A Peptide-binding Motif for I-A^{g7}, the Class II Major Histocompatibility Complex (MHC) Molecule of NOD and Biozzi AB/H Mice

By Leonard C. Harrison,* Margo C. Honeyman,* Sylvie Trembleau,‡ Silvia Gregori,‡ Fabio Gallazzi,‡ Petra Augstein,* Vladimir Brusic,* Juergen Hammer,‡ and Luciano Adorini‡

*From *the Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital, Parkville 3050, Australia; and ‡Roche Milano Ricerche, 1-20132, Milano, Italy*

Summary

The class II major histocompatibility complex molecule I-A^{g7} is strongly linked to the development of spontaneous insulin-dependent diabetes mellitus (IDDM) in non obese diabetic mice and to the induction of experimental allergic encephalomyelitis in Biozzi AB/H mice. Structurally, it resembles the HLA-DQ molecules associated with human IDDM, in having a non-Asp residue at position 57 in its β chain. To identify the requirements for peptide binding to I-A^{g7} and thereby potentially pathogenic T cell epitopes, we analyzed a known I-A^{g7}-restricted T cell epitope, hen egg white lysozyme (HEL) amino acids 9–27. NH_2- and COOH-terminal truncations demonstrated that the minimal epitope for activation of the T cell hybridoma 2D12.1 was M12-R21 and the minimum sequence for direct binding to purified I-A^{g7} M12-Y20/K13-R21. Alanine (A) scanning revealed two primary anchors for binding at relative positions (p) 6 (L) and 9 (Y) in the HEL epitope. The critical role of both anchors was demonstrated by incorporating L and Y in poly(A) backbones at the same relative positions as in the HEL epitope. Well-tolerated, weakly tolerated, and nontolerated residues were identified by analyzing the binding of peptides containing multiple substitutions at individual positions. Optimally, p6 was a large, hydrophobic residue (L, I, V, M), whereas p9 was aromatic and hydrophobic (Y or F) or positively charged (K, R). Specific residues were not tolerated at these and some other positions. A motif for binding to I-A^{g7} deduced from analysis of the model HEL epitope was present in 27/30 (90%) of peptides reported to be I-A^{g7}–restricted T cell epitopes or eluted from I-A^{g7}. Scanning a set of overlapping peptides encompassing human proinsulin revealed the motif in 6/6 good binders (sensitivity = 100%) and 4/13 weak or non-binders (specificity = 70%). This motif should facilitate identification of autoantigenic epitopes relevant to the pathogenesis and immunotherapy of IDDM.

Non obese diabetic (NOD)[1] mice develop autoimmune, T cell–mediated destruction of pancreatic islet β cells and are a model of human insulin-dependent diabetes mellitus (IDDM) (1). In common with humans who develop IDDM, NOD mice have immune responses to islet autoantigens such as insulin and glutamic acid decarboxylase (GAD). In addition, they share a structurally similar class II MHC molecule associated with disease susceptibility. This molecule, I-A^{g7}, has a β chain sequence otherwise found only in Biozzi AB/H mice that are susceptible to chronic relapsing experimental allergic encephalomyelitis (CR-EAE) (2). It is characterized by a non-Asp residue at position 57 (3), as in the β chain of the HLA-DQ molecules associated with human IDDM (4). The capacity of these unique class II molecules to bind and present peptides to autoreactive T cells could be critical in the development of IDDM and CR-EAE.

Although amino acid motifs for peptides that bind to individual class I and some class II MHC molecules have been well defined (5, 6), the rules that govern binding of peptides to I-A^{g7} are still unclear. Reich et al. (7) eluted and sequenced several naturally processed peptides from I-A^{g7} and concluded that binding may require an acidic residue in the COOH terminus of the peptide. Carrasco-Marin et al. (8) found that I-A^{g7} either on the surface of antigen-presenting cells or in SDS-PAGE after its purification was unstable and that the binding of known I-A^{g7}–restricted T cell

[1] *Abbreviations used in this paper:* CR-EAE, chronic relapsing experimental allergic encephalomyelitis; DOC, sodium deoxycholate; EAE, experimental allergic encephalomyelitis; GAD, glutamic acid decarboxylase; HEL, hen egg white lysozyme; IDDM, insulin-dependent diabetes mellitus; MOG, myelin oligodendrocyte glycoprotein; MSA, mouse serum albumin; NOD, nonobese diabetic; OGP, octyl-β-D-glucopyranoside; p, position; PEG, polyethylene glycol; PLP, proteolypid protein; RT, room temperature.

J. Exp. Med. © The Rockefeller University Press • 0022-1007/97/03/1013/09 $2.00
Volume 185, Number 6, March 17, 1997 1013–1021

epitopes or the peptides eluted by Reich et al. (7) was difficult or impossible to demonstrate. This led them to hypothesize that weak peptide binding by I-A^{g7} militated against elimination of autoreactive T cells in the NOD mouse. Amor et al. (9) investigated the fine specificity of peptides from myelin oligodendrocyte glycoprotein (MOG) or proteolipid protein (PLP) for the induction of experimental allergic encephalomyelitis (EAE) in Biozzi AB/H mice and suggested a core motif for I-A^{g7} binding peptides.

In this study, we used the I-A^{g7}–restricted T cell epitope, hen egg white lysozyme (HEL) amino acids 9–27, as a template with which to analyze the amino acid sequence of peptides that bind to purified, native I-A^{g7} and activate a T cell hybridoma. This has enabled us to define general rules that identify most known I-A^{g7} binding peptides.

