Peptide Binding Specificity of HLA-DR4 Molecules: Correlation with Rheumatoid Arthritis Association

By Juergen Hammer, Fabio Gallazzi, Elisa Bono, Robert W. Karr,‖ Jeanmarie Guenot,§ Paola Valsasnini, Zoltan A. Nagy,‡ and Francesco Sinigaglia

From Roche Milano Ricerche, I-20132 Milan, Italy; the ‡Departments of Inflammation/Autoimmune Diseases and §Physical Chemistry, Hoffmann-La-Roche Inc., Nutley, New Jersey 07110; and the ‖Department of Immunology, G.D. Searle & Co., St. Louis, Missouri 63198

Summary

We have investigated whether sequence 67 to 74 shared by β chains of rheumatoid arthritis (RA)–associated HLA-DR molecules imparts a specific pattern of peptide binding. The peptide binding specificity of the RA-associated molecules, DRB1*0401, DRB1*0404, and the closely related, RA nonassociated DRB1*0402 was, therefore, determined using designer peptide libraries. The effect of single key residues was tested with site-directed mutants of DRB1*0401. The results have demonstrated striking differences between RA-linked and unlinked DR allotypes in selecting the portion of peptides that interacts with the 67–74 area. Most differences were associated with a single amino acid exchange at position 71 of the DR β chain, and affected the charge of residues potentially contacting position 71. The observed binding patterns permitted an accurate prediction of natural protein derived peptide sequences that bind selectively to RA-associated DR molecules. Thus, the 67-74 region, in particular position 71, induces changes of binding specificity that correlate with the genetic linkage of RA susceptibility. These findings should facilitate the identification of autoantigenic peptides involved in the pathogenesis of RA.

Susceptibility to rheumatoid arthritis (RA)[1] is specifically associated with the class II MHC alleles *DRB1*0401*, *DRB1*0404*, and *DRB1*0101* (1, 2). Interestingly, the DRβ chains encoded by these genes possess a "shared epitope" formed by a short stretch of amino acids (at positions 67–74) that is highly conserved among RA-associated molecules (3, 4). Since DRB*0402, a closely related molecule not associated with RA, differs from the RA-linked DRB1*0401 and DRB1*0404 molecules only in the 67–74 region (at positions 67, 70, and 71), this part of the molecule is likely to be critical for disease association.

Polymorphic MHC residues can affect the immune system in at least two different ways. First, they can alter the specificity of the peptide binding site (5) and thus control the array of antigenic peptides presented to T cells. Second, they influence the selection of TCR specificities during the development of the T cell repertoire (6, 7). The role of the 67–74 region in RA association can be explained on a similar basis, since autoimmunity is considered to underlie the pathogenesis of RA. Thus, RA-associated DR allotypes, because of their similar peptide binding specificity, may be the ones capable of presenting certain autoantigenic peptides critical for the development of disease. Alternatively, this region may be required for the selection of TCR capable of recognizing the pathogenic self peptide(s).

In an attempt to gain insight into the mechanism of RA association, we have studied the influence of the 67–74 region on the peptide binding specificity of RA associated and nonassociated DR molecules, using site-directed mutants of class II molecules, designer peptide libraries, and high throughput binding assays (8, 9). The results reported here demonstrate a striking correlation between binding specificity and disease association, depending largely on a single amino acid position 71 of the β chain.

Materials and Methods

Preparation of HLA-DR Molecules

DR molecules were isolated from human lymphoblastoid cell lines BSM (DRB1*0401), 2046 (DRB1*0402), and MT (DRB1*0404), and L cell transfectants as described (10). The following L cell transfectants were used: L 565H.1 (DRB1*0401) and site-directed mutants L 286.10 where K was exchanged to E at position 71 of DRB1*0401 (0401/K71E) and L 335.2 where K was exchanged to R at position 71 of DRB1*0401 (0401/K71R) (11). Site-directed

[1] *Abbreviations used in this paper:* NMP, *N*-methyl-2 pyrrolidone; RA, rheumatoid arthritis; RT, room temperature; p, peptide position.

mutagenesis of DRB1*0401 cDNA was performed by PCR overlap extension method, and the mutants were sequenced to confirm the presence of the mutation and absence of misincorporation.

Peptide Synthesis

The selectively randomized peptide libraries, the set of peptide position 1 (p1) anchored designer peptides and peptides with natural protein-derived sequences were synthesized with a multiple peptide synthesizer (model 396; Advanced ChemTech, Louisville, KY) using Fmoc chemistry and solid phase synthesis (12).

Selectively Randomized Peptide Libraries. 19 selectively randomized peptide libraries GZAAXAXAAX were synthesized where each one carries a different natural amino acid residue (except Cys) at relative p1 (Z), p4, p6, and p9 were randomized (X) in all sublibraries and p2, p3, p5, p7, and p8 were occupied by Ala residues (A). The synthesis was performed as follows: 3 g resin was equally distributed into 19 reaction vessels preceding the synthesis of a random peptide position. Each Fmoc natural amino acid residue (except Cys) was individually coupled using quadruple coupling procedures. After the synthesis of the Ala spacer and before the next random position, the resin was recovered from the reaction vessels, mixed, and redistributed. Mixing was omitted after the synthesis of the Z position.

P1-anchored Designer Peptides. A set of p1-anchored designer peptides was synthesized as described (9). It consists of the basic peptide GFKAAAAAAA-NH_2 where all natural amino acid residues (except Cys) were single substituted at p3–p7, and F corresponds to p1. The sequence of the basic peptide is the result of several optimization experiments with the purpose of reaching an IC_{50} value within a given binding assay, which allows one to measure 100-fold decreases and increases of the IC_{50} after the introduction of anchor and inhibitory residues, respectively. The peptides were analyzed by analytical reverse-phase HPLC and the peptides used in this study were routinely ≥75% pure.

Biotinylation of Peptides. Peptides were specifically NH_2-terminal biotinylated using the Advanced ChemTech 396 robotic system as follows: 50 mg of resin with NH_2-terminal deprotected peptides were suspended and washed at least three times in *N*-methyl-2-pyrrolidone (NMP). A two-fold excess of biotin-XX-*n*-hydroxysuccinimide over the free NH_2 termini of the resin in 1 ml NMP was added, and the suspension was incubated for 30 min with agitation. The biotinylation step was repeated once. Subsequently, the resin was washed three times with NMP, five times with methanol, and dried in vacuum. The biotinylated peptides were deprotected and cleaved from the resin with 95% TFA, 2.5% thioanisol, 1.25% distilled water, and 1.25% phenol.

