



Association of HLA-DQ Heterodimer Residues –18β and β57 With Progression From Islet Autoimmunity to Diabetes in the Diabetes Prevention Trial–Type 1

Diabetes Care 2022;45:1610–1620 | <https://doi.org/10.2337/dc21-1628>

Lue Ping Zhao,^{1,2} Jay Skyler,³ George K. Papadopoulos,^{4S} Alberto Pugliese,³ James Antonio Najera,² George P. Bondinas,⁵ Antonis K. Moustakas,⁵ Ruihan Wang,⁶ Chul-Woo Pyo,⁶ Wyatt C. Nelson,⁶ Daniel E. Geraghty,⁶ and Åke Lernmark⁷

OBJECTIVE

The purpose was to test the hypothesis that the HLA-DQαβ heterodimer structure is related to the progression of islet autoimmunity from asymptomatic to symptomatic type 1 diabetes (T1D).

RESEARCH DESIGN AND METHODS

Next-generation targeted sequencing was used to genotype HLA-DQA1-B1 class II genes in 670 subjects in the Diabetes Prevention Trial–Type 1 (DPT-1). Coding sequences were translated into DQ α- and β-chain amino acid residues and used in hierarchically organized haplotype (HOH) association analysis to identify motifs associated with diabetes onset.

RESULTS

The opposite diabetes risks were confirmed for HLA *DQA1*03:01-B1*03:02* (hazard ratio [HR] 1.36; $P = 2.01 \times 10^{-3}$) and *DQA1*03:03-B1*03:01* (HR 0.62; $P = 0.037$). The HOH analysis uncovered residue –18β in the signal peptide and β57 in the β-chain to form six motifs. DQ*VA was associated with faster (HR 1.49; $P = 6.36 \times 10^{-4}$) and DQ*AD with slower (HR 0.64; $P = 0.020$) progression to diabetes onset. VA/VA, representing *DQA1*03:01-B1*03:02* (DQ8/8), had a greater HR of 1.98 ($P = 2.80 \times 10^{-3}$). The DQ*VA motif was associated with both islet cell antibodies ($P = 0.023$) and insulin autoantibodies (IAAs) ($P = 3.34 \times 10^{-3}$), while the DQ*AD motif was associated with a decreased IAA frequency ($P = 0.015$). Subjects with DQ*VA and DQ*AD experienced, respectively, increasing and decreasing trends of HbA_{1c} levels throughout the follow-up.

CONCLUSIONS

HLA-DQ structural motifs appear to modulate progression from islet autoimmunity to diabetes among at-risk relatives with islet autoantibodies. Residue –18β within the signal peptide may be related to levels of protein synthesis and β57 to stability of the peptide-DQab trimolecular complex.

Type 1 diabetes (T1D) is a chronic autoimmune disease resulting in the selective loss of pancreatic β-cells and leads to insulin dependence for life. It is one of the most common chronic diseases among children and young adults. Individuals with

¹Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA

²School of Public Health, University of Washington, Seattle, WA

³Diabetes Research Institute and Division of Endocrinology, Diabetes and Metabolism, University of Miami Miller School of Medicine, Miami, FL

⁴Laboratory of Biophysics, Biochemistry, Biomaterials and Bioprocessing, Faculty of Agricultural Technology, Technological Educational Institute of Epirus, Arta, Greece

⁵Department of Food Science and Technology, Faculty of Environmental Sciences, Ionian University, Argostoli, Kefalonia, Greece

⁶Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA

⁷Department of Clinical Sciences, Lund University Clinical Research Centre, Skåne University Hospital, Malmö, Sweden

Corresponding authors: Lue Ping Zhao, lzhao@fredhutch.org, and Åke Lernmark, ake.lernmark@med.lu.se

Received 13 August 2021 and accepted 7 March 2022

This article contains supplementary material online at <https://doi.org/10.2337/figshare.19525624>.

§G.K.P. has been retired from Technological Educational Institute (TEI) of Epirus, Arta, Greece since 1 September 2018. The affiliation is given for identification purposes only. As of 1 October 2018, the TEI of Epirus has been absorbed by the University of Ioannina. The respective department is now called Department of Agriculture.

© 2022 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

a positive family history have substantially increased risk (1), and genetic factors contribute considerably to familial aggregation of T1D (2–4). HLA class II genes contribute >50% of the inherited genetic susceptibility compared with >70 non-HLA genetic loci that individually contribute modestly to the genetic risk (5,6). HLA proteins function as antigen-presenting molecules and are key restricting elements of adaptive T and B responses. The importance of HLA also derives from studies highlighting associations of certain HLA types in secondary prevention studies to prevent progression to clinical onset of T1D, which was noted in subjects positive for HLA-DR4 but not in those positive for HLA-DR3 (7).

