

# Natural ligand motifs of H-2E molecules are allele specific and illustrate homology to HLA-DR molecules

Hansjörg Schild, Ulrike Grüneberg<sup>2</sup>, Georg Pougialis, Hans-Joachim Wallny<sup>1</sup>, Wieland Keilholz, Stefan Stevanović and Hans-Georg Rammensee

Department of Tumovirus-Immunology, German Cancer Research Institute, Im Neuenheimer Feld 242, 69120 Heidelberg, Germany

<sup>1</sup>The Basel Institute for Immunology, Grenzacherstr. 487, 4005 Basel, Switzerland

<sup>2</sup>Present address: Human Immunogenetics Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

**Keywords:** class II proteins, epitope, MHC, peptide, ligand motif, T<sub>H</sub> cell

## Abstract

**Motifs of peptides naturally associated with H-2E<sup>k</sup> and E<sup>d</sup> molecules were determined by (I) pool sequencing of natural ligand mixtures and (II) sequencing of individual natural ligands followed by their alignment to the basic motif suggested by pool sequencing. The data reveal nine amino acid motifs with interaction sites at relative positions P1, P4, P6 and P9, with specificities that are identical at some but different at other anchor positions between E<sup>d</sup> and E<sup>k</sup> motifs, illustrating the different requirements for peptides to be presented by these two MHC molecules. The anchors with the most restricted specificity are P1 and P9. P1 is aliphatic for E<sup>k</sup> and predominantly aromatic for E<sup>d</sup>. P9 is positively charged for both molecules. P4 and P6 show a totally different amino acid preference between E<sup>k</sup> and E<sup>d</sup> ligand motifs. An alignment of E<sup>d</sup> and E<sup>k</sup> protein sequences to the recently reported HLA-DR1 pocket residues is in agreement with observed anchor residues in E<sup>k</sup> and E<sup>d</sup> motifs, thus confirming the predicted similarity of mouse class II E molecules with human DR molecules. Furthermore, this alignment was extended to the putative pockets of class II E<sup>b</sup> and E<sup>s</sup> molecules, and allowed, together with sequence information of previously identified natural ligands of E<sup>b</sup> and E<sup>s</sup> molecules, a prediction of their respective motifs. The information obtained by this study should be useful to identify putative class II E epitopes in proteins and to design peptides for blocking class II E molecules.**

## Introduction

The analysis of peptides eluted from MHC class I molecules revealed clear ligand motifs, typically characterized by a certain peptide length and two anchor positions that have to be occupied by closely related amino acids (1). However, such motifs are less apparent with ligands eluted from class II MHC molecules because these peptides vary extensively in their length, complicating the alignment of isolated natural ligands (2–6). Therefore, most of the information of class II binding specificity is derived from peptide binding studies using mutations of known T cells epitopes (7–11) or by *in vitro* selection of peptides derived from phage libraries (12). These studies have the disadvantage of being biased for peptide binding only and do not account for influences of antigen

processing. A good example for possible discrepancies between ligand and binding motifs is the H-2K<sup>b</sup> motif. Whereas all isolated K<sup>b</sup> ligands carry a hydrophobic residue at P9, no preference for certain amino acids at this position is detectable in peptide binding studies (13). Furthermore, in experiments using mutated synthetic peptides, time and money allow the study of only a limited number of variant peptides. Therefore, complementary effects between different anchor positions might stay undetected, since not all combinatorial possibilities can be studied. Recent experiments using pool sequencing of natural ligands in combination with sequencing and alignment of individual peptides showed that this approach can be used to identify motifs of class II-associated peptides (14–

17). The validity of these motifs is supported by the recently published three-dimensional structure of DR1 molecules (18,19). The structure revealed five pockets, each of which has a certain specificity for peptide ligand side chains which is in agreement with the peptide binding motif for DR1 molecules derived by either peptide binding studies (12) or pool sequencing analysis (14). In order to gain more insight into H-2E motifs and especially to understand the different requirements for peptide binding by H-2E<sup>k</sup> and H-2E<sup>d</sup> molecules, we analyzed natural peptide ligands of these two molecules, using pool and individual peptide sequencing. Even though H-2E<sup>k</sup> and H-2E<sup>d</sup> molecules have a high sequence homology, the obtained ligand motifs show remarkable differences explaining why most peptides which bind to H-2E<sup>k</sup> molecules do not bind to H-2E<sup>d</sup> molecules and vice versa. The differences in the obtained motifs are also reflected in the putative pocket structures based on the HLA-DR molecule. This implies, in addition to the high sequence homology, also a high structural homology between H-2E and HLA-DR molecules. The resulting E<sup>k</sup> and E<sup>d</sup>-specific motifs were used as a basis to attempt predictions on H-2E<sup>s</sup> and E<sup>b</sup> motifs.

## Methods

### Cell lines and cell culture

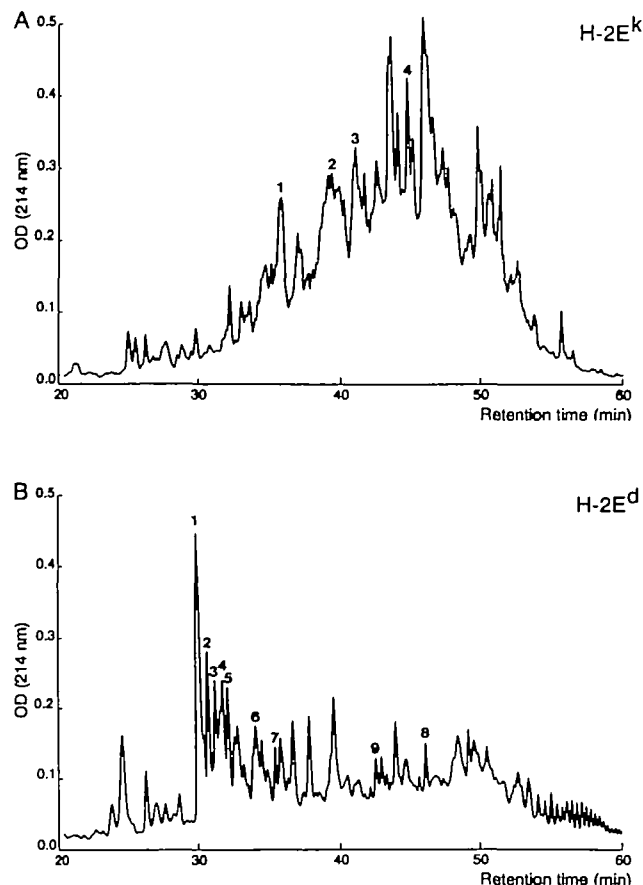
As a source for E<sup>k</sup> molecules, CH27 cells or wt-Ek cells (20) were used. E<sup>d</sup> molecules were purified from A20 cells (21) or LQ-9E10 cells (provided by Dr D. Vidović). For bulk culture, cells were grown in 2 l glass roller bottles (Schott Duran, Frankfurt, Germany) at 37°C with up to 1.5 l RPMI 1640 medium supplemented with 3.5% FCS, 3.5% newborn calf serum, glutamine, penicillin/streptomycin and 2-mercaptoethanol. Cells were split every 3–4 days. For harvest, (typically 24–28 l) cell suspensions were pelleted with a continuous flow rotor (Beckman, Munich, Germany). Cell pellets were either used directly for peptide extraction or stored at –70°C.

