Cell-surface MHC density profiling reveals instability of autoimmunity-associated HLA

Hiroko Miyadera,1,3 Jun Ohashi,2 Åke Lernmark,4 Toshibo Kitamura,5 and Katsushi Tokunaga1

1Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. 2Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Chiba, Japan. 3Faculty of Medicine, University of Tsukuba, Ibaraki, Japan. 4Department of Clinical Sciences, Lund University/CRC, Skåne University Hospital SUS, Malmö, Sweden. 5Division of Cellular Therapy and Division of Stem Cell Signaling, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Sciences, The University of Tokyo, Tokyo, Japan.

Polymorphisms within HLA gene loci are strongly associated with susceptibility to autoimmune disorders; however, it is not clear how genetic variations in these loci confer a disease risk. Here, we devised a cell-surface MHC expression assay to detect allelic differences in the intrinsic stability of HLA-DQ proteins. We found extreme variation in cell-surface MHC density among HLA-DQ alleles, indicating a dynamic allelic hierarchy in the intrinsic stability of HLA-DQ proteins. Using the case-control data for type 1 diabetes (T1D) for the Swedish and Japanese populations, we determined that T1D risk–associated HLA-DQ haplotypes, which also increase risk for autoimmune endocrinopathies and other autoimmune disorders, encode unstable proteins, whereas the T1D–protective haplotypes encode the most stable HLA-DQ proteins. Among the amino acid variants of HLA-DQ, alterations in 47α, the residue that is located on the outside of the peptide-binding groove and acts as a key stability regulator, showed strong association with T1D. Evolutionary analysis suggested that 47α variants have been the target of positive diversifying selection. Our study demonstrates a steep allelic hierarchy in the intrinsic stability of HLA-DQ that is associated with T1D risk and protection, suggesting that HLA instability mediates the development of autoimmune disorders.

Introduction
HLA (also known as the MHC in other vertebrates) proteins present self- and non-self peptides to the T cell receptor (TCR) to maintain self-tolerance and adapted immunity (ref. 1 and Figure 1A). Certain HLA-DR-DQ haplotypes, such as DR3-DQA1*05–DQB1*02:01 (DR3–DQ2.5) and DR4–DQA1*03–DQB1*03:02 (DR4–DQ8.3) in Europeans, confer a risk for autoimmune diseases, including type 1 diabetes (T1D), celiac disease, and autoimmune endocrinopathy (2–7). In East Asian populations (Japanese) DR9–DQA1*03–DQB1*03:03 (DR9–DQ9.3) and DR4–DQA1*03–DQB1*04:01 (DR4–DQ4.3) are the major risk factors for T1D (8–10) and other autoimmune endocrinopathies (11) (Table 1 and see Supplemental Figure 1 for abbreviations for the DQA1–DQB1 haplotypes [DQ haplotype]; supplemental material available online with this article; doi:10.1172/JCI74961DS1). Despite accumulating genetic evidence, the mechanism through which particular HLA alleles confer risk for autoimmune diseases has not been fully uncovered.

The HLA-autoimmunity association has been generally explained by the allelic differences in self-epitope presentation (see, for instance, refs. 12, 13). However, the binding affinity and specificity in the MHC-self-epitope interaction are highly variable. Studies of T1D, multiple sclerosis, and other autoimmune disorders have found that the disease-relevant self-epitopes interact with the MHC with high or low affinity or in kinetically unfavorable registers (14–29). Some of these self-epitopes bind to both the disease risk and neutral/protective allele products (i.e., promiscuous binders) (15, 30–32). Although self-epitope presentation is critical in autoimmune pathogenesis, the above findings suggest that additional factors may also contribute to the allele-specific disease risk.

In the 1990s, Kwok’s and Unanue’s groups and other researchers reported that the T1D risk alleles of HLA-DQ and murine I-A encode SDS unstable proteins (33–36). The SDS stability measures the migration of non-boiled MHC class II (MHC II) on SDS-PAGE (37) and was initially regarded as an indicator of peptide occupancy. It was later found that SDS stability reflects the stabilization of the peptide-MHC (pMHC) at the P1 and P9 pockets and at the extended peptide residues (38–45). In some of these and other studies, however, SDS stability was not affected by the peptide-binding affinity (41, 42, 46) and was maintained through the peptide-independent stabilization (46). The mechanism of SDS stability, and hence its relevance to the MHC protein function, has remained controversial.

The stability of the pMHC is maintained through the heterodimerization of the α and β subunits and peptide presentation (Supplemental Figure 2). The interaction of the peptide side chain atoms with MHC stabilizes the pMHC in a peptide-specific manner and has been extensively analyzed (1). In this study, we focused on the possibilities that the MHC stability might differ intrinsically among the alleles and that this stability may be associated with autoimmunity. The intrinsic stability of the MHC protein in this study refers to the MHC stability that is formed through the α/β assembly and peptide main chain interactions. The contribution of both the polymorphic and nonpolymorphic residues in the heterodimerization and peptide main chain interactions suggests that MHC stability might differ intrinsically among alleles. However, it has not been possible to measure the intrinsic stability of MHC protein.
MHC protein. \( \Delta MHC \) was then used to analyze the relationship between the intrinsic stability of MHC protein and autoimmune disease risk. \( \Delta MHC \) measures the combined outcomes of the heterodimer assembly, cell-surface transport, and turnover, but not the chemical or physical stability of the MHC protein. However, for simplicity, \( \Delta MHC \) is used as an equivalent to the protein stability in this article.

In this study, we identified an allelic diversity in the intrinsic stability of HLA-DQ that has been maintained through evolution and is associated with genetic risk for T1D. Our study provides a new framework through which to interpret the HLA-autoimmunity association profiles and uncover the mechanism of autoimmunity.

or to demonstrate its allelic differences, because the pMHC is usually stabilized through both the peptide main chain and side chain interactions.