The development of robust autoantibody markers to multiple autoantigens and natural history studies of longitudinal cohorts have uncovered the chronic nature of islet autoimmunity. A combination of autoantibody and metabolic testing led to T1D being classified as stage 1 in the presence of two or more autoantibodies and the absence of dysglycemia, stage 2 as progression to dysglycemia, and stage 3 as diabetes (8). In many, but not all, subjects, autoantibody seroconversion may occur in the first years of life primarily in those positive for DR4-DQ8, with insulin autoantibodies (IAAs) as the first-appearing autoantibody category, while in those positive for DR3-DQ2, it may occur later, with GAD65 autoantibodies as the first-appearing autoantibody category (9–11). Unless conducted from birth, autoantibody screening of individuals at increased risk of T1D because of family relation to a proband, ascertained genetic risk, or both, will typically identify those with autoantibody positivity in whom the time of seroconversion is unknown. This is the case of the Diabetes Prevention Trial–Type 1 (DPT-1) (12,13), data from which are analyzed in here. Relatives who tested positive for autoantibodies (retrospectively classified as stage 1 and 2) were followed with periodic metabolic evaluation to estimate β -cell function and the development of stage 3 T1D. Progression to clinical symptoms was highly variable, taking from a few months to several years (14). It has been hypothesized that genetic mechanisms underlying the initiation and progression differ (10,15–17). Because most of past T1D genetic studies were case-control based, we have limited

understanding of the genetic mechanisms that may influence T1D progression.

Preclinical stages 1 and 2 of T1D represent opportunities for secondary prevention of T1D (13,14,18,19). A follow-up investigation of DPT-1 showed that *HLA-DQB1*03:02* and *HLA-DQB1*03:01* were differentially associated with faster or slower progression, respectively (20). In a related study of DPT-1 subjects, the 5 year risk of progression according to HLA-DQ types was confirmed (21) to also include earlier reports of strong protection afforded by the *HLA-DQB1*06:02* allele occasionally found among relatives with autoantibody positivity (22). Furthermore, lack of metabolic progression and a limited autoantibody response was reported among at-risk relatives enrolled in the TrialNet Pathway to Prevention study (23). We also found evidence that the HLA-DQ2/8 genotype was associated with fast progression and young age at clinical diagnosis, but the DQ2 association was limited among young people (24).

HLA is critically important to host immunity, and further studies of associations with disease progression remain of great importance. Such knowledge could inform risk assessment and guide enrollment and stratification of subjects in natural history and prevention studies. For example in most prevention trials, the presence of the T1D-protective *HLA-DQB1*06:02* is used as an exclusion criterion. Genetic risk factors, together with the autoantibody profile, could improve the assessment of disease progression and recruit at-risk subjects to prevention studies. Finally, learning at least some of the molecular mechanisms by which HLA molecules are involved in the progression of islet autoimmunity from asymptomatic to symptomatic T1D (or progression hereafter) may provide new leads into therapeutic agents that could target specific immunological processes (i.e., stopping or slowing the autoimmunity against β -cells). Indeed, dissecting the contributions of HLA genes to T1D progression could affect the secondary prevention of T1D, a main motivation of the current project, using DNA resources collected by the DPT-1.

RESEARCH DESIGN AND METHODS

DPT-1

The DPT-1 screened first- and second-degree relatives of patients with T1D to

identify those expressing islet cell antibodies (ICAs). Relatives positive for ICA who did not carry the T1D-protective *HLA-DQB1*06:02* were randomly assigned to placebo or treatment, as previously described. After excluding participants with no DNA samples or those with poor-quality DNA, 670 subjects were included in the current investigation. Among them, 316 had a projected T1D risk of >50% over 5 years and were randomly assigned to a parenteral insulin regimen or to placebo; 354 subjects had a projected risk of 25–50% over 5 years and were randomly assigned to oral insulin and placebo, as previously described (20,25).

Risk Factors and Clinical Outcome

The DPT-1 collected clinical and demographic factors from each subject, including race, sex, age, treatment type, risk level, and use of parenteral or oral insulin. The distributions for all 670 subjects were calculated (Supplementary Table 1). All subjects were followed until the clinical diagnosis of T1D or termination of the trial, and progression from islet autoimmunity to T1D was quantified through a censored time to T1D. Incidences in four combinations with two risk levels and with treatment/placebo were calculated (Supplementary Fig. 1). Associations of all risk factors with time to T1D were evaluated using a Cox regression model (Supplementary Table 1). Furthermore, distributions of race, sex, treatment type, and use of insulin were largely comparable between subjects in the low- and high-risk groups (Supplementary Table 2). Both groups had comparable age distributions (Supplementary Fig. 2).

Repeatedly Measured Biomarkers

Subjects were evaluated at 6 month intervals. During each visit, they had blood glucose, ICA, IAA, and HbA_{1c} levels assessed and underwent a 2 h oral glucose tolerance test, during which glucose and C-peptide levels were measured at –10, 0, 30, 60, 90, and 120 min.

Genotyping HLA Using Next-Generation Targeted Sequencing Technology

HLA typing was performed using the ScisGo HLA v6 Kit (Cisco Genetics Inc., Seattle, WA), following the kit protocol. Briefly, the method uses an amplicon-based two-stage PCR, followed by sample pooling and sequencing using

a MiSeq v2 PE500 Kit (Illumina, San Diego, CA). The protocol yielded three-field coverage of *HLA-DQA1* and *HLA-DQB1*. Phase within each gene was determined in part by overlapping sequences for HLA class I and a database lookup table for HLA class II (26). Since next-generation targeted sequencing (NGTS) of DNA nucleotides of selected exons was originally used to infer *DQA1* and *DQB1* alleles (27), we determined *HLA-DQA1* and *HLA-DQB1* genotypes at the high resolution of ≥ 6 digits based on the IMGT HLA nomenclature (https://raw.githubusercontent.com/ANHIG/IMGTHLA/Latest/alignments/DQB1_prot.txt; released July 2019; accessed July and October 2019). The physical positions of individual amino acids in the α -chain are referenced from codon $\alpha 1$ to $\alpha 232$ and in the β -chain from codon $\beta 1$ to $\beta 237$. We adopted the numbering system first suggested in 1998 (28) and improved in 2007 (29) that essentially allows for structural equivalence among residues of various MHC II alleles, regardless of gene locus or species, and based on the structure of the first published MHC II allele HLA-DR1 (30,31). There were also 23 and 32 residues in the signal peptides of *DQA1* and *DQB1*, respectively, and their positions were referenced as positions -23 and -32 to position -1 . By convention, the first amino acid of the mature functional chain is numbered 1.