DR-Peptide Binding Assays

Peptide side chain scanning was performed either with a high-flux DR-peptide competition assay as described (8, 9), or with an ELISA-based high-flux competition assay, especially developed to detect peptide binding to DR molecules isolated from L cell transfectants. The ELISA-based assay was performed as follows. Peptides were dissolved in DMSO at a concentration of 5 mM, diluted 1:4 in PBS, and transferred into the (A)-row wells of 96-well plates (dilution plates). 90 μl of DMSO/PBS (1:4) was added to the (B)-(F) row wells of the dilution plates and the peptides were simultaneously diluted by serially transferring 10 μl aliquots from row (A) to row (B)-(F). 10 μl of each well was transferred to the corresponding wells of a second set of 96-well plates (binding plates). 10 μl of NH_2-terminal biotinylated peptide in PBS (0.5–5 μM) was added to each well of the binding plate. Subsequently, 30 μl of affinity-purified DR4 antigens (~0.2–1 μg protein) in 5 μl volume and 25 μl binding buffer (2% NP40, 2 mM PMSF, 2 mM EDTA, 20 μg/ml each of soybean trypsin inhibitor, antipain, pepstatin, leupeptin, and chymostatin, 0.15 M NaCl, and 50–100 mM citrate phosphate buffer, pH 7) was added to each well of the binding plates and incubated for ≥24 h at room temperature (RT). The ELISA plates were prepared as follows. 50 μl per well of 10 μg/ml anti-DR mAb (L243) was used to coat the plates. The wells were washed twice with distilled water, blocked with 200 μl blocking buffer (PBS with 0.05% Tween 20, 1 mM EDTA, 0.25% BSA fraction V, and 0.05% NaN_3) for at least 2 h at RT and washed twice with distilled water. 25 μl PBS/0.15% Tween 20 was added to each well of the binding plates. Subsequently, 50 μl of each well was transferred to the corresponding wells of the ELISA plates and incubated for ≥2 h at RT. After washing three times with PBS/0.05% Tween 20, 50 μl Streptavidin-alkaline-phosphatase in PBS/0.05% Tween 20 was added to each well and incubated for 30 min at RT. The plates were then washed twice with distilled water, once with PBS/0.05% Tween 20, and twice again with distilled water. 200 μl per well of 1 mg/ml *p*-nitrophenylphosphate in 50 mM $NaCO_3$ and 50 mM $MgCl_2$ was added and incubated for ~3 h at 37°C. The amount of substrate hydrolyzed was assessed with a spectrophotometer (405 nm). The radiolabeled (conventional high-flux assay [9]) or biotinylated (ELISA-based assay) peptides used in this study were either flu hemagglutinin peptide HA 307-319 (for DRB1*0401 only) or the "universal DR4" (UD4) peptide, YPKFVKQNTLKAA, designed to bind to all DR4 allotypes with high affinity.

*Modeling of the DRB*0402 Binding Site*

The DRB1*0402 peptide model was constructed from the DRB1*0101-HA peptide crystal structure model (13) by substituting the DRB1*0402 amino acid residues at positions 9, 11, 13, 26, 28, 30, 31, 37, 67, 70, 71, and 86 on the DRB1*0101 β-subunit and replacing the residue side chains in the HA peptide with Ac-FKAWATAAA-NH_2 such that the Phe in the modeled peptide occupied the Tyr site in the HA peptide. The coordinates for the peptide backbone were not altered from the crystallographic positions. This model was minimized to a 0.1 Å root mean square (RMS) energy gradient using the AMBER 4.0 molecular mechanics and dynamics software package (14) with the all atom force field (15) and a distance-dependent dielectric model. Minimization was followed by 150 picoseconds of molecular dynamics at 300 K during which the entire ligand and the altered amino acid residues on the protein were allowed to move according to the classical laws of motion. An additional minimization to a 0.01 Å RMS energy gradient yielded the final structure. The refined model was displayed using the GRASP software package (16).

Results

*Similar Residues at the Main Anchor Positions Are Required for Peptide Binding to DRB1*0401, and DRB1*0404, and DRB1*0402.* The RA-associated DR4 subtypes, DRB1*0401, and DRB1*0404, differ only at amino acid positions 67, 70, and 71 from the nonassociated subtype DRB1*0402 (Table 1). As shown by a model of the DRB1*0402/DRA*0101 peptide binding site (Fig. 1), this difference is located such that it may not alter the specificity of DR4 subtypes for anchor residues at p1 and p6, although it is predicted to affect residues at p4, p5, and p7. We tested this hypothesis by com-

Table 1. *Amino Acid Differences between the DR Molecules Used in this Study*

	DRβ residues at position			
HLA-DR	67	70	71	86
DRB1*0401	L	Q	K	G
DRB1*0404	–*	–	R	V
DRB1*0402	I	D	E	V
DRB1*0401/K71R‡	–	–	R	–
DRB1*0401/K71E§	–	–	E	–

* Sequence identity with DRB1*0401.
‡ Site-directed mutants of Lys to Arg (K71R), and
§ Lys to Glu (K71E) at position 71 of DRB1*0401.

paring the peptide binding specificity of DRB1*0404 and DRB1*0402. To determine the p1 anchor specificity, we used selectively randomized peptide libraries with relative anchor positions 4, 6, and 9 randomized and p1 fixed to 1 of 19 amino acid residues as described in Materials and Methods. By this method, the influence of a particular peptide sequence on binding is minimized, and thus, the effect of a single residue at p1 can be tested in a manner largely independent of the remaining sequence. The results defining p1 anchor specificity are shown in Fig. 2. It is clear that Ile, Leu, Met, Phe, and Val can serve as p1 anchor residues for both DR allotypes. This pattern is almost identical with the p1 specificity of DRB1*0401 and DRB1*0101 (17). The difference is that residues Trp and Tyr can serve as anchors for DRB1*0401 and DRB1*0101, but they are not accepted by either DRB1*0404 or DRB1*0402. This correlates with a Gly to Val exchange at position 86 of the corresponding DRβ chains that is known to alter the hydrophobic pocket accommodating the p1 anchor (13).

