound peptides) to an 88% alignment in DR4 (7/8 peptides), with no set of allele-specific peptides completely conforming to a strict amino acid position preference (Table 8). It is noteworthy that the distance between the index position (*i*) and the preferred residue (↓) varied between alleles, implying that this spacing may in part contribute to allele-specific peptide binding. The promiscuous self-peptides, demonstrated to bind to all six DR alleles, also varied in alignment conformity from 60 to 100% (the worst alignment was again observed with DR2). Thus, it appears that class II molecules govern peptide binding by subtle differences in peptide composition rather than by strictly defined side chain requirements. This conclusion is supported by the identification of antigenic peptides capable of binding to multiple DR alleles (23, 33–39) as well as the promiscuous self-peptides identified in this report.

As an alternative to using synthetic peptides and inferring that these findings are reflective of peptide binding in vivo, we have chosen to analyze those peptides that have been processed and bound to MHC molecules within cells. Thus, in addition to the identification of the bound peptides, much can be learned about the repertoire of presented antigen and the machinery involved in antigen processing. In contrast to the peptides bound to class I MHC proteins, class II–bound peptides do not have specific side chain requirements at precise distances from the termini, suggesting that the overall binding energy is provided by the sum of cooperative interactions between the peptide backbone and the cleft and does not involve the peptide termini per se. Although allele specificity surely exists, the identification of promiscuous self-peptides (capable of binding to multiple alleles) suggests that rigid allele-specific motifs for class II molecules do not exist, thus permitting a broad binding specificity. Perhaps this promiscuous binding both modulates the presentation of foreign antigens during an immune response and also broadens T cell tolerance to self-peptides during thymic development. Finally, the substantial amount of MHC-related bound peptide (both in terms of relative abundance and the presentation of peptides derived from the three major isotypes found among class I and II molecules) must be considered when interpreting alloreactivity and may have significant implications for transplantation immunology.

We thank Mary Gordy and Neil Wright for technical assistance and Hidde Ploegh for critical reading of the manuscript.

This research was supported by the Irvington Institute for Medical Research (R. M. Chicz and R. G. Urban), the Juvenile Diabetes Foundation International (J. C. Gorga), the International Human Frontier Science Program Organization (D. A. A. Vignali), and by National Institutes of Health grant CA-47554 (J. L. Strominger).

Address correspondence to Roman M. Chicz, Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138. Joan C. Gorga's present address is the Division of Immunogenetics, Department of Pediatrics, University of Pittsburgh School of Medicine, Children's Hospital of Pittsburgh, Pittsburgh, PA 15213.

Received for publication 4 February 1993 and in revised form 29 March 1993.

References

1. Rothbard, J.B., and M.L. Gefter. 1991. Interactions between immunogenic peptides and MHC proteins. *Annu. Rev. Immunol.* 9:527.
2. Bjorkman, P.M., 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 (Lond.).* 329:506.
3. Garrett, T.P.J., M.A. Saper, P.J. Bjorkman, J.L. Strominger, and D.C. Wiley. 1989. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature (Lond.).* 342:692.
4. 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 (Lond.).* 353:321.
5. 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 (Lond.).* 348:213.
6. Falk, K., O. Rotzschke, S. Stevanovic, G. Jung, and H. Rammensee. 1991. Allele-specific motifs revealed by sequencing of