Haplotype Association Analysis With Progression

HLA-DQA1 and *HLA-DQB1* are in strong linkage disequilibrium, which can be used to infer a pair of haplotypes, known as diplotypes. Of 670 subjects, DQ haplotypes were inferred unambiguously for 667 (32). Of the remaining three subjects, DQ haplotypes were inferred with high confidence (corresponding posterior probability 0.94). We denote DQ diplotype as HLA-DQ or DQ hereafter, and each diplotype DQ has two DQ haplotypes H1/H2, each of which can take one of K haplotypes (h_1, h_2, \dots, h_k). To assess the haplotype association with progression, we converted DQ diplotypes to a vector of K haplotype counts ($c_{i1}, c_{i2}, \dots, c_{ik}$), in which each $c_{ik} = 0, 1$, and 2 corresponded to the number of observed haplotypes h_k observed on the i^{th} subject. The

haplotypic association analysis assessed the association of haplotype counts with the time to T1D using a Cox proportional hazards regression model (33), which may be modeled through a hazard function and was written as

$$\lambda(t|c_{ik}, x_i) = \lambda_0(t) \exp(\beta_k c_{ik} + \gamma' x_i),$$

where $\lambda_0(t)$ is the baseline hazard function estimated from the data, the coefficient β_k quantifies the haplotypic association, and $\gamma' x_i$, a linear combination of covariates, is included for adjustment.

Hierarchically Organized Haplotype Association Analysis

Each DQ haplotype was coded by a sequence of amino acid residues. To uncover progression-associated residues, we applied a hierarchically organized haplotype (HOH) approach to cluster all DQ haplotypes based on similarities of their sequences and which clusters of DQ haplotypes were identified. Focusing on a cluster of multiple DQ haplotypes that have variable associations with progression HOH identified those polymorphic residues within the cluster after excluding monomorphic residues or nearly monomorphic ones. HOH further assessed the residue association within the cluster, through haplotyping residue and cluster, and selected those residues that were significantly associated with progression. The residue association was modeled through the Cox regression model described above. Multiple residues in phase form motifs and their associations with progression can be assessed through the same Cox regression model.

Generalized Estimating Equation Analysis of Repeatedly Measured Biomarkers

To functionally validate motif associations, we assessed their associations with protein-based biomarkers (ICA, IAA, and HbA_{1c}) as well as with β -cell function (fasting blood glucose and C-peptide) that were measured repeatedly during the cohort study. Let $y_i(t_{ij})$ denote the level of biomarker measured at time t_{ij} for the i^{th} subject, we modeled the biomarker association with the genetic polymorphisms, after adjusting for the temporal trend, through the linear regression for the mean biomarker value:

$$E[y_i(t_{ij})|c_{ik}] = \eta_k + \nu_k c_{ik} + \rho_k t_{ij},$$

where $\nu_k c_{ik}$ quantified the genetic association of the k^{th} motif with the coefficient ν_k , the temporal trend of the biomarker under the adjustment $\rho_k t_{ij}$, and intercept η_k . Recognizing that repeatedly measured biomarkers, denoted by vector $Y_i = [y_i(t_{i1}), y_i(t_{i2}), \dots, y_i(t_{i3})]'$, had intraindividual correlations, we used a generalized estimating equation (GEE) to estimate coefficients (ν_k, η_k, ρ_k), along with estimated SEs, Wald statistics, and P values (34–37). Since the temporal trend ($\rho_k t_{ij}$) may vary with the genetic polymorphisms, we modified the above linear regression model by including an interaction term into the average level of the biomarker, which may be written as

$$E[y_i(t_{ij})|c_{ik}] = \eta_k + \nu_k c_{ik} + \rho_k t_{ij} + \delta_k c_{ik} t_{ij},$$

where the coefficient δ_k quantified the allele-specific linear trend. GEE was used to estimate all coefficients as described above.

Statistical Analysis

All analyses in this project relied on statistical functions and packages in R. To calculate distances between all unique sequences pairwise, we used the “stringdist” function with the Levenshtein distance and then applied the “hclust” function to perform the hierarchical clustering with the agglomeration method of ward.D2. For haplotyping *DQA1* and *DQB1*, we applied the “haplo.em” function to deduce DQ diplotypes. After recoding DQ haplotypes, we performed the Cox regression analysis by the “coxph” function. Finally, we used the “geeglm” function to perform GEE association analysis.

To minimize the influence of skewed ICA, IAA, and HbA_{1c} levels, we applied a bidirectional power transformation so that transformed values within normal and abnormal ranges had symmetric distributions. The threshold values for ICA, IAA, and HbA_{1c} are chosen to be 10 JDFU, 100 units/mL, and 6.5%, respectively.