To detect the potential allelic differences in the intrinsic stability of the MHC protein, we used an alternative approach to the conventional stability assays. Specifically, instead of analyzing protein stability itself, we measured the biological outcome, the cell-surface expression of MHC protein. We quantified the amount of cell-surface MHC in engineered conditions and confirmed, through the use of mutagenesis and the model peptides, that the level of cell-surface MHC protein density (referred to herein as the \( \Delta MHC \)) reflects the intrinsic stability of the MHC protein. \( \Delta MHC \) was then used to analyze the relationship between the intrinsic stability of MHC protein and autoimmune disease risk. \( \Delta MHC \) measures the combined outcomes of the heterodimer assembly, cell-surface transport, and turnover, but not the chemical or physical stability of the MHC protein. However, for simplicity, \( \Delta MHC \) is used as an equivalent to the protein stability in this article.

In this study, we identified an allelic diversity in the intrinsic stability of HLA-DQ that has been maintained through evolution and is associated with genetic risk for T1D. Our study provides a new framework through which to interpret the HLA-autoimmunity association profiles and uncover the mechanism of autoimmunity.
Table 1. Associations of DR-DQ haplotypes with autoimmune and other immune disorders

<table>
<thead>
<tr>
<th>DR-DQA1-DQB1 haplotype</th>
<th>Susceptible</th>
<th>Protective</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR3-DQ2.5 (DRB1<em>03-DQA1</em>05-DQB1*02)</td>
<td>Celiac disease</td>
<td>selective IgA deficiency</td>
</tr>
<tr>
<td>DR4-DQ8.3 (DRB1<em>04-DQA1</em>03-DQB1*03:02)</td>
<td>APS type II</td>
<td>Celiac disease</td>
</tr>
<tr>
<td>DR5-DQ9.3 (DRB1<em>09-DQA1</em>03-DQB1*03:03)</td>
<td>TID</td>
<td>TID</td>
</tr>
<tr>
<td>DR4-DQ4.3 (DRB1<em>04-DQA1</em>03-DQB1*04:01)</td>
<td>APS type III</td>
<td>Microscopic polyangiitis</td>
</tr>
<tr>
<td>DR15-DQ6002 (DRB1<em>15-DQA1</em>02-DQB1*06:02)</td>
<td>Narcolepsy</td>
<td>APS types II and III</td>
</tr>
</tbody>
</table>

aHaplotypes commonly observed in Europeans. bHaplotypes commonly observed in the Japanese population. cTID<sub>06</sub> is SDS stable (43). dTID<sub>06</sub> is SDS unstable (43).

Results

Measurement of ΔMHC. We generated an MHC expression system using fibroblasts (NIH3T3, murine embryonic fibroblasts) as expression hosts and GFP as an internal control. MHC II is expressed in a functionally intact form in fibroblasts (47, 48) but in inappropriately paired or unassembled forms, the α and β subunits are retained during intracellular transport and are degraded (49, 50). The observation that cell-surface MHC expression on fibroblasts is altered through the gain and loss of hydrogen bond(s) (H-bond) between the MHC and the peptidase (51) indicates that subtle changes in the net stability of the pMHC can be detected using cell-surface MHC protein expression levels. H2-DM, which is expressed in antigen-presenting cells (APCs) and stabilizes pMHC, may be absent in this expression system.

We established HLA-DQ<sub>B1</sub>-stable cells using the retrovirus vector pMXs-puro and the packaging cell line PLAT-E (52, 53). We then transduced the HLA-DQ<sub>B1</sub>-stable cells with a retrovirus containing pMXs-IG/DQAI (Figure 1B). Using a graded concentration of retrovirus particles, it was possible to express both the HLA-DQ<sub>B1</sub> and GFP at several different levels (Figure 1C). Cell-surface HLA-DQ<sub>B1</sub> and cytosolic GFP were measured by flow cytometry using the pan-HLA II β mAb (WR18). The mean fluorescence intensity (MFI) for both the MHC [MFI (MHC)] and the GFP [MFI (GFP)] showed good linear correlation (R<sup>2</sup> = 0.9). The increase in MFI (MHC) relative to MFI (GFP) (slope in Figure 1D), which indicates the amount of cell-surface MHC normalized to GFP, was calculated for each HLA-DQ<sub>B1</sub> allelic pair and was designated as ΔMHC (Figure 1, C and D, and Supplemental Figure 3, A and B). To minimize interassay variation, ΔMHC was normalized to ΔMHC for the DQA1*01:02-DQB1*06:02 haplotype product (DQ0602), which is highly SDS stable (36) and showed one of the highest ΔMHC values among the tested alleles. Hereafter, the ΔMHC values that were normalized to the ΔMHC of DQ0602 are indicated in the figures unless otherwise specified. Representative ΔMHC assay data are presented in Supplemental Figure 4, A and B. ΔMHC was measured for the major HLA-DQ<sub>B1</sub> alleles in worldwide populations and in their possible trans combinations, given that the trans DQA1-DQB1 pair forms heterodimers (54), and certain trans combinations are associated with autoimmunity (55–59). In this study, the HLA allele and haplotype protein products are indicated using the nontypical version of the gene name (e.g., DQ0602 represents the DQ0602 haplotype product).

Figure 2A shows the ΔMHC profile for HLA-DQ<sub>B1</sub>. ΔMHC varied by nearly 100-fold among the HLA-DQ<sub>B1</sub> alleles. Consistent with earlier work (60–62), HLA-DQA1 and DQB1 alleles of the same evolutionary sublineage (63) expressed HLA-DQ on the cell surface (Figure 2, A and B). These sublineages are referred to herein as the subgroups DQ2/3/4 and DQ5/6. HLA-DQA1*02, *03, and *05 and certain DQB1*06 alleles also expressed HLA-DQ<sub>B1</sub> on the cell surface (Figure 2A). HLA-DQ<sub>B1</sub> cell-surface expression was not detectable in the absence of HLA-DQA1 or in the presence of the incompatible HLA-DQA1 alleles (Supplemental Figure 5, A–C, and Supplemental Table 1). The HLA-DQ<sub>B1</sub> cell-surface expression pattern and the assembly of the DQ<sub>B1</sub> and DQ<sub>B1</sub> subunits were confirmed using stable insect cells (Drosophila melanogaster S2) (Supplemental Figure 6, A–D).