self-peptides eluted from MHC molecules. *Nature (Lond.).* 351:290.
7. 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 (Lond.).* 353:326.
8. Hunt, D.F., R.A. Henderson, J. Shabanowicz, K. Sakaguchi, H. Michel, N. Sevilir, A. Cox, E. Appella, and V.H. Engelhard. 1992. Characterization of peptides bound to class I MHC molecule HLA-A2.1 by mass spectrometry. *Science (Wash. DC).* 255:1261.
9. Guo, H.-C., T.S. Jardetzky, T.P.J. Garrett, W.S. Lane, J.L. Strominger, and D.C. Wiley. 1992. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature (Lond.).* 360:364.
10. Sadegh-Nasseri, S., and R.N. Germain. 1991. A role for peptide in determining MHC class II structure. *Nature (Lond.).* 353:167.
11. Jensen, P.E. 1992. Long-lived complexes between peptide and class II major histocompatibility complex are formed at low pH with no requirement for pH neutralization. *J. Exp. Med.* 176:793.
12. Rudensky, A.Y., P. Preston-Hurlburt, S.-C. Hong, A. Barlow, and C.A. Janeway. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature (Lond.).* 353:622.
13. 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 (Wash. DC).* 256:1817.
14. Nelson, C.A., R.W. Roof, D.W. McCourt, and E.R. Unanue. 1992. Identification of the naturally processed form of hen egg white lysozyme bound to the murine major histocompatibility complex class II molecule I-A^k. *Proc. Natl. Acad. Sci. USA.* 89:7380.
15. Chicz, R.M., R.G. Urban, W.S. Lane, J.C. Gorga, L.J. Stern, D.A.A. Vignali, 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 (Lond.).* 358:764.
16. Gorga, J.C., P.J. Knudsen, J.A. Foran, J.L. Strominger, and S.J. Burakoff. 1986. Immunochemically purified DR antigens in liposomes stimulate xenogeneic cytolytic T cells in secondary in vitro cultures. *Cell. Immunol.* 103:160.
17. Matsuyama, T., J. Schwenzer, J. Silver, and R.J. Winchester. 1986. Structural relationships between the DR beta 1 and DR beta 2 subunits in DR4, 7, and w9 haplotypes and the DRw53 (MT3) specificity. *J. Immunol.* 137:934.
18. Gorga, J.C., V. Horejsi, D.R. Johnson, R. Raghupathy, and J.L. Strominger. 1987. Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J. Biol. Chem.* 262:16087.
19. Gorga, J.C., J.H. Brown, T. Jardetzky, D.C. Wiley, and J.L. Strominger. 1991. Crystallization of HLA-DR antigens. *Res. Immunol.* 142:401.
20. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680.
21. Lane, W.S., A. Galat, M.W. Harding, and S.L. Schreiber. 1991. Complete amino acid sequence of the FK506 and rapamycin binding protein, FKBP, isolated from calf thymus. *J. Protein Chem.* 10:151.
22. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403.
23. Busch, R., G. Strang, K. Howland, and J.B. Rothbard. 1990. Degenerate binding of immunogenic peptides to HLA-DR proteins on B cell surfaces. *Int. Immunol.* 2:443.
24. 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.* 147:189.
25. Springer, T.A., J.F. Kaufman, L.A. Siddoway, D.L. Mann, and J.L. Strominger. 1977. Purification of HLA-linked B lymphocyte alloantigens in the immunologically active form by preparative sodium dodecyl sulfate-gel electrophoresis and studies on their subunit association. *J. Biol. Chem.* 252:6201.
26. Dornmair, K., B. Rothenhausler, and H.M. McConnell. 1989. Structural intermediates in the reactions of antigenic peptides with MHC molecules. *Cold Spring Harbor Symp. Quant. Biol.* 54:409.
27. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387.
28. Bachvaroff, R.J., F. Miller, and F.T. Rapaport. 1980. Appearance of cytoskeletal components on the surface of leukemia cells and of lymphocytes transformed by mitogens and Epstein-Barr virus. *Proc. Natl. Acad. Sci. USA.* 77:4979.
29. Cheng, Y.-S.E., C.E. Patterson, and P. Staeheli. 1991. Interferon-induced guanylate-binding proteins lack an N(T)KXD consensus motif and bind GMP in addition to GDP and GTP. *Mol. Cell. Biol.* 11:4717.
30. Bourne, H.R. 1988. Do GTPases direct membrane traffic in secretion? *Cell.* 53:669.
31. Colombo, M.I., L.S. Mayorga, P.J. Casey, and P.D. Stahl. 1992. Evidence of a role for heterotrimeric GTP-binding proteins in endosome fusion. *Science (Wash. DC).* 255:1695.
32. DeNagel, D.C., and S.K. Pierce. 1992. A case for chaperones in antigen processing. *Immunol. Today.* 13:86.
33. Sinigaglia, F., M. Guttinger, J. Kilgus, D.M. Doran, H. Matile, H. Etlinger, A. Trzeciak, D. Gillessen, and J.R.L. Pink. 1988. A malaria T-cell epitope recognized in association with most mouse and human MHC class II molecules. *Nature (Lond.).* 336:778.
34. Guttinger, M., P. Romagnoli, L. Vandel, R. Meloen, B. Takacs, J.R.L. Pink, and F. Sinigaglia. 1991. HLA polymorphism and T cell recognition of a conserved region of p190, a malaria vaccine candidate. *Int. Immunol.* 3:899.
35. Lamb, J.R., A.D.M. Rees, V. Bal, H. Ikeda, D. Wilkinson, R.R.P. De Vries, and J.B. Rothbard. 1988. Prediction and identification of an HLA-DR-restricted T cell determinant in the 19-kDa protein of Mycobacterium tuberculosis. *Eur. J. Immunol.* 18:973.
36. Roche, P.A., and P. Cresswell. 1990. High-affinity binding of an influenza hemagglutinin-derived peptide to purified HLA-DR. *J. Immunol.* 144:1849.
37. O'Sullivan, D., J. Sidney, E. Appella, L. Walker, L. Phillips, S.M. Colon, C. Miles, R.W. Chesnut, and A. Sette. 1990. Characterization of the specificity of peptide binding to four DR haplotypes. *J. Immunol.* 145:1799.
38. O'Sullivan, D., J. Sidney, M.-F. Del Guercio, S.M. Colon, and A. Sette. 1991. Truncation analysis of several DR binding epitopes. *J. Immunol.* 146:1240.
39. O'Sullivan, D., T. Arrhenius, J. Sidney, M.-F. Del Guercio, M. Albertson, M. Wall, C. Oseroff, S. Southwood, S.M. Colon, F.C.A. Gaeta, and A. Sette. 1991. On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. *J. Immunol.* 147:2663.