### Antibodies

Immunoprecipitations were performed with 14-4-4S antibodies (22) specific for H-2E<sup>k,d,r,s</sup>. Hybridomas were grown in roller culture. Supernatant was filtered and the pH adjusted to 7.5. The antibodies were further purified on a Protein A-Sepharose column (Pharmacia, Freiburg, Germany) according to standard protocols.

### MHC precipitation and peptide extraction

This was done as described (18). Briefly, pellets (20–40 ml) were lysed in 1% (v/v) NP-40, 1 mM phenylmethylsulphonyl fluoride, 0.02% (w/v) pepstatin, 0.02% (w/v) leupeptin and 0.02% (w/v) aprotinin (Boehringer, Mannheim, Germany). The lysate was ultracentrifuged for 60 min at 86,000 *g*. The supernatant was passed over Sepharose beads (Pharmacia) coupled with glycine and then over Sepharose beads coupled with 14-4-4S antibodies. Peptides were acid-released from the loaded beads by treatment with 0.1% (v/v) trifluoroacetic acid (TFA). In order to remove antibody molecules and E $\alpha$ - and E $\beta$ -chains from the peptides, the solution was passed through ultrafiltration membranes (Centricon 10; Amicon, Witten, Germany) of 10,000 Da exclusion size. The filtrate



**Fig. 1.** HPLC separation of H-2E ligands. Peptides eluted from immunoprecipitated H-2E<sup>k</sup> (A) and H-2E<sup>d</sup> (B) molecules were separated by reversed-phase HPLC. The numbers indicate peaks that were sequenced individually. The sequences are listed in Tables 2 and 4. Ten percent of the volume of the numbered peaks was also used for pool sequencing. Peaks appearing in the glycine control (not shown) were omitted from the pools.

was separated on a reversed-phase HPLC column ( $\mu$ RPC C2/C18; 2.1  $\times$  100 mm; Pharmacia), using a SMART system (Pharmacia). Eluent A was 0.1% (v/v) TFA in water, eluent B was 0.08% (v/v) TFA in 80% acetonitrile/water. All eluents used were of HPLC grade (Riedel de Haen, Frankfurt, Germany). The gradient used was the same as previously described (14). Peptide fractions were pooled, whereby dominant single peaks were used only to 10%. Pools and dominant single peaks were sequenced separately.

### Peptide analysis

Pools and single peaks were sequenced by Edman degradation on a pulsed-liquid protein sequencer 476A (Applied Biosystems, Weiterstadt, Germany) as described (23). Cys was not modified and therefore not detectable.

### Interpretation of pool sequencing

In the raw data tables, amino acids are ordered according to their hydrophathy (24). Ordering the amino acids in this way facilitates perception of certain physico-chemical groups clustering in cycles of Edman degradation. We use two arbitrary

**Table 1.** Pool sequencing of H-2E<sup>k</sup> extracted peptides

Cycle	Amino acid residues (in pmol)																			
	I Ile	V Val	L Leu	F Phe	W Trp	M Met	A Ala	G Gly	T Thr	S Ser	Y Tyr	P Pro	H His	Q Gln	N Asn	E Glu	D Asp	K Lys	R Arg	
1	25.4	42.2	41.8	14.2	4.1	3.8	28.3	33.1	25.8	8.3	10.6	4.2	21.9	12.3	7.6	21.2	14.1	17.2	2.9	
2	13.6	34.2	26.8	11.6	4.4	<b>16.0</b>	<u>41.8</u>	22.3	10.3	7.5	6.5	<b>140.9</b>	7.5	14.4	<b>15.3</b>	22.1	<u>15.8</u>	19.4	2.3	
3	<u>18.6</u>	43.7	23.2	7.2	4.5	7.6	<u>22.8</u>	<u>30.9</u>	<b>24.9</b>	7.3	6.1	88.6	8.3	19.6	13.8	<u>23.7</u>	<u>20.8</u>	17.2	<b>7.4</b>	
4	<u>31.1</u>	29.4	<b>59.4</b>	7.3	4.8	<b>17.4</b>	<u>26.3</u>	<u>23.5</u>	15.6	5.3	<b>12.1</b>	40.6	<b>16.8</b>	<u>22.3</u>	<u>14.4</u>	<u>31.1</u>	<u>25.1</u>	14.2	7.5	
5	<b>101.7</b>	<b>74.9</b>	67.4	8.4	4.8	6.6	14.9	10.9	10.5	4.8	7.0	26.0	16.4	13.1	8.0	16.9	23.3	16.0	6.6	
6	55.5	40.6	<u>37.2</u>	<b>28.1</b>	4.5	6.7	<u>19.4</u>	<u>13.3</u>	<u>13.8</u>	<u>8.6</u>	8.7	13.0	16.1	<u>18.8</u>	<b>14.3</b>	<b>50.5</b>	<u>24.7</u>	<u>20.5</u>	<u>8.9</u>	
7	<b>91.5</b>	<u>51.2</u>	29.3	<u>38.9</u>	3.6	10.0	<u>21.6</u>	<u>15.5</u>	4.9	<u>10.2</u>	5.6	9.5	14.6	<u>11.8</u>	8.4	29.8	<u>18.3</u>	<u>16.8</u>	<u>9.3</u>	
8	<u>63.3</u>	<u>41.3</u>	<u>27.0</u>	<u>35.8</u>	2.9	<u>8.1</u>	<u>29.7</u>	<u>12.5</u>	4.7	<u>16.3</u>	5.2	9.7	11.3	<b>29.4</b>	10.1	23.8	<u>17.3</u>	<u>18.3</u>	<u>9.5</u>	
9	49.2	34.3	26.4	<u>18.4</u>	2.3	<b>32.7</b>	<u>22.4</u>	12.7	5.1	9.9	3.8	<u>14.3</u>	8.0	<u>31.5</u>	11.6	18.0	12.5	<u>24.4</u>	<b>22.6</b>	
10	28.1	26.5	<b>40.0</b>	14.5	1.9	24.0	<u>30.8</u>	<u>14.5</u>	5.0	<u>10.6</u>	3.6	<u>14.6</u>	<u>11.5</u>	<u>38.7</u>	<b>19.3</b>	14.2	9.9	14.0	18.5	
11	24.7	19.8	<u>54.0</u>	9.5	1.7	19.6	<u>20.9</u>	<u>14.5</u>	3.3	7.2	4.3	<u>13.0</u>	11.9	<u>19.3</u>	<u>14.8</u>	12.9	7.9	<b>30.1</b>	<u>18.5</u>	
12	20.7	15.8	<u>36.7</u>	<u>11.9</u>	1.5	15.3	<u>20.9</u>	<u>20.4</u>	2.3	4.8	3.8	<u>14.9</u>	7.8	<u>15.9</u>	<b>27.6</b>	13.1	7.6	<u>36.3</u>	<b>28.8</b>	
13	14.2	11.0	<u>28.7</u>	<u>8.7</u>	1.3	13.8	<u>12.9</u>	<u>27.6</u>	1.3	3.6	2.9	<u>11.6</u>	6.0	8.5	<u>20.1</u>	6.9	6.3	<b>57.8</b>	<u>32.8</u>	
14	10.3	<u>19.0</u>	23.8	6.8	1.1	12.1	10.0	<u>22.6</u>	2.6	3.6	1.9	9.3	4.8	6.4	13.4	7.0	7.3	<u>41.2</u>	<u>35.9</u>	
15	7.8	<u>12.7</u>	15.5	4.9	1.0	11.8	7.2	<u>18.3</u>	2.1	2.7	2.8	8.3	3.2	6.2	9.4	5.5	8.4	<u>41.1</u>	<u>27.9</u>	
16	<b>15.3</b>	7.7	10.2	4.8	0.9	7.6	5.0	<u>13.2</u>	2.9	2.5	4.6	8.1	3.9	5.8	6.1	4.3	7.1	28.4	20.0	