Among the major DQ haplotypes, DQ6002 and DQ9.2 showed the highest ΔMHC, whereas the ΔMHC of the DQA1*01:04-DQB1*05:01 (DQ5001) product was below the threshold (Figure 2, A and C, and Supplemental Table 1). The conserved hierarchy in ΔMHC among the DQ5/6 (DQA1*01:02 > *01:01, *01:03 > *01:04 and DQB1*06:02 > *06:01, *06:03 > *06:04 > *05:03 > *05:02 > *05:01) (Figure 2A and Supplemental Figure 7, A–D) indicates that polymorphic variants in each subunit act independently of the variants in the other subunit in the regulation of ΔMHC. In DQ2/3/4, ΔMHC decreased in the order of DQA1*02 > *03, *05 > 06 and DQB1*03:01, *03:03, *04 > 02, *03:02, except for the high ΔMHC value of DQA1*03-DQB1*02 (DQ2.3) (Figure 2A and Supplemental Figure 7, E–H). The mechanism that stabilizes DQ2.3 was not identified in this study. The hierarchy in ΔMHC for the major DQ haplotype products was similar to the hierarchy in the SDS stability (DQ6002 > DQ6003 > DQ6004, DQ7.3, DQ5001, DQ8.3, DQ5.5) that was measured using the cell lysates of DQ-homologous B lymphoblastoid cell lines (B-LCLs) (56). SDS stability may be sensitive to the variants at 57β, given that DQ9.3 (ΔMHC = 0.12, carries Asp57β) is SDS stable (43), whereas DQ2.2 (ΔMHC = 0.4–0.6, carries non-Asp57β) is reported as SDS unstable (43).

The mAb WR18 stains diverse HLA-DQ (64, 65), -DR, and -DP allele products (H. Miyadera, unpublished observations), indicating that the mAb WR18 recognizes a common epitope on HLA II β.
The ΔMHC profile was reproducible with pan-HLA II β mAbs BL-IA/6, TDR31.1 (Supplemental Figure 8, A–C), and IVA-12 (H. Miyadera, unpublished observations). The ΔMHC assay was not successful in the B cell lines due to the low efficiency of retroviral transduction (H. Miyadera, unpublished observation).

The effects of peptides on ΔMHC. To confirm that ΔMHC reflects the net stability of the MHC protein, ΔMHC was measured in the presence of high- and low-affinity peptides using the DQB1*06:02 peptide fusion constructs (designed according to ref. 66) (see Methods). ΔMHC increased 3.4-fold in the presence of insulin B_{1-15} a strong binder of DQ0602 (30), and 1.5-fold in the presence of the class II–associated invariant chain peptide (CLIP)_{81-107} (Figure 3A).

ΔMHC decreased 0.6-fold in the presence of insulin B_{9-23}, a weak binder of DQ0602 (30), as well as in the presence of an artificial negative control peptide (GGSGGSGGSGGS) (GGS peptide) (Figure 3A), with which the interaction of the MHC protein with the peptide side chain may be limited. These data confirm that ΔMHC reflects the net stability of the pMHC.

The effect of endogenous peptides on the ΔMHC profile (Figure 2A) was estimated using DQB1-GGS peptide fusion constructs.
We analyzed the effect of CLIP on ΔMHC using the DQ81–CLIP 81–107 fusion constructs. The fusion of CLIP 81–107 greatly increased ΔMHC for certain alleles including DQ2.5 (Figure 3D), which is consistent with the high affinity of CLIP 94–104 for DQ2.5 (68). The variable effects of CLIP 81–107 on ΔMHC of the tested alleles indicate that the ΔMHC profile (Figure 2A) was not affected by CLIP-mediated stabilization.

Collectively, the ΔMHC profile (Figure 2A) does not appear to be biased by endogenous peptides, invariant chain, or CLIP. These observations support the possibility that the ΔMHC profile represents the allelic hierarchy in intrinsic HLA-DQ protein stability.

Polymorphic residues that regulate the ΔMHC of DQ5/6. We next searched for polymorphic sites that regulate ΔMHC. Supplemental Figures 11 and 12 show the pairwise comparisons of ΔMHC and the polymorphisms between the representative alleles (Supplemental Figure 11, A and B, and Supplemental Figure 12, A and B). For the highly conserved DQ5/6 subgroup, the responsible residues were identified through mutagenesis. For the highly polymorphic DQ2/3/4 subgroup, the major regulators of ΔMHC were identified through association analysis (see Supplemental Figure 13, A and B, for the amino acid sequence alignments of HLA-DQ).
Figure 4. Association of amino acid variants in DQ2/3/4 with \( \Delta MHC \). (A) The HLA-DQ heterodimers in DQ2/3/4 are organized in the order of their \( \Delta MHC \) values. Error bars represent the SEM. (B and C) Polymorphic variants in HLA-DQA1 (B) and -DQB1 (C) and their association with \( \Delta MHC \). Numbers indicate the amino acid residues. Residue numbers for the \( \alpha2, \beta2, \) transmembrane, and cytosolic domains are shaded in brown. Variants identical to DQA1*02:01-DQB1*04:01 (magenta) and other variants (white or gray). The association between each amino acid variant and \( \Delta MHC \) was analyzed. The lowest P values at each site are indicated on the left with asterisks (2-tailed t test). The association table is presented in Supplemental Table 2.
In DQ5/6, the HLA-DQA1*01 alleles differ at Asp/Gly2α, Tyr/Phe25α, Gln/Glu34α, and Lys/Arg47α. Of these, variants at 2α, 25α, and 34α, but not at 41α, significantly altered ΔMHC (Supplemental Figure 14, A–D). Tyr/Phe25α also altered ΔMHC between DQA1*04:01 and *06:01 (Supplemental Figure 14E). Asp2α, which is located outside of the peptide-binding groove, stabilized the HLA-DQ protein, possibly through the formation of interdomain H-bonds (Supplemental Figure 14, F and G). Tyr25α and Gln34α stabilized the HLA-DQ through the interdomain/intersubunit H-bonds and interactions with the peptide main chain (Supplemental Figure 14, H and I). The hierarchy in ΔMHC among DQA1*01 alleles is perfectly explained by the stabilizing and destabilizing effects of 2α, 25α, and 34α (Supplemental Figure 14J).