40. Sette, A., A. Vitiello, P. Farness, J. Furze, J. Sidney, J.M. Claverie, H.M. Grey, and R. Chesnut. 1991. Random association between the peptide repertoire of A2.1 class I and several different DR class II molecules. *J. Immunol.* 147:3893.
41. Bjorkman, P.J., and P. Parham. 1990. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu. Rev. Biochem.* 59:253.
42. Shimizu, Y., and R. DeMars. 1989. Production of human cells expressing individual transferred HLA-A, -B, -C genes using an HLA-A, -B, -C null human cell line. *J. Immunol.* 142:3320.
43. Jardetzky, T.S., J.C. Gorga, R. Busch, J. Rothbard, J.L. Strominger, and D.C. Wiley. 1990. Peptide binding to HLA-DR1: a peptide with most residues substituted to alanine retains MHC binding. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:1797.
44. Krieger, J.I., R.W. Karr, H.M. Grey, W.-Y. Yu, D. O'Sullivan, L. Batovsky, Z.-L. Zheng, S.M. Colon, F.C.A. Gaeta, J. Sidney, M. Albertson, M.-F. Del Guercio, R.W. Chesnut, and A. Sette. 1991. Single amino acid changes in DR and antigen define residues crucial for peptide-MHC binding and T cell recognition. *J. Immunol.* 146:2331.
45. Hammer, J., B. Takacs, and F. Sinigaglia. 1992. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. *J. Exp. Med.* 176:1007.
46. Newcomb, J.F., and P. Cresswell. 1993. Characterization of endogenous peptides bound to purified HLA-DR molecules and their absence from invariant chain-associated $\alpha\beta$ dimers. *J. Immunol.* 150:499.
47. Unanue, E.R., and J.-C. Cerottini. 1989. Antigen presentation. *FASEB (Fed. Am. Soc. Biol. Organ.) J.* 3:2496.
48. Jin, Y., J.W.-K. Shih, and I. Berkower. 1988. Human T cell response to the surface antigen of hepatitis B virus (HBsAg). Endosomal and nonendosomal processing pathways are accessible to both endogenous and exogenous antigen. *J. Exp. Med.* 168:293.
49. Weiss, S., and B. Bogen. 1989. B lymphoma cells process and present their endogenous immunoglobulin to major histocompatibility complex-restricted T cells. *Proc. Natl. Acad. Sci. USA.* 86:282.
50. Eisenlohr, L.C., and C.J. Hackett. 1989. Class II major histocompatibility complex-restricted T cells specific for a virion structural protein that do not recognize exogenous influenza virus. Evidence that presentation of labile T cell determinants is favored by endogenous antigen synthesis. *J. Exp. Med.* 169:921.
51. Nuchtern, J.G., W.E. Biddison, and R.D. Klausner. 1990. Class II MHC molecules can use the endogenous pathway of antigen presentation. *Nature (Lond.).* 343:74.
52. Thomas, D.B., J. Hodgson, P.F. Riska, and C.M. Graham. 1990. The role of endoplasmic reticulum in antigen processing. N-glycosylation of influenza hemagglutinin abrogates $CD4^+$ cytotoxic T cell recognition of endogenously processed antigen. *J. Immunol.* 144:2789.
53. Chen, B.P., A. Madrigal, and P. Parham. 1990. Cytotoxic T cell recognition of an endogenous class I HLA peptide presented by a class II HLA molecule. *J. Exp. Med.* 172:779.
54. Weiss, S., and B. Bogen. 1991. MHC class II-restricted presentation of intracellular antigen. *Cell.* 64:767.
55. Moreno, J., D.A.A. Vignali, F. Nadimi, S. Fuchs, L. Adorini, and G.J. Hammerling. 1991. Processing of an endogenous protein can generate MHC class II-restricted T cell determinants distinct from those derived from exogenous antigen. *J. Immunol.* 147:3306.