Amino acids are ordered according to their hydrophathy (24). **Bold** numbers indicate increases of >50% as compared with the pmol amounts of the same PTH-amino acid at the previous position. Underlined numbers indicate smaller increases or decreases lower than expected from tailing effects.

levels of significance for evaluation of data. An increase of the amount of a given PTH amino acid at a given cycle of >50% as compared with the previous cycle is considered highly significant as indicated in **bold**. An increase <50% is considered to reflect a significant signal and is indicated by underlining the respective numbers. In some instances, a decrease smaller than the lag effect usually found for a given amino acid is taken as indication that some of the peptides in the pool have this amino acid. Such values are also underlined.

## Results

### Analysis of H-2E<sup>k</sup> eluted peptides

CH27 cells, expressing E<sup>k</sup> molecules, were grown to high cell numbers, lysed with detergent and E<sup>k</sup> molecules were precipitated with solid-phase bound mAb 14-4-4S. Bound MHC molecules were treated with TFA to elute associated peptides and the peptide containing supernatants were separated by reversed-phase HPLC (Fig. 1A). Peptides eluting in large distinct peaks were sequenced individually by Edman degradation. The remaining fractions were pooled and sequenced as such. As observed with pool sequencing of other class II ligands (14,15), a strong Pro signal was found at cycle 2, followed by clusters of hydrophobic, polar and charged residues (Table 1) that reflect the anchor positions used by the ligands to bind to the class II molecule (14,15). The appearance of anchor residues in clusters stretching over three cycles rather than as a sharp peak confined to one cycle is probably caused by variation in the number of amino acids between the N-terminus and the first anchor position. One evident cluster appearing after the Pro in cycle 2 is observed in cycles 4 and 5 consisting of aliphatic amino acids like Val, Leu and Ile. These are considered to represent the first anchor position. The Tyr signal in cycle 4 is sharp

and therefore unlikely to be part of an anchor residue. After this, a cluster of hydrophobic amino acids, including Phe, but also small and polar amino acids like Ser, appears in cycles 7 and 8, considered to represent the second anchor position. The strong and unclustered Glu signal in cycle 6 is part of the sequence of peptide 3 in Table 2 and is therefore not considered as part of an anchor. In cycles 9 and 10, a significant increase of the polar residue Gln becomes evident, with still some Asn and Ala contribution, forming the putative third anchor position. Cycles 11–14 reflect the putative fourth anchor position, which is dominated by strong Lys and Arg signals, with some contribution of Gly. This information, together with the alignment of peptides derived from single peak sequences, leads to the basis for the detailed E<sup>k</sup> motif in Table 2, with four anchors at relative positions 1, 4, 6 and 9. Position 1 and 9 seem to be the most stringent ones, allowing mainly aliphatic residues at position 1 and positive charged residues at position 9. The anchors at positions 4 and 6 and the correlation of all anchor residues to the predicted E<sup>k</sup> pockets according to the DR1 crystal structure will be discussed later.

### Analysis of E<sup>d</sup> eluted peptides

E<sup>d</sup> associated peptides were eluted from large batches of A20 or LQ-9E10 cells using the same antibody and procedure as described for E<sup>k</sup> bound peptides (Fig. 1B). Again, a very strong Pro signal can be seen in cycle 2 (Table 3). After this, hydrophobic residues, including the large, aromatic amino acids, form the cluster interpreted to reflect the first anchor position in cycles 4–6. From cycles 6 to 14 strong signals for Lys and Arg dominate with the most significant values in cycles 11–13. The very strong Gln signals in cycles 7 and 8 are most likely caused by dominant single peptides, which also contributed to the strong Lys signals in cycles 8 and 9. In addition to the residues thought to reflect the main anchors,

**Table 2.** H-2E<sup>k</sup> motif and ligands

	Relative position									Peak no.															
	1	2	3	4	5	6	7	8	9																
Anchor residues	I			I		Q			K																
	L			L		N			R																
	V			V		Δ																			
	F			F		G																			
	Y			S																					
	W																								
Natural ligands																									
β <sub>2m</sub> 42-56	H	P	P	H	I	E	I	Q	M	L	K	N	G	K	K					a					
HSC70 234-248			D	N	R	M	V	H	F	I	A	E	F	K	R	K				a					
M.ser.alb. 347-361			T	P	T	L	V	E	A	A	R	N	L	G	R	V	G			a					
C'inhib. 41-55			V	N	K	E	I	Q	N	A	V	Q	G	V	K	H	I			a					
ER60 448-461			G	F	P	T	I	Y	F	S	P	Δ	N	K	K	L				b					
Unknown				I	P	L	I	M	L	I	N	K	A	R	N	K	A	E		b					
α1-Antitryp. 397-410			Y	D	R	N	T	K	S	P	L	F	V	G	K	V				1					
Unknown							F	A	E	E	G	T	L	K	K	A	A	V	H	Y	4				
(Human) dead box protein					L	H	L	G	Y	L	P	N	Q	L	F	R					3				
Bovine fetuin 342-	I	P	G	G	P	V	R	L	C	P	G	R	I	R							2				
T cell epitopes <sup>c</sup>																									
MCC 91-103				R	A	D	L	I	A	Y	L	K	Q	A	T	K									
PCC 91-104				R	A	D	L	I	A	Y	L	K	Q	A	T	A	K								
λrep 12-26	L	E	D	A	R	R	L	K	A	I	Y	E	K	K	K										
SWMb 66-78				V	T	V	L	T	A	L	G	Δ	I	L	K	K									
EqMb 69-77							L	T	A	L	G	G	I	L	K										
MoMb 69-77							L	T	A	L	G	G	I	L	K										
MoHb 68-76							I	T	A	E	N	E	G	L	K										
HEL 1-18				K	V	F	G	R	C	E	L	A	A	Δ	M	K	R	H	G	L	D				
HEL 81-96			S	A	L	L	S	S	D	I	T	A	S	V	N	C	A	K							
HEL 108-119									W	V	A	W	R	N	R	C	K	G	T	D					
SNase 51-70						V	E	K	Y	G	P	E	A	S	Δ	F	T	K	K	M	V	E	N	A	K
SNase 81-100	R	T	D	K	Y	G	R	G	L	A	Y	I	Y	Δ	D	G	K	M	V	N					
SNase 121-140				H	E	Q	H	L	R	K	S	E	Δ	Q	A	K	K	E	K	L	N	I	W		
LLO 218-226									I	A	K	E	G	T	A	F	K								

<sup>a</sup>Sequences adapted from ref. 5.<sup>b</sup>Sequences obtained from extraction not shown in Fig. 1 (A).<sup>c</sup>T cell epitopes are from refs 10 and 25.**Table 3.** Pool sequencing of H-2E<sup>d</sup> extracted peptides

