Divergent Motifs but Overlapping Binding Repertoires of Six HLA-DQ Molecules Frequently Expressed in the Worldwide Human Population

John Sidney, Amiyah Steen, Carrie Moore, Sandy Ngo, Jolan Chung, Bjoern Peters and Alessandro Sette

*J Immunol* 2010; 185:4189-4198; Prepublished online 1 September 2010; doi: 10.4049/jimmunol.1001006

http://www.jimmunol.org/content/185/7/4189

Supplementary Material  http://www.jimmunol.org/content/suppl/2010/09/01/jimmunol.1001006.DC1

Why *The JI*?

- **Rapid Reviews! 30 days** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Speedy Publication!** 4 weeks from acceptance to publication

*average

References  This article cites 86 articles, 28 of which you can access for free at: http://www.jimmunol.org/content/185/7/4189.full#ref-list-1

Subscription  Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

Permissions  Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts  Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
Divergent Motifs but Overlapping Binding Repertoires of Six HLA-DQ Molecules Frequently Expressed in the Worldwide Human Population

John Sidney, Amiyah Steen, Carrie Moore, Sandy Ngo, Jolan Chung, Bjoern Peters, and Alessandro Sette

Knowledge of the binding repertoires and specificities of HLA-DQ molecules is somewhat limited and contradictory, partly because of the scarcity of reports addressing some of the most common molecules and possibly because of the diversity of the techniques used. In this paper, we report the development of high-throughput binding assays for the six most common DQ molecules in the general worldwide population. Using comprehensive panels of single substitution analogs of specific ligands, we derived detailed binding motifs for DQA1*0501/DQB1*0301, DQA1*0401/DQB1*0402, and DQA1*0101/DQB1*0501 and more detailed motifs for DQA1*0501/DQB1*0201, DQA1*0301/DQB1*0302, and DQA1*0102/DQB1*0602, previously characterized on the basis of sets of eluted ligands and/or limited sets of substituted peptides. In contrast to what has previously been observed for DR and DP molecules, DQ motifs were generally less clearly defined in terms of chemical specificity and, strikingly, had little overlap with each other. However, testing a panel of peptides spanning a set of Phleum pratense Ags, and panels of known DQ epitopes, revealed a surprisingly significant and substantial overlap in the repertoire of peptides bound by these DQ molecules. Although the mechanism underlying these apparently contradictory findings is not clear, it likely reflects the peculiar mode of interaction between DQ (and not DR or DP) molecules and their peptide ligands. Because the DQ molecules studied are found in >85% of the general human population, these findings have important implications for epitope identification studies and monitoring of DQ-restricted immune responses. The Journal of Immunology, 2010, 185: 4189–4198.

The online version of this article contains supplemental material.

La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037

Received for publication March 29, 2010. Accepted for publication July 20, 2010.

This work was supported by National Institutes of Health-National Institute of Allergy and Infectious Diseases Contracts HHSN266200400006C, HHSN272200900044C, HHSN272200900042C, and HHSN272200700048C (all to A.S.).

Address correspondence and reprint requests to Dr. Alessandro Sette, La Jolla Institute for Allergy and Immunology, 9420 Athena Circle, La Jolla, CA 92037. E-mail address: alex@liai.org

The online version of this article contains supplemental material.

Abbreviations used in this paper: ARB, average relative binding affinity; DQA, DQ α; DQB, DQ β; SAAS, single amino acid substitution; SF, specificity factor.

Copyright © 2010 by The American Association of Immunologists, Inc. 0022-1767/10/$16.00

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1001006

In the case of HLA class I, previous studies have demonstrated the existence of MHC supertypes, which define sets of A and B class I molecules associated with largely overlapping peptide-binding repertoires (9–12). In the context of HLA class II, four different loci must be considered: DRB1, DRB3/4/5, DP, and DQ. Several studies have suggested the existence of class II supertypes encompassing several DR and DP specificities that, as with class I supertypes, describe sets of molecules sharing largely overlapping peptide-binding repertoires (13–29). Some epitope and peptide-binding specificity overlap has also been reported in the case of two HLA-DQ molecules (26).

But in comparison, DQ-associated motifs have been less extensively characterized. Indeed, the vast majority of studies characterizing DQ-binding specificity have been targeted at DQ2 (DQA1*0501/DQB1*0201) (26, 30–34) and DQ8 (DQA1*0301/DQB1*0302) (26, 30, 35–37), which have been associated with susceptibility to celiac disease and type 1 diabetes mellitus (38–42). Studies characterizing the peptide-binding specificities of other common DQ molecules have been more limited in both breadth and detail (43–47). Thus, with these limitations, it has not been possible to discern whether a general mode of DQ binding can be defined, as has been done for both DR and DP molecules. That the existing data might not be enough to extrapolate a general DQ supertypic binding specificity is indicated from studies analyzing the structure of DQ molecules and IA molecules (the murine ortholog of DQ) (37, 47–52). These structural studies have shown that in several instances, interactions between amino acid side chains of the peptide ligands and the peptide-binding pockets of the MHC contribute relatively little to the overall binding energy. Instead, binding may, in these cases, be more dependent on peptide backbone–MHC interactions.