Materials and Methods

Purification of I-A^{g7}. I-A^{g7} protein was affinity-purified from detergent lysates of 4G4.7 B cell hybridoma cells by desorption from OX-6 mouse monoclonal antibody. The 4G4.7 B cell hybridoma was derived by polyethylene glycol (PEG)-induced fusion of NOD mouse T cell–depleted splenocytes with the HAT-sensitive A20.2J lymphoma line (10). OX-6 is a mouse monoclonal IgG1 antibody against an invariant determinant of rat Ia, which also recognizes I-A^{g7} but not I-A^{d} (11, 12). Approximately 15 mg of OX-6 antibody was first bound to 4 ml of protein A–Sepharose 4 Fastflow (Pharmacia, Uppsala, Sweden) and then chemically cross-linked to the protein A with dimethyl pimelimidate dihydrochloride (Sigma Chemical Co., St. Louis, MO) in sodium borate buffer, pH 9.0. After 60 min at room temperature (RT), the reaction was quenched by incubating the Sepharose in 0.2 M ethanolamine, pH 8.0, for 60 min at RT. The suspension was washed thoroughly in PBS and stored in PBS, 0.02% sodium azide (NaN_3).

4G4.7 cells were harvested by centrifugation, washed in PBS, resuspended at 10^8 cells/ml of lysis buffer, and then allowed to stand at 4°C for 120 min. The lysis buffer was 0.05 M sodium phosphate, pH 7.5, containing 0.15 M NaCl, 1% (vol/vol) NP-40 detergent and the following protease inhibitors: 1 mM phenylmethylsulphonyl fluoride, 5 mM ϵ-amino-*n*-caproic acid and 10 μg/ml each of soybean trypsin inhibitor, antipain, pepstatin, leupeptin and chymotrypsin. Lysates were cleared of nuclei and debris by centrifugation at 27,000 *g* for 30 min and stored as such if not immediately processed further. To the postnuclear supernatant was added 0.2 vol of 5% sodium deoxycholate (DOC). After mixing at 4°C for 10 min, the supernatant was centrifuged at 100,000 *g* at 4°C for 120 min, carefully decanted, and filtered through a 0.45-μm nylon membrane. The lysate of 5×10^{10} 4G4.7 cells was gently mixed overnight at 4°C with 4 ml of OX-6–protein A–Sepharose, and the suspension then poured into a column and washed with at least 50 vol each of buffers A, B, and C. Buffer A was 0.05 M Tris, pH 8.0, 0.15 M NaCl, 0.5% NP-40, 0.5% DOC, 10% glycerol, and 0.03% NaN_3; buffer B was 0.05 M Tris, pH 9.0, 0.5 M NaCl, 0.5% NP-40, 0.5% DOC, 10% glycerol, and 0.03% NaN_3; buffer C was 2 mM Tris, pH 8.0, 1% octyl-β-d-glucopyranoside (OGP), 10% glycerol, and 0.03% NaN_3. Bound I-A^{g7} was eluted with 50 mM diethylamine HCl, pH 11.5 in 0.15 M NaCl, 1 mM EDTA, 1% OGP, 10% glycerol, and 0.03% NaN_3, and immediately neutralized with 1 M Tris.

Peptide Synthesis. Peptides were synthesized with a multiple peptide synthesizer (model 396; Advanced ChemTech, Louisville, KY) using Fmoc chemistry and solid phase synthesis on Rink Amide resin. All acylation reactions were effected with a threefold excess of activated Fmoc amino acids, and a standard coupling time of 20 min was used. Each Fmoc amino acid was coupled at least twice. Cleavage and side chain deprotection was achieved by treating the resin with 90% trifluoroacetic acid, 5% thioanisole, 2.5% phenol, 2.5% water. The indicator peptide for the binding assay was biotinylated before being cleaved from resin by coupling two 6-aminocaproic acid spacers on the NH_2 terminus and one biotin molecule sequentially, using the above-described procedure. Individual peptides were analyzed by reverse-phase HPLC and those used in this study were routinely ≥85% pure.

T Cell Hybridoma. Hybridoma 2D12.1 was generated by PEG-induced fusion of HEL-immune lymph node cells from a NOD mouse with the TCR-α/β–negative variant of the BW5147 thymoma, as described previously (13). Reactivity of 2D12.1 to HEL peptides was assayed by incubating 2.5×10^5 NOD spleen cells and HEL peptides (0.3 nM to 10 μm) with 5×10^4 T hybridoma cells/well. Culture medium was RPMI 1640 supplemented with 10% FCS, 2 mM l-glutamine, 50 μg/ml gentamicin, and 50 μm 2-mercaptoethanol. After 24 h of culture, 50 μl of supernatants were transferred to culture wells containing 10^4 IL-2–responsive CTLL-2 cells. During the final 4 h of a 24-h culture, CTLL-2 cells were pulsed with 1 μCi [^{3}H]thymidine. Thymidine incorporation was measured by scintillation spectrometry. The concentration of peptide that caused 50% of maximum stimulation is referred to as SC_{50}.