Cycle	Amino acid residues (in pmol)																		
	I Ile	V Val	L Leu	F Phe	W Trp	M Met	A Ala	G Gly	T Thr	S Ser	Y Tyr	P Pro	H His	Q Gln	N Asn	E Glu	D Asp	K Lys	R Arg
1	41.7	44.4	109.2	12.5	3.4	5.5	50.4	44.6	21.9	31.8	15.8	5.7	5.1	11.4	3.1	46.2	23.6	13.2	15.9
2	27.1	38.3	45.5	11.8	2.8	5.3	<b>126.3</b>	31.8	16.9	17.9	12.4	<b>57.1</b>	4.6	<b>58.1</b>	<b>6.2</b>	64.7	<u>35.4</u>	16.9	8.9
3	17.8	27.1	<b>141.4</b>	<b>27.0</b>	3.2	<b>10.0</b>	50.6	<b>55.7</b>	13.1	14.4	<b>28.8</b>	31.7	5.7	31.5	5.8	<u>103.7</u>	<u>29.3</u>	16.0	7.7
4	<u>23.0</u>	19.4	83.4	<u>35.0</u>	4.8	<u>11.5</u>	<b>114.6</b>	31.3	11.3	9.9	<b>55.1</b>	25.1	<b>9.1</b>	19.2	5.5	<u>87.1</u>	27.8	15.2	<b>20.0</b>
5	<u>33.3</u>	<b>45.1</b>	<u>96.3</u>	<u>49.4</u>	<b>33.2</b>	5.5	39.0	22.7	12.5	7.0	51.5	22.3	11.6	15.3	4.5	41.9	21.4	12.4	15.7
6	<u>34.5</u>	27.2	45.6	27.9	18.5	5.5	24.4	13.9	11.4	9.9	<u>61.5</u>	<u>23.0</u>	<u>14.6</u>	<u>21.8</u>	<b>8.6</b>	34.7	19.4	<b>40.9</b>	<b>48.2</b>
7	<b>54.3</b>	<u>30.9</u>	36.9	18.6	19.5	4.4	20.5	<b>28.4</b>	8.0	<b>21.3</b>	22.4	21.7	15.0	<b>64.8</b>	8.4	<u>35.6</u>	19.6	29.7	38.7
8	41.6	<u>34.5</u>	28.0	17.1	<u>13.4</u>	<u>6.6</u>	20.5	16.1	7.9	11.1	22.1	18.2	12.4	<u>55.2</u>	<u>8.6</u>	<u>27.6</u>	16.5	<b>108.7</b>	<u>47.3</u>
9	<u>23.0</u>	<u>45.6</u>	<b>87.8</b>	<b>32.0</b>	7.4	<u>7.6</u>	21.6	14.8	8.5	<b>19.4</b>	14.9	14.8	<b>30.6</b>	27.0	6.9	17.2	14.5	81.4	<u>47.6</u>
10	<b>47.3</b>	<u>34.8</u>	47.7	13.0	4.4	4.7	14.4	<u>17.1</u>	<u>10.0</u>	12.6	15.1	13.1	<u>50.4</u>	17.4	6.8	<u>29.4</u>	11.5	<u>38.9</u>	<u>60.0</u>
11	<u>38.2</u>	27.2	39.2	11.6	2.9	3.8	12.4	<b>46.6</b>	8.7	9.3	10.7	<u>17.7</u>	<u>35.7</u>	<u>18.7</u>	6.9	16.1	11.0	<b>54.7</b>	<u>75.0</u>
12	<u>47.5</u>	<u>27.5</u>	<u>30.6</u>	10.2	2.3	3.0	<u>17.5</u>	<u>31.5</u>	6.8	7.1	10.5	<u>20.7</u>	<u>43.2</u>	<u>20.6</u>	6.3	<u>29.4</u>	9.9	<u>52.7</u>	<u>77.7</u>
13	<u>31.3</u>	<u>24.6</u>	<b>56.9</b>	8.4	2.2	3.5	<u>15.3</u>	<u>21.8</u>	<u>10.2</u>	4.3	<u>12.4</u>	<u>18.2</u>	<u>18.0</u>	<u>11.4</u>	<u>8.7</u>	<u>20.6</u>	7.8	<u>67.7</u>	<u>79.9</u>
14	17.3	<u>13.8</u>	29.3	6.1	1.9	2.4	13.7	17.9	8.4	4.7	4.1	12.7	16.2	8.3	<u>7.2</u>	18.3	8.2	<u>58.6</u>	42.7
15	13.9	9.3	15.8	4.4	1.6	1.7	11.2	15.4	6.9	5.1	5.8	11.2	15.1	7.5	7.0	18.0	7.9	<u>38.5</u>	43.5
16	10.5	8.0	8.9	4.0	1.3	1.6	8.5	11.8	5.8	4.6	5.2	9.4	13.9	6.2	5.8	16.5	8.5	<u>32.8</u>	41.2

Organization of data as in Table 1.

**Table 4.** H-2E<sup>d</sup> motif and ligands

	Relative position									Peak no.														
	1	2	3	4	5	6	7	8	9															
Anchor residues	<b>W</b>			<b>K</b>		<b>I</b>			<b>K</b>															
	<b>Y</b>			<b>R</b>		<b>L</b>			<b>R</b>															
	<b>F</b>			<b>I</b>		<b>V</b>																		
	<b>I</b>					<b>G</b>																		
	<b>L</b>																							
	<b>V</b>																							
Natural ligands																								
IL-2R. $\gamma$ chain 168-182	S	Q	L	E	L	R	<b>W</b>	K	S	<b>R</b>	H	<b>I</b>	K	E	<b>R</b>	2,5,7								
170-182				L	E	L	<b>W</b>	K	S	<b>R</b>	H	<b>I</b>	K	E	<b>R</b>	1,4								
Apo-E prec. 222-236				E	R	A	E	A	<b>W</b>	R	Q	<b>K</b>	<b>L</b>	H	G	<b>R</b>	3							
223-236				R	A	E	A	<b>W</b>	R	Q	<b>K</b>	<b>L</b>	H	G	<b>R</b>	L	1							
Unknown				A	Q	<b>F</b>	M	<b>W</b>	<b>I</b>	<b>I</b>	R	K	<b>R</b>	I	Q	L	P	8						
Similar apolipoprotein B	S	L	D	E	H	<b>Y</b>	H	<b>I</b>	<b>R</b>	V	H	L	V	<b>K</b>			6							
Human 2211-2224																								
<i>C. elegans</i> cDNA homolog 74-87	G	Q	F	Y	<b>F</b>	L	<b>I</b>	<b>R</b>	K	R	<b>I</b>	H	L	R			9							
Actin B 8-				L	V	<b>V</b>	D	N	G	S	<b>G</b>	M	C	<b>K</b>	A	G	F	#						
T cell epitopes <sup>a</sup>																								
Ig lambda 91-108							A	L	W	F	R	N	H	F	V	F	G	G	T	K	V	T	V	
SWM 102-118	K	Y	L	E	F	I	S	E	A	<b>I</b>	<b>I</b>	H	V	L	H	S	<b>R</b>							
SWM 132-146				N	K	A	L	E	L	F	R	<b>K</b>	D	<b>I</b>	A	A	<b>K</b>	Y						
HEL 108-119							W	V	A	W	<b>R</b>	N	R	C	<b>K</b>	G	T	D						
SNase 112-129							A	Y	V	Y	<b>K</b>	P	N	N	T	H	E	Q	H	L	R	K	S	E
FLU PR/8 HA 109-119							S	S	<b>F</b>	E	R	F	E	<b>I</b>	F	P	<b>K</b>							
$\lambda$ rep 12-26																								
HIV-1 gag p17 17-28	L	E	D	A	R	R	L	K	A	<b>I</b>	Y	E	K	K	<b>K</b>									
							E	K	<b>I</b>	R	L	<b>R</b>	P	<b>G</b>	G	K	<b>K</b>	K						

