High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations

(peptide-HLA-DR binding/M13 random peptide library/binding motif)

JUERGEN HAMMER^{†‡}, CHARLES BELUNIS[†], DAVID BOLIN[§], JOANNE PAPADOPOULOS[†], ROBERT WALSKY[‡], JACQUELINE HIGELIN[¶], WALEED DANHO[§], FRANCESCO SINIGAGLIA[‡], AND ZOLTAN A. NAGY[†]

[†]Department of Inflammation and Autoimmune Diseases, and [§]Department of Molecular Sciences, Hoffman-La Roche Inc., 340 Kingsland Street, Nutley, NJ 07110-1199; [‡]Roche Milano Ricerche, via Olgettina 58, I-20132 Milan, Italy; and [¶]Department of Pharmaceutical Research New Technologies, F. Hoffman-La Roche Ltd., CH-4002 Basel, Switzerland

Communicated by Sidney Udenfriend, December 27, 1993 (received for review August 31, 1993)

ABSTRACT We have previously identified four anchor positions in HLA-DRB1*0101-binding peptides, and three anchors involved in peptide binding to DRB1*0401 and DRB1*1101 molecules, by screening of an M13 peptide display library (≈ 20 million independent nonapeptides) for DR-binding activity. In this study, high stringency screening of the M13 library for DRB1*0401 binding has resulted in identification of three further anchor positions. Taken together, a peptide-binding motif has been obtained, in which six of seven positions show enrichment of certain residues. We have demonstrated an additive effect of anchors in two different ways: (i) the addition of more anchors is shown to compensate for progressive truncation of designer peptides; (ii) the incorporation of an increasing number of anchors into 6- or 7-residue-long designer peptides is shown to result in a gradual increase of binding affinity to the level of 13-residue-long high-affinity epitopes. The anchor at relative position 1 seems to be obligatory, in that its substitution abrogates binding completely, whereas the elimination of other anchors results only in partial loss of binding affinity. The spacing between anchors is critical, since their effect is lost by shifting them one position toward the N or C terminus. The information born out of this study has been successfully used to identify DR-binding sequences from natural proteins.

Each allelic form of the highly polymorphic major histocompatibility complex (MHC) molecules can bind a large array of different peptides to ensure T-cell-mediated immunity to the antigenic universe. Thus, provided that peptides bind to a given MHC molecule in similar or identical conformation (1–4), the sequencing of a sufficiently large pool of bound peptides should reveal the rules of MHC/peptide interaction. Recently, two different approaches have been described for the isolation and characterization of MHC-bound peptide pools. (i) Endogenously processed peptides were eluted from MHC molecules and sequenced in pool or individually. This approach resulted in binding motifs for class I MHC molecules (5–7). However, its application for MHC class II molecules proved to be difficult, because of the limited number of independent sequences identified and the high degree of length heterogeneity of bound peptides (8–11). (ii) Phage-expressed large (≈ 20 million independent sequences) random peptide libraries were screened for MHC binding. Sequencing of the bound peptides has revealed allele-specific binding motifs for three different HLA-DR molecules (12, 13). Here, we report a further refinement of the phage library screening, which allowed us to identify additional DRB1*0401-binding anchors. We also demonstrate that the anchor residues in class II binding peptides are additive and

position specific, and, in combination, enable short peptides to bind tightly to DR molecules.

MATERIALS AND METHODS

Preparation of HLA-DR Molecules. DR molecules isolated from cell lines BSM, PREISS (DRB1*0401), HOM-2, LG-2 (DRB1*0101), and SWEIG (DRB1*1101) were purified and biotinylated as described (12, 14).