I-A^{g7} Peptide-binding Assay. Peptides were dissolved at 10 mM in DMSO and diluted into 20% DMSO/PBS for assay. Indicator I-A^{g7} binding peptide, HEL 10–23, was synthesized with a biotin molecule and two spacer residues at the NH_2 terminus. Approximately 200 nM of this biotinylated HEL peptide and each test peptide in seven concentrations ranging from 50 μM to 50 pM,

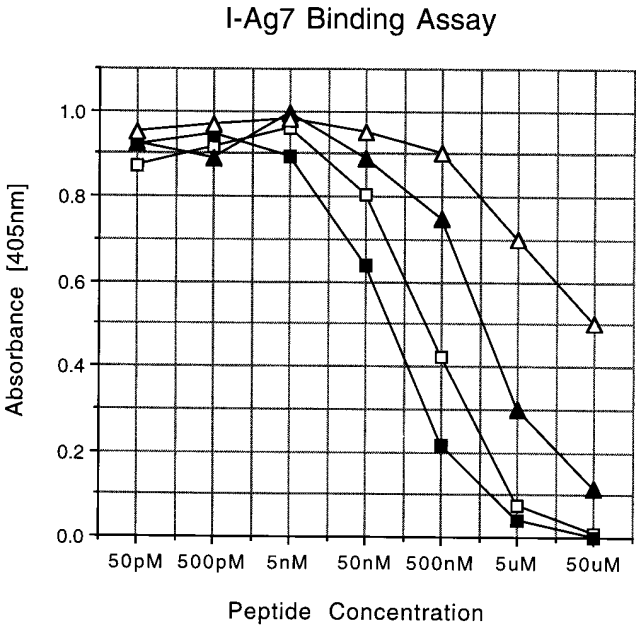

Figure 1. Examples of competition between biotinylated HEL peptide (amino acids 10–23) and unlabeled peptides for binding to purified I-A^{g7}, measured by ELISA (see Materials and Methods). Unlabeled HEL 10–23 (□) was used as an internal control in each 96-well plate assay; ■, good binder (IC_{50} 100 nM); ▲, weak binder (IC_{50} 2000 nM); △, non-binder (IC_{50} 50,000 nM).

Table 1. *Truncation Analysis of I-A^{g7}–restricted HEL 9–27 Epitope*

HEL	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Binding to I-A^{g7} IC_{50}	Activation of T cell hybridoma SC_{50}
																				nM	
9–27	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	350	60
10–27		A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	300	84
11–27			A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	300	84
12–27				M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	300	380
13–27					K	R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	300	3,000
14–27						R	H	G	L	D	N	Y	R	G	Y	S	L	G	N	2,000	>10,000
15–27							H	G	L	D	N	Y	R	G	Y	S	L	G	N	20,000	>10,000
16–27								G	L	D	N	Y	R	G	Y	S	L	G	N	>50,000	>10,000
17–27									L	D	N	Y	R	G	Y	S	L	G	N	>50,000	>10,000
18–27										D	N	Y	R	G	Y	S	L	G	N	>50,000	>10,000
19–27											N	Y	R	G	Y	S	L	G	N	>50,000	>10,000
9–26	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L	G		300	60
9–25	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S	L			300	140
9–24	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y	S				200	72
9–23	A	A	A	M	K	R	H	G	L	D	N	Y	R	G	Y					150	180
9–22	A	A	A	M	K	R	H	G	L	D	N	Y	R	G						300	240
9–21	A	A	A	M	K	R	H	G	L	D	N	Y	R							300	360
9–20	A	A	A	M	K	R	H	G	L	D	N	Y								150	>10,000
9–19	A	A	A	M	K	R	H	G	L	D	N									15,000	>10,000
9–18	A	A	A	M	K	R	H	G	L	D										>15,000	>10,000
9–17	A	A	A	M	K	R	H	G	L											>50,000	>10,000
12–23				M	K	R	H	G	L	D	N	Y	R	G	Y					250	800
12–22				M	K	R	H	G	L	D	N	Y	R	G						600	Not done
12–21				M	K	R	H	G	L	D	N	Y	R							1,000	1,200
12–20				M	K	R	H	G	L	D	N	Y								1,250	>10,000
13–23					K	R	H	G	L	D	N	Y	R	G	Y					200	Not done
13–22					K	R	H	G	L	D	N	Y	R	G						250	Not done
13–21					K	R	H	G	L	D	N	Y	R							5,000	>10,000
13–20					K	R	H	G	L	D	N	Y								30,000	>10,000
14–23						R	H	G	L	D	N	Y	R	G	Y					500	Not done
14–22						R	H	G	L	D	N	Y	R	G						3,000	Not done

were coincubated with ~200 ng of I-A^{g7} protein in U-bottomed polypropylene 96-well plates (Costar Serocluster, Costar Corp., Cambridge, MA) in binding buffer at RT. The binding buffer was 6.7 mM citric phosphate, pH 7.0, with 0.15 M NaCl, 2% NP-40, 2 mM EDTA, and the protease inhibitors as used in the lysis buffer. After a minimum of 24 h, each incubate was transferred to the corresponding well of an ELISA plate (Nunc Maxisorp, Nunc, Roskilde, Denmark) containing prebound OX-6 antibody (5 μg/ml overnight at 4°C, followed by washing). After incubation at RT for at least 2 h, and washing, bound biotinylated peptide–I-A^{g7} complexes were detected colorimetrically at 405 nm after reaction with streptavidin–alkaline phosphatase and paranitrophenolphosphate. Competition binding curves were plotted and the affinity of peptide for I-A^{g7} was expressed as an inhibitory concentration 50 (IC_{50}), the concentration of peptide required to inhibit the binding of bio-HEL 10–23 by 50%.