We then compared the effect of different side chains at p6 on peptide binding to the three DR4 subtypes. These experiments were performed using a set of p1-anchored designer peptides where all natural amino acid residues (except Cys) were single substituted at relative p6 as described in Materials and Methods. As shown in Fig. 3, *a* and *b*, the same p6 anchor residues, Ser, Thr, Asn, and Val, were identified for all three subtypes. We observed only one significant difference between DRB1*0402 and the two RA-associated subtypes, namely that Arg at p6 did not influence binding to the former, but inhibited binding to the latter (see below). We also performed anchor shifting experiments, and demonstrated identical spacing of the two major anchor positions (p1 and p6) for the three DR allotypes (Fig. 3 *c*). These findings suggest a similar overall conformation of peptides bound to RA-associated and nonassociated DR4 subtypes.

Position 71 of the DRβ Chain Induces Major Differences of Peptide Binding Specificity between DR4 Subtypes. The overall similar conformation of peptides bound to the three DR4 subtypes permitted us to systematically scan p3 to p7 for the

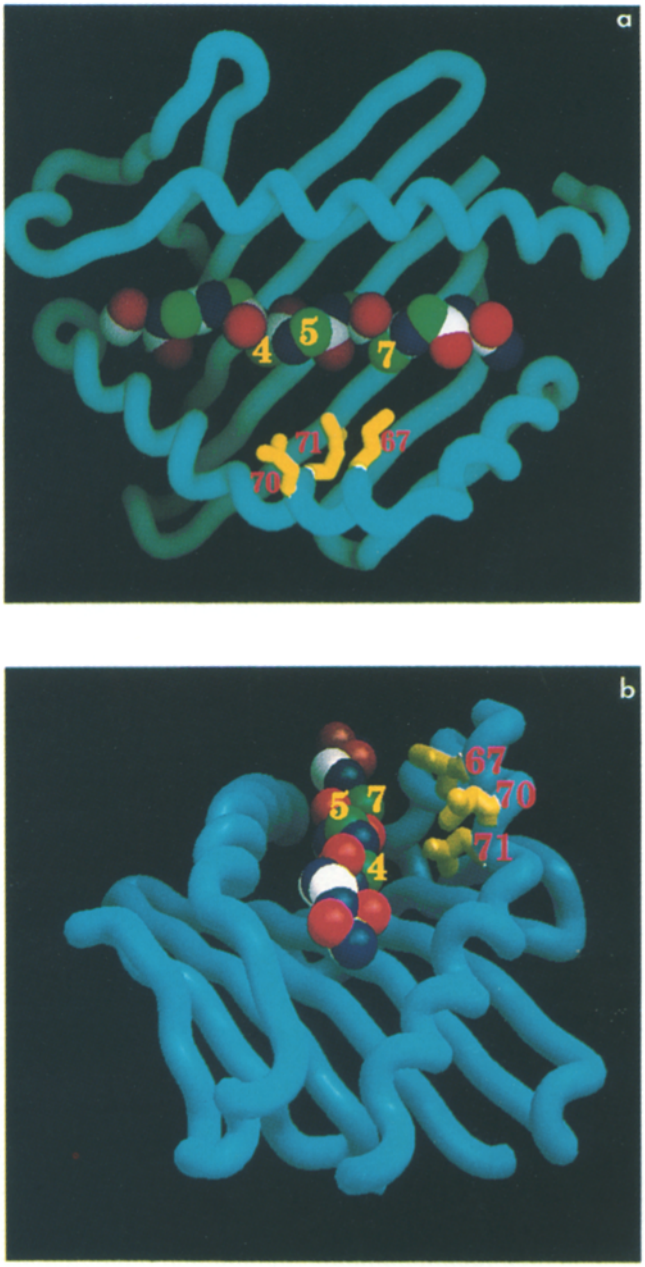

Figure 1. The DRB1*0402 peptide model viewed from above the binding site (*a*), and from the side of position 1 of the bound peptide (*b*). The protein Cα backbone of the binding site is shown schematically as a tube diagram (*cyan*) with the α-subunit helix above the β-subunit helix (*a*), and the α-subunit helix left and β-subunit helix right (*b*). The amino acid substitutions distinguishing DRB1*0402 from the RA-linked DR4 molecules at position 67, 70, and 71, respectively, are clustered near the center of the β-subunit helix (*yellow*). Note that the center of the β-subunit helix is adjacent to the middle of the bound peptide (*a*) and it bulges up above the bound peptide (*b*), offering sites of interaction with side chains branching out of the Cα atoms (*green*) at p4, p5, and p7 of the peptide (p6, when occupied by a small anchor residue such as Thr, is buried in the binding site, and thus not seen). The backbone of the bound peptide (Ac-FKAWATAA-NH2) is shown as spheres (*blue*, N; *red*, O; *white*, carbonyl C; *green*, Cα).

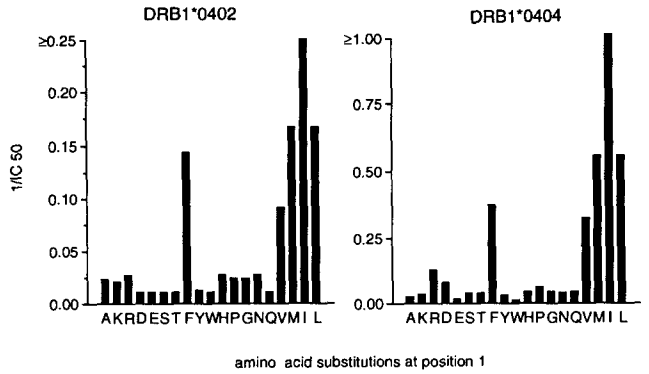

Figure 2. p1 anchor residues defined by 19 selectively randomized designer peptide libraries. Each of the 19 peptide libraries carries a different natural amino acid residue at relative p1, and three selectively randomized p4, p6, and p9 (see Materials and Methods). Data are expressed as reciprocal IC_{50} values from the DR-peptide competition assay using radiolabeled UD4 as a reference peptide.