There were multiple comparisons involving multiple alleles, genotypes, residues, and outcomes in various analyses. To control for multiple comparisons in the initial allelic association analysis, we calculated the false-positive error rates (38,39) (i.e., q -value, complementary to the P value). Once DQ allelic associations

were established, additional analyses were aimed at characterizing associations and, thus, produced *P* values without multiple comparisons adjustments.

Molecular Simulation and Depiction of HLA-DQ Molecules

The simulation of the structure of HLA-DQ7cis (*DQA1*03:01-B1*03:01*) in complex with antigenic peptides had been performed as previously described (40–42) using antigenic peptides shown experimentally to bind to the purified molecule (41). The crystal structure of the very closely related molecule HLA-DQ8 (*DQA1*03:01-B1*03:02*) that confers susceptibility to T1D was used as the base molecule (42). Depictions of simulated and known structures were performed using the DSViewer Pro 6.0 and WebLab Viewer 3.5 programs (Accelrys, San Diego, CA). Details of the depictions are provided in the respective figures.

RESULTS

HLA-DQA1 and HLA-DQB1 Haplotypes Are Associated With Progression

By centering on a *DQA1-B1* haplotype, we counted the number of DQ haplotypes observed in the data set, which took a value of 0, 1, or 2, corresponding to the number of haplotype copies. Regressing the haplotype count with time to onset of T1D allowed us to estimate hazard ratio (HR), SE, *z*-score, *P* value, and *q*-value (Supplementary Table 3). The sixth haplotype (*DQA1*03:01-B1*03:02*) had 446 copies and was significantly associated with faster progression (HR 1.36; $P = 2.01 * 10^{-3}$) (highlighted green in Supplementary Table 3). On the other hand, the ninth haplotype (*DQA1*03:03-B1*03:01*), with 79 copies, was significantly associated with slower progression (HR 0.62; $P = 0.037$) (highlighted red in Supplementary Table 3). All 16 rare haplotypes, with <10 observed copies, were merged into a rare haplotype category (Supplementary Table 3). All DQ haplotype-specific association analyses were adjusted for disease severity and age because these two were significantly associated with progression and could possibly be confounders (Supplementary Table 1).

Subjects With DQ8.1 Have Variable Associations With Disease Progression

The nucleotide sequence of each DQ haplotype was converted to a sequence of amino acid residues. We then evaluated distances between each pair of 29 DQ haplotypes (including 16 rare haplotypes) based on their amino acid sequence similarity measure and performed hierarchical cluster analysis over all DQ haplotypes. The fan representation of DQ haplotype clusters with rare haplotypes, neutral haplotypes ($P > 0.05$), risk association, and resistant association are illustrated in Fig. 1A. As expected, similar DQ haplotypes were clustered together because of their sequence similarities. Six haplotypes, belonging to DQ8.1, were clustered together, and were referred accordingly. Interestingly, the DQ8.1 cluster included one risk haplotype, one resistant haplotype, two neutral haplotypes, and two rare haplotypes, despite that they had highly similar residue sequences and protein structures.

Residues –18β and β57 Among Subjects With DQ8.1 Associate With Disease Progression

Leaving out two rare haplotypes and extracting all polymorphic residues across four DQ8.1 haplotypes, we observed four residues (–18β, α157, –18β, and β57), which retained the same haplotype associations (Table 1). We then assessed the association of each residue, within the DQ8.1 cluster through HOH and found that residues –18β and β57 were significantly associated with time to T1D ($P = 0.0052$ and 0.018 , respectively) (Table 1). In contrast, the residues on the α-chain were not significantly associated with progression.

Focusing on the two residues –18β and β57, we observed six motifs (AA, AD, VA, VD, VS, VV) in the cohort (Table 1). Interestingly, only three motifs (AD, VA, and VD) were present on DQ8.1 haplotypes. Motif AD was associated with resistance to disease progression (HR 0.62; $P = 0.033$), while VA was associated with an increased risk to progression (HR 1.47; $P = 1.94 * 10^{-4}$). Motif VD tended to be associated with faster progression, but the elevated risk was not significant because of small sample size ($P = 0.169$). Surprisingly, all six motifs were observed on other haplotypes than DQ8.1, but none were associated with progression. When

pooling them together, motifs AD and VA retained comparable associations to those observed among subjects with DQ8.1 (Table 1).

To build the linkage of the identified motifs with common DQ nomenclature, we list all corresponding DQ haplotypes that share the same motif (Supplementary Table 4). For example, the risk motif VA corresponds to the *DQA1*03:01-B1*03:02*, *DQA1*03:03-B1*03:02*, and *DQA1*05:05-B1*03:02* haplotypes; the first two are within the DQ8.1 cluster, while the third represents a rare haplotype. The resistant motif AD included seven different haplotypes: Two were within the DQ8.1 cluster and the others within rare haplotypes. In contrast, the motif AA included *DQA1*05:01-B1*02:01* (DQ2.5), *DQA1*02:01-B1*02:02* and *DQA1*03:03-B1*02:02* (DQ2.2), and a rare haplotype. Their haplotype frequencies are shown among subjects with DQ8.1 and subjects without DQ8.1 haplotypes in the first and second columns of Supplementary Table 4.