<sup>a</sup>T cell epitopes adapted from ref. 24.

a cluster of aliphatic and small amino acids can be observed in cycles 7-12. After combining this pool sequencing information with the result of sequencing individual peptides and the alignment of T cell epitopes, a ligand motif for E<sup>d</sup> molecules becomes evident which has the requirement for basic residues at relative position 9 in common with the E<sup>k</sup> motif (Table 4). The other anchor residues, however, seem to differ between E<sup>k</sup> and E<sup>d</sup> motifs. Whereas aromatic residues are not dominant at the first anchor position of E<sup>k</sup> ligands, the former are favored in E<sup>d</sup>-associated peptides. Also interesting is the high frequency of positively charged residues found at position 4 (cycles 6-9), which is not observed in peptides eluted from E<sup>k</sup> molecules. On the other hand, polar residues like Gln and Asn, which form the anchor at position 6 (cycles 9-12) for E<sup>k</sup> ligands, are not significantly represented at this position in E<sup>d</sup> bound peptides. Instead, aliphatic residues can be found here. More details of the motif and differences between E<sup>k</sup> and E<sup>d</sup> will be discussed below, especially in light of the pocket compositions predicted for E<sup>k</sup> and E<sup>d</sup> molecules.

**Discussion**

In this study, we have analyzed natural peptide ligands eluted from mouse class II E molecules in order to identify allele-specific motifs. Pool sequencing analysis of peptides eluted from E<sup>k</sup> and E<sup>d</sup> molecules revealed clusters of amino acids that aligned well with sequences of individual peptides, and allowed us to identify anchors at relative positions 1, 4, 6 and 9. In the case of the E<sup>k</sup> motif, our results fit well to the motif

derived by binding studies with 220 substitutions of the E<sup>k</sup>-restricted T cell epitope moth cytochrome c (MCC) 93-103 (10,11), where positions 1, 6 and 9 were identified to be critical for binding. However, the anchor position at P4 could not be identified in those studies, probably because of a weaker contribution to the overall affinity of the peptide and the fact that in peptide binding studies with mutated peptides only one position can be studied at one time. Optimal amino acid selections at several anchor positions can therefore mask the influence of suboptimal amino acid selections at a single anchor position. This might also explain why many replacements were tolerated at the P1, P6 and P9 anchor positions in the peptide binding assay, whereas our sequence analysis showed a more stringent selection of amino acids among natural ligands, especially at relative positions 1 and 9. At position 1, we find the aliphatic amino acids Val, Leu and Ile to be clearly favored over the aromatic residues. At position 9, the basic residues Lys and Arg dominate with a small contribution of Gly. Relative position 6 is mainly occupied by large, polar residues like Asn or Gln, but smaller polar groups like Ser and Thr, in addition to small hydrophobic residues and Ala, are also found. For this anchor position, peptide binding studies and pool sequencing are in perfect agreement. We find an additional anchor at relative position 4 which is mainly formed by aliphatic and small polar amino acids, but also Phe. The identification of P4 as an anchor position is in agreement with the results obtained by Reay *et al.* (10). When testing MCC peptide substitutions for the influence on TCR interactions they found that only substitutions

with V, L, I and F, which we observe at P4, preserved optimal T cell stimulatory capacity. All 24 E<sup>k</sup> restricted natural ligands and T cell epitopes in Table 2 follow this ligand motif with at least three optimal amino acids at the anchor positions P1, P4, P6 and P9, with P1 and P9 being the most stringent positions.

The analysis of E<sup>d</sup> eluted peptides revealed anchors at identical positions, but interestingly with different amino acid preferences at positions 1, 4 and 6. At relative position 1, bulky aromatic residues are more prominent than the smaller hydrophobic amino acids, indicating that the corresponding pocket in the E<sup>d</sup> molecule might be more spacious. The anchor at relative position 4 shows, besides weak signals for Val and Ile, a very strong preference for Arg and Lys that is not observed with peptides eluted from E<sup>k</sup> molecules. A contribution of positively charged amino acids to the P4 anchor of E<sup>d</sup> ligands has been observed previously (30). Sequencing cycles 9–11, which correspond to the third pocket at relative position 6, show an increase of signals for aliphatic residues and Gly, but no significant contribution of polar amino acids like Gln and Asn. This finding might indicate that polar interactions with the E<sup>d</sup> molecule at this position are not favored, as opposed to interactions of peptides with E<sup>k</sup> molecules. The anchor at position 9 seems to be identical to the E<sup>k</sup> motif and is therefore again dominated by Lys and Arg residues. Fourteen out of the 15 E<sup>d</sup>-restricted natural ligands and T cell epitopes follow this proposed ligand motif with the optimal amino acid selection at at least three of the four anchor positions with, analogous to the E<sup>k</sup> motif, the most stringent anchors at P1 and P9.

Since mouse class II E molecules are predicted to have a high structural homology to human DR molecules and the three-dimensional structure of the DR1 is known, we tried to correlate E<sup>k</sup> and E<sup>d</sup> motifs obtained from pool and individual peptide sequencing with the composition of the putative pockets in class II E molecules. The residues predicted to form the first pocket in E molecules differ from each other and from DR1 only at position  $\beta$ 86. DR1 molecules have a Gly at this position (27), E<sup>k</sup> molecules a Phe and E<sup>d</sup> molecules a Ser (28). DR1 binding peptides show a preference for aromatic residues at the first anchor position, which is also the case for E<sup>d</sup> binding peptides, where  $\beta$ 86 is also occupied by a small amino acid. Peptides from E<sup>k</sup> molecules, on the other hand, have a strong preference for aliphatic residues, which can be well explained by the dramatic reduction of the pocket volume by the bigger side chain  $\beta$ 86Phe. A similar influence has been observed for peptides presented by DR5, DR4Dw10 and DR4Dw14 where position  $\beta$ 86 is occupied by Val (14,17).

The second pocket, which accommodates P4 in peptides binding to DR1 molecules, has a more degenerate specificity than the first pocket, allowing aliphatic and also polar residues, but does not favor positively charged amino acids. The residues that are polymorphic between DR1 and E<sup>k</sup> and predicted to be part of the second pocket are  $\beta$ 78 (Tyr → Val),  $\beta$ 13 (Phe → Ser),  $\beta$ 74 (Ala → Glu) and  $\beta$ 71 (Arg → Lys) (26,27). These changes are in agreement with the appearance of aliphatic and polar residues at P4 of E<sup>k</sup> ligands. The change of Ala to Glu at  $\beta$ 74 can also explain why a basic residue at P4 in some E<sup>k</sup>-restricted T cell epitopes can be tolerated.