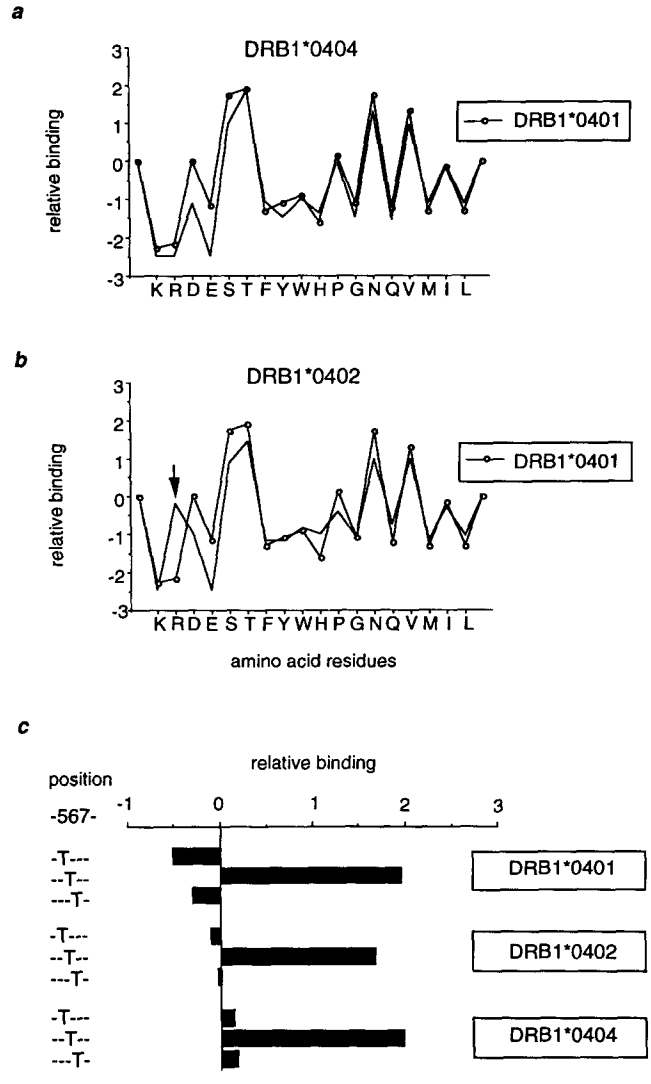

Figure 3. Similar specificity (*a* and *b*) and spacing (*c*) of the p6 anchor for peptide binding to DRB1*0401, DRB1*0404, and DRB1*0402. The effect of 19 amino acid side chains on binding to DRB1*0404 (*a*), and DRB1*0402 (*b*) was determined at relative anchor position 6. P6 anchor shifting (*c*) was performed to test spacing between p1 and p6. Data were determined by side chain scanning using GFKAAAAAAA based, single-substituted, p1-anchored peptides, and compared with side chain scanning results with DRB1*0401 using YASAAAAAA based peptides (*Y* = p1) (9). The data in (*a–c*) are expressed as relative binding, derived as follows: IC_{50} values were normalized by dividing them with IC_{50} values of Ala substitutions. The logarithm of reciprocal, normalized IC_{50} values was plotted [log ($IC_{50\text{-}Ala}/IC_{50\text{-}Exp}$)]. Data are derived from the high-flux DR-peptide competition assay, using radiolabeled HA 307-319 (DRB1*0401) and UD4 (DRB1*0402 and DRB1*0404) as reference peptides.

Table 2. *Results of Side Chain Scanning on DRB1*0402*

G	F	K	X	X	X	X	X	A	A
	p1	p2	p3	p4	p5	p6	p7	p8	p9
A			0	0	0	0	0		
D			−0.9*	**−2.3‡**	**−2.3**	−1.0	−2.1		
E			−1.1	**−2.3**	−1.3	−2.5	−1.4		
F			−0.3	+0.3	−0.3	−1.2	+0.9		
G			−0.9	−0.8	−1.4	−1.1	−1.3		
H			−0.2	+1.2	−0.2	−1.0	+0.5		
I			+0.5	+0.1	−0.1	−0.3	+0.5		
K			−0.6	**+0.1**	+0.3	−2.5	0		
L			+0.2	−0.6	−0.1	−1.0	+1.0		
M			+0.6	+0.6	−0.1	−1.2	+0.8		
N			−0.5	−0.4	−1.2	+1.0	+0.6		
P			−0.7	−1.3	+0.1	−0.4	−1.0		
Q			−0.6	−0.4	−0.2	−0.8	+1.1		
R			−0.4	**+1.0**	+0.5	**−0.2**	**+1.7**		
S			−0.4	−0.9	−0.2	+0.9	−0.4		
T			−0.6	−0.5	−0.1	+1.5	+0.1		
V			0	−0.6	0	+1.0	+0.2		
W			−0.6	**+1.6**	−0.6	−0.8	**+1.4**		
Y			0	−0.5	−0.3	−1.2	+0.9		

* Data are expressed as relative binding, i.e., the logarithm of the reciprocal of Ala-normalized IC_{50} values (log [$IC_{50\text{-}Ala}/IC_{50\text{-}Exp}$]). Side chain scanning was performed using a set of p1-anchored designer peptides where all amino acid residues, except Cys, were single substituted at relative positions 3–7. GFKAAAAAAA-NH_2 was the basic peptide where *F* corresponds to relative position 1. The average IC_{50} of the basic peptide was 6 μM. The radiolabeled UD4 was used as a reference peptide in the high-flux DR-peptide competition assay.