Results and Discussion

I-A^{g7} Purification and Binding Assay. Approximately 2 mg of protein, estimated by Coomassie blue binding (Bio Rad Protein assay), was purified from 5×10^{10} 4G4.7 cells. In SDS-PAGE, the majority (>95%) of the protein was resolved as two bands of molecular weight ~33,000 and ~28,000 that correspond to the α and β subunits, respec-

Table 2. *Effect of Selected Amino Acid Substitutions on Binding and T Cell Activation of HEL 10–22**

Substituted amino acid	HEL 10–22												
	10 A	11 A	12 M	13 K	14 R	15 H	16 G	17 L	18 D	19 N	20 Y	21 R	22 G
ALA (A)	−	−	600‡	300	300	300	400	30,000	150	150	30,000	800	350
	−	−	46§	8,000	>10,000	>10,000	>10,000	>10,000	>10,000	42	10,000	2,200	45
ASP (D)	950	950	950	950	7,500	1,750	2,000	7,500	−	2,000	7,000	750	150
	48	48	60	1,200	800	600	2,000	1,600	−	1,100	1,100	1,300	200
LYS (K)	160	750	500	−	900	800	1,500	7,000	600	7,000	900	350	1,500
	200	100	60	−	200	>10,000	4,750	3,000	3,000	3,500	4,750	200	250
PRO (P)	1,300	1,300	1,300	1,300	2,500	7,000	2,500	10,000	300	450	10,000	350	650
	240	240	400	9,000	9,000	>10,000	>10,000	>10,000	>10,000	>10,000	8,000	1,000	5,000
TYR (Y)	275	350	125	250	900	150	750	3,500	100	>50,000	−	300	300
	270	155	310	>10,000	>10,000	>10,000	4,000	4,900	>10,000	>10,000	−	>10,000	>10,000
LEU (L)	300	400	550	1,100	750	450	1,400	−	75	1,450	1,550	500	1,500
	270	400	>10,000	>10,000	10,000	>10,000	>10,000	−	>10,000	>10,000	>10,000	5,370	350
GLN (Q)	450	450	600	900	550	400	700	12,000	200	2,000	2,000	700	900
	25	47	38	2,250	>10,000	>10,000	475	>10,000	>10,000	>10,000	>10,000	650	142

*For natural, unsubstituted HEL 10–22, the binding affinity for I-A^{g7} (IC_{50}; mean ± SD) was 295 ± 72 nM (n = 15) and the dose (SC_{50}) for activation of the T cell hybridoma was 52 ± 23 nM (n = 6).
‡Binding to I-A^{g7}: IC_{50} (nM).
§Activation of the T cell hybridoma 2D12.1: SC_{50} (nM).

tively, of mouse class II MHC molecules (data not shown). The competition binding assay with purified I-A^{g7} was sensitive and specific (Fig. 1), and highly reproducible; in 15 separate assays the mean ± SD of the IC_{50} for competition between biotinylated and unlabeled HEL 10–23 was 295 ± 72 nM.

Carrasco-Marin et al. (8) were unable to demonstrate direct binding of HEL 11–25 to purified I-A^{g7} and proposed that I-A^{g7} was inherently unstable. We found that purified I-A^{g7} stored at −70°C for more than 1 yr reproducibly bound HEL 10–23 with high affinity. Therefore, our results do not support their hypothesis that I-A^{g7} is inherently unstable, which they postulated would impair its ability to bind and induce tolerance to autoreactive peptides.

Truncation Analysis of HEL 9–27. Peptides representing sequential truncations of HEL 9–27, from either the NH_2 or COOH-terminus, were each assayed in parallel for binding to I-A^{g7} and for their ability to activate the 2D12.1 hybridoma. Inspection of these data (Table 1) reveals that the minimum T cell epitope is M12-R21, and the minimum binder is M12-Y20 or K13-R21.

Effect of Selected Substitutions on Binding and Bioactivity of HEL 12–22. Substitution of alanine (A) at each position in HEL 12–22 (Table 2) had no significant effect on binding, with the sole exceptions of positions L17 and Y20. Substitution at either of these two positions virtually abolished binding. On the other hand, while having no effect on binding, substitutions by A at K13, R14, H15, G16, and D18, and to a lesser extent at R21, abolished T cell activation. Removal of R21 (see Table 1) abolished T cell activation. Further substitutions of representative amino acids (D, K, P, Y, L, Q) at each position (Table 2) revealed varying levels of tolerance of specific residues/positions for binding (see below) and generally confirmed the results of the alanine substitutions on T cell activation. On the basis of these results, we can deduce that most residues in the minimal T cell epitope HEL 12–21 have TCR contacts and that two, L17 and Y20, are essential for binding to I-A^{g7} (Fig. 2).

Anchor Residues for Peptide Binding. The critical roles of L17 and Y20 in the HEL epitope, as model anchor residues for binding to I-A^{g7}, was demonstrated with poly(A) pep-

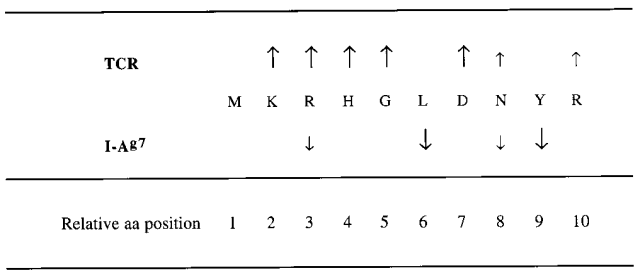

Figure 2. Minimal T cell epitope, HEL 12(M)–21(R), showing TCR contact residues and L and Y primary anchors for binding to I-A^{97} at relative positions 6 and 9, and positions 3 and 8 at which specific residues are also not tolerated for binding.

Table 3. *Binding to I-A^{g7} of L- and Y-substituted Poly(A) Peptides*

Peptide	Binding IC_{50} (nM)
* K A A A A A A A A	25,000
K A A A A L A A A A	25,000
K A A A A A A Y A	20,000
K A A A A L A A Y A	50
A A A A A L A A Y A	50
A A A A A Y A A L A	12,000
A A A A A L A A A Y A	12,000
A A A A A A L A Y A	12,000
A A A A A A L A A Y	25
A A A A A L A A Y	200
A A A A L A A A Y A	600
A A A A L A A A Y A A	150
A A A L A A A Y A A	400
A A A L A A A Y A	20,000
A A L A A A Y A A A A A	6,000
A A L A A A Y A	30,000
L A A A Y A A A A A	35,000

*K added at NH_2 terminus to improve solubility.