In this paper, we report the development of high-throughput binding assays for the six most common HLA-DQ molecules in...
MHC molecules were then eluted with 50 mM diethylamine in 0.15 M NaCl. The effectiveness of depletion steps were monitored by SDS-PAGE and bicinchoninic acid assay. The eluates were concentrated by centrifugation in Centriprep 30 microconcentrators at 2000 g for 30 min. Tris (pH 6.8) was added to the eluate to reduce the pH to 8. Eluates were loaded on phosphocellulose (Phospho-S, Whatman, Maidstone, UK) columns equilibrated with PBS, and 2-column volumes of PBS containing 0.4% (w/v) 2-octylglucoside (pH 11.5). A 1/26 volume of 2.0 M NaOH was added to the eluate to reduce the pH to 8. Eluates were concentrated by centrifugation in Centriprep 30 microconcentrators at 2000 g. Lysates were then passed through 0.45 μm filters and cleared of nuclei and debris by centrifugation at 10,000 g for 20 min. For affinity purification, columns of inactivated Sepharose CL-4B and protein A-Sepharose were used as precolumns. HLA-DR, -DQ, and -DP molecules were captured by repeated passage of lysates over LBS1 (anti–HLA-DR), SPV-L3 (anti–HLA-DQ), and B7/21 (anti–HLA-DP) columns. After two to four passages of the lysate, Ab columns were washed with 10-column volumes of 10 mM Tris-HCl (pH 8) with 1% (v/v) Nonidet P-40, 2-column volumes of PBS, and 2-column volumes of PBS containing 0.4% (w/v) n-octylglucoside. MHC molecules were then eluted with 50 mM diethylamine in 0.15 M NaCl containing 0.4% (w/v) n-octylglucoside (pH 11.5). A 1/26 volume of 2.0 M NaOH was added to the eluate to reduce the pH to 8. Eluates were concentrated by centrifugation in Centriprep 30 microconcentrators at 2000 g (Amicon, Beverly, MA). Protein purity, concentration, and the effectiveness of depletion steps were monitored by SDS-PAGE and bicinchoninic acid assay.

**Results**

The peptide-binding repertoire (R) of each MHC molecule (i) was defined as the set of the peptides that bind that molecule with an affinity of ≥1000 nM. The relationship between two molecules has been measured by determining cross-reactivity and repertoire overlap. Cross-reactivity is defined as the fraction of peptides that bind one MHC molecule that also bind a second or (R AND R[i])/R ≥ 100%. Repertoire overlap is defined as the fraction of peptides binding either molecule that bind both (i.e., repertoire overlap = [R AND R[i]])/R ≥ 100%). Population coverage was calculated as described previously (57). Gene frequencies for each HLA allele were calculated from population frequencies obtained from the National Database for Histocompatibility (IBM Corporation, Bethesda, MD [58]). Phenotypic frequencies were calculated using the binomial distribution formula: phenotypic frequencies = 1 – (1 – \( \frac{1}{2} \) \( g \))^n. To obtain total potential population coverage, no linkage disequilibrium was assumed.

Analysis of the single amino acid substitution binding data were performed essentially as described previously for analysis of positional scanning combinatorial library data (29, 50, 60). Briefly, the IC50 value of each substituted peptide was standardized relative to the geometric mean IC50 nM value of the wild-type peptide. Relative binding values were then represented in a 20 aa by position matrix. Next, an average (geometric) relative binding affinity (ARB) was calculated for each position, encompassing all 20 possible residues. Finally, for each position, the ratio of the ARB for the entire library to the position-specific ARB was derived. We have denominated this ratio, which derives from the ratio of the ARB for the entire library to the position-specific ARB, as the specificity factor (SF). As calculated, positions with the highest specificity will have the highest SF value. Primary anchor positions were then defined as those with an SF ≥ 2.4. This criterion identifies positions where the majority of residues are associated with significant decreases in binding capacity.

In selecting the specific peptide used as the basis for probing the binding specificity of each molecule in detail, we considered several different criteria before selecting one specific candidate. These criteria included, to afford the greatest sensitivity, the capacity of the wild-type peptide to bind with high affinity. We also privileged in our selection, as much as possible, peptides that had been previously reported as recognized by T cell responses and, secondarily, those reported as endogenously bound ligands, rather than those only described as “binders.” Finally, we also balanced the need to select peptides that would, on paper, be less problematic from a synthesis standpoint; thus, peptides with M or C residues, for example, were typically passed over as candidates in favor of another peptide with similar affinity for the respective HLA molecule.

Because of their extensive polymorphism, different HLA class II types are expressed at different frequencies in different ethnicities.
Both DQα (DQA) and DQβ (DQB) chains are polymorphic. On the basis of HLA crystal structures and sequence analyses, it is hypothesized that variation in both the DQA and DQB chains appear in regions likely to affect binding specificity, unlike the cases of DR and DP (37, 47, 52, 61–66). The frequency of specific DQ molecules resulting from expression of DQA and DQB chain allelic variants in various ethnicities, as reported at DbMHC [National Center for Biotechnology Information (58)], were compiled, with the goal of selecting a panel of DQ heterodimers representing the most common molecules overall encountered in major populations worldwide and thereby affording coverage of the majority of the human population, irrespective of ethnicity (Table I).

Accordingly, we selected a panel of six different HLA-DQ molecules (DQA1*0501/DQB1*0301, DQA1*0301/DQB1*0302, DQA1*0102/DQB1*0602, DQA1*0401/DQB1*0402, DQA1*0101/DQB1*0501, and DQA1*0501/DQB1*0201). This panel encompasses all DQ molecules with a frequency of >15% in at least three major populations for which frequency data are currently available, including Europe, North America, South America, Southeast Asia, Southwest Asia, and Sub-Saharan Africa. Furthermore, each molecule is found with an average phenotypic frequency across these populations of ≥10%. In aggregate, the panel covers ~85% of the average population, with equally high and balanced population coverage, ranging from 72 to 98%, for the populations listed on Table I.