For DQB1*06, we analyzed the effect of polymorphic residue on ΔMHC through mutagenesis at the sites that differed between DQB1*06:02 and *06:04 (9β, 30β, 57β, and 87β) (Supplemental Figure 14, K and L). The replacement of Tyr30β with His30β in DQB1*06:02 and of Tyr9β with Phe9β in DQB1*06:04 decreased ΔMHC. However, the substitution of Phe9β with Tyr9β in the presence of Tyr30β (in DQB1*06:02) and the substitution of His30β with Tyr30β in the presence of Tyr9β (in DQB1*06:04) did not affect ΔMHC (Supplemental Figure 14, K and L), indicating that Tyr9β and Tyr30β act complementarily to increase the ΔMHC. Tyr9β forms an H-bond with Asn72α, which interacts with the peptide main chain (Ser7p O and Glu9p N) in DQ8.3 (ref. 69 and Supplemental Figure 14N). The higher ΔMHC of DQB1*06:02 compared with that of DQB1*06:03 suggests a greater stabilizing effect of Phe9β-Tyr30β than of Tyr9β-His30β (Supplemental Figure 11B). Asp57β mediates an H-bond/salt bridge with Arg79α and H-bonds with Ala10p N and Tyr37β (ref. 70 and Supplemental Figure 14O). Arg70β, which is predicted to interact with the p6 residue (71), increased ΔMHC relative to Gly70β in DQB1*06:04, but not in DQB1*06:02 (Supplemental Figure 14, K and L). Tyr/Phe87β did not affect ΔMHC (Supplemental Figure 14, K and L). Collectively, the hierarchy in ΔMHC among DQB1*06:02, *06:03, and *06:04 is shaped mainly through the variants at 9β, 30β, and 57β, the residues that alter the intrinsic stability of the MHC protein.

The large difference in ΔMHC between DQB1*05 and *06 is consistent with their difference in SDS stability (36) and was replicated with pan–HLA II β mAb IVA-12 and the DQw1-specific mAb Genox3.53 (H. Miyadera, unpublished observations). We determined that 14β (Leu in DQB1*05 and Met in DQB1*06), which projects its side chain toward the interface of the α1 and α2/β2 domains, is one of the responsible variants that diversifies ΔMHC between DQB1*05 and *06 (H. Miyadera, unpublished observations). Other variants that also contribute to the low ΔMHC value of DQB1*05 have not been fully elucidated in this study.

Polymorphic residues that regulate the ΔMHC of DQ2/3/4. To identify the residues that affect the ΔMHC value of DQ2/3/4, we conducted an association analysis between each amino acid variant and ΔMHC (Figure 4, A–C). In HLA-DQA1, variants at 47α
The Journal of Clinical Investigation

ResearCh article

Figure 6. Evolutionary diversification of the stability regulatory sites. (A) Neighbor-joining tree for HLA-DQA1 showing the amino acid substitutions at 2a and 47a. (B and C) d-n-ds for codons in HLA-DQA1 (14 alleles) (B), and ape DQA1 (20 alleles) (C). The x axis indicates the amino acid residue numbers starting at 18a and ending at 79a. The y axis indicates the d-n-ds at each codon, as calculated by the Nei-Gojobori method (124) using SNAP (126) (http://www.hiv.lanl.gov). The codons with a positive d-n-ds (P < 1 × 10−7, calculated for pn and ps) are plotted in orange. (D) Location of the amino acid residues that showed a positive d-n-ds (P < 1 × 10−7) in both human and ape DQA1. See Supplemental Tables 3 and 4 for the association tables.

(Lys vs. Cys and Gln), 52a (His vs. Arg), and 54a (Leu vs. Phe) were most strongly associated with ΔMHC (P = 4.4 × 10−11 by 2-tailed t test) (Figure 4B and Supplemental Table 2). In the structures of DQ2.5 and DQ8.3 (69, 72), 47a is located at the interface of the α1 and α2/β domains and appears to affect heterodimer stability, Arg52a, which does not participate in interdomain interaction, and Phe54a, which projects its side chain toward the outer surface or to the β1 domain, does not seem to explain the lower ΔMHC values of DQA1*03 and *05 compared with those of DQA1*02.

47a, which encodes Arg, Lys, Gln, and Cys, is the most variable site in HLA-DQA1 (Supplemental Figure 13A). Unlike typical MHC polymorphisms, 47a is located outside the peptide-binding groove and TCR-recognition surface (Figure 5A and Supplemental Figure 14F). In DQA1*01:02, Arg47a forms extensive H-bonds with the α2 domain (70), which may be partially maintained by Lys47a (in DQA1*02), but not by Gln47a in DQ8.3 (69, 73) or by Cys47a in DQ2.5 (ref. 72 and Figure 5, B–D). The substitution of Cys47a with Lys47a in DQ2.5 increased the ΔMHC value by 4.1-fold (from 0.10 to 0.40) (Figure 5E), accounting for approximately 69% to 86% of the ΔMHC of DQ2.2 (ΔMHC = 0.46–0.57). The substitution of Gln47a with Lys47a in DQ8.3 increased the ΔMHC value by 2.1-fold (from 0.09 to 0.18) (Figure 5E), accounting for approximately 39% of the ΔMHC of DQ8.2 (DQA1*02:01-DQB1*03:02) (ΔMHC = 0.46). Mutagenesis at 47a in the least stable DQA1*06:01 confirmed greater stabilizing effects of Lys47a and Arg47a compared with those of Gln47a and Cys47a (Figure 5F). Therefore, DQ2.2 and DQ9.2, which carry Lys47a, are intrinsically more stable than the heterodimers formed by DQA1*03–*06.

In HLA-DQB1*02, *03, and *04, none of the variants were associated with ΔMHC (Figure 4C and Supplemental Table 2). Non-Asp/Asp57β is responsible for the differences in the ΔMHC within DQB1*03, but not the entire DQ2/3/4 that is attributable to HLA-DQA1 (Figure 5G).