‡ The DRB1*0402 scanning values that differ from both DRB1*0401 (at least 100-fold) and DRB1*0404 (at least 10-fold) are in boldface. Significant differences between DRB1*0402 and the RA-associated subtypes were confirmed by several independent experiments using different sets of peptides.

induction of differences in binding specificity. As shown by the model in Fig. 1, this part of the peptide is likely to be influenced by the 67–74 area of the binding site. The results of side chain scanning for DRB1*0402 are summarized in Table 2. Similar scanning was performed for DRB1*0401 (9),

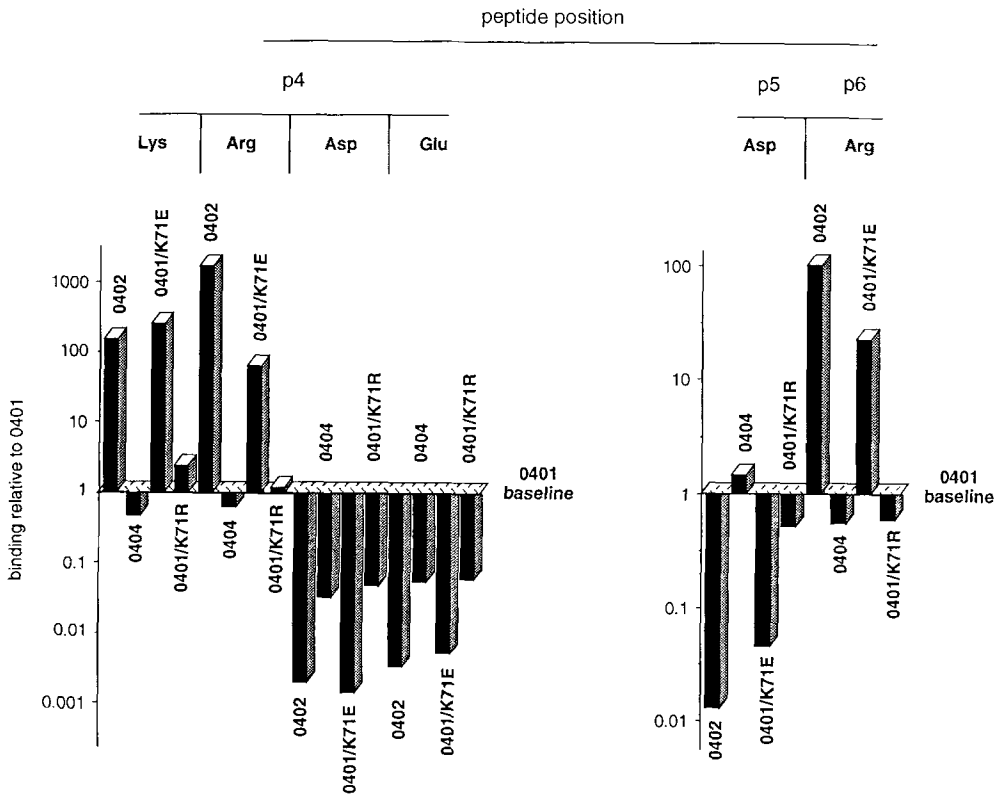

Figure 4. Differences in peptide binding specificity of RA-associated and nonassociated DR4 subtypes, and the effect of position β 71. The effect of the charged residues (in boldface) in Table 2 is shown for p4, p5, and p6. Relative binding data were derived as in Fig. 3, and plotted in comparison to binding data of DRB1*0401 as the baseline. GFK-AAAAAAA was used as basic peptide for all DR subtypes and mutants. Data were either derived from the DR-peptide competition assay, using radiolabeled UD4 as reference peptide, or from the ELISA-based DR-peptide competition assay (DR molecules from wild-type and mutant transfectants), using biotinylated UD4 as a reference peptide.

and DRB1*0404 (data not shown). Significant differences between DRB1*0402 and the two RA-associated subtypes were observed in the effect of charged residues. Most of these effects were seen at p4, but some were also identified at p5–p7. The nature of these charge effects is shown in Fig. 4. The most striking difference is that peptides with positively charged residues, Lys or Arg, at p4 or p6 bind 100–1,000-fold better to DRB1*0402 than to DRB1*0401 and DRB1*0404. Conversely, peptides with negative residues, Asp or Glu, at p4 or p5 bind 100–500-fold less well to DRB1*0402 than to DRB1*0401, and ~50-fold less well than to DRB1*0404. A site-directed mutant molecule that differs from DRB1*0401 only in a single Lys to Glu exchange at position 71 of the β chain exhibits a binding specificity similar to that of DRB1*0402 (Fig. 4). Thus, exchange of the basic Lys to the acidic Glu at position 71 seems to account alone for most major differences in binding specificity between RA-associated and nonassociated DR4 allotypes. We have also observed a smaller but significant difference between the two RA-associated molecules. Namely, peptides with a negative residue (Asp or Glu) at p4 bound 20–50-fold better to DRB1*0401 than to DRB1*0404. Mutating position 71 Lys of DRB1*0401 to the Arg found in DRB1*0404 resulted in decreased binding of these peptides to levels comparable to the binding by DRB1*0404 (Fig. 4). Thus, position 71 of the β chain is responsible also for the minor difference between the two RA-linked allotypes. We found three additional significant differences between DRB1*0402 and the other two DR4 subtypes, i.e., peptides with Trp at p4 or p7, or Arg at p7 bound to the former 10–100-fold better than to the latter (Table 2). Since these effects could not be unequivocally associated with position 71 (data not shown), additional amino acid differences between these molecules (see in Table 1) may be responsible for them.

Prediction of Natural Protein Sequences Selectively Binding to RA-associated DR4 Molecules. To address the biological significance of the observed differences in binding specificity, we asked whether it is possible to predict protein regions selectively binding to RA-associated DR4 subtypes. For this purpose, we selected four proteins implicated in the pathogenesis of arthritis (18–21). Of 3,518 nonapeptide frames screened using the selection criteria in Fig. 5, only 14 frames were predicted to bind to DRB1*0401 and failed to bind to DRB1*0402. In the selection, we used a relatively low binding score for DRB1*0401 (corresponding to a low affinity cutoff; 9) to include most frames, potentially binding to this molecule. To evaluate our prediction, all peptides with scores ≥ 2 (9) and p1 anchor residues used by all DR4 subtypes (see above) were synthesized and tested for binding to DRB1*0401. 25 of them, with a high binding affinity to DRB1*0401 comparable to that of HA 307-319 (± fivefold, data not shown), were selected for comparative binding studies (peptide sequences listed in Table 3). Eight of them were predicted to bind selectively to RA-associated DR4 subtypes according to the rules in Fig. 5, that is, they were expected to bind poorly to DRB1*0402. As shown in Fig. 6 *a*, these 8 peptides did bind poorly to DRB1*0402, whereas 15 of 17 peptides predicted to be nonselective DR4 binders were found