Establishment of high-throughput HLA-DQ-binding assays

Previous studies have described the development of quantitative peptide-binding assays based on the use of purified MHC for each of the targeted specificities, specifically DQA1*0501/DQB1*0201 (26, 30–34), DQA1*0501/DQB1*0301 (43), DQA1*0301/DQB1*0302 (26, 30, 35, 36), DQA1*0401/DQB1*0402 (44), DQA1*0101/DQB1*0501 (45), and DQA1*0102/DQB1*0602 (46). However, the specific assays used different methodologies, including gel filtration, MHC capture and cell-bound MHCs, and also different readouts, such as biotinylation, fluorescence, and radioactivity. Because the sensitivity of these different assay systems ranged over several orders of magnitude, we sought to establish assays based on a single high-sensitivity platform.

Binding assays based on the inhibition of binding of a radio-labeled high-affinity ligand affords high-throughput capacity and high sensitivity, often in the 25 nM range for class II. To establish assays for the six molecules selected, we assembled a panel of known DQ epitopes, endogenous ligands, or ligands known to bind specific DQ molecules with high affinity. These peptides were radiolabeled and tested in direct binding assays for their capacity to bind purified DQ molecules, as described in Materials and Methods. For each DQ molecule, a ligand providing a strong and specific signal was identified (Table II). Ligand specificity was further verified by assessing capture with HLA class II locus-specific Abs (Supplemental Fig. 2).

To optimize the assays, the effects of pH, temperature, and duration of incubation and MHC capture times were investigated. A representative analysis is shown in Supplemental Fig. 2. In general, optimal assay conditions were different from those used for HLA-DR but similar to those for HLA-DP assays (Table II) (15, 29, 53, 54). Specifically, DQ assays required, in most cases, incubation for 72 h at 37°C and an extended Ab capture step of up to 24 h. As with DR and DP, optimal pHs were in the 5–7 range, depending on the specific molecule. These optimal conditions were used in all subsequent competitive inhibition of binding assays. Assay sensitivities (IC50 of the unlabeled ligand) were in the 12–33 nM range.

Following the biochemical validation of the assays described above, we next sought to examine the correlation between assay specificity and the binding capacity of known HLA-restricted epitopes. For this purpose, we queried the Immune Epitope Database (http://www.iedb.org [67]) and selected a panel of 25 different epitopes and endogenous ligands (representing 27 different restriction events) presented by one or more of the six common DQ specificities selected in the current study (Table III). As shown, and similar to the case for HLA-DR and -DP (15, 29), in 81% (22 of 27) of the cases a DQ-restricted epitope or endogenous ligand bound its corresponding restricting molecule with a binding affinity of 1000 nM or better; ~50% of the cases were associated with affinities of 50 nM or better.

Definition of peptide-binding motifs for HLA DQA1*0501/DQB1*0201, DQA1*0301/DQB1*0302, and DQA1*0102/DQB1*0602 molecules

Next, we used sets of single amino acid substitution (SAAS) analogs of prototype ligands to define specific binding motifs. ARB matrices were compiled on the basis of measured binding affinities, as described in Materials and Methods. As in previous studies (29, 59, 60), for each position an SF was calculated, and primary anchor positions were defined as those with an SF ≥ 2.4. At each main anchor, preferred residues were defined as those with ARB values within 5-fold of the optimal residue. In a first set of experiments, we

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Europe</th>
<th>North Africa</th>
<th>North America</th>
<th>Other</th>
<th>South America</th>
<th>Southwest Asia</th>
<th>Sub-Saharan Africa</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>0501/0301</td>
<td>29.5</td>
<td>24.0</td>
<td>50.3</td>
<td>40.4</td>
<td>76.2</td>
<td>38.9</td>
<td>26.2</td>
<td>40.8</td>
</tr>
<tr>
<td>0301/0302</td>
<td>12.7</td>
<td>10.2</td>
<td>67.6</td>
<td>44.8</td>
<td>24.0</td>
<td>0.0</td>
<td>4.7</td>
<td>23.4</td>
</tr>
<tr>
<td>0102/0602</td>
<td>16.3</td>
<td>15.0</td>
<td>1.5</td>
<td>9.6</td>
<td>0.0</td>
<td>11.9</td>
<td>34.0</td>
<td>12.6</td>
</tr>
<tr>
<td>0401/0402</td>
<td>4.8</td>
<td>5.7</td>
<td>20.7</td>
<td>26.4</td>
<td>29.5</td>
<td>3.2</td>
<td>21.9</td>
<td>16.0</td>
</tr>
<tr>
<td>0101/0501</td>
<td>16.8</td>
<td>18.9</td>
<td>1.5</td>
<td>15.1</td>
<td>0.0</td>
<td>15.7</td>
<td>25.9</td>
<td>13.4</td>
</tr>
<tr>
<td>0501/0201</td>
<td>19.1</td>
<td>17.0</td>
<td>2.7</td>
<td>8.1</td>
<td>0.0</td>
<td>17.2</td>
<td>12.6</td>
<td>10.9</td>
</tr>
<tr>
<td>0201/0201</td>
<td>16.7</td>
<td>20.3</td>
<td>1.9</td>
<td>10.5</td>
<td>2.0</td>
<td>14.2</td>
<td>9.7</td>
<td>10.8</td>
</tr>
<tr>
<td>0101/0503</td>
<td>2.2</td>
<td>3.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>10.3</td>
<td>0.0</td>
<td>2.3</td>
</tr>
<tr>
<td>0102/0502</td>
<td>13.6</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>11.5</td>
<td>0.0</td>
<td>4.0</td>
</tr>
<tr>
<td>0103/0601</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.4</td>
<td>0.0</td>
<td>8.0</td>
<td>0.0</td>
<td>1.5</td>
</tr>
<tr>
<td>03/*0302</td>
<td>0.0</td>
<td>8.7</td>
<td>0.0</td>
<td>0.0</td>
<td>14.3</td>
<td>22.2</td>
<td>0.0</td>
<td>6.5</td>
</tr>
<tr>
<td>0103/0603</td>
<td>13.2</td>
<td>2.1</td>
<td>0.0</td>
<td>3.4</td>
<td>0.0</td>
<td>7.2</td>
<td>5.6</td>
<td>4.5</td>
</tr>
<tr>
<td>0301/0402</td>
<td>0.0</td>
<td>0.0</td>
<td>2.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>0102/0604</td>
<td>3.9</td>
<td>8.7</td>
<td>0.0</td>
<td>3.4</td>
<td>0.0</td>
<td>6.8</td>
<td>5.1</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Table I. Phenotypic frequencies of HLA-DQ molecules