Table 4. *Effects of Amino Acid Substitutions on Binding of HEL 12–22 to I-A^{g7}*

	Relative position (HEL 12–22)										
	1	2	3	4	5	6	7	8	9	10	11
	M	K	R	H	G	L	D	N	Y	R	G
Well-tolerated residues (IC_{50} <1,000 nM)	A D K L Q Y	A D Q Y	A K L Q Y	A K L Q Y	A Q Y	I M V	A E K L P Q R Y	A G P S	F K R	A D K L P Q Y	A D P Q Y
Weakly tolerated residues (IC_{50} 1,000–10,000 nM)	P	L P	D G P W S	D P	D K L P	D F H K N T Y		D E F K L Q R	D H L M N Q W		K L
Nontolerated (IC_{50} >10,000 nM)			E F			A E G P Q R S W		W Y	A E G I P S T V		

tides (Table 3). The nonbinding poly(A) peptide, KAA AAAAAA, was converted to a super binder simply by incorporating L and Y at the same relative positions as in the HEL epitope. Either residue alone was not sufficient. Binding was reconstituted only when these two residues were appropriately spaced and in the correct order. In addition, this approach reveals the importance of the frame or context of the anchor residues. The LAAY sequence must be flanked by at least two As, an absence of which at the COOH terminus can be compensated for by at least three As on the NH_2 terminus, but not vice versa. This suggests that binding of these specific residues within the I-A^{g7} groove requires stabilization by hydrogen bonding from nonspecific flanking residues, in particular at the NH_2 terminus. For the purpose of further analysis, the relative positions (p) of L and Y in the HEL epitope 12–21 are designated p6 and p9.

Effect of Multiple Substitutions at p6 and p9. In addition to the selected substitutions at all positions (see Table 2), we investigated the effect on binding of all possible substitutions (except labile cysteine) at p6 or p9. A single residue substitution was classified as well tolerated, weakly tolerated, or nontolerated according to a threshold on its IC_{50} value: well tolerated, <1,000 nM; weakly tolerated, 1,000–10,000 nM; nontolerated, ⩾10,000 nM. Although somewhat arbitrary, this classification corresponds to generally accepted notions of good binders, moderate binders, and weak to non-binders. The results, combined with those from Table 2, are presented in Table 4. Optimally, p6 is a large, hydrophobic residue (L, I, M, V), whereas p9 is aromatic and hydrophobic (Y, F) or positively charged (K, R). Most amino acids are not well tolerated at these anchor positions. Additionally, specific amino acids are not tolerated at other positions, namely F and E at p3 and W and Y at p8. This information allowed us to propose and test minimum rules for a motif for I-A^{g7} binding peptides. These were that a binder must have the following: (*a*) two well-tolerated residues or one well-tolerated and one weakly-tolerated residue at anchor positions p6 and p9, (*b*) no nontolerated residues at positions p3 and p8, and (*c*) at least two residues flanking p6 and p9, or at least three residues NH_2-terminal of p6.

Motifs in Peptides Known or Deduced to Bind I-A^{g7}. Relatively few peptides containing sequences that might bind to I-A^{g7} have been reported in the literature. They include peptides that stimulate I-A^{g7}–restricted T cells or T cell hybridomas, compete for antigen presentation to T cell hybridomas, induce EAE in Biozzi AB/H mice, or have been eluted from I-A^{g7} (listed in Table 5). It should be noted