Basic residues are not favored in DR1 binding peptides (19), presumably because of unfavorable electrostatic interactions with  $\beta$ 71 Arg, which  $\beta$ 74 Glu in E<sup>k</sup> molecules might compensate for. The second pocket in E<sup>d</sup> molecules has polymorphic changes at positions  $\beta$ 13 (Phe → Ser),  $\beta$ 74 (Ala → Ser),  $\beta$ 70 (Gln → Glu) and  $\beta$ 71 (Arg → Ala) (26,27), giving a more negative character to this pocket. This might explain why, in addition to some aliphatic residues, positively charged amino acids are favored at P4 of E<sup>d</sup> bound peptides, whereas negatively charged residues are not observed. Similar effects have been observed with the composition of P4 residues in peptides released from DR4Dw4 and DR4Dw10 molecules (17), where  $\beta$ 71 is changed from Lys to Glu. The increase of the Glu signals in cycles 10 and 12 in Table 3 is not considered significant for an anchor, since it can be read as part of the sequence of the IL-2R  $\gamma$  chain peptide (Table 4).

The third pocket identified in DR1 molecules makes contact to the amino acid at relative position 6 of the peptide and prefers small residues. The polymorphic residues predicted to be responsible for the specificity of this pocket in E<sup>k</sup> molecules are  $\beta$ 13 (Phe → Ser) and  $\beta$ 11 (Leu → Cys), which should give a polar character to this pocket. That this prediction is correct can be seen from the strong increase of the signals for polar residues, like Gln and Asn in cycles 8–11, building the cluster for P6 in E<sup>k</sup> bound peptides (Table 1). A small increase for Ser and Thr is observed as well. That the bigger polar residues are favored over the small ones can be explained by the expected increase in the pocket volume caused by the Leu → Cys change. In the corresponding pocket in E<sup>d</sup> molecules,  $\beta$ 11 Cys is changed to Val which should make this pocket less polar and more hydrophobic. Indeed, the pool sequencing signals for polar residues in E<sup>d</sup> bound peptides show no significant increase, whereas strong signals for Val, Leu and Ile can be observed in cycles 8–11 (Table 3), suggesting that P6 in E<sup>d</sup> ligands are preferentially occupied by aliphatic residues.

The fourth pocket described for DR1 is very degenerate and no specificity could be detected at relative position 7 for DR1 binding peptides and natural ligands (12,14,19). Similarly, the alignment of single peptides and pool sequence information revealed no specific interactions at P7 for E<sup>k</sup> bound peptides. Of interest might be that the alignment of E<sup>d</sup> bound peptides shows a high frequency of positively charged amino acids at P7 which probably contributed to the very broad signals for Lys and Arg (cycles 6–14) in the pool sequencing analysis. Whether or not these signals reflect an anchor residue at P7 can not be confirmed by the putative nature of the predicted corresponding pocket, because H-2E molecules differ from DR molecules at  $\beta$ 65 where Lys is replaced by Pro, which is expected to alter the pocket structure dramatically. The pocket number 5 for P9 of DR1 ligands is reported to have a preference for aliphatic residues, allowing aromatic and polar residues as well (12,14). Both E<sup>k</sup> and E<sup>d</sup> molecules differ in their predicted pocket only in one position,  $\beta$ 9. DR1 molecules have Trp and all H-2E  $\beta$ -chains have a Glu (28). This difference is reflected in a strong clustering of positively charged residues in cycles 12–14 in the pool sequencing analysis of E<sup>k</sup> and E<sup>d</sup> eluted peptides (Tables 1 and 3). Weak signals for aliphatic residues are found as well, whereas P9 is clearly dominated by Arg and

**Table 5.** Relations between HLA-DR and predicted H-2E pocket composition and preference for ligand residues

Class II molecule	Pocket for relative position			
	P1	P4	P6	P9
DRB1*0101 residues	86 Gly, 85 Val	13 Phe, 70 Gln, 71 Arg, 74 Ala, 78 Tyr	11 Leu, 13 Phe	9 Trp
ligand specificity	<b>Y, V, L, F, I, A, M, W</b>	<b>L, A, I, V, M, N, Q</b> ; no positive charge	<b>A, G, S, T, P</b>	<b>L, A, I, V, F, Y</b>
H-2E <sup>k</sup> β chain polymorphic residues	86 Phe, 85 Ile	13 Ser, 70 Gln, 71 Lys, 74 Glu, 78 Val	11 Cys, 13 Ser	9 Glu
Ligand specificity	<b>I, L, V, F, Y, W</b>	<b>I, L, F, S, A</b> ; R, K, D, E allowed	<b>Q, N, S, T, A, H, R, E</b>	<b>K, R, G</b>
H-2E <sup>d</sup> β chain polymorphic residues	86 Ser, 85 Ile	13 Ser, 70 Glu, 71 Ala, 74 Ser, 78 Tyr	11 Val, 13 Ser	9 Glu
ligand specificity	<b>W, Y, F, I, L, V</b>	<b>R, K, I, V</b> ; no negative charge	<b>I, L, V, G</b>	<b>K, R</b>
H-2E <sup>b</sup> β chain polymorphic residues	86 Ser, 85 Ile	13 Ser, 70 Gln, 71 Lys, 74 Glu, 78 Val	11 Cys, 13 Ser	9 Glu
predicted ligand specificity	<b>W, Y, F, I, L, V</b>	<b>I, L, F, S, A</b> ; R, K, D, E allowed	<b>Q, N, S, T, A, H, R, E</b>	<b>K, R</b>
H-2E <sup>s</sup> β chain polymorphic residues	86 Leu, 85 Ile	13 Ser, 70 Gln, 71 Arg, 74 Ala, 78 Val	11 Ser, 13 Ser	9 Glu
predicted ligand specificity	<b>I, L, V, F, Y, W</b>	<b>L, I, V, T, S, N, Q</b> ; no positive charge	<b>Q, N, S, T, A, H, R, E</b>	<b>K, R</b>

Lys. A similar preference of basic amino acids at the P9 anchor of isolated ligands is observed for the HLA-DRB5\*0101 molecule, where the corresponding pocket has a negative character (29). Also interesting is the finding that some E<sup>d</sup> ligands seem to have the amino acid that interacts with the fifth pocket of the MHC molecule shifted from P9 to P8 (Table 4), whereas E<sup>k</sup> ligands show in some cases a shift from P9 to P10 (Table 2). Similar observations have been made for ligands of class I MHC molecules as well (1). Whether or not this reflects differences in the pocket structures caused by polymorphic residues can only be determined by comparing the crystal structure of both molecules. The influences of the predicted pockets of E<sup>k</sup> and E<sup>d</sup> molecules on the selection of the amino acids at the anchor positions 1, 4, 6 and 9 are again summarized in Table 5.