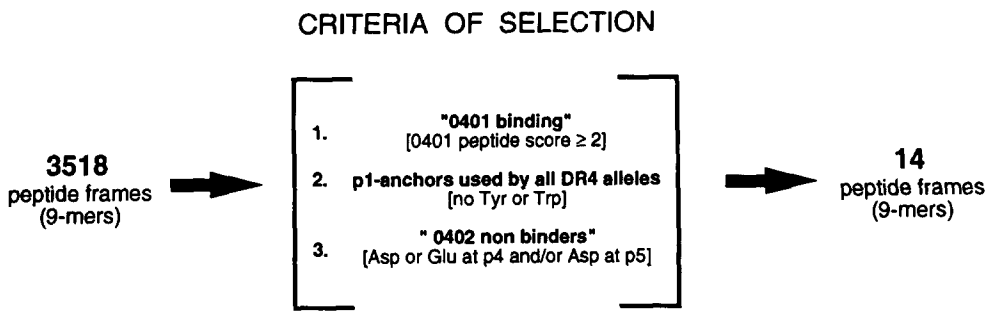

Figure 5. Criteria used to identify natural protein-derived peptides that selectively bind to RA-associated DR4 molecules. The peptide score is the sum of the effects exerted by side chains at each position, in comparison to Ala, on binding of the peptide. This value is derived from side chain scanning data for DRB1*0401, and correlates with binding affinity (9). The 3518 nonamers screened were from cartilage link protein, procollagen alpha 1 (II) chain precursor, hsp65 of Mycobacterium tuberculosis, and fillagrin (up to residue 1243).

Table 3. *High Affinity DRB1*0401-binding Natural Peptides with Predicted Selectivity for RA-associated DR4 Molecules*

No.	Selected peptides		Predicted selective binding
		123456789*	
1.	Link 61–73	AKFYRDPTAFGSG§	+‖
2.	Link 170–182	YNLNFHEAQQAAL	–¶
3.	Link 249–261	DVFAFTSNFNGRF	–
4.	Link 268–280	TKLTYDEAVQAAL	+
5.	Collagen 13–25	LTLLVAAVLRAQG	–
6.	Collagen 1165–1177	DPLQYMRADQAAG	–
7.	Collagen 1168–1180	QYMRADQAAGGLR	+
8.	Collagen 1306–1318	MTFLRLLSTEGSQ	–
9.	Collagen 1307–1319	TFLRLLSTEGSQN	–
10.	Collagen 1340–1352	KALLIQGSNDVEI	–
11.	Collagen 1350–1362	VEIRAEGNSRFTY	+
12.	hsp65 181–193	FGLQLELTEGMRF	+
13.	hsp65 189–201	EGMRFDKGYISGY	+
14.	hsp65 216–228	PYILLVSSKVSTV	–
15.	hsp65 284–296	KAMLQDMAILTGG	+
16.	hsp65 509–521	GLFLTTEAVVADK	–
17.	Fillagrin 9–21	YILFAHLVNKVHI	–
18.	Fillagrin 34–46	FAIINLFNEYSKK	–
19.	Fillagrin 491–503	GSFLYQVSTHEQS	–
20.	Fillagrin 492–504	SFLYQVSTHEQSE	–
21.	Fillagrin 644–656	SSFSQDRDSQAQS	+
22.	Fillagrin 815–827	GSFLYQVSTHEQS	–
23.	Fillagrin 816–828	SFLYQVSTHEQSE	–
24.	Fillagrin 1139–1151	GSFIYQVSTHEQS	–
25.	Fillagrin 1140–1152	SFIYQVSTHEQSS	–

* Numbers indicate the position and the p1-frames of the selected peptides.
‡ Peptides were selected according to criteria 1 and 2 in Fig. 5, and shown experimentally to have affinities ± fivefold that of HA 307–319. Peptide sequences were derived from the proteins described in Fig. 5.
§ Cys residues were replaced by Ala residues in all peptides.
‖ Predicted to bind selectively to RA-associated DR4 subtypes (Fig. 5).
¶ Predicted to be nonselective DR4 binders according to Fig. 5.