DQ Associations With Progression Are Restricted to Subjects at High Risk for Progressing to T1D

The DPT-1 began screening on 15 February 1994, and the first subject was randomly assigned on 31 December 1994. This study was conducted before the implementation of the current disease staging (8). DPT-1 investigated whether parenteral insulin could prevent T1D among high-risk subjects, most of whom were at what is currently classified as stage 2. The intervention with oral insulin was tested in subjects with an estimated 25–50% 5 year risk of T1D. Current staging would rather place these individuals in a stage 1 category. We thus performed the stratified association analysis among high- and low-risk subjects, adjusting for their age (Table 2). Interestingly, motifs AD and VA retained their significant resistant and risk associations with progression among the high-risk subjects (HR 0.57 and 1.70; $P = 0.010$ and $7.09 * 10^{-5}$, respectively) in the parenteral insulin arm, but while the direction remained, the associations were not significant ($P = 0.383$ and 0.381 , respectively) in the oral insulin arm.

Genotypic Associations of DQ Motifs With Progression Among High-Risk Subjects

Recognizing that DQ associations were restricted to high-risk participants, we

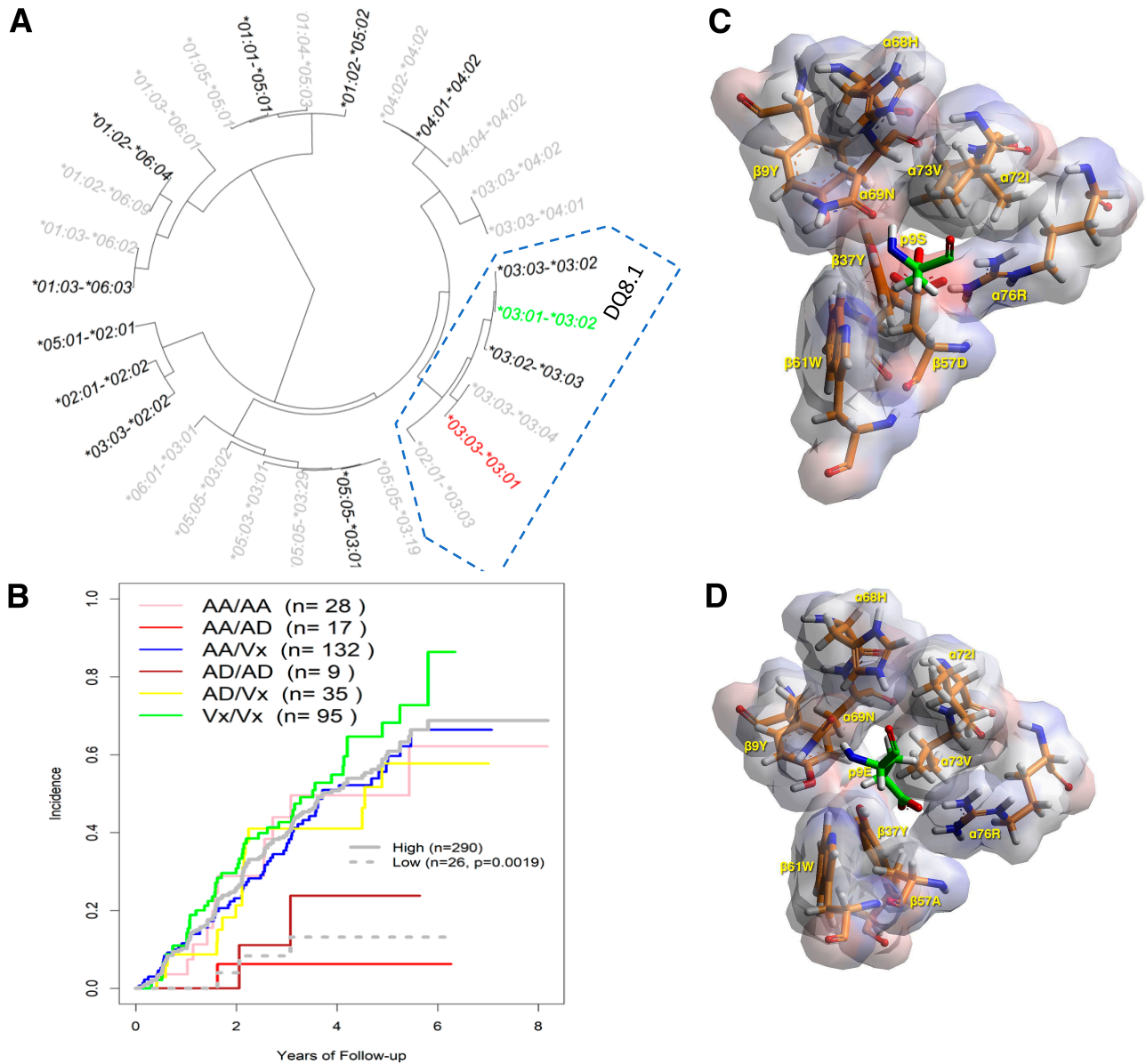


Figure 1—A: Fan representation of DQA1-DQB1 HOHs based on their residue sequences and their associations with progression accelerating progression (green), decelerated progression (red), averaged progression (black), and uncommon haplotypes (gray). A dashed line identifies a cluster of DQ8.1 haplotypes. B: Estimated T1D incidence curves of six HLA-DQ genotypes: AA/AA, AA/AD, AA/Vx, AD/AD, AD/Vx, and Vx/Vx at amino acid residue positions –18β and β57. High (AA/AA, AA/Vx, AD/Vx, and Vx/Vx; n = 290) differs from low (AA/AD and AD/AD; n = 26; P = 0.0019). C: Detailed close-up T-cell receptor view of p9 from the modeled structure of HLA-DQ7 (A1*03:01-B1*03:01) in complex with the insulin A2–14 peptide (40,41). This allele has an identical peptide-binding motif as DQA1*03:03-B1*03:01 because of identical antigen-binding α1β1 domains. The prominence of the β57Asp-α76Arg interaction is evident. The HLA-DQ7 residues are depicted in stick form (nitrogen, blue; oxygen, red; carbon, orange; sulfur, yellow; hydrogen, white) with transparent surface according to atomic charge (positive, blue; negative, red; neutral, gray; partial positive or negative charges shown by shading in between). The bound antigenic peptide is in stick form (same atom convention except that carbon is green). In addition to the specified salt bridge, these two residues form hydrogen bonds with proximal amide and carbonyl groups of the antigenic peptide backbone (not shown for clarity). This interaction, in addition to the bulky β37Tyr, render the respective p9 as shallow; hence, only small polar/apolar residues, such as p9Ser here, are accommodated. D: Detailed T-cell receptor view of p9 of HLA-DQ8 (A1*03:01-B1*03:02) in complex with the insulin B11-23/Gly peptide from the respective crystal structure (1jk8.pdb), as previously reported (42). The β57Ala residue, present in this molecule, cannot