Table 5. *Motif in Reported I-A^{g7} Binding Peptides*

Protein	Peptide	Method of Identification	Reference	Motif
HEL (9–29)	AAAMKRHG**L**DN**Y**RGYSLGNW	T cell hybridoma stimulation	14	+
HEL (90–104)	SV<u>N</u>CA**K**K**I**VS<u>D</u>GNGM	T cell stimulation	15	+
h/m GAD65 (524–543)	SRLSKVAPVIKAR**M**ME**Y**GTT	″ ″	16	+
h GAD65 (509–528)	IPPSLRY**L**LD<u>N</u>EERMSRLSK	″ ″	17	+
h GAD65 (247–266)	NMYAM**M**IAR<u>F</u>KM**F**PEVKEKG	″ ″	″	+
OVA (323–339)	ISQAVHAAHAEINEAGR	″ ″	18	−
λ Repressor protein (12–26)	LEDARRLKA**I**YE**K**KK	″ ″	″	+
In serum albumin (560–574)	KPKATAEQLKTVMDD	Elution and sequencing	7	−
Transferrin (55–68)	GHNYV<u>T</u>AI**R**NQQEG	″ ″	″	+
hRNPA1 (44–59)	VVMRDPQT**K**RS**R**GFGF	″ ″	″	+
hRNPA24B1 (51–66)	VVMRDPAS**K**RS**R**GFGF	″ ″	″	+
hRNPA24B1 (31–43)	ETT**e**ES**L**RN**Y**YEQ	″ ″	″	+
r MOG (8–22)	PGYP**I**RA<u>L</u>VGDEQED	Induction of CR-EAE	9	+
m PLP (56–70)	DYEYLINV**I**HA**F**QYV	″ ″	″	+
m MBP (12–35)	YLATASTMDHARHG<u>F</u>LP**R**HRDTGI	″ ″	″	+
Heat shock protein (437–460)	VLGGGCALLRC**I**PA<u>L</u>DSLTPANED	T cell hybridoma stimulation	8	+
Carboxypeptidase H (362–382)	KNS**L**IN**Y**LEQIHRGV**K**GF**V**R	″ ″	″	+
Carboxypeptidase H (440–464)	FSPAVGVDFELES<u>F</u>SE**R**KEEEKEEL	″ ″	″	+
Staphylococcal nuclease (61–79)	FTKK**M**VE<u>N</u>AKKIEVEFDK	T cell immunogen and/or T cell	19	+
Equine myoglobin (131–153)	MTKA**L**EL**F**RNDIAA**K**YKELGFQG	Hybridoma competition		
m α-1 antitrypsin (148–163)	LSQAVHKAV**L**TI**D**ETG	″ ″	″	+
m laminin b1 chain precursor (1594-1612)	MVKEALEEAEKAQVAAEKA	″ ″	″	−
m prostate secretory glycoprotein precursor (63–76)	FE<u>N</u>RK**R**IEPVLIRK	″ ″	″	+
m myoglobin (131–153)	MSKA**L**EL**F**RNDIAA**K**YKLGFQG	″ ″	″	+
m insulin B chain (9–23)	SHL**V**EA**L**YLVCGERG	T cell stimulation	20	+
m TCR Vβ8.2 (38–60)	DTG<u>H</u>GL**R**L<u>I</u>H<u>Y</u>S**Y**GAGSTEKGDI	″ ″	21	+
m TCR Vβ6 (38–60)	DSG<u>K</u>GL**R**L**I**Y**Y**S**I**TE<u>N</u>DLQKGDL	″ ″	″	+
m myoglobin (110–121)	I**I**IE**V**<u>L</u>K**K**R**H**SG	″ ″	″	+
m myoglobin (69–78)	LTALG<u>T</u>IL**K**K	″ ″	″	+
Ribosomal S30 peptide (75–96)	KVHGSLARAGK**V**RG<u>Q</u>TP<u>K</u>VA**K**Q	″ ″	″	+
m MOG (35–55)	MEVGWYRSPFSR**V**V<u>H</u>L**V**R<u>N</u>GK	Induction of CR-EAE	22	+

Residues shown to be well tolerated at the p6 or p9 anchor positions (see Table 4) are bolded; weakly tolerated are underlined; nontolerated are bolded in lower case. h, human; m, mouse; r, rat.

that apart from the present study and that of Carrasco-Marin et al. (8), binding has not been determined by direct peptide interaction with purified I-A^{g7}, but either by elution from I-A^{g7}, competition with peptides that activate I-A^{g7}–bearing T cell hybridomas or induction of EAE. The motif we have defined correctly identifies 27/30 (90%) of the published sequences (Table 5). Interestingly, we found that one of these sequences, mouse serum albumin 560–574, that does not contain the motif, did not bind to I-A^{g7} (data not shown).

Two groups have suggested putative motifs for peptides that bind to I-A^{g7}. Reich et al. (7) found that several peptides eluted from I-A^{g7} had an acidic residue at the COOH terminus. Their data also indicated that this residue was separated by three from a basic residue. Whereas basic residues are major p9 anchors in our motif, an acidic residue at the COOH terminus is not a uniform feature of other peptides deduced (Table 5) or shown (Table 6) to bind to I-A^{g7}. However, it is conceivable that in some cases, e.g., OVA 323–339 (see Table 5), a COOH-terminal acidic residue could compensate for a nontolerated residue at p9. Amor et al. (9) described a possible motif shared by encephalitogenic peptides in the Biozzi AB/H (I-A^{g7}) mouse. It contained hydrophobic (I or L), basic (K, R, or H), a small T cell contact (A or G) and large hydrophobic (L or F) residues within a 6– to 7–amino acid core. They studied the effect of K substitutions on the immunogenicity of phospholipid protein 56–70 (see Table 5), in which they had deduced a core sequence, NVIHAFQ, necessary for the induction of EAE. This sequence contains our motif I (p6) and F (p9). K substitutions at I, H, A, or F completely abolished the ability of the peptide to induce EAE. We would have predicted abolition of binding by the K substitution at I or A and, by analogy with HEL (see Table 2), a significant reduction in T cell activation by K substitution at H or F. Thus, the features of this encephalitogenic motif are contained within the expanded and generalized motif we have described.

Presence of Motif in Overlapping Peptides from Human Proinsulin. We tested peptides overlapping by four residues and spanning the entire sequence of human proinsulin for binding to I-A^{g7}, and inspected them for presence of the binding motif (Table 6). All six (100%) good binders contained a motif. However, a motif was present in 4/13 (30%) weak or non-binders. Clearly, the motif rules do not fully account for the effects of residue combinations or flanking sequences. Proinsulin 5–19 has a well-tolerated V at p6 and a weakly tolerated L at p9 yet did not bind, but when this anchor pair moves towards the NH_2 terminus in the following 9–23 sequence, the peptide becomes a binder. Proinsulin 45–59 has a well-tolerated L at p6 and a weakly tolerated L at p9 and binds with high affinity, but when this anchor pair moves towards the NH_2 terminus in the following 49–63 sequence, the affinity of the peptide decreases. Human proinsulin 17–31 has a weakly tolerated Y at p6 and a well-tolerated K at p9, yet does not bind. Although this anchor pair is close to the COOH terminus, this does not preclude other peptides, e.g., human proinsulin 65–79, from binding with high affinity. However, human proinsulin 17–31 has a positively charged p9/COOH terminus, whereas the other binding peptides are generally neutral or tend to be acidic. Reich et al. (7) noted a bias towards acidic residues at the COOH termini of peptides they eluted from I-A^{g7}. When the anchor pair in this peptide moves towards the NH_2 terminus in the following 21–35 sequence with an acidic COOH terminus, the peptide becomes a binder. Thus, in a small set of unbiased peptides, the motif appears to have high sensitivity and some degree of specificity. A similar degree of specificity was found for an I-E^k motif by correlating binding with the presence of the motif in a panel of ~150 peptides (23).