Construction and High Stringency Screening of the M13 Peptide Library. The M13 peptide library was constructed as described (12). The library consisted of 20 million independent phages displaying 9 random amino acids flanked by glycine spacers, located at the N-terminal part of phage protein III. Approximately 20 billion library phages were incubated with 50 pmol of biotinylated DRB1*0401 in binding buffer [150 mM NaCl/1 mM EDTA/1 mM phenylmethylsulfonyl fluoride/0.2% Nonidet P-40 (NP-40)/80 mM citrate/phosphate, pH 6] for 48 hr at room temperature. Bovine serum albumin (BSA)-blocked streptavidin on 4% beaded agarose was added and incubated for 10 min. The M13 phage/DRB1*0401 complexes were purified by washing the solid phase eight times with 200 volumes of wash buffer (50 mM Tris-HCl, pH 7.5/150 mM NaCl/1 mM EDTA/0.2% NP-40), once with 60 volumes of pre-elution buffer (70 mM citrate/phosphate, pH 3/500 mM NaCl) for 10 min, and once again with 200 volumes wash buffer. The phages were eluted with 60 volumes of elution buffer (0.1 M glycine hydrochloride, pH 2.2/1 mg of BSA per ml) for 10 min and neutralized with 2 M Tris base. The phage eluate was concentrated and washed with 150 mM NaCl/50 mM Tris-HCl (pH 7.5) by ultradialysis (Microsep Microconcentrators, 30-kDa cutoff, Filtron Technology, Northborough, MA). Enrichment of phages was measured as described (12). The screening procedure was repeated five times. The amplified eluate of the third, fourth, and fifth round was used to isolate and sequence individual M13 clones.

Peptide Synthesis and DR-Peptide Binding Assay. Peptides were prepared by solid-phase synthesis (15) and purified by preparative reversed-phase HPLC; their homogeneity was confirmed by analytical HPLC, amino acid analysis, and fast atom bombardment mass spectrometry. Competition-type assays with affinity-purified DR molecules, ¹²⁵I-labeled indicator peptides, and unlabeled competitor peptides were used as described (12, 16). DR-bound and unbound labeled peptide was separated by the spin column method (16). In assays with large sample numbers, the DR molecules were captured by biotinylated anti-DR antibody to streptavidin-coated agarose beads, and the solid phase was counted directly (J. Hammer,

FIG. 1. Anchor positions defined by high stringency screening of an M13 phage-expressed random peptide library. (A) To define a high stringency screening condition, 10^9 phages displaying the NNX peptide known to bind strongly DRB1*0101 and DRB1*0401 were mixed with 10^{10} M13mp19 background phages, and the mixture was allowed to bind to DRB1*0101 (Upper) or DRB1*0401 (Lower) as described (12). After eight washes in binding buffer (see text), the

unpublished data). The average 50% inhibitory concentrations (IC_{50}) of labeled vs. unlabeled indicator peptides were as follows: 130 nM of flu hemagglutinin peptide (HA) 307–319 for DRB1*0101, 220 nM of the same peptide for DRB1*0401, and 490 nM of tetanus toxoid peptide (TT) 831–843 for DRB1*1101.

RESULTS

New Anchor Residues in DRB1*0401-Binding Peptides Revealed by High Stringency Screening of an M13 Peptide Library. Using the phage display technology (17–19), we have previously described a binding motif for DRB1*0401 consisting of an aromatic or aliphatic hydrophobic residue at position (p) 1, a hydrophobic residue (typically Met or Ala) at p4, and Thr or Ser at p6 (13). To identify possible additional anchors, we tried to increase the stringency of phage library screening. We used in these studies the model bacteriophage M13NNX that binds specifically to both DRB1*0101 and DRB1*0401 (12). As shown in Fig. 1A, whereas a pH 3 wash removed most background (M13mp19) and only some M13NNX phages from DRB1*0101, it eluted both background and M13NNX phages almost completely from DRB1*0401. Thus, the selection of phages remaining DRB1*0401-bound after a pH 3 wash introduced more stringency into the screening. An M13 peptide library of 20 million independent phage recombinants (12, 13) was then screened in five cycles of DRB1*0401 binding, preelution, elution, and amplification. After the third, fourth, and fifth cycles, 181 phages were isolated and the inserts were sequenced (data not shown). Seventy percent of independent sequences contained aromatic residues at p1 and Thr or Ser at p6—i.e., the expected residues at the two dominant anchor sites (13). After alignment based on p1 and p6 anchors, 53% of sequences had the expected hydrophobic residues (Met, Ala) at p4. In addition, enrichments of Arg at p2 (47%), Gly and Ala at p3 (43%), and Leu at p7 (47%) were observed (Fig. 1B). Altogether, amino acid preferences could be demonstrated at six of seven positions.