In both DQ5/6 and DQ2/3/4, the hierarchies in ΔMHC are generated mainly through stabilization and destabilization at the polymorphic sites that mediate the intersubunit or interdomain interactions or the interaction with the peptide main chain. These findings demonstrate that the major factor that determines the ΔMHC is the intrinsic stability of the MHC protein. Based on these and earlier findings, the ΔMHC values in Figure 2A were used as indicators of the intrinsic stability of HLA-DQ.

Evolutional divergence at 47a. The extensive variation at the 47a stability regulatory site (Supplemental Figure 13A) suggests that 47a may have been the target of positive natural selection, as has been observed for the peptide-binding sites (74, 75). The variants at 47a may have appeared before or at the time of the divergence of the HLA-DQA1 sublineages (Figure 6A). To determine whether positive selection has been operating at 47a, we compared the number of nonsynonymous substitutions per nonsynonymous sites (dN) with the number of synonymous substitutions per synonymous sites (dS)
for each codon in the DQA1 of human, ape, and other mammals. The comparison of \( dn \) with \( ds \) detects a positive selection operating on each codon or gene. Under a null hypothesis of selective neutrality, the \( dn = ds \) is expected (ref. 76 and references therein).

In human and ape DQA1, the codons for 69\( \alpha \) showed the highest \( dn-ds \) value (\( dn-ds = 1.573, P = 5.17 \times 10^{-12} \) in human DQA1; \( P \) value calculated by the Wilcoxon signed-rank test for the proportion of synonymous differences per synonymous site [ps] and the proportion of nonsynonymous differences per nonsynonymous site [pn]; see Methods). The codons for 47\( \alpha \) showed a \( dn-ds \) value that was one of the highest in DQA1 (\( dn-ds = 0.443, P = 1.63 \times 10^{-11} \) in human DQA1; \( dn-ds = 0.664, P = 1.43 \times 10^{-23} \) in ape DQA1) (Figure 6, B and C, and Supplemental Tables 3 and 4), indicating that 47\( \alpha \) has been subjected to positive diversifying selection. The codons for 47\( \alpha \) in swine DQA1 also showed a significantly positive \( dn-ds \) value (H. Miyadera, unpublished observations). The other residues that showed significantly positive \( dn-ds \) values (\( P < 1.0 \times 10^{-7} \)) in both human and ape DQA1 were 25\( \alpha \), 34\( \alpha \), 40\( \alpha \), 45\( \alpha \), 48\( \alpha \), 69\( \alpha \), and 76\( \alpha \) (Figure 6, B–D, and Supplemental Tables 3 and 4). Of these, 25\( \alpha \) and 34\( \alpha \) participate in peptide-binding and protein stability regulation, and 69\( \alpha \) and 76\( \alpha \) for each codon in the DQA1 of human, ape, and other mammals.
The composition of the DR-DQ haplotypes differs greatly between the Swedish and Japanese populations (Figure 7A); however, the frequency of the DQ haplotypes is similar in the 2 populations when the haplotypes are subgrouped by ΔMHC (Figure 7B). The DQ haplotypes with low ΔMHC values (ΔMHC < 0.2) are predominant in T1D cases, whereas those with high ΔMHC values (ΔMHC > 0.8) are present at higher frequencies in controls than in T1D cases (Figure 7B). The steep hierarchy in ΔMHC among the DQ haplotypes appears to be correlated with the risk and protection for T1D (Figure 7C and Supplemental Figure 15, A and B). Indeed, a clear inverse relationship exists between ΔMHC and susceptibility to T1D (estimated by the odds ratio [OR]) for the majority of DQ haplotypes, including those haplotypes that are most predisposing (DQ2.5, DQ8.3, DQ9.3, and DQ4.3) and protective (DQ602, DQ603, and DQ9.2). In the absence of DQ0501 and DQ0503 (in the Swedish and Japanese populations) and DRB1*04:06-DQ8.3 (in the Japanese population), which encode unstable HLA-DQ proteins and are neutral or protective of T1D, the rankings in ΔMHC and in the ORs were inversely associated (P = 4.3 × 10⁻⁴ [Swedish]; P = 0.03 [Japanese], Spearman’s rank correlation test) (Figure 7B). Although the CIs for the ORs are large, these data confirm an overall inverse relation between ΔMHC and genetic risk for T1D. The protective phenotype of the DRB1*04:06-DQ8.3 haplotype may be ascribed to DRB1*04:06 constituting the peptide-binding pockets. The significant excess of dn versus ds at 47α was not due to positive selection operating on antigen-binding sites, because the variants at 47α were not in linkage disequilibrium with the variants at the peptide-binding sites, such as 25α and 69α. The signature of positive selection at 47α independently of antigen-binding sites suggests that HLA-DQ has been evolving in favor of diversification in protein stability, in addition to increasing variations in the peptide-binding spectrum.

Autoimmune-susceptible DQ haplotypes encode unstable proteins. We next analyzed the relationship between ΔMHC and genetic risk for autoimmunity using the ΔMHC profile (Figure 2A) and case-control data for T1D in Swedish (77, 78) and Japanese populations (10). T1D is caused by the autoimmune destruction of insulin-producing β cells in the pancreas. Both HLA-DR and HLA-DQ confer a predisposition to T1D. The association of DR-DQ haplotypes with T1D has been extensively studied, and the risk hierarchy among the haplotypes has been established (59, 79, 80). In European and African-American populations, DR3-DQ2.5 and DR4-DQ8.3 are the risk haplotypes for T1D (59, 79–81). In the Japanese population, DR9-DQ9.3 and DR4-DQ4.3 are most strongly associated with T1D (8–10). DR9-DQ9.3 confers a risk for T1D in the Filipino and Korean populations (8, 82, 83). DQ602 confers protection against T1D in Swedes and Japanese, among other populations (ref. 79, Figure 7A, and Supplemental Table 5).
The Journal of Clinical Investigation

Figure 9. Hypothetical mechanisms of the HLA-autoimmunity association. (A) Peptide specificity model. This model postulates that the peptide-binding spectrum of the MHC determines the genetic association of HLA with autoimmunity. According to this model, the HLAs that are able to present the disease-relevant peptides (magenta) confer a risk for autoimmunity. Colors indicate the disease-relevant (magenta) and irrelevant peptides (white). (B) MHC stability model. This model postulates that intrinsically unstable MHC proteins (magenta), which form unstable MHC–self-epitope complexes through presentation of diverse self-peptides and are more likely to form unstable pMHC than intrinsically stable MHC, confer a risk for autoimmunity.