Phenotypic frequencies derived from data available at DbMHC (National Center for Biotechnology Information).
Table II. Optimized conditions for HLA-DQ binding assays

<table>
<thead>
<tr>
<th>Molecule (DQA1*/DQB1*)</th>
<th>Serological Ag</th>
<th>Preferred Cell Lines&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Source&lt;sup&gt;b&lt;/sup&gt;</th>
<th>pH</th>
<th>Assay (h)</th>
<th>Sensitivity (IC&lt;sub&gt;50&lt;/sub&gt; nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0501/0201</td>
<td>DQ2.3</td>
<td>VAVY</td>
<td>KPLIIAEDVEGEY</td>
<td>E</td>
<td>5.5</td>
<td>72</td>
</tr>
<tr>
<td>0501/0301</td>
<td>DQ3.1 (7)</td>
<td>Herluf</td>
<td>YAIIIIAAHAAHAAAAHAA</td>
<td>A</td>
<td>7</td>
<td>72</td>
</tr>
<tr>
<td>0301/0302</td>
<td>DQ3.2 (8)</td>
<td>Preiss</td>
<td>EEDIIIFIPQEEY</td>
<td>B</td>
<td>5</td>
<td>48</td>
</tr>
<tr>
<td>0401/0402</td>
<td>DQ4</td>
<td>OLL</td>
<td>EEDIIIFIPQEEY</td>
<td>B</td>
<td>5</td>
<td>72</td>
</tr>
<tr>
<td>0101/0501</td>
<td>DQ5(1)</td>
<td>LQG</td>
<td>AAHSAAFPEDLVFQVEY</td>
<td>D</td>
<td>7</td>
<td>72</td>
</tr>
<tr>
<td>0102/0602</td>
<td>DQ6(1)</td>
<td>MGAR</td>
<td>AAATAGTIVYYAFAA</td>
<td>C</td>
<td>7</td>
<td>72</td>
</tr>
</tbody>
</table>

MHC purification and capture assays for all alleles were performed as described in Materials and Methods using the SPVL3 mAb. All assays were performed at 37°C in a final Nonidet P-40 concentration of 0.15%, and with enough labeled peptide (0.1–1 nM) to provide 15,000 cpm radioactivity. Binding of the radiolabeled ligand is determined follow a 24-h capture of MHC molecules on Ab-coated plates.

The cell line used for the majority of experiments is shown. Alternative lines include Sweig (0501/0301), Yar and 145b (0301/0302), AMAI (0102/0602), RSH (0401/0402), and MAT, Cox, Steinlin, and QBL (0501/0201).

<sup>a</sup>A, ROIV reiterative nonnatural ligand; B, *Homo sapiens* CD20 249; C, nonnatural analog of *H. sapiens* GAD65 334; D, influenza nucleoprotein 335; and E, *Mycobacterium tuberculosis* 65 kDa hsp 32.

Table III. HLA-DQ ligands and epitopes bind their restricting molecule with high affinity