Since natural ligands of E<sup>s</sup> and E<sup>b</sup> molecules have been identified (3,5), we also attempted pocket and motif predictions (lower part of Table 5) and alignment of peptides for these class II molecules (Table 6). The first predicted pocket of the E<sup>s</sup> molecule differs from DR1, E<sup>k</sup> and E<sup>d</sup> molecules at position β86. Leu at this position should decrease the pocket volume compared with Gly in DR1 and Ser in E<sup>d</sup> molecules, therefore allowing, analogous to the E<sup>k</sup> motif, mainly aliphatic residues in P1. This is observed in four out of the five peptide isolated from E<sup>s</sup> molecules. E<sup>b</sup> molecules are identical to E<sup>d</sup> molecules with regard to the amino acids participating in the first predicted pocket. Aromatic amino acids should be favored at P1, as we observed for E<sup>d</sup> molecules. Indeed, all three natural E<sup>b</sup> ligands carrying Tyr at P1. The pocket for P4 in E<sup>s</sup> molecules differs from DR1 at β78 (Tyr → Val) and β13 (Phe → Ser). This should still allow the binding of aliphatic and polar residues, and should disfavor the presence of positively charged residues, very similar to the amino acids observed at P4 for DR1 molecules (14). All six E<sup>s</sup> eluted peptides fulfill this requirement. The second pocket of E<sup>b</sup> molecules is identical to that of E<sup>k</sup> molecules, where P4 of the peptides is occupied by hydrophobic, polar but also charged residues. The small number of known E<sup>b</sup> eluted peptides have residues at P4 which belong to these groups of amino acids. The third pocket, accommodating P6 of the

peptide, differs in both E<sup>s</sup> and E<sup>b</sup> molecules from DR1 at position β11, where Leu is replaced by Ser (E<sup>s</sup>) or Cys (E<sup>b</sup>). This exchange should give the pocket, as in E<sup>k</sup> molecules, more room and a more polar character, allowing polar and slightly bigger amino acids compared with DR1. The fourth pocket for P7, which is very degenerate in DR1 molecules, is probably similar in E<sup>s</sup> molecules and differs in E<sup>b</sup> molecules only at β71, where Arg is replaced by Lys. This exchange is not expected to alter the pocket specificity, resulting in a very degenerate specificity at P7 of E<sup>s</sup> and E<sup>b</sup> binding peptides.

The fifth predicted pocket of E<sup>s</sup> and E<sup>b</sup> molecules is identical to E<sup>d</sup> and E<sup>k</sup>, which differ from DR1 at β9 where Trp is changed to Glu. Therefore, analogous to E<sup>d</sup> and E<sup>k</sup>, a basic residue is expected at this position. The natural E<sup>b</sup> and E<sup>s</sup> ligands can be easily aligned according to this prediction. This consideration on predicted E<sup>b</sup> and E<sup>s</sup> motifs should be useful for designing peptides for binding assays, in order to verify or modify the prediction. In support of our predicted similarity of E<sup>k</sup> and E<sup>s</sup> ligand motifs is the finding of Marrack *et al.* (11) who could isolate the same β<sub>2</sub>-microglobulin (β<sub>2</sub>m) peptide as a natural ligand from E<sup>k</sup> and E<sup>s</sup> molecules.

We show that sequencing of peptide pools and individual ligands provides detailed information on allele-specific motifs of E<sup>k</sup> and for the first time, E<sup>d</sup> ligands, allowing to understand the different requirements for peptides to bind to these two H-2E molecules. The motifs can be aligned to pockets in class II E molecules, predicted according to the crystal structure of the HLA-DR1 molecule. The polymorphisms between the E<sup>k</sup> and E<sup>d</sup> molecules at the 'pocket' residues are reflected in differences within the peptide motifs. The latter indicate a very high structural similarity of mouse class II E molecules with human DR molecules. The anchor at relative position 9 is conserved between all class II E binding peptides [as suggested by Leighton *et al.* (7)], whereas P1, P4, P6 and possibly P7 are different and responsible for the allelic differences. Since all anchor residues contribute to binding, a suboptimal amino acid selection at one anchor position may be compensated by an optimal fit at another anchor residue. This might explain why some peptides can bind to more than one but not all H-2E molecules. A good example

**Table 6.** Alignment of natural ligands and T cell epitopes to the predicted H-2E<sup>s</sup> and E<sup>b</sup> motifs

H-2E <sup>s</sup> motif	Relative position																					
	1	2	3	4	5	6	7	8	9													
Predicted anchor residues	I		L		Q		K															
	V		I		N		R															
	L		Y																			
Natural ligands <sup>a</sup>																						
Carboxypeptidase A 44-				L	Y	V	L	K	I	G	K	K	D	G								
β <sub>2</sub> m 42-	H	P	P	H	I	E	I	Q	M	L	K	N	G	K	K							
H-2L <sup>d</sup> 160-			E	G	E	C	V	E	W	L	H	R	Y	L	K	N	G					
β-Actin 286-	M	Q	K	E	I	T	A	L	A	P	S	T	M	K	I	K	I	I				
Substance P receptor 255-261				C	T	F	A	I	C	W	L	P	F	H	V	F	F	L				
HSP60 478-492		E	G	S	L	I	V	E	K	I	M	Q	S	S	S	E						
T cell epitopes <sup>b</sup>																						
MCC 95-103		I	A	Y	L	K	Q	A	T	K												
H-2E <sup>b</sup> motif																						
	1	2	3	4	5	6	7	8	9													
Predicted anchor residues	W		L		Q		K															
	F		I		N		R															
	Y		F		Δ																	
				Y																		
Natural ligands <sup>c</sup>																						
MuLV env protein 454-469	S	P	S	Y	V	Y	H	Q	E	E	R	R	A	K	Y	K						
MuLV env protein 454-475	S	P	S	Y	V	Y	H	Q	E	E	R	R	A	K	Y	K	R	E	P	V	S	L
MuLV env protein 454-467	S	P	S	Y	V	Y	H	Q	E	E	R	R	A	K								
BSA 141-153				G	K	Y	L	Y	E	I	Δ	R	R	H	P	Y	F	Y				
Unknown	X	P	Q	S	Y	L	I	H	E	X	X	X	I	S								
T cell epitopes <sup>b</sup>																						
SNase 81-100	R	T	D	K	Y	G	R	G	L	A	Y	I	Y	Δ	D	G	K	M	V	N		
MCC 95-103									I	A	Y	L	K	Q	A	T	K					

<sup>a</sup>Sequences adapted from ref. 5.<sup>b</sup>T cell epitopes adapted from ref. 22.<sup>c</sup>Sequences adapted from ref. 3.

for this is the MCC 95-103 peptide. It is reported to bind to E<sup>k,b,s</sup> molecules (26). P1, 4, 6 and 9 have the optimal amino acid selection for binding to E<sup>k</sup> and E<sup>s</sup> molecules and are nearly optimal for E<sup>b</sup>, which should prefer aromatic instead of aliphatic residues at P1. E<sup>d</sup> molecules, on the other hand, prefer, in addition to aromatic amino acids at P1, basic residues at P4 and aliphatic residues at P6. MCC 95-103 does not fulfil any of these requirements and, because of this, might not bind to E<sup>d</sup> molecules.

Our analysis shows that structural and sequence information together can be used to predict allele-specific motifs. The four H-2E motifs we have described or predicted in this paper can be used to identify ligands within proteins supposed to be recognized during immune responses against infectious agents or in autoimmune diseases. The identification of such ligands will be helpful for designing mouse models of active immune intervention since the design of both subunit vaccines or MHC antagonist requires a detailed knowledge of anchor positions and other general rules for peptide presentation by MHC molecules.

### Acknowledgements

We thank S. Faath for technical assistance and Dr Damir Vidovic for the cell line LQ-9E10 cells. This work was supported by a grant from the Bundesminister für Forschung und Technologie and by the Deutsch Forschungsgemeinschaft (Leibnizprogramm). The Basel Institute of Immunology was founded and is supported by F. Hoffmann-La Roche & Co. Ltd, Basel, Switzerland.