Additive Effect of Anchor Residues. DR-binding peptides derived from endogenously processed proteins are 13–25 residues long (10, 11) and, judged by the binding motif information available to us (12, 13), they contain typically two or three anchor residues. The length requirement for class II MHC binding may thus reflect a compensation by main chain interactions (4, 20) for the missing binding energy that could be provided by side chain interactions of additional anchor residues. To investigate this possibility, we have performed a series of truncations on designer peptides containing different numbers of anchors. As shown in Fig. 2A, the loss of binding activity due to decreasing length could be compensated for by increasing the number of anchors. Furthermore, the data in Fig. 2B demonstrate that the binding affinity of hexa- and heptapeptides could be increased by adding more anchors, up to the affinity of 13- to 14-residue-long epitopes (e.g., HA 307–319 and TT830–843). The results also show that Arg at p2, identified as an anchor for DRB1*0401 by high stringency phage library screening, also exerts a positive effect on peptide binding to DRB1*0101 and DRB1*1101. Altogether, a clear additive effect of anchor

remaining phages were eluted as follows: fraction 1, binding buffer; fraction 2, 70 mM citrate/phosphate, pH 3/500 mM NaCl ("pH 3 wash"); fraction 3, binding buffer; fraction 4, pH 3 wash; fractions 5–7, elution buffer (pH 2.2) (see text). Each fraction was of 1 ml volume. Phage numbers were determined by plating the mixture on 5-bromo-4-chloro-3-indolyl β -D-galactoside indicator plates (12). (B) Enrichment of particular residues at positions 2, 3, and 7 (compared to position 5) in peptides from DRB1*0401 selected phages isolated after high stringency screening conditions. Peptides from independent phages were aligned according to p1 and p6 anchors.

FIG. 2. Additive effects of anchor residues. (A) Increasing number of anchors compensates for affinity loss due to truncation. Relative binding values were derived from the reciprocal of 50% inhibitory concentration (IC_{50}) of each peptide. (B) Construction of high-affinity-binding short peptides by anchor addition. Binding of high-affinity natural epitopes HA 307–319 and TT831–843 is included for comparison.

residues is evident from these experiments. Two exceptions to this rule were noted. (i) Short peptides with Tyr at p1 and Ala at p3 and p6 failed to bind, suggesting that Ala anchors may not contribute to the binding energy. (ii) Removal of the p1 anchor could not be compensated for by the sum of all remaining anchors.

The Anchor at p1 Is Essential for HLA-DR/Peptide Interaction. Peptides with p1 truncated or substituted to Ala fail to bind to DR molecules, even if they contain all additional anchor residues (Fig. 2B and unpublished data). This observation suggests an obligatory role for the p1 anchor. Because of its importance as a potential starting point in several allele-specific motifs, we analyzed the effect of 18 amino acid substitutions at p1. The results in Fig. 3 demonstrate that only the aromatic residues Tyr, Phe, and Trp, identified by the M13 screening procedure (13), and, to a smaller extent, hydrophobic residues Leu, Ile, Val, and Met enable peptide binding to DRB1*0101 and DRB1*0401. Thus, a hydrophobic

anchor at p1 is essential for peptide binding to at least the two DR molecules investigated.

The Effect of Anchor Residues Is Position Specific. To investigate whether exact spacing between two anchor residues was crucial for class II binding, or some flexibility might be permitted as in the case of certain class I binding peptides (3), we tested various designer peptides, with anchor residues shifted one position toward either the N or the C terminus, for DR binding. As shown in Fig. 4, peptides with the p2 or p4 anchor shifted had significantly lower binding affinity, and both anchors out of phase caused an almost complete loss of binding. Shifting of the p6 anchor had different effects depending on the DR allotype. In DRB1*0101, shifting of p6 Gly did not decrease binding affinity. This may be explained by the fact that another small residue, such as Ala in the substituted peptides, can also serve as anchor for DRB1*0101 (12). Alternatively, this anchor may not contribute to the binding energy of peptides, as also suggested by data in Fig. 2B. In

FIG. 3. Amino acid residues accepted at anchor position 1. Data are expressed in terms of relative binding (see in Fig. 2). For comparison, the binding of HA 307-319 is given.

contrast, shifting of the allele-specific p6 anchor, Thr in DRB1*0401 and Arg in DRB1*1101 binding peptides, caused a significant decrease of binding affinity. Thus, the latter

FIG. 4. Effect of anchor shifting on peptide binding to three allelic DR molecules. Shifted residues are given in boldface type. Filled columns indicate the binding of peptides with an anchor residue or with a residue mediating a negative effect at p6.

anchor residues did contribute to binding affinity, and their effect was lost upon shifting. The results also demonstrated that inappropriate residues at the p6 site—e.g., Leu or Glu for DRB1*0101 and Arg for DRB1*0401—had a position-specific negative effect on peptide binding.