The simplest mechanism by which an unstable MHC protein might confer autoimmune disease risk would be the contribution of MHC protein instability to the incomplete thymic negative selection. This possibility is consistent with the established concept of autoimmunity, in which the formation of an unstable pMHC...
by low-affinity peptides permits thymic escape of self-reactive T cells (99, 100). Upon weak self-epitope presentation, intrinsically unstable MHC may form MHC–self-epitope complexes in the low-stability range and confer a risk for autoimmunity. The ability of intrinsically unstable MHC to form unstable pMHC with diverse self-epitopes, including the promiscuous binders, may explain the association of TID risk DR-DQ haplotypes with a variety of autoimmune disorders (Table 1), as well as the involvement of promiscuous self-epitopes in allele-specific disease pathogenesis.

The intrinsically unstable MHC protein may be expressed either at a basal level or maximum level depending on the availability of high-affinity peptides and accessory molecules. The intrinsically unstable MHC protein is also expected to preferentially present high-affinity peptides on the cell surface. In contrast, intrinsically stable MHC may be expressed relatively constantly on the cell surface and present the peptides with wider affinity ranges. The potential variation in the expression patterns of each MHC type might affect the outcomes of thymic selection, peripheral activation, and subset development of T cells (101–105), processes that are controlled by cell-surface MHC density. The actual mechanism that links MHC instability to autoimmunity should involve diverse immunological processes and could be highly complex.

The $\Delta MHC$ profile provides a clue for dissecting the mechanism of the $DQ$ haplotype association with TID risk. The high risk for TID of $DQ2.5/DQ8.3$ heterozygotes (58, 59, 79, 80) may be ascribed to the instability of $DQA1^05/DQB1^03:02$ (DQ8.5) ($\Delta MHC = 0.05$) (Figure 7C and Supplemental Figure 15, A and B) and its unique peptide-binding spectrum (106, 107).

$DQ0602$ and $DQ2.2$ may confer protective effects through abundant or sustained MHC expression. The high stability of $DQ0602$ and its ability to bind diabetogenic self-epitopes (30–32) are compatible with the proposed mechanisms of protection, such as the thymic deletion of self-reactive T cells (108) and the “affinity model” or “determinant capture” (106, 109, 110). The protective associations of $DRB1^15:01$–$DQB1^06:02$ with TID, autoimmune polyglandular syndrome (APS) types II and III, and selective IgA deficiency (Table 1) suggest a shared mechanism of protection among these disorders.

The neutral and protective phenotypes of $DQ0501$, $DQ0503$, $DQ7.3$, and $DQ7.5$, which encode unstable HLA-DQ proteins (Figure 7D and Supplemental Figure 15B), are not explained by the $\Delta MHC$ value. As established for the $DRB1^04$–$DQ8.3$ haplotypes (10, 59, 79–81), the predisposing and protective phenotypes of DR-DQ haplotypes can be largely influenced by the $D$ allele. The instability of HLA-DQ could be the condition that permits the linked DR to exert pathogenic or protective effects. The increased risk for TID of $DQ8.3$/$DQ0501$ and $DQ8.3$/$DQ0604$ heterozygotes (79, 80, 111, 112) ($\Delta DR4/DR13$ in the Japanese population; refs. 8-10), which do not generate trans DQ heterodimers (Figure 2A), also remains difficult to explain by the existing hypothesis.

The dependency of MHC II on the invariant chain, CLIP, and HLA-DM are proposed as the risk factors for TID (113–117). The affinity of HLA-DQ with CLIP, appears to be variable among the alleles (Figure 3D). The dependency of HLA-DQ on HLA-DM may also be variable due to polymorphisms at the 47α–56α segment (117). It will be intriguing to speculate how these interrelated parameters and their combinations might affect TID risk.

In this study, $\Delta MHC$ was used as an indicator of the intrinsic stability of the MHC protein. Several lines of direct and indirect evidence, including the mutagenesis studies, have confirmed that the intrinsic stability of the MHC protein is the major factor in determining $\Delta MHC$. In fibroblasts, the stabilization of MHC II promotes rapid transport of pMHC into the Golgi apparatus (118), and the MHC II proteins that weakly interact with peptides are protease sensitive and are degraded in the endosomes (50). In B cells, unassembled HLA-DR subunits can be degraded rapidly during intracellular transport (119, 120). Presumably, intrinsically less stable MHC proteins that are inefficient in both α and β subunit assembly and in the formation of the pMHC may generate a substantial amount of unassembled or unfolded subunits that are sensitive to proteolysis, resulting in reductions in both the number and lifetime of the pMHCs that reach the cell surface. However, because $\Delta MHC$ is measured by the cell-based assay, the possibility that $\Delta MHC$ is biased by unknown cellular components cannot be completely ruled out.

It should also be noted that the $\Delta MHC$ profile, which is measured in an engineered condition, does not represent the expression profile of HLA-DQ on professional APCs. As observed in B-LCLs, the HLA-DQ alleles that showed low $\Delta MHC$ values can be expressed at a high level (36), possibly through transcriptional upregulation and/or stabilization by high-affinity peptides.

Our study provides new insights into the molecular evolution of the MHC, which has been explained by diversification of the peptide-binding spectrum (ref. 121 and references therein). The DQ haplotypes that encode unstable and stable proteins are maintained at a high frequency in a variety of populations (allele-frequencies.net; ref. 122), indicating that the alleles encoding unstable HLA-DQ heterodimers are not selectively disadvantageous. Presumably, both the stable and unstable MHC may have functional advantages against important pathogens. The stability of HLA-DR may also be diverse (123); however, the scale of variation may be limited by the monomorphic nature of HLA-DRA.

Collectively, our study reveals an additional layer of functional hierarchy among the HLA alleles that has been generated through evolution and is associated with autoimmunity. These findings complement and extend the existing model of the HLA-autoimmunity association and suggest a mechanistic basis of autoimmune susceptibility.