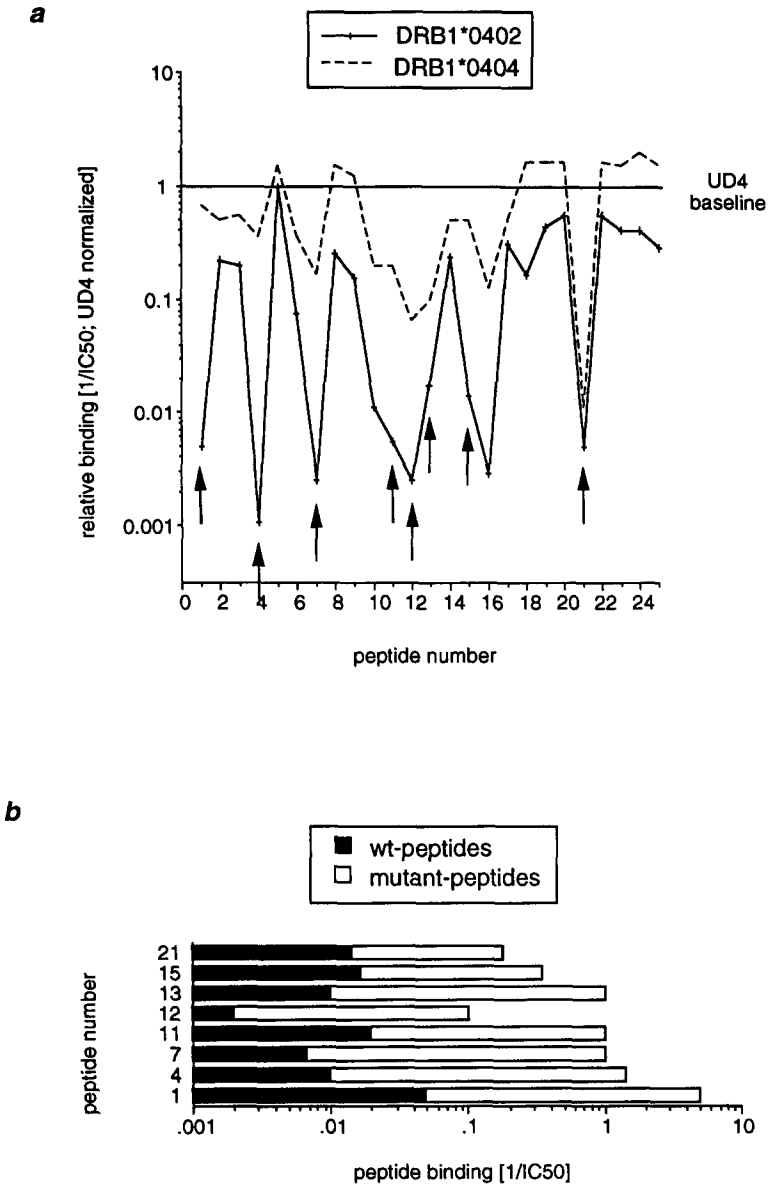

Figure 6. Binding of the selected DRB1*0401-binding peptides (shown in Table 3) to RA-associated DRB1*0404 and nonassociated DRB1*0402 molecules (*a*), and the effect of removing negative peptide residues on DRB1*0402 binding (*b*). (*a*) The relative binding data, derived from reciprocal IC_{50} values, were based on the binding values of the UD4 peptide ($IC_{50} \sim 0.1$ μM) which was set as 1. Arrows indicate peptides predicted to bind selectively to RA-associated DR4 subtypes, i.e., to bind poorly to DRB1*0402. Peptide numbers are defined in Table 3. Data were derived from the DR-peptide competition assay, using radiolabeled UD4 as a reference peptide. (*b*) In the mutant peptides, the negative residues Asp and Glu were exchanged to Ala at p4 (and at p5 of peptides 4 and 11, because of slight negative effects of E and G at p5). Binding data were derived from reverse IC_{50} values from the ELISA-based DR-peptide competition with biotinylated UD4 as reference peptide.

to be nonselective. Moreover, exchange of the critical p4 residues to Ala in the peptides predicted to bind selectively to RA-associated subtypes increased their binding affinity ~100-fold for the RA nonassociated molecule DRB1*0402 (Fig. 6 *b*). These data confirm the role of position 71 of the DRβ chain in determining peptide binding specificity, and demonstrate the potential applicability of the concept in the identification of autoantigenic peptides.

Discussion

The study presented here, together with our previous work (9, 17, 22, 23), establishes the rules of peptide binding to DR molecules associated with susceptibility to RA. There are three major anchor and inhibitory positions involved in binding, at p1, p4, and p6 of the p1-frame. At p1, a hydrophobic anchor is required for binding to all DR molecules tested so far. However, there is a difference in fine specificity at this position, namely Trp and Tyr that can serve as anchors for DRB1*0401 and DRB1*0101, are not accepted by either DRB1*0404 or DRB1*0402. This difference correlates with a Gly/Val dimorphism at position 86 of DRβ chains. Since position 86 forms part of a hydrophobic pocket accommodating the p1 anchor (13), the position 86 dimorphism was expected to have a direct effect on peptide binding. Clearly, this dimorphism cannot influence RA-linkage, since both residues occur in RA-associated (as well as nonassociated) molecules. However, if one assumes that a common set of peptides has to be presented by different DR molecules for RA induction, such a set cannot comprise peptides with Trp or Tyr at p1. Another important anchor at p6 has been previously shown to influence allele specificity of peptide binding (23). This anchor is shared by all DR4 molecules tested, ir-

respective of RA association. Although the p6 specificity does not correlate with genetic linkage, the preponderance of DR4 subtypes in RA indicates that the DR4-specific p6 anchor, Thr, Ser, Asn, or Val, is probably an important component of disease-inducing peptides.

The critical peptide position for RA association has turned out to be p4. The p4 anchor appeared pseudo-monomorphic in previous studies, because Met was the preferred residue for three different DR alleles, although allele-dependent differences in fine specificity have been noted (22, 23). The use of a set of p1-anchored designer peptides in a recent study (9) and in the present experiments has revealed a striking difference in the effect of charged residues at p4. Namely, a negative charge is accepted and a positive charge is not accepted by the RA-associated DR4 subtypes, and the opposite is the case for the nonassociated DRB1*0402. A similar charge effect operates also at p5 (Asp). Substitution experiments on natural protein-derived peptide sequences have further confirmed these results. Most peptides with negative charges at p4 bind poorly to the nonassociated DRB1*0402. However, they lose their selective binding to RA-associated DR4 subtypes when Asp or Glu is replaced by Ala at p4. Thus, a negative charge at p4 and/or p5 may be a distinctive feature of the putative pathogenic peptides involved in RA. Only two exceptions were noted where charge effects do not explain the poor binding to DRB1*0402 (peptides 10 and 16 in Fig. 6 *a*). This may be due to peptide sequence-specific parameters, not detected by the side chain scanning.

As shown by modeling, p4 seems to anchor in the vicinity of the RA-associated 67–74 region of the β chain. In support of this is our result that a Lys to Glu exchange at position 71 can change most RA-associated binding features into the RA nonassociated pattern. Our data also imply a minor role for residues 67 and 70 in peptide binding, which, however, remains to be investigated in detail. Thus, the residue at position 71 seems to exert a major effect on the RA-associated binding pattern.

Our results do not directly address the association of RA with DRB1*0101. Although this molecule also shares the 67–74 sequence with RA-associated DR4 subtypes, it differs from the latter at several amino acid positions within the peptide binding site. Considering also that the DRB1*0101 linkage has a partially different ethnic distribution (24), it is reasonable to assume that the disease involves a different set of peptides, although the same charge effect may apply around p4.

Altogether, the results strongly support the hypothesis that selective binding of peptides is the major mechanism underlying MHC association of RA. It should be pointed out, however, that our data do not necessarily argue against a possible role of the 67–74 region in T cell repertoire selection. In fact, the important residues of this region are located in the center of the α helix that protrudes above the plane of the binding site, and is therefore highly likely to be contacted by TCRs. In general, the selection of complementary TCRs by MHC molecules may be a prerequisite for the recognition of all antigens presented by the selecting molecule in the periphery, including the antigens involved in disease. However, the specific question of whether or not the 67–74 region is the only structure capable of selecting pathogenic T cells involved in RA is not addressed by our studies.

The results presented herein provide a sufficiently accurate basis for the prediction of T cell epitopes of any protein implicated as a potential autoantigen in RA. The peptides corresponding to such sequences can then be tested directly for their role as autoantigens.