Identification of DR-Binding Peptides from Natural Proteins by Using the Anchors Defined. Considering the synergistic nature and position specificity of class II anchors, it is reasonable to assume that peptides with several anchors in frame are more likely to bind to a particular class II molecule than peptides lacking most or all anchors in frame. To investigate this hypothesis, we compared for DRB1*0101 binding a set of natural protein-derived sequences with three or four anchor residues in frame to another set with one potential p1 or no anchor. Twelve of the 13 peptides with anchors bound to DRB1*0101, 10 of them with affinities comparable to that of HA 307–319 (Fig. 5). In contrast, all 11 peptides lacking two anchors in frame failed to bind. Thus, the rules described here seem to apply to natural proteins. However, the observed variation of binding affinities and the failure of one peptide with four anchors to bind (probably due to the aliphatic p1 residue together with negative effects at p4 and elsewhere) suggest that additional rules remain to be identified for a precise, quantitative epitope prediction.

FIG. 5. Identification of good DRB1*0101-binding and nonbinding peptides from natural protein sequences based on anchor combinations. The sequences were taken from TT (21), human glutamic acid decarboxylase (22), and human cartilage link protein (23). In the synthetic peptides, Cys residues of natural sequences were replaced by Ala. Data are expressed as relative binding, with the level of HA 307–319 binding indicated. Uppercase letters in the binding motif denote anchor residues identified by phage library screening, and lowercase letters are anchor residues defined by synthetic peptide chemistry. Predicted anchor residues in peptide sequences are given in boldface type.

DISCUSSION

Class II MHC anchors, by our definition, represent position-specific enrichments of certain amino acid residues in a pool of DR-binding peptides selected from a large, phage-expressed random peptide library. By these criteria, we have demonstrated here three new anchor positions in peptides binding to DRB1*0401—namely, Arg at p2, Ala or Gly at p3, and Leu at p7. Of these at least one, Arg at p2, seems to be shared by at least two additional DR allotypes, DRB1*0101 and DRB1*1101. Taken together with our previous studies (12, 13), DRB1*0401-binding heptapeptides can thus contain up to six anchor positions.

Mechanistically, anchors can be of two different types: “type 1” anchors that contribute to the binding energy of the peptide and “type 2” ones that probably reflect steric constraints only. Most anchors fall into the first category. This category can be subdivided into two—namely, obligatory and auxiliary type 1 anchors. Obligatory is in our hands the aromatic-hydrophobic residue at p1, whose exchange or truncation leads to a practically complete loss of peptide binding to DRB1*0101 and 0401 (Fig. 3 and unpublished data). The importance of this position has also been pointed out by others (20, 24, 25). The p1 anchor probably fits into the recently identified hydrophobic pocket of the DR-binding site (4), and this may account for its obligatory nature. It is present in several DR-binding motifs. All other type 1 anchors appear to be auxiliary, in that either of them can be omitted without complete loss of peptide binding. The type 2 anchors are represented by Ala or Gly at p6 for DRB1*0101 and possibly also at p3 for DRB1*0401. These residues do not seem to contribute significantly to the binding energy of peptides (Figs. 2B and 4). However, residues with bulky side chains at p6 decrease binding affinity (Fig. 4, and unpublished observation), suggesting steric constraints in this part of the binding site. Such anchor sites, therefore, appear to be sensitive to negative effects. Since the p6 anchor shows allelic variation (13), steric constraint seems to be one of the mechanisms whereby allele specificity is imposed on peptide binding to DR molecules.

It appears that at least two or three correctly spaced anchors, as defined in our present and previous studies (12, 13), are present in DR-restricted epitopes (Fig. 5, and refs. 24 and 25) as well as in naturally processed DR-binding peptides (10, 11), suggesting an important role for these anchors in the biological function of DR molecules. The only apparent exception to this rule is the poly(Ala) peptide reported by Jardetzky *et al.* (20) that exhibits good binding to DRB1*0101 via a single aromatic type 1 anchor. However, according to our data with poly(Ala) peptides, main chain interactions may play an important role in binding (Fig. 2). In natural peptides, these interactions may be weakened by negative effects similar to those shown in Fig. 4, and, as a result, natural sequences with a single aromatic anchor do not bind (e.g., peptides in the lower panel of Fig. 5).