### Abbreviations

α1-antitryp.	α1-antitrypsin
β <sub>2</sub> m	β <sub>2</sub> -microglobulin
Apo-E prec.	apolipoprotein E precursor
BSA	bovine serum albumin
C' inhib.	C' cytotoxic inhibitor
EqMb	equine myoglobin
ER60	ER resident protease 60 kDa
Flu PR/8 HA	influenza A PR/8/34 hemagglutinin
HEL	hen egg lysozyme
HIV-1 gag	HIV-1 gag protein
HLA	human lymphocyte antigen
HSC70	heat-shock cognate protein 70



LLO	listeriolysin O
λrep	λ repressor
MCC	moth cytochrome c
MoMb	mouse myoglobin
MuLV env	murine leukemia virus envelope protein
PCC	pigeon cytochrome c
ser. alb	serum albumin
SNase	staphylococci nuclease
SWMb	sperm whale myoglobin
TFA	trifluoroacetic acid

## References

- Rammensee, H.-G., Falk, K. and Rötzschke, O. 1993. Peptides naturally presented by MHC class I molecules. *Annu. Rev. Immunol.* 11:213.
- Rudensky, A. Y., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. and Janeway, C. A. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature* 353:622.
- Rudensky, A. Y., Preston-Hurlburt, P., Al-Ramadi, B.K., Rothbard, J. and Janeway, C. A. 1992. Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature* 359:429.
- Chicz, R. M., Urban, R. G., Gorga, J. C., Vignali, D. A. A., Lane, W. S. and Strominger, J. L. 1993. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J. Exp. Med.* 178:27.
- Freed, J. H. and Marrack, P. 1993. Tissue-specific expression of self peptides bound by major histocompatibility complex class II molecules. In Sette, A., ed., *Naturally Processed Peptides*, p. 88. Karger, Basel.
- Hunt, D. F., Michel, H., Dickinson, T. A., Shabanowitz, J., Cox, A. L., Sakaguchi, K., Appella, E., Grey, H. M. and Sette, A. 1992. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-A<sup>d</sup>. *Science* 256:1817.
- Leighton, J., Sette, A., Sidney, J., Appella, E., Ehrhardt, C., Fuchs, S. and Adorini, L. 1991. Comparison of structural requirements for interaction of the same peptide with I-E<sup>k</sup> and I-E<sup>d</sup> molecules in the activation of MHC class II-restricted T cells. *J. Immunol.* 147:198.
- Sette, A., Adorini, L., Appella, E., Colon, S. M., Miles, C., Tanaka, S., Ehrhardt, C., Doria, G., Nagy, Z. A., Buus, S. and Grey, H. M. 1989. Structural requirements for the interaction between peptide antigens and I-E<sup>d</sup> molecules. *J. Immunol.* 143:3289.
- Sette, S., Buus, S., Colon, S., Miles, S. and Grey, H. M. 1989. Structural analysis of peptides capable of binding to more than one Ia antigen. *J. Immunol.* 142:35.
- Reay, P. A., Kantor, R. M. and Davis, M. M. 1994. Use of global amino acid replacements to define the requirements for MHC binding and T cell recognition of moth cytochrome c (93–103). *J. Immunol.* 152:3946.
- Marrack, P., Ignatowicz, L., Kappler, J. W., Boymel, J. and Freed, J. H. 1993. Comparison of peptides bound to spleen and thymus class II. *J. Exp. Med.* 178:2173.
- Hammer, J., Takacs, B. and Sinigaglia, F. 1992. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. *J. Exp. Med.* 176:1007.
- Matsumura, M., Fremont, D. H., Peterson, P. and Wilson, I. A. 1992. Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* 257:927.
- Falk, K., Rötzschke, O., Stevanović, S., Jung, G. and Rammensee, H. G. 1994. Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules. *Immunogenetics* 39:230.
- Malcherek, G., Falk, K., Rötzschke, O., Rammensee, H.-G., Stevanović, S., Gnau, V., Jung, G. and Melms, A. 1993. Natural peptide ligand motifs of two HLA molecules associated with myasthenia gravis. *Int. Immunol.* 5:1229.
- Vorreck, F. A., van de Poel, A., Termijtelen, A., Amons, R., Drijfhout, J.-W. and Koning, F. 1994. Identification of an HLA-DQ2 peptide binding motif and HLA-DPw3-bound self-peptide by pool sequencing. *Eur. J. Immunol.* 24:375.
- Friede, T., Gnau, V., Jung, G., Keilholz, W., Stevanović, S. and Rammensee, H.-G. 1995. Natural ligand motifs to closely related HLA-DR4 molecules and implications for differential disease association, submitted.
- Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L. and Wiley, D. C. 1993. Three-dimensional structure of the human class-II histocompatibility antigen HLA-DR1. *Nature* 364:33.
- Stern, L. J., Brown, J. H., Jardetzky, T. S., Gorga, J. C., Urban, R. G., Strominger, J. L. and Wiley, D. C. 1994. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 368:215.
- Schild, H., Mavaddat, N., Litzenberger, C., Ehrlich, E. W., Davis, M. M., Bluestone, J. A., Matis, L., Draper, R. K. and Chien, Y.-H. 1994. The nature of major histocompatibility complex recognition by γδ T cells. *Cell* 76:29.
- Kim, K. J., Kanellopoulos-Langevin, C., Mervin, R. M., Sachs, D. H. and Asofsky, R. 1979. Establishment and characterization of BALB/c lymphoma lines with B cell properties. *J. Immunol.* 122:549.
- Ozato, K., Mayer, N. and Sachs, D. H. 1980. Hybridoma cell lines secreting monoclonal antibodies to mouse H-2 and Ia antigens. *J. Immunol.* 124:533.
- Stevanović, S. and Jung, G. 1993. Multiple sequence-analysis—pool sequencing of synthetic and natural peptide libraries. *Anal. Biochem.* 212:212.
- Kyte, J. and Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105.
- Stevanović, S. and Rammensee, H.-G. 1994. The structure of T-cell epitopes. In van Regenmortel, M. H. V., ed., *Structure of Antigens*, vol. 3. CRC Press, Boca Raton, FL.
- Hedrick, S. M., Matis, L. A., Hecht, T. T., Samelson, L. E., Longo, D. L., Heber-Katz, E. and Schwartz, R. H. 1982. The fine specificity of antigen and Ia determinant recognition by T cell hybridoma clones specific for pigeon cytochrome c. *Cell* 30:141.
- Marsh, S. G. E. and Bodmer, J. G. 1992. HLA class nucleotide sequences, 1992. *Eur. J. Immunogenet.* 20:83.
- Figuerola, F. and Klein, J. 1986. The evolution of MHC class II genes. *Immunol. Today* 7:78.
- Vogt, A. B., Kropshofer, H., Kalbacher, H., Kalbus, M., Rammensee, H.-G., Coligan, J. E. and Martin, R. 1994. Ligand motifs of HLA-DRB5\*0101 and DRB1\*1501 molecules delineated from self-peptides. *J. Immunol.* 153:1665.
- Bogeu, B. and Lambris, J. D. 1989. Minimum length of an idiotypic peptide and a model for its binding to a major histocompatibility-complex class II molecule. *EMBO J.* 8:1947.