We thank Saskia De Koster and René de Vries (University Hospital, Leiden, The Netherlands) for providing DR-homozygous EBV-B cell lines; Larry Stern (Massachusetts Institute of Technology, Cambridge, MA) and Don Wiley (Harvard University, Cambridge, MA) for the DRB1*0101-HA peptide crystal structure coordinates; and Anthony Nicholls (Columbia University, New York, NY) for the GRASP software package.

This work was supported in part by National Institutes of Health grant AI-27214.

Address correspondence to Dr. J. Hammer, Roche Milano Ricerche, Via Olgettina 58, I-20132 Milan, Italy.

Received for publication 1 September 1994 and in revised form 19 December 1994.

References

1. Nepom, G.T., P. Byers, C. Seyfried, L.A. Healey, K.R. Wilske, D. Stage, and B.S. Nepom. 1989. HLA genes associated with rheumatoid arthritis. *Arthritis Rheum.* 32:15–21.
2. Nepom, G.T., and H. Erlich. 1991. MHC class-II molecules and autoimmunity. *Annu. Rev. Immunol.* 9:493–525.
3. Gregersen, P.K., J. Silver, and R.J. Winchester. 1987. The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis. Rheum.* 30:1205–1213.
4. Hiraiwa, A., K. Yamanaka, W.W. Kwok, E.M. Mickelson, S. Masewicz, J.A. Hansen, S.F. Radka, and G.T. Nepom. 1990. Structural requirements for recognition of the HLA-Dw14 class II epitope — a key HLA determinant associated with rheumatoid arthritis. *Proc. Natl. Acad. Sci. USA.* 87:8051–8055.
5. Roy, S., M.T. Scherer, T.J. Briner, J.A. Smith, and M.L. Gefter. 1989. Murine MHC polymorphism and T cell specificities. *Science (Wash. DC).* 244:572–575.
6. Kappler, J., N. Roehm, and P. Marrack. 1987. T cell tolerance by clonal elimination in the thymus. *Cell.* 49:273–280.
7. Teh, H.S., P. Kisielow, B. Scott, H. Kishi, Y. Uematsu, H.

Blüthmann, and H. von Boehmer. 1988. Thymic major histocompatibility complex antigens and the $\alpha\beta$ T-cell receptor determine the CD4/CD8 phenotype of T cells. *Nature (Lond.).* 335:229–233.

8. Hammer, J., and F. Sinigaglia. Techniques to identify the rules governing class II MHC–peptide interaction. *In* MHC: A Practical Approach. G. Butcher and N. Fernandez, editors. Oxford University Press, Oxford, UK. In press.
9. Hammer, J., E. Bono, F. Gallazzi, C. Belunis, Z. Nagy, and F. Sinigaglia. 1994. Precise prediction of major histocompatibility complex class II–peptide interaction based on peptide side chain scanning. *J. Exp. Med.* 180:2353–2358.
10. Sinigaglia, F., P. Romagnoli, M. Guttinger, B. Takacs, and R.J.L. Pink. 1991. Selection of T cell epitopes and vaccine engineering. *Methods Enzymol.* 203:370–386.
11. Fu, X.-T., C.P. Bono, S.L. Woulfe, C. Swearingen, N.L. Summers, F. Sinigaglia, A. Sette, B.D. Schwartz, and R.W. Karr. Pocket 4 of the HLA-DR (α,β1*0401) molecule is a major determinant of T cell recognition of peptide. *J. Exp. Med.* 181:915–926.
12. Barany, G., and R.B. Merrifield. 1980. Solid-phase peptide synthesis. *In* The Peptides. Vol. 2. E. Gross and J. Meienhofer, editors. Academic Press, New York. 1–284.
13. Stern, L.J., J.H. Brown, T.S. Jardetzky, J.C. Gorga, R.G. Urban, J.L. Strominger, and D.C. Wiley. 1994. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature (Lond.).* 368:215–221.
14. Pearlman, D.A., D.A. Case, J.C. Caldwell, G.S. Seibel, P. Weiner, and P.A. Kollmann. 1991. AMBER 4.0. University of California, San Francisco.
15. Weiner, S.J., P.A. Kollman, D.T. Nguyen, and D.A. Case. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.
16. Nicholls, A., and B. Honig. 1991. GRASP 1.0. Columbia University, New York.
17. Hammer, J., C. Belunis, D. Bolin, J. Papadopoulos, R. Walsky, J. Higelin, W. Danho, F. Sinigaglia, and Z.A. Nagy. 1994. High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. *Proc. Natl. Acad. Sci. USA.* 91:4456–4460.
18. Mikecz, K., T.T. Glant, and A.R. Poole. 1987. Immunity to cartilage proteoglycans in BALB/c mice with progressive polyarthritis and ankylosing spondylitis induced by injection of human cartilage proteoglycan. *Arthritis Rheum.* 30:306–318.
19. Trentham, D.E., A.S. Townes, and A.H. Kang. 1977. Autoimmunity to type II collagen: an experimental model of arthritis. *J. Exp. Med.* 146:857–868.
20. Gaston, J.S.H., P.F. Life, P.J. Jenner, M.J. Colston, and P.A. Bacon. 1990. Recognition of a Mycobacteria-specific epitope in the 65-kD heat-shock protein by synovial fluid–derived T cell clones. *J. Exp. Med.* 171:831–841.
21. Simon, M., E. Girbat, M. Sebbag, V. Gomès-Daudrix, C. Vincent, G. Salama, and G. Serre. 1993. The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis. *J. Clin. Invest.* 92:1387–1393.
22. Hammer, J., B. Takacs, and F. Sinigaglia. 1992. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. *J. Exp. Med.* 176:1007–1013.
23. Hammer, J., P. Valsasini, K. Tolba, D. Bolin, J. Higelin, B. Takacs, and F. Sinigaglia. 1993. Promiscuous and allele-specific anchors in HLA-DR-binding peptides. *Cell.* 74:197–203.
24. Schiff, B., Y. Mizrachi, S. Orgad, M. Yaron, and I. Gazit. 1982. Association of HLA-Aw31 and HLA-DR1 with adult rheumatoid arthritis. *Ann. Rheum. Dis.* 41:403–404.