Although the core region length of class II MHC binding peptides is ~13 residues (24), analysis of binding motifs (25) indicates that only nine residues within the core region are essential for binding. The motif we describe conforms with this requirement. It is simplistic in the sense that each residue is assumed to contribute to binding independently of other residues and, when located at a given position, to contribute the same amount to binding even within different sequences. The rules that govern binding to class II MHC molecules are more complex and will have to take account of interactions between residues. Nevertheless, by defining tolerated and nontolerated residues for binding at

Table 6. *Overlapping Human Proinsulin Peptides Tested for Binding to I-A^{g7}*

	Peptide sequence	IC_{50} (nM)	Motif
(1–15)	F **V** N Q H L *A* G S H L **V** E A L	7,000	−
(5–19)	H L *A* G S H L **V** E A L Y L V *A*	30,000	+
(9–23)	S H L **V** E A L Y L V *A* G E R G	1,000	+
(13–27)	E *A* L Y L V *A* G E R G F F Y T	4,000	−
(17–31)	L V A G E R G F F Y T P **K** T R	>50,000	+
(21–35)	E R G F F Y T P **K** T R **R** E A E	1,000	+
(25–39)	F Y T P **K** T R **R** E A E D L Q V	1,000	+
(29–43)	K T R **R** E A E D L Q V G Q V E	7,000	−
(33–47)	E A E D L Q V G Q V E L G G G	>50,000	−
(37–51)	L Q V G Q V E L G G G P G A G	>50,000	−
(41–55)	Q V E L G G G P G A G S L Q P	6,000	−
(45–59)	G G G P G A G S **L** Q P L A L E	200	+
(49–63)	G A G S **L** Q P L A L E G S L Q	1,000	+
(53–67)	**L** Q P L A L **e** G S **L** Q K **R** G I	25,000	−
(57–71)	A L **e** G S **L** Q K **R** G I V E Q *A*	15,000	−
(61–75)	S **L** Q K **R** G I V E Q *A* *A* T S I	12,000	−
(65–79)	R G I V E Q *A* *A* T S **I** *A* S L Y	400	+
(69–83)	E Q *A* *A* T S **I** *A* S **L** Y Q L E N	12,000	+
(73–86)	T S **I** *A* S **L** Y Q **L** E N **Y** *A* N	10,000	+

Residues well tolerated at the p6 or p9 anchor positions (see Table 4) are bolded; weakly tolerated are underlined; nontolerated are bolded in lower case. Cysteines have been substituted by alanine (*A* in italics).

key positions in HEL 12–22, we appear to have unearthed general rules that identify a large majority of known binders to I-A^{g7}, and discriminate most non-binders.

The high sensitivity of the motif for reported I-A^{g7} binders or T cell epitopes is remarkable, but the utility of the motif will depend on its specificity, i.e., its absence in non-binders. Specificity was 70% for the peptides in Table 6, but the database is small. Even if somewhat degenerate in its present form, the motif should considerably narrow the search for possible binders. The I-A^{g7} binding assay we have described is robust and will enable the database of binders and non-binders to be enlarged progressively to further validate the motif. Experiments to fine tune the motif by using peptide libraries are in progress. Just as a motif for human class II DR4 (*0401) binding peptides was applied to scan candidate autoantigen proteins in rheumatoid arthritis for potential epitopes (26), so also might the motif for I-A^{g7} binding be applied to identify potential autoepitopes for IDDM in NOD mice and CR-EAE in Biozzi AB/H mice.

L.C. Harrison and M.C. Honeyman were supported by the National Health and Medical Research Council of Australia and a Diabetes Interdisciplinary Research Program grant from the Juvenile Diabetes Foundation International. P. Augstein was supported by Deutscher Akademischer Austauschdienst, Bonn. Secretarial assistance was provided by Margaret Thompson.

Address correspondence to L.C. Harrison, The Walter and Eliza Hall Institute of Medical Research, Post Office, Royal Melbourne Hospital, Parkville, 3050, Australia.

Received for publication 18 November 1996 and in revised form 8 January 1997.