at all substitute for any of the interactions of β57Asp and provides more space for p9; rather, a p9Glu anchor, as seen here, can well serve through a salt bridge to α76Arg to promote interactions between the antigenic peptide and the DQαβ molecule. This rendering has been rotated 20° around the x-axis (top side toward the reader, bottom side away from the reader) with respect to its conventional view (all β-sheet floor segments nearly parallel to the plane of the page/screen) to better reveal the p9Glu-α76Arg interactions. The same atom and surface color conventions are used as in panel C.

next investigated their genotypic associations. Six motifs formed 17 genotypes, and only 9 of them had >10 copies (Supplementary Table 5). The genotypic

association results, in which all rare genotypes (with <10 copies) were grouped into a single composite genotype, showed that the genotype AA/AD,

with only 17 copies, was significantly resistant to progression (HR 0.08; P = 0.014). On the other hand, VA/VA, observed among 35 subjects,

Table 1—Results from HOH analysis

	<i>n</i>	Estimated coefficient	HR	SE	z-score	<i>P</i> *	DQ haplotype
Motif							
MAVA	446	0.31	1.36	0.10	3.09	2.01E-03	*03.01-*03.02
MDVA	70	0.31	1.36	0.20	1.58	1.15E-01	*03.03-*03.02
TDVD	10	0.64	1.90	0.45	1.42	1.57E-01	*03.02-*03.03
MDAD	79	−0.48	0.62	0.23	−2.09	3.69E-02	*03.03-*03.01
Residue							
−6αT		0.53	1.70	0.45	1.17	2.43E-01	
α157D		−0.16	0.85	0.15	−1.10	2.72E-01	
−18βA		−0.64	0.53	0.23	−2.80	5.17E-03	
β57D		−0.49	0.61	0.21	−2.37	1.78E-02	
DQ8.1 motif							
Carriers							
AA							
AD		−0.48	0.62	0.23	−2.13	3.28E-02	
VA		0.38	1.47	0.10	3.73	1.94E-04	
VD		0.62	1.86	0.45	1.37	1.69E-01	
VS							
VV							
Noncarriers							
AA		−0.05	0.95	0.11	−0.45	6.52E-01	
AD		−0.43	0.65	0.24	−1.78	7.46E-02	
VA		−0.07	0.94	1.00	−0.06	9.48E-01	
VD		−0.10	0.91	0.19	−0.50	6.15E-01	
VS		−0.42	0.65	0.50	−0.84	4.01E-01	
VV		−0.17	0.84	0.15	−1.10	2.70E-01	
Entire cohort							
AA		−0.09	0.91	0.11	−0.88	3.76E-01	
AD		−0.45	0.64	0.17	−2.71	6.70E-03	
VA		0.40	1.49	0.10	3.88	1.06E-04	
VD		0.03	1.03	0.18	0.14	8.87E-01	
VS		−0.28	0.76	0.51	−0.55	5.81E-01	
VV		−0.17	0.85	0.15	−1.08	2.80E-01	

Extracted haplotype association results for DQ8.1 haplotypes with all polymorphic residues (−6α, α157, −18β, and β57); individual residues among DQ8.1 carriers; and motifs of −18β and β57 with progression among DQ8.1 carriers, noncarriers, and all subjects. **P* value in individual haplotypes. When *P* < 0.05, the corresponding haplotype (or residue or motif) with HR >1 is taken as a risk haplotype, and, otherwise, is a resistant haplotype.

was significantly associated with progression (HR 1.98; *P* = 0.0028). Interestingly, the T1D delaying effect of the composite rare genotypes appeared to

be significant (*P* = 0.045). Note that three genotypes VA/VA, VA/VD, and VA/VV shared the same association direction, and they would be grouped into a

single composite genotype Vx because VA, VD, VS, and VV shared the same residue −18βV. The six genotype-specific incidence curves shown in Fig. 1B

Table 2—Results from motif −18β and β57 association analysis among subjects in the DPT-1 parenteral insulin and oral insulin arms