Methods

Measurement of $\Delta MHC$. Full-length cDNAs for HLA-DQA1 and DQB1 were cloned from HLA-typed cell lines or peripheral blood samples from healthy individuals. The cDNAs were inserted into the retroviral vectors pMXs-puro or pMXs-IG (52) with an EcoRI site and a Kozak sequence in the 5′ terminus of the pMXs-puro and pMXs-IG vectors, respectively. The first nucleotide of the second amino acid of HLA-DQA1 and -DQB1 was changed to guanine to introduce a Kozak sequence. The HLA-DQA1-Strep-tag II and HLA-DQB1-His-tag were inserted into the pMXs-puro and pMXs-IG vectors, respectively, using EcoRI and NorI sites. To generate the retroviruses containing pMXs-puro/DQBl and pMXs-IG/DQBl, approximately $0.5 \times 10^6$ PLAT-E cells (53) were transfected with 1.5 μg plasmid and 4.5 μl Fu-GENE reagent (Roche Diagnostics), according to the manu-
The Journal of Clinical Investigation

RESEARCH ARTICLE

Fabricant’s instructions. HLA-DQβ1-stable cells were established through the transduction of NIH3T3 cells with a retrovirus containing pMXs-puro/DQβ1 and selection with puromycin (6 μg/ml). To measure ΔMHC, the HLA-DQβ1-stable cells were seeded at approximately 2 x 10^5 cells per well in a 12- or 24-well plate and cultured overnight. The cells were transduced with a retrovirus containing pMXs-IG/DQα1 using 5–60 μl PLAT-E medium. Forty-eight hours after transduction, GFP and cell-surface MHC (in GFP-positive cells) levels were measured by flow cytometry (EPICS-XL; Beckman Coulter) using anti–HLA II β mAb (WR18) (65) (MorphoSys AG) or isootype control (mouse IgG2α [6H3]; Medical & Biological Laboratories Co. Ltd.) with phycocerythrin-conjugated anti-mouse IgG (Rockland Immunchemcials or Southern Biotechnology Associates Inc.). The MFI (GFP) was defined as the MFI for GFP-positive cells minus MFI for GFP-negative cells. The MFI (MHC) was defined as the MFI for anti–HLA II β [WR18] minus the MFI for the isotype control of the GFP-positive cells (Supplemental Figure 3A). The MFI (GFP) and MFI (MHC) were plotted to calculate the increase in the MHC relative to GFP (ΔMHC) (Figure 1D and Supplemental Figure 3B). To minimize interassay variation, the ΔMHC for each DQ heterodimer was normalized with the ΔMHC value of DQ0602, which was measured on the same day (Figure 1D and Supplemental Figure 4, A and B). The lower detection threshold of ΔMHC was set at 0.005 (after normalization to DQ0602). The assay was performed more than 3 times for each HLA-DQA1 and -DQB1 allelic pair. Site-directed mutagenesis was performed with QuickChange II (Agilent Technologies), according to the manufacturer’s instructions. The following anti-HLA mAbs were purchased: BL-IA/6 (Santa Cruz Biotechnology, Inc.) and TDR31.1 (Ancell Corporation). The mAb IVA12 was an anti-HLA mAbs purchased: BL-IA/6 (Santa Cruz Biotechnology, Inc.) and TDR31.1 (Ancell Corporation). The mAb IVA12 was a anti-HLA II mAb (65) (MorphoSys AG) or isootype control (mouse IgG2α–FITC) (Beckman Coulter). 3′ Rapid amplification of cDNA ends (3′-RACE) was performed with the 3′-Full RACE Core Set (Takara Bio) using total RNA (1 μg) from the stable S2 cells as a template. Total RNA was primed for reverse transcription at 42°C with the Oligo dT-3 Sites Adaptor Primer (Takara Bio). 3′-RACE PCR was performed according to the manufacturer’s instructions. Specifically, the internal primers for HLA-DQA1 (p-50: 5′-ACTCTACCGCTGCCATAC-3′ and HLA-DQB1 (p-48: 5′-ACGGTGTGCAGACGTATACAGTGATATCGACCAGTTC-3′) and HLA-DQB1 (p-48: 5′-ACGGTGTGCAGACGTATACAGTGATATCGACCAGTTC-3′) and HLA-DQB1 were used as the 5′ primers, and the Oligo dT-3 Sites Adaptor Primer was used as the 3′ primer. The Drosophila ribosomal protein rp49 was amplified as a positive control with the primers -266 (5′-ATGACCATCGCCGGCACTAC-3′) and p-267 (5′-TGTTATCCGGACAGTTC-3′). For the purification of HLA-DQ proteins and Western blotting, the stably transfected S2 cells were harvested 48 hours after induction and stored at ~80°C. The cells were thawed on ice and lysed in 30 mM sodium phosphate buffer (pH 8.0, 150 mM NaCl, 0.5% NP-40) in the presence of a protease inhibitor cocktail (Sigma-Aldrich) and Benzonase Nuclease (EMD Millipore, Merck KGaA) for 1 hour at 4°C. The lysate was centrifuged at 2,000 g for 10 minutes at 4°C, and the supernatant was subjected to purification. Purification by Strep-tag II and His-tag was performed with a Strep-Tactin Spin Column (IBA GmbH) and MagExtractr (Toyobo Co., Ltd.), respectively, according to the manufacturers’ instructions. For Western blotting, the elution fractions were subjected to denaturation and boiling in the presence of SDS (2%) and reducing agents. Strep-Tactin-AP (IBA GmbH) and anti-His6 mAb-HRP (Roche Diagnostics) were used for the detection of DQα and DQβ, respectively.