References

1. Kikutani, H., and S. Makino. 1992. The murine autoimmune diabetes model: NOD and related strains. *Adv. Immunol.* 51: 285–322.
2. Liu, G.Y., D. Baker, S. Fairchild, F. Figueroa, R. Quartey-Papafio, M. Tone, D. Healey, A. Cooke, J.L. Turk, and D.C. Wraith. 1993. Complete characterization of the expressed immune response genes in Biozzi AB/H mice: structural and functional identity between AB/H and NOD A region molecules. *Immunogenetics*. 37:296–300.
3. Acha-Orbea, H., and H.O. McDevitt. 1987. The first external domain of the nonobese diabetic mouse class II I-A β chain is unique. *Proc. Natl. Acad. Sci. USA*. 84:2435–2439.
4. Tait, B.D., and L.C. Harrison. 1991. Overview: the major histocompatibility complex and insulin-dependent diabetes mellitus. *Baillieres Clin. Endocrinol. & Metab.* 5:211–228.
5. Rammensee, H.-G. 1995. Chemistry of peptides associated with MHC class I and class II molecules. *Curr. Opin. Immunol.* 7:85–96.
6. Sinigaglia, F., and J. Hammer. 1994. Defining rules for the peptide–MHC class II interaction. Curr. Opin. Immunol. 6: 52–56.
7. Reich, E.-P., H. von Grafenstein, A. Barlow, K.E. Swendon, K. Williams, and C.A. Janeway, Jr. 1994. Self peptides isolated from MHC glycoproteins of non-obese diabetic mice. *J. Immunol.* 152:2279–2288.
8. Carrasco-Marin, E., J. Shimizu, O. Kanagawa, and E.R. Unanue. 1996. The class II MHC I-A^{g7} molecules from non-obese diabetic mice are poor peptide binders. *J. Immunol.* 156:450–458.
9. Amor, S., J.K. O'Neill, M.M. Morris, R.M. Smith, D.C. Wraith, N. Groome, and D. Baker. Encephalitogenic epitopes of myelin basic protein, proteolipid protein, and myelin oligodendrocyte glycoprotein for experimental allergic encephalomyelitis induction in Biozzi AB/H (H-2A^{g7}) mice share an amino acid motif. *J. Immunol.* 156:3000–3008.
10. Kappler, J., J. White, D. Wegmann, E. Mustain, and P. Marrack. 1982. Antigen presentation by Ia^+ B cell hybridomas to H-2-restricted T cell hybridomas. *Proc. Natl. Acad. Sci. USA*. 79:3604–3608.
11. O'Reilly, L.A., P.R. Hutchings, P.R. Crocker, E. Simpson, T. Lund, D. Kioussis, F. Takei, J. Baird, and A. Cooke. 1991. Characterization of pancreatic islet cell infiltrates in NOD mice: effect of cell transfer and transgene expression. *Eur. J. Immunol.* 21:1171–1180.
12. Chosich, N., and L.C. Harrison. 1993. Suppression of diabetes mellitus in the non-obese diabetic (NOD) mouse by an autoreactive (anti-T-A^{nod}) islet-derived CD4+ T cell line. *Diabetologia*. 36:716–721.
13. Adorini, L., A. Sette, S. Buus, H.M. Grey, M. Darsley, P.V. Lehmann, G. Doria, Z.A. Nagy, and E. Appella. 1988. Interaction of an immunodominant epitope with Ia molecules in T-cell activation. *Proc. Natl. Acad. Sci. USA*. 85:5181–5185.
14. Hurtenbach, U., E. Lier, L. Adorini, and Z.A. Nagy. 1993. Prevention of autoimmune diabetes in non-obese diabetic mice by treatment with a class II major histocompatibility complex–blocking peptide. *J. Exp. Med.* 177:1499–1594.
15. Deng, H., R. Apple, M. Clare-Salzler, S. Trembleau, D. Mathis, L. Adorini, and E. Sercarz. 1993. Determinant capture as a possible mechanism of protection afforded by major histocompatibility complex class II molecules in autoimmune disease. *J. Exp. Med.* 178:1675–1680.
16. Chen, S.-L., P.J. Whiteley, D.C. Freed, J.B. Rothbard, L.B. Peterson, and L.S. Wicker. 1994. Responses of NOD congenic mice to a glutamic acid decarboxylase–derived peptide. *J. Autoimmunity*. 7:635–641.
17. Kaufman, D.L., M. Clare-Salzler, J. Tian, T. Forsthuber, G.S.P. Ting, P. Robinson, M.A. Atkinson, E.E. Sercarz, A.J. Tobin, and P.V. Lehmann. Spontaneous loss of T-cell tolerance to glutamic acid decarboxylase in murine insulin-dependent diabetes. *Nature (Lond.).* 366:69–72.
18. Smilek, D.E., C.B. Lock, and H.O. McDevitt. 1990. Antigen recognition and peptide-mediated immunotherapy in autoimmune disease. *Immunol. Rev.* 118:37–71.
19. Vaysburd, M., C. Lock, and H. McDevitt, 1995. Prevention

of insulin-dependent diabetes mellitis in nonobese diabetic mice by immunogenic but not by tolerated peptides. *J. Exp. Med.* 182:897–902.

20. Daniel, D., and D.R. Wegmann. 1996. Protection of nonobese diabetic mice from diabetes by intranasal or subcutaneous administration of insulin peptide B-(9–23). *Proc. Natl. Acad. Sci. USA*. 93:956–960.
21. Ridgway, W.M., M. Fasso, A. Lanctot, C. Garvey, and C.G. Fathman. 1996. Breaking self-tolerance in nonobese diabetic mice. *J. Exp. Med.* 183:1657–1662.
22. Bernard, C.C.A., T.G. Johns, A. Slavin, M. Ichikawa, C. Ewing, J. Liu, and J. Bettadapura. 1996. Myelin oligodendrocyte glycoprotein: a novel candidate autoantigen in multiple sclerosis. *J. Mol. Med.* In press.
23. Leighton, J., A. Sette, J. Sidney, E. Appella, C. Ehrhardt, S. Fuchs, and L. Adorini. 1991. Comparison of structural requirements for interaction on the same peptide with I-E^k and I-E^d molecules in the activation of MHC class II–restricted T cells. *J. Immunol.* 147:198–204.
24. Jardetzky, T.S., J.B. Brown, J.C. Gorga, L.J. Stern, R.G. Urban, J.L. Strominger, and D.C. Wiley. 1996. Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II–like conformation for bound peptides. *Proc. Natl. Acad. Sci. USA*. 93:734–738.
25. Rammensee, H.G., T. Friede, and S. Stevanovic. 1995. MHC ligands and peptide motifs: first listing. *Immunogenetics*. 41:178–228.
26. Hammer, J., F. Gallazzi, E. Bono, R.W. Karr, J. Guenot, P. Valsasnini, Z.A. Nagy, and F. Sinigaglia. 1995. Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association. *J. Exp. Med.* 181:1847–1855.