Motif	Parenteral insulin arm						Oral insulin arm				
	Estimated coefficient	HR	SE	z	<i>P</i> *		Estimated coefficient	HR	SE	z	<i>P</i> *
AA	−0.13	0.88	0.13	−0.94	3.48E-01		0.03	1.03	0.18	0.15	8.79E-01
AD	−0.56	0.57	0.22	−2.56	1.04E-02	Resistant	−0.22	0.80	0.26	−0.87	3.83E-01
VA	0.53	1.70	0.13	3.97	7.09E-05	Risk	0.15	1.16	0.17	0.88	3.81E-01
VD	0.02	1.02	0.24	0.07	9.41E-01		0.03	1.03	0.28	0.12	9.07E-01
VS	−0.47	0.63	1.00	−0.47	6.41E-01		−0.20	0.82	0.59	−0.34	7.36E-01
VV	−0.22	0.80	0.20	−1.10	2.72E-01		−0.11	0.89	0.24	−0.47	6.39E-01

**P* value and *q*-value for false-positive error rate in individual haplotypes. When *P* < 0.05, the corresponding haplotype (or residue or motif) with HR >1 is taken as a risk haplotype, and, otherwise, is a resistant haplotype.

Table 3—Results from association analysis of motif -18β and $\beta 57$ with repeatedly measured ICA, IAA, and HbA_{1c}

Motif	Freq	ICA				IAA				HbA _{1c}			
		Coeff	SE	z	P	Coeff	SE	z	P	Coeff	SE	z	P
Main													
AA	421	−0.04	0.04	−1.07	2.83E-01	−0.08	0.06	−1.43	1.52E-01	0.02	0.03	0.70	4.85E-01
AD	151	−0.02	0.06	−0.40	6.88E-01	−0.21	0.08	−2.45	1.45E-02	−0.05	0.04	−1.34	1.79E-01
VA	519	0.09	0.04	2.28	2.28E-02	0.17	0.06	2.93	3.34E-03	0.04	0.03	1.50	1.35E-01
V ⁺	249	−0.05	0.05	−1.10	2.71E-01	0.03	0.06	0.53	5.93E-01	−0.04	0.03	−1.41	1.60E-01
Trends													
AA	421	0.01	0.04	0.31	7.55E-01	0.05	0.02	2.28	2.25E-02	0.01	0.02	0.50	6.17E-01
AD	151	−0.02	0.05	−0.34	7.33E-01	0.01	0.02	0.34	7.35E-01	−0.04	0.01	−2.98	2.87E-03
VA	519	0.03	0.04	0.75	4.53E-01	−0.02	0.02	−0.99	3.24E-01	0.03	0.01	2.51	1.21E-02
V ⁺	249	−0.04	0.05	−0.79	4.31E-01	−0.04	0.02	−1.79	7.31E-02	−0.03	0.02	−1.57	1.17E-01

Main genetic effects and linear trend effects (estimated mean, SE, Wald statistic, and *P* value). Note: V⁺ = VD, VS, and VV. When *P* < 0.05, the corresponding motif with HR > 0 increases the level autoantibody, and, otherwise, depresses the autoantibody level. Coeff, coefficient.

reveal that the incidence curves of four genotypes (AA/AA, AA/Vx, AD/Vx, Vx/Vx) are intertwined, except that the Vx/Vx curve tended to have an upward trajectory. By contrast, AA/AD and AD/AD incidence curves were substantially below the four other genotypes. Comparing the aggregate of the top four incidence curves (high, gray line in Fig. 1B) with the aggregate of the bottom two incidence curves (low, gray dotted line in Fig. 1B) reveals that they are significantly different (*P* = 0.0019).

DQB1 Motifs Are Associated With Levels of Autoantibodies ICA and IAA and HbA_{1c}

The DPT-1 measured ICA and IAA as well as HbA_{1c} repeatedly during the follow-up, and the available data offered an opportunity to assess whether identified DQ motifs were associated with these biomarkers. As typical immune responses, these biomarkers had a skewed distribution with a long right tail along with some extreme outliers (Supplementary Fig. 3). Through bidirectional power transformation, these values were transformed to have a bimodal distribution in which the first modality corresponded to normal values and the second to the elevated levels.

The first analysis was to assess whether biomarker levels were associated with counts of each DQ motif through GEE analysis to account for multiple measurements for each subject. On each DQ motif, the estimated coefficient quantified the average change of biomarker levels with the incremental of a single observed DQ motif. From GEE, the association statistics included SE, z-score, and *P*

value. We chose to group three less common motifs (VD, VS, VV) as V⁺, leaving VA as a separate motif for assessment (Table 3). Indeed, the motif DQ*VA was associated with elevated ICA and IAA levels (coefficients 0.09 and 0.17; *P* = 0.023 and 0.0033, respectively), but not with HbA_{1c} levels (*P* = 0.135). On the other hand, the motif DQ*AD was associated with lower IAA levels (coefficient −0.21; *P* = 0.015), with no associations with ICA or HbA_{1c} (*P* = 0.688 and 0.179, respectively). Note that each association analysis was adjusted for the linear temporal trend over the follow-up by incorporating times of observations as a covariate in the GEE analysis.

The next analysis investigated whether DQ motifs influenced the linear trend of biomarker levels over time during follow-up (Table 3). Again, we used GEE to estimate the coefficient that quantified the change in the linear trend, along with other association statistics. Interestingly, DQ*AD was significantly associated with the trend of decreasing values of HbA_{1c} during the follow-up (coefficient −0.04; *P* = 0.0029), while DQ*VA was associated with the increasing trend of HbA_{1c} values (coefficient 0.03; *P* = 0.0121). The motif DQ*AA marginally contributed to the positive incline of IAA levels over time (coefficient 0.05; *P* = 0.0225).

The Residue at $\beta 57$ Has Significant Structural Properties, While the Role of -18β Is a Complete Unknown

Residue $\beta 57$ is part of pocket 9 (p9), the last of the five pockets that contribute to the binding of antigenic peptides to HLA-DQ heterodimers (Fig. 1C and D). As this

residue is polymorphic, each different amino acid at this position leads to a radically different anchoring residue at p9: $\beta 57$ Asp ensures a small aliphatic anchor at p9 (Fig. 1C), while $\beta 57$ Ala/Ser preferentially leads to an acidic anchor at this pocket, interacting with $\alpha 76$ Arg (42) (Fig. 1D and Supplementary Fig. 4A and B). The two alleles have four differing residues in the antigen-binding groove and two outside it, with the first four residues having known important functional roles (Supplementary Table 6). In the HLA-DQ2 series of alleles, the spaciousness of this pocket places the preference for aromatic anchors (43). The expression of the *DQB1*03:01* and *DQB1*03:02* alleles was investigated at the mRNA and protein level where it was found that the former showed nearly twofold higher expression at both levels (44,45). This was due to polymorphisms in the 5' upstream region that allowed for better binding of transcription factors in *DQB1*03:01* than in *DQB1*03:02*. HLA class II expression levels are crucial in the magnitude of the immune response (46). The possible influence of the -18β residue, or any other residue of the signal peptide, in protein expression of HLA-DQ proteins has not been investigated to our knowledge. Structural studies of mammalian ribosomes docked onto the signal recognition particles (SRPs), with a newly synthesized protein containing a signal peptide emerging from the ribosomal exit tunnel, did not show any particular features of the signal peptide, as is necessary for improved expression of the given protein (47). Mutational studies of preprorenin and preproinsulin have

revealed mutations at positions near –18 of the signal peptide as important (Supplementary Fig. 5).

CONCLUSIONS

Leveraging clinical data and biospecimens stored from DPT-1, we obtained high-resolution *HLA-DQA1* and *HLA-DQB1* genotypes using NGTS to assess genetic associations with T1D progression. We uncovered two residues, –18 β and β 57, that were associated with T1D progression to explain the association of DQ molecules carrying specific amino acid residues in those unique positions. The major finding that DQ*AD was significantly associated with resistance and DQ*VA with faster progression is novel, especially as both associations were restricted to those DPT-1 subjects enrolled because they had an estimated 5 year risk of >50%. Our findings that DQ*AD was associated with lower IAA levels and DQ*VA was associated with elevated ICA and IAA levels underscore the possible functional importance of the –18 β and β 57 residues in the overall disease process. This was further supported by the genotypic association analysis demonstrating that subjects with AA/AD and AD/AD had exceptionally slow progression to T1D. Taken together, subjects with these genotype-based motifs should be considered to be excluded from secondary prevention clinical trials, especially if the emphasis is on prevention among high-risk individuals. Indeed, future therapeutic interventions in subjects with these different genotypes may have to be tailored differently. It cannot be excluded that molecular or genetic targeting could be designed specifically for subjects with VA/VA and AA/AD. Equivalently, our observation that subjects with DQ*AD were found in seven DQ haplotypes (*DQA1*01:03-B1*06:01*, *A1*03:03-B1*03:01*, *A1*05:03-B1*03:01*, *A1*05:05-B1*03:01*, *A1*05:05-B1*03:19*, *A1*05:05-B1*03:29*, and *A1*06:01-B1*03:01*) and DQ*AA in four DQ haplotypes (*DQA1*02:01-B1*02:02*, *A1*03:03-B1*02:02*, *A1*03:03-B1*03:04*, and *A1*05:01-B1*02:01*) underscores the usefulness of identifying these amino acid residue motifs, as the –18 β and β 57 residues somehow contribute to disease progression associated with lower IAA levels (Supplementary Table 7). The identification of these motifs and their

association with progression provides a novel approach that may advance prediction and risk assessment. Importantly, the NGTS to identify these residues also in rare DQ haplotypes could be applied toward distinguishing slow from rapid progressors.

The strength of our study is based on the DPT-1 (12,13) that was conducted among first- and second-degree relatives of subjects with T1D. ICA and IAA positivity were used as inclusion criteria, which today would represent stage 1 or stage 2 subjects (8). On the front of structure-function relationships of HLA-DQ as the main genetic component of disease susceptibility, we note that again the HLA-DQ8 (*DQA1*03:01-B1*03:02*) haplotype was associated with faster progression to disease. By contrast, the HLA-DQ7 (*DQA1*03:03-B1*03:01*) haplotype was associated with slower progression, as noted previously (20). Their difference at the most crucial position of β 57, with HLA-DQ8 having an Ala residue and DQ7 an Asp residue, determines many diverse properties of the respective HLA-DQ $\alpha\beta$ heterodimer (40). Several T1D autoantigenic epitopes have been linked to HLA-DQ8 with the cloning of specific CD4⁺ T cells; however, this is not the case for HLA-DQ7. There have only been identifications of candidate epitopes with strong binding, but no specific and restricted CD4⁺ T cell lines or clones of effector T cells or regulatory T (Treg) cells were isolated and associated with T1D (41). The comparatively more stable pDQ7, in no small part due to β 57Asp, would be expected