56. Jaraquemada, D., M. Marti, and E.O. Long. 1990. An endogenous processing pathway in vaccinia virus-infected cells for presentation of cytoplasmic antigens to class II-restricted T cells. *J. Exp. Med.* 172:947.
57. Malnati, M.S., M. Marti, T. LaVaute, D. Jaraquemada, W. Biddison, R. DeMars, and E.O. Long. 1992. Processing pathways for presentation of cytosolic antigen to MHC class II-restricted T cells. *Nature (Lond.).* 357:702.
58. 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.
59. Liu, Z., Y.-K. Sun, Y.-P. Xi, P. Harris, and N. Suciu-Foca. 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.
60. Rudensky, A.Y., P. Preston-Hurlburt, B.K. Al-Ramadi, J. Rothbard, and C.A. Janeway, Jr. 1992. Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature (Lond.).* 359:429.
61. Rothbard, J.B., and W.R. Taylor. 1988. A sequence pattern common to T cell epitopes. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:93.
62. Donermeyer, D.L., and P.M. Allen. 1989. Binding to Ia protects an immunogenic peptide from proteolytic degradation. *J. Immunol.* 142:1063.
63. Mouritsen, S., M. Meldal, O. Werdelin, A. Stryhn Hansen, and S. Buus. 1992. MHC molecules protect T cell epitopes against proteolytic destruction. *J. Immunol.* 149:1987.
64. Ishioka, G.Y., A.G. Lamont, D. Thomson, N. Bulbow, F.C.A. Gaeta, A. Sette, and H.M. Grey. 1992. MHC interaction and T cell recognition of carbohydrates and glycopeptides. *J. Immunol.* 148:2446.
65. Harding, C.V., D.S. Collins, J.W. Slot, H.J. Geuze, and E.R. Unanue. 1991. Liposome-encapsulated antigens are processed in lysosomes, recycled, and presented to T cells. *Cell.* 64:393.
66. Collins, D.S., E.R. Unanue, and C.V. Harding. 1991. Reduction of disulfide bonds within lysosomes is a key step in antigen processing. *J. Immunol.* 147:4054.
67. Van Noort, J.M., J. Boon, A.C.M. Van der Drift, J.P.A. Wagenaar, A.M.H. Boots, and C.J.P. Boog. 1991. Antigen processing by endosomal proteases determines which sites of sperm-whale myoglobin are eventually recognized by T cells. *Eur. J. Immunol.* 21:1989.
68. Cresswell, P. 1985. Intracellular class II antigens are accessible to transferrin-neuraminidase conjugates internalized by receptor-mediated-endocytosis. *Proc. Natl. Acad. Sci. USA.* 82:8188.
69. Neefjes, J.J., V. Stollorz, P.J. Peters, H.J. Geuze, and H.L. Ploegh. 1990. The biosynthetic pathway of MHC class II but not class I molecules intersects the endocytic route. *Cell.* 61:171.
70. Peters, P.J., J.J. Neefjes, V. Oorschot, H.L. Ploegh, and H.J. Geuze. 1991. Segregation of MHC class II molecules from MHC class I molecules in the Golgi complex for transport to lysosomal compartments. *Nature (Lond.).* 349:669.
71. Knott, T.J., R.J. Pease, L.M. Powell, S.C. Wallis, S.C. Rall, Jr., T.L. Innerarity, B. Blackhart, W.H. Taylor, Y. Marcel, R. Milne et al. 1986. Complete protein sequence and identification of structural domains of human apolipoprotein B. *Nature (Lond.).* 323:734.
72. Harding, C.V., and E.R. Unanue. 1990. Quantitation of antigen-processing cell MHC class II/peptide complexes necessary for T-cell stimulation. *Nature (Lond.).* 346:574.
73. Demotz, S., H.M. Grey, and A. Sette. 1990. The minimal number of class II MHC-antigen complexes needed for T cell activation. *Science (Wash. DC).* 249:1028.