<table>
<thead>
<tr>
<th>Binding Assay</th>
<th>Peptide</th>
<th>Source</th>
<th>Reported Restriction</th>
<th>Binding (IC&lt;sub&gt;50&lt;/sub&gt; nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQA1<em>0501/DQB1</em>0201</td>
<td>AFILGDNLFPKV</td>
<td>Betula pendula Bet v 1-A</td>
<td>DQB1*0201</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>KPVSKMRMATPLMQLAP</td>
<td>Homo sapiens Li chain</td>
<td>DQB1*0201</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>EEEVMTPADLDDFD</td>
<td>Herpesvirus UL48</td>
<td>DQB1*0201</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>YQSYGPSGQYTHEFD</td>
<td><em>H. sapiens</em> DQA</td>
<td>DQB1*0201</td>
<td>270</td>
</tr>
<tr>
<td></td>
<td>TEDQAMEDIQMEAESIS</td>
<td><em>Bos taurus</em> a-S1 casein</td>
<td>DQB1*0201</td>
<td>439</td>
</tr>
<tr>
<td>DQA1<em>0501/DQB1</em>0301</td>
<td>DVKFPGGGQIVGVYVLPRR</td>
<td>Hepatitis C polyprotein</td>
<td>DQB1*0301</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>MGDDYGVLACAIATHAKIRD</td>
<td>Dermatophagoides</td>
<td>DQB1*0301</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>HGSPECIPHRKGPKFQLEAV</td>
<td><em>D. pteronyssinus</em> Der p 2</td>
<td>DQB1*0301</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>EYLNKIQNSLSTEWSCSVT</td>
<td><em>Plasmodium falciparum</em> CS</td>
<td>DQB1*0301</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>VLPVQCHGPSEPCIHRR</td>
<td><em>D. pteronyssinus</em> Der p 2</td>
<td>DQB1*0301</td>
<td>3,898</td>
</tr>
<tr>
<td>DQA1<em>0301/DQB1</em>0302</td>
<td>KDILEDERAADVTC</td>
<td><em>H. sapiens</em> DRB1*0402</td>
<td>DQB1*0302</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>CDGERPTLAFLDQVM</td>
<td><em>H. sapiens</em> GAD65</td>
<td>DQB1*0302</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>DSNMINSSNNVMDIEFFFEK</td>
<td>P. falciparum 101 kDa</td>
<td>DQB1*0302</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>EEEVMTPADLDDFD</td>
<td>Herpesvirus UL48</td>
<td>DQB1*0302</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>DLLCAYSEIFGTNISKEHD</td>
<td>Gallus gallus ovomucoid</td>
<td>DQB1*0302</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>RMMEYGIIMTVSYQPL</td>
<td><em>H. sapiens</em> GAD65</td>
<td>DQB1*0302</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>FDRSTKVIDHFYPNE</td>
<td><em>H. sapiens</em> GAD65</td>
<td>DQB1*0302</td>
<td>3,957</td>
</tr>
<tr>
<td></td>
<td>TPTEKDIECARVNH</td>
<td><em>H. sapiens</em> B2-m</td>
<td>DQB1*0302</td>
<td>29,516</td>
</tr>
<tr>
<td>DQA1<em>0401/DQB1</em>0402</td>
<td>KDILEDERAADVTC</td>
<td><em>H. sapiens</em> DRB1*0402</td>
<td>DQB1*0401</td>
<td>45</td>
</tr>
<tr>
<td>DQA1<em>0101/DQB1</em>0501</td>
<td>SDDELPYIDPNMEPV</td>
<td>Herpesvirus nuc</td>
<td>DQ5</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>PLYRYLOGSFSSHVL</td>
<td><em>C. trachomatis</em> enolase</td>
<td>DQB1*0501</td>
<td>2,455</td>
</tr>
<tr>
<td></td>
<td>EELKSLNISVQAYQA</td>
<td><em>C. trachomatis</em> CT579</td>
<td>DQB1*0501</td>
<td>3,531</td>
</tr>
<tr>
<td>DQA1<em>0102/DQB1</em>0602</td>
<td>INEPTAAIAAGYGLDR</td>
<td><em>H. sapiens</em> HSP70</td>
<td>DQB1*0602</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>PLYATGRLSQQMLPMPPM</td>
<td>Herpesvirus UL48</td>
<td>DQB1*0602</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>NNYGSTIEGLLD</td>
<td>Herpesvirus VP16</td>
<td>DQB1*0602</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>RGYFKMRGKSSMRS</td>
<td>Influenza A hemagglutinin</td>
<td>DQB1*0602</td>
<td>369</td>
</tr>
<tr>
<td></td>
<td>NFRDAKACVHVGSDLK</td>
<td><em>H. sapiens</em> ATPase</td>
<td>DQB1*0602</td>
<td>379</td>
</tr>
</tbody>
</table>

<sup>IC</sup><sub>50</sub> < 1000 nM are in bold.
for DQA1*0301/DQB1*0302 (26, 30, 35, 36), as well as a crystal structure (37), which place this specificity at P9 of the binding core. Unlike previous studies, however, we did not identify significant influences at other positions.

Finally, in the case of DQA1*0102/DQB1*0602, positions T7, V9, and A12 of an analog of the GAD65 334–344 peptide were identified as main anchor positions (Table IV, Supplemental Table III). On the basis of a previously published DQB1*0602 crystal structure (47), as well as a previously published motif (46), we aligned these positions to correspond to a P4, P6, and P9 main anchor spacing. Position 4 was found to be associated with a preference for the small residues T, A, and S but also the small polar residue N and the hydrophobic/aliphatic residues M, V, L, and I. Position 6 was associated with a preference for the hydrophobic/aliphatic residues I, M, and V, and A. As with position 4, position 9 has a preference for small residues, where A, S, and G were found to be associated with the highest binding. This motif is largely in agreement with previous reports (46, 47) of DQA1*0102/DQB1*0602 binding specificity.

Taken together, the data in the present section have provided detailed motifs for three common HLA-DQ molecules. These motifs are largely compatible with previous reports, thereby validating the relevance of our approach.

**Definition of peptide-binding motifs for HLA DQA1*0501/DQB1*0301, DQA1*0401/DQB1*0402, and DQA1*0301/DQB1*0302 molecules**

Next, we defined the peptide-binding specificity of the remaining DQ molecules, for each of which no detailed motif has been reported to date. Panels of SAAS of prototype molecule-specific binders were tested, and the data were analyzed as described above.

In the case of DQA1*0501/DQB1*0301, positions G7, G8, and I10 of the hepatitis C virus polyprotein 21–35 peptide (sequence DVYKFPGGGGVGNY) were identified as main anchors, with SF of 2.6, 4.2, and 2.8, respectively (Table IV, Supplemental Table IV). To comport with the spacing of the majority of class II motifs, we aligned these as the P3, P4, and P6 positions of the canonical class II core-binding region. Thus, in this case, no position corresponding to the prototypic aromatic/aliphatic P1 anchor, characteristic of nearly all class II MHC molecules, could be identified. This is similar to what was described above for DQB1*0602, associated with main anchors at P4, P6, and P9, and also reminiscent of the case of H-2 IAα1 molecules (1A is the murine locus corresponding to the human DQ locus), which is associated with P4 and P6 main anchors. In position 3, residues representing a broad chemical specificity were preferred and included aromatic (Y, W, and F), small (G, A, C, and S), basic (R and H), and polar (Q) residues. At position 4, the small residues A, G, and P, as well as the small acidic residues C and S, were preferred. Finally, the preference in position 6 was for the small residues P, A, and T, but also the hydrophobic/aliphatic residues V, I, and L, and the basic residue R. Interestingly, this motif is not dissimilar from that described previously for DQA1*0301/DQB1*0301, where main anchors spaced two positions apart were identified, with each having a preference for small residues (68).

What we believe is a unique anchor spacing was identified in the case of DQA1*0401/DQB1*0402, where A7 and D14 of human HSP70 173–187 (sequence INEPTAAAIAYGLDR) were identified as main anchors, with SF of 3.0 and 2.9, respectively (Table IV, Supplemental Table V). If A7 is taken to be a P1 anchor, the resulting anchor pattern would correspond to a P1–P8 spacing. Alternatively, aligning D14 with P9, to correspond to the pattern observed for DQA1*0301/DQB1*0302, would place A7 as a P2 anchor, delineteating a P2–P9 spacing. Neither of these alignments has, to the best of our knowledge, been described for HLA class II. In Table IV, we have chosen the latter, P2–P9, alignment to reflect similarities in the binding repertoires of DQA1*0401/DQB1*0402 and DQA1*0301/DQB1*0302 (see below). In position 2, preferred residues were small (V, A, C, G, S, and T), aromatic (Y, W, and F), or aliphatic/hydrophobic (M and L). In this respect, the specificity at P2 may be more precisely described as the avoidance of acidic (D and E) and basic (H and K) residues. Position 9 is characterized by a preference for acidic residues (E and D), although the large polar residue Q and the aromatic residue F, were also allowed.

Finally, SAAS of FcεR 104–119 (sequence SQDLELSWNLGLQAY) identified positions E5, W8, and L10 as the main anchor contacts for DQA1*0101/DQB1*0501, with SF of 5.6, 53, and 4.8, respectively (Table IV, Supplemental Table VI). These residues correspond to a P1–P4–P6 main anchor spacing pattern. At P1 the most preferred residues were Y, V, E, Q, and R, representing a somewhat diverse chemical specificity. At P4, which based on SF would appear to be the dominant anchor position, the aromatic residues W and F were preferred. Finally, the aliphatic/hydrophobic residues V, L, I, and M were preferred in position 6.

**Summary of motifs associated with DP peptide-binding specificity**

Taken together, the data in the previous sections defined comprehensive quantitative binding motifs for each of the six most common HLA DQ specificities in the general population. A summary of the motif associated with each of the six different common DQ molecules, as defined above, is presented in Table IV.

On average, ~2.3 anchor positions were defined for each DQ molecule, similar to the cases of DR and DP. This anchor rate perhaps reflects the general structural similarities of these molecules and that they share a similar number of peptide-binding pockets (37, 47, 51, 52, 62, 63, 69). However, it is also apparent that, in general, motifs for DQ molecules are less sharply defined than for DR and DP molecules, consistent with the hypothesis that peptide anchor–MHC pocket interactions are less crucial for DQ binding. For example, in a recent report defining the motifs for five common DP molecules using the same methodology, DP anchor positions were associated with higher SF values (average, 8.95; median, 9.0) than DQ (average, 4.7; median, 4.8; excluding the unusually selective P4 anchor of DQA1*0101/DQB1*0501). Similarly, DP anchor pre-

---

**Table IV. Summary HLA DQ motifs**

<table>
<thead>
<tr>
<th>DQ Molecule</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P6</th>
<th>P9</th>
<th>P10</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQA1<em>0501/DQB1</em>0301</td>
<td>FWV</td>
<td>YWGA</td>
<td>FRC</td>
<td>SQH</td>
<td>GPACS</td>
<td>APIV</td>
<td>RLT</td>
</tr>
<tr>
<td>DQA1<em>0501/DQB1</em>0301</td>
<td>FWV</td>
<td>YWGA</td>
<td>FRC</td>
<td>SQH</td>
<td>GPACS</td>
<td>APIV</td>
<td>RLT</td>
</tr>
<tr>
<td>DQA1<em>0301/DQB1</em>0302</td>
<td>YVEQ R</td>
<td>VYATSCMMLGF</td>
<td>WMIF</td>
<td>PWQRACED</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQA1<em>0401/DQB1</em>0402</td>
<td>VYATSCMMLGF</td>
<td>WMIF</td>
<td>PWQRACED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQA1<em>0301/DQB1</em>0302</td>
<td>VYATSCMMLGF</td>
<td>WMIF</td>
<td>PWQRACED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DQA1<em>0102/DQB1</em>0602</td>
<td>VYATSCMMLGF</td>
<td>WMIF</td>
<td>PWQRACED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ferences could typically be defined by a consistent chemical specificity, usually reflecting a requirement for aromatic and/or hydrophobic residues. For some DQ anchors, in contrast, preferences represented a wide, and sometimes even antithetical, range of chemical specificities, perhaps more reflective of the fact the overriding determinant is that certain residues are forbidden, rather than a direct preference.

When the motifs defined for the various DQ molecules are compared, each molecule is associated with what we believe is a unique spacing of main anchor residues, and differing specificity for the various positions is generally noted. These results, indicating significant divergence in the mode and specificity of binding, are very much in contrast to what has been shown for the most common HLA-DR (15) and -DP (29) molecules, where molecules of each locus could be aligned with a nearly identical main anchor pattern and specificity. In fact, it is difficult to identify even pairs of DQ molecules with similar motifs. For example, although DQA1*0501/DQB1*0201 and DQA1*0101/DQB1*0501 have a preference for aromatic residues with similar motifs. For example, although DQA1*0501/DQB1*0201

In fact, it is difficult to identify even pairs of DQ molecules with similar motifs. For example, although DQA1*0501/DQB1*0201 and DQA1*0101/DQB1*0501 have a preference for aromatic residues at position 1, their specificities at position 6, where DQA1*0501/DQB1*0201 prefers acidic residues and DQA1*0101/DQB1*0501 prefers aliphatic/hydrophobic residues, are highly divergent. Perhaps the closest pairings involve DQA1*0501/DQB1*0301 and DQA1*0102/DQB1*0602, which share some overlap in their preferences for small or hydrophobic residues in positions 4 and 6, or DQA1*0301/DQB1*0302 and DQA1*0401/DQB1*0402, which share preferences for acidic residues in position 9. However, even in these cases, any overlap in specificity is not readily apparent.

Taken together, the present data indicate that, unlike previous observations in the context of HLA-DR and -DP, a common DQ-supertypic specificity is not readily apparent.

Unexpected DQ cross-reactivity revealed by testing overlapping peptide and DQ epitope sets for binding to the various DQ allelic molecules

The data presented above demonstrated that the various DQ molecules considered have rather loose peptide-binding motifs, and the defined motifs differ significantly from each other. This would seem to imply that these differences would also translate into largely divergent peptide-binding repertoires. Thus, we next asked whether significant cross-reactivity would be observed with a large panel of random peptides.

To address this point, we tested a panel of 425 nonredundant peptides for their capacity to bind each DQ molecule. The set of peptides consisted of 15-mer, overlapping by 10 residues, spanning the entire sequences of the P. pratense 1, 2, 3, 4, 5, 6, 7, 11, 12, and 13 pollen Ags, which are implicated in allergic reaction to timothy grass.

Approximately 28% of all possible DQ-binding events were associated with an affinity of 1000 nM or better (range, 15–64%), an overall rate not dissimilar to what has been previously observed with the same set of peptides for HLA-DR and -DP molecules [Sidney et al. (29) and J. Sidney and A. Sette, unpublished observations]. But somewhat surprisingly, given the differences in motifs, almost half of the peptides bound multiple DQ molecules, indicating that there is more overlap in the repertoires than random association. Altogether, 107 (25%) of the peptides bound two of the different DQ molecules tested and 45 (11%) bound three DQ molecules. Forty-seven (11%) peptides were very promiscuous DQ binders, binding more than half of the DQ molecules tested with an affinity of 1000 nM or better. This included 39 (9%) that bound four or five and eight (2%) that bound all six DQ molecules tested. By comparison, on the basis of the rates of binding to each individual molecule in our panel, it would have been expected that 17 peptides binding four or more molecules would have been identified. The high rate of cross-reactivity was found to be statistically significant ($\chi^2: p < 0.00001$).

Consideration these observations, we examined the panel of HLA-DQ–restricted epitopes and endogenous ligands described above in Table III for their capacity to cross-react among different DQ molecules. Interestingly, an even higher level of cross-reactivity was found than with the panel of P. pratense overlapping peptides as 48% ($n = 12$) of known epitopes and ligands bound four or more of the six DQ molecules with an affinity of 1000 nM or better (Fig. 1B).

Taken together, these data reveal that despite apparently largely unrelated peptide-binding motifs, a surprising degree of cross-reactivity exists among DQ specificities. This is especially apparent when known DQ-restricted epitopes are considered. However, the patterns of overlap do not suggest that cross-reactive DQ binding is, in general, a function of a common, supertypic mode of binding.

**Discussion**

We have studied the peptide-binding specificity of the six most common HLA-DQ α/β heterodimers present in the worldwide population, namely DQA1*0501/DQB1*0301, DQA1*0301/DQB1*0302, DQA1*0102/DQB1*0602, DQA1*0401/DQB1*0402, DQA1*0101/DQB1*0501, and DQA1*0101/DQB1*0201. These molecules were found to be associated with rather divergent peptide-binding motifs, as defined using a uniform MHC-binding assay platform. Surprisingly,
Although in general terms our results are in good agreement with most of the results described in the literature, there are some discrepancies or interesting differences. For example, in the context of P1 of DQ2, we have found that Asp and Glu are accommodated barely better than Arg and Lys, whereas based on published structures, it could be inferred that acidic residues would be accommodated far better than basic ones. It is also possible that proline at P1 might have had a better ARB had the peptide been extended by one residue on its N terminus to engage all possible hydrogen bonds. Also, still in the case of DQ2, the CLIP used and shown in this paper to bind with an affinity of 30 nM was previously reported to bind with ~1000-fold less affinity. This latter discrepant result can perhaps be explained by the optimized binding assay conditions that have been used in this report, which may allow to more readily detect high-affinity binding. Finally, in the case of DQ8, at p9 it was noted that R is associated with equal anchor potency as D, whereas previous binding experiments from other laboratories (26, 30, 35, 36), as well as a crystal structure (37), have shown that D/E have equal potency at this pocket and R/K/H are excluded. Also, at p4 our results show Y to be slightly better or worse accommodated as E, D, Q, and N when crystallography and previous binding experiments have shown decent accommodation for F/Y and poor for D/E and Q/N. These apparent discrepancies might reflect differences in methodology and/or effects associated with the specific ligands used in the experiments. It is also possible that some of the paradoxical or divergent preference patterns observed are the result of the use of longer peptides, which may use alternative binding frames.

Comparatively, the motifs we have described for HLA-DQ were surprisingly “loose,” with few main anchor positions associated with dominant effects on binding. Similarly, in several cases, the range of residues preferred at specific anchor positions represented a relatively diverse set of chemical specificities. These observations suggest that perhaps DQ binding is less contingent upon strong peptide–MHC pocket interactions but instead is more dependent upon backbone interactions. These observations are consistent with previous studies analyzing the structure of DQ molecules and IA molecules (the murine ortholog of DQ) (37, 47–52). Those studies pointed out that in several instances the peptide-binding pockets of the MHC were either not engaged by the amino acid side chains of the peptide ligands, or when engaged, the interaction appeared to contribute relatively little to the overall binding energy. Conversely, much of the binding energy appeared to take place with lateral interaction between the peptide and the α helices demarcating the peptide-binding groove. This is in contrast with most common DR and DP molecules, where anchor residues engaging the main MHC peptide-binding pockets contribute much of the binding affinity, and clear anchors and allele-specific binding motifs can be more easily discerned.

Several publications from Stern’s group (72–75) using DR1 and the HA306–318 peptide showed that the exquisite fitting of an aromatic residue at p1 determines peptide affinity for a DR-b86G+ allele. By contrast, in DQ alleles, two or more different pockets can be of importance by being selective, whereas other pockets may be less restrictive. Indeed, the lateral interactions are important, and the strict conservation of the MHC class II residues interacting with the peptide backbone testifies to their importance. In this paper, the experiments suggested the presence of the p10 shelf for DQ2, also shown by the Stern group (76) for DR1, and by crystallography for DQ2 (64) with P9 being unoccupied.

Despite some differences, the peptide-binding mode of DQ molecules shares fundamental characteristics with DR and DP molecules. Overall, each of the DQ molecules tested bound between 14 and 63% of the peptides tested at the 1000 nM level. These levels of
binding are consistent, both in terms of frequency and affinity, with what we have observed with other HLA class II molecules encoded in the DR and DP loci (13, 15, 29, 77, 78). Furthermore, >80% of known DQ-restricted epitopes bound their respective restricting MHC molecule with an affinity of 1000 nM or better. The binding rates and affinities detected are similar to what was previously described for HLA-DR and -DP epitopes (15, 29) and suggest that 1000 nM is a generally applicable threshold for biologically relevant binding in the context of HLA class II.

When the overlap in peptide-binding repertoire between different DQ molecules was examined by testing a panel of peptides spanning a set of P. pratense Ags and panels of known DQ epitopes, a significant and substantial overlap was revealed. In fact, several peptides were identified that have the capacity to bind all six of the DQ molecules tested with affinities of 100 nM or better (see, for example, Fig. 1, Supplemental Fig. 1E, 1H, 1I). As noted above, a relatively minor role of peptide-binding pockets and a more prominent role of lateral interactions provide the most likely explanation for this otherwise apparent paradox. On average, ~25% of the repertoires of any pair of the DQ molecules studied overlapped, which significantly exceeds what would be expected at random. The biological relevance of the binding overlap is emphasized by the fact that ~48% of the known DQ epitopes were found to be promiscuous DQ binders, having the capacity to bind four or more of the six DQ molecules tested.

It is also possible that cross-reactivity between different molecules is due to the ability of peptides to use different frames, depending on the specific molecule. That this is possible in the case of DQ has been shown previously (26) and suggested in modeling studies (66). Furthermore, this type of cross-reactivity is perhaps related to the previously reported phenomenon of epitope/motif clustering, or “hot spots” (79–85). It has been reported that certain protein regions are targets of T cells restricted by multiple HLA specificities that recognize overlapping, yet distinct, peptides. It has been hypothesized that this clustering occurs significantly more frequently than would be expected by random chance. The available data suggest that it may be possible to identify highly promiscuous helper T cell epitopes capable of mediating activation of T cells restricted by multiple HLA molecules either within a locus (DR, DQ, or DP) or even across multiple loci.

Previous studies in the DR system have shown that promiscuous binders can bind in similar registers and that indeed pan-DR–binding peptides can be engineered (86, 87). It is further possible that the same peptide can bind different HLA class II in different registers. In either case, such promiscuous binders, of obvious interest for vaccine applications, could be detected by either a “brute force” approach, namely scanning pathogenic or allergenic proteins and performing relevant binding experiments, and/or by using multiple allele specific bioinformatic predictions (71, 88, 89).

In conclusion, in the current study, we describe assays and peptide-binding motifs for the most common DQ molecules. These motifs are in general agreement with those described previously for DQA1*0501/DQB1*0201, DQA1*0301/DQB1*0302, and DQA1*0102/DQB1*0602 (26, 30–32, 34–36, 46) but also represent more detailed quantification of the specificity of these alleles. Motifs for DQA1*0501/DQB1*0301, DQA1*0401/DQB1*0402, and DQA1*0101/DQB1*0501 are novel, to the best of our knowledge, at least in terms of the level of their detail. Our experiments reveal discordant motifs but an unexpected high degree of overlap between the peptide-binding repertoires of these molecules. We anticipate that the data presented in this paper will facilitate molecular studies identifying and investigating in more detail the specificity of DQ-restricted epitopes, which despite their contribution to HLA class II responses and potentially crucial roles in association with several autoimmune disorders, including celiac disease and type 1 diabetes mellitus (38–42), have thus far received much less attention than their DR-restricted counterparts.

Acknowledgments
We thank Carla Oseroff and Howard Grey for helpful discussions.

Disclosures
The authors have no financial conflicts of interest.

References