Evolutionary analysis. The dn-ds for each codon was calculated in a pairwise manner by the Nei-Gojobori method (124) with Jukes-Cantor correction (125) using SNAP (126) (http://www.hiv.lanl.gov). The amino acids 18a–79a of the DQA1 for humans (HLA-DQA1, 14 alleles) and apes (Gogo-DQA1 [Gorilla gorilla], Pattr-DQA1 [Pan troglodytes], Hyla-DQA1 [Hylobates lar], and Poppy-DQA1 [Pongo pygmaeus], 20 alleles), which represent all of the alleles that are registered in the IMGT/HLA and IMGT/IPD databases (EMBL-EBI) (127-130) and which were nonredundant in the 18a–79a region (i.e., at least 1 codon was different from the others), were used for the analyses. The codons with a deletion or insertion (55a and 56a in the human and ape DQA1) were excluded from the analysis. For the sequence pairs with codons encoding Met or Trp (48a, 66a, and 76a in human and ape DQA1), the ps was assumed to be 0 and was used to calculate the ds. For sequence pairs showing a ps greater than or equal to 0.75 (32a, 36a, 60a, 71a, and 75a in human and ape DQA1), the ds was not calculated (at these codons, the ps was greater than the pns, and...
there were no codons with a \( p \geq 0.75 \) (Supplemental Tables 3 and 4). To assess the possibility of positive selection (i.e., diversifying selection), the differences between the \( ps \) and \( pn \) as well as between the \( ds \) and \( dn \) were statistically tested for each codon using the Wilcoxon signed-rank test. A neighbor-joining phylogenetic tree was constructed for the nucleotide sequences of exon 2 of the HLA-DQA1 alleles using the alleles with frequencies greater than 0.02 in worldwide populations (131). Molecular Evolutionary Genetics Analysis (MEGA) version 4 software (132) was used for the analysis.

Association analysis between the \( \Delta \text{MHC} \) and genetic risk for T1D. The case-control data for T1D in the Swedish population (77) were collected and genotyped by the T1D component of the 13th International Histocompatibility Workshop (78) and were deposited in the dbMHC database (http://www.ncbi.nlm.nih.gov/projects/gv/mhc/).

For the Japanese population, published case-control data (10), which were collected and genotyped by the Committee on T1D of the Japanese Diabetes Society, were used for the analysis. All of the T1D subtypes in the Japanese population (acute, fulminant, and slowly progressive) (10) were combined. The HLA-DQA1 alleles were manually estimated from the DRB1-DQB1 haplotype data based on the DRB1-DQA1-DQB1 haplotype frequency in the general European-American (133) and Japanese (134) populations. In the Japanese population, all of the DRBI*08:02-DQB1*03:02 haplotypes were assumed to carry HLA-DQA1*03. The rare DRBI-DQBI haplotypes and DRBI-DQBI haplotypes that are linked to more than 2 HLA-DQA1 alleles were excluded from the analysis. Closely related alleles, such as HLA-DQA1*01:04 and *01:05; HLA-DQA1*05:01, *05:03, and *05:05; and HLA-DQBI*04:01 and *04:02, were assumed to have identical \( \Delta \text{MHC} \) values. Because not all of the HLA-DQA1*03 and -DQB1*02 alleles were genotyped at a 4-digit resolution in the case-control data, the \( \Delta \text{MHC} \) values of DQA1*03 and of DQB1*02 were assumed to be the average of the \( \Delta \text{MHC} \) values for DQA1*03:01, *03:02, and *03:03, and of DQB1*02:01 and *02:02, respectively, and were used for the association analyses. For the analysis of the association between the amino acid variants and genetic risk for T1D, the variants at 160\( \alpha \), which differ between HLA-DQA1*03:01 and *03:03, and at 135\( \beta \), which differ between HLA-DQB1*02:01 and *02:02, were excluded from the analysis.

Statistics. Allelic differences in the \( \Delta \text{MHC} \) values as well as associatiions between the amino acid variants and \( \Delta \text{MHC} \) were analyzed by a 2-tailed \( t \) test. Differences between \( ps \) and \( pn \) and between \( ds \) and \( dn \) were analyzed by the Wilcoxon signed-rank test. Associations between the DQ haplotype frequency and susceptibility to T1D were analyzed by the \( \chi^2 \) test with Bonferroni’s correction. Spearman’s rank test was used to evaluate associations between the ranking in the \( \Delta \text{MHC} \) and ORs. A \( \chi^2 \) test was used to analyze associations between the amino acid variants and genetic risk for T1D. A \( P \) value of less than 0.05 was considered statistically significant.

Study approval. The use of cDNAs and cell lines from healthy individuals was approved by the ethics committee of the Graduate School of Medicine, The University of Tokyo, and written informed consent was obtained from all participants.

Acknowledgments
We acknowledge W.W. Kwok for his early work and valuable comments, without which this work could not have been achieved. We thank S. Harada and K. Kita for their valuable advice. We thank E.D. Mellins and E. Mignot for discussions. We are grateful to the T1D component of the 13th International Histocompatibility Workshop; the dbMHC Database (a public database funded by the NIH); and the Committee on T1D of the Japanese Diabetes Society for the case-control data. We thank T. Yabe for the HLA-genotyped cells and S. Kawai for the IVA12 mAb. We are grateful to the following groups for permission to use structural data in the figures: DQ2.5 (1s9v), C.-Y. Kim, H. Quarsten, C. Khosla, and L. M. Solidi; DQ8.3 (2nna), K.N. Henderson, J.A. Tye-Din, J. Rossjohn, and R. P. Anderson; and DQ602 (1uvq), C. Siebold, B.E. Hansen, E.Y. Jones, and L. Fugger. We thank the RIKEN Cell Bank for the NIH3T3 cell line. This work was funded by grants from the Japan Society for the Promotion of Science (JSPS) KAKENHI (22133008, to K. Tokunaga; 22133006 and 18770106, to H. Miyadera; and 23133502, to J. Ohashi).

Address correspondence to: Hiroko Miyadera, Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Phone: 81.3.5841.3653; E-mail: miyadera-h@umin.net.

Jun Ohashi’s present address is: Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan.


26. Leviisetti MG, Lewis DM, Suri A, Unanue ER. Weak proinsulin peptide-major histocompatibil-


87. Suri A, Vidavsky I, van der Drift K, Kanagawa O, Gross ML, Unanue ER. In APCs, the autologous peptides selected by the diabetogenic I-Ag7 molecule are unique and determined by the amino acid changes in the P9 pocket. *J Immunol.* 2002;168(3):1223–1227.