74. Suzuki, G., Y. Kawase, S. Koyasu, I. Yahara, Y. Kobayashi, and R.H. Schwartz. 1988. Antigen-induced suppression of the proliferative response of T cell clones. *J. Immunol.* 140:1359.
75. Webb, S., C. Morris, and J. Sprent. 1990. Extrathymic tolerance of mature T cells: clonal elimination as a consequence of immunity. *Cell.* 63:1249.
76. Kawabe, Y., and A. Ochi. 1991. Programmed cell death and extrathymic reduction of $V\beta8^+$ $CD4^+$ T cells in mice tolerant to Staphylococcus aureus enterotoxin B. *Nature (Lond.).* 349:245.
77. Rocha, B., and H. von Boehmer. 1991. Peripheral selection of the T cell repertoire. *Science (Wash. DC).* 251:1225.
78. Miller, J.F.A.P., and G. Morahan. 1992. Peripheral T cell tolerance. *Annu. Rev. Immunol.* 10:51.
79. Anderson, M.S., and J. Miller. 1992. Invariant chain can function as a chaperone protein for class II major histocompatibility molecules. *Proc. Natl. Acad. Sci. USA.* 89:2282.
80. Roche, P.A., and P. Cresswell. 1990. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature (Lond.).* 345:615.
81. Roche, P.A., and P. Cresswell. 1991. Proteolysis of the class II-associated invariant chain generates a peptide binding site in intracellular HLA-DR molecules. *Proc. Natl. Acad. Sci. USA.* 88:3150.
82. Claesson-Welsh, L., and P.A. Peterson. 1985. Implications of the invariant gamma-chain on the intracellular transport of class II histocompatibility antigens. *J. Immunol.* 135:3551.
83. Bakke, O., and B. Dobberstein. 1990. MHC class II-associated invariant chain contains a sorting signal for endosomal compartments. *Cell.* 63:707.
84. Lotteau, V., L. Teyton, A. Peleraux, T. Nilsson, L. Karlsson, S.L. Schmid, V. Quaranta, and P.A. Peterson. 1990. Intracellular transport of class II MHC molecules directed by invariant chain. *Nature (Lond.).* 348:600.
85. Teyton, L., D. O'Sullivan, P.W. Dickson, V. Lotteau, A. Sette, P. Fink, and P.A. Peterson. 1990. Invariant chain distinguishes between the exogenous and endogenous antigen presentation pathways. *Nature (Lond.).* 348:39.
86. Peterson, M., and J. Miller. 1992. Antigen presentation enhanced by the alternatively spliced invariant chain gene product p41. *Nature (Lond.).* 357:596.
87. Sette, A., S. Ceman, R.T. Kubo, K. Sakaguchi, E. Appella, D.F. Hunt, T.A. Davis, H. Michel, J. Shabanowitz, R. Rudersdorf, H.M. Grey, and R. DeMars. 1992. Invariant chain peptides in most HLA-DR molecules of an antigen-processing mutant. *Science (Wash. DC).* 258:1801.
88. Riberdy, J.M., J.R. Newcomb, M.J. Surman, J.A. Barbosa, and P. Cresswell. 1992. HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature (Lond.).* 360:474.
89. Holmes, N., and P. Parham. 1985. Exxon shuffling in vivo can generate novel HLA class I molecules. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2849.
90. 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 (Wash. DC).* 257:927.
91. Zhang, W., A.C.M. Young, M. Imarai, S.G. Nathenson, and J.C. Sacchettini. 1992. Crystal structure of the major histocompatibility complex class I $H\text{-}2K^b$ molecule containing a single viral peptide: implications for peptide binding and T-cell receptor recognition. *Proc. Natl. Acad. Sci. USA.* 89:8403.
92. Silver, M.L., H.-C. Guo, J.L. Strominger, and D.C. Wiley. 1992. Atomic structure of a human MHC molecule presenting an influenza virus peptide. *Nature (Lond.).* 360:367.
93. Kropshofer, H., H. Max, C.A. Muller, F. Hesse, S. Stevanovic, G. Jung, and H. Kalbacher. 1992. Self-peptide released from class II HLA-DR1 exhibits a hydrophobic two-residue motif. *J. Exp. Med.* 175:1799.
94. Geluk, A., A.A.M. Janson, J. Wouter Drijfhout, R.H. Meloen, R.R.P. De Vries, and T.H.M. Ottenhoff. 1992. Functional analysis of DR17(DR3)-restricted mycobacterial T cell epitopes reveals DR17-binding motif and enables the design of allele-specific competitor peptides. *J. Immunol.* 149:2864.
95. Sidney, J., C. Oseroff, S. Southwood, M. Wall, G. Ishioka, F. Korning, and A. Sette. 1992. DRB1*0301 molecules recognize a structural motif distinct from the one recognized by most $DR\beta1$ alleles. *J. Immunol.* 149:2634.
96. 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 (Lond.).* 332:845.