Using synthetic peptide chemistry, we have demonstrated here two important features of anchor residues—namely, that they must be precisely spaced and that they act in an additive manner. The latter characteristic applies primarily to type 1 anchors, and it is reflected in gradual increase of the binding affinity of peptides. These two properties of anchors strongly suggest that the conformation of DR-bound peptides should be very similar. Furthermore, the observed additive effect permits the construction of hexa- and heptapeptides with high binding affinity. Finally, we have provided an example to show that the motif information from our studies can be

useful in predicting DR-binding sequences from natural proteins. We think that once both positive and negative effects on binding are understood, the combined information can be incorporated into an algorithm enabling precise epitope prediction.

We thank Drs. W. Haas and M. Steinmetz for criticism and suggestions, Dr. F. Falcioni and Mr. P. Familletti for help in cell cultures, Dr. K. Tolba for help in DNA sequencing, and Ms. G. Ugolini and Ms. A. Wolfer for expert technical assistance.

- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1991) *Nature (London)* **353**, 321–325.
- Zhang, W., Young, A. C. M., Imarai, M., Nathenson, S. G. & Sacchettini, J. C. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8403–8407.
- Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257**, 919–927.
- Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L. & Wiley, D. C. (1993) *Nature (London)* **364**, 33–39.
- Falk, K., Roetzschke, O., Stevanovic, S., Jung, G. & Ramensee, H.-G. (1991) *Nature (London)* **351**, 290–296.
- Van Bleek, G. M. & Nathenson, S. G. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11032–11036.
- Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E. & Engelhard, V. H. (1992) *Science* **255**, 1261–1263.
- Rudensky, A. Y., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. & Janeway, C. A. (1991) *Nature (London)* **353**, 622–627.
- Hunt, D. F., Michel, H., Dickinson, T. A., Shabanowitz, J., Cox, A. L., Sakaguchi, K., Appella, E., Grey, H. M. & Sette, A. (1992) *Science* **256**, 1817–1820.
- Chicz, R. M., Urban, R. G., Lane, W. S., Gorga, J. C., Stern, L. J., Vignali, D. A. A. & Strominger, J. L. (1992) *Nature (London)* **358**, 764–768.
- Chicz, R. M., Urban, R. G., Gorga, J. C., Vignali, D. A. A., Lane, W. S. & Strominger, J. L. (1993) *J. Exp. Med.* **178**, 27–47.
- Hammer, J., Takacs, B. & Sinigaglia, F. (1992) *J. Exp. Med.* **176**, 1007–1013.
- Hammer, J., Valsasnini, P., Tolba, K., Bolin, D., Higelin, J., Takacs, B. & Sinigaglia, F. (1993) *Cell* **74**, 197–203.
- Sinigaglia, F., Romagnoli, P., Guttinger, M., Takacs, B. & Pink, J. R. L. (1992) *Methods Enzymol.* **203**, 370–386.
- Barany, G. & Merrifield, R. B. (1980) in *The Peptides*, eds. Gross, E. & Meienhofer, J. (Academic, New York), Vol. 2, pp. 1–284.
- Scheirle, A., Takacs, B., Kremer, L., Marin, F. & Sinigaglia, F. (1992) *J. Immunol.* **149**, 1994–1999.
- Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
- Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
- Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dover, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
- Jardetzky, T. S., Gorga, J. C., Busch, R., Rothbard, J. B., Strominger, J. L. & Wiley, D. C. (1990) *EMBO J.* **9**, 1797–1803.
- Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E. & Niemann, H. (1986) *EMBO J.* **5**, 2495–2502.
- Karlsen, A. E., Hagopian, W. A., Grubin, C. E., Dube, S., Distech, C. M., Adler, D. A., Barmeier, H., Mathewes, S., Grant, F. J., Foster, D. & Lenmark, A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8337–8341.
- Dudhia, J. & Hardingham, T. E. (1990) *Nucleic Acids Res.* **18**, 1292.
- Hill, C. M., Hayball, J. D., Allison, A. A. & Rothbard, J. B. (1991) *J. Immunol.* **147**, 189–197.
- O'Sullivan, D., Arrhenius, T., Sidney, J., Del Guercio, M.-F., Albertson, M., Wall, M., Oseroff, C., Southwood, S., Colon, S. M., Gaeta, F. C. A. & Sette, A. (1991) *J. Immunol.* **147**, 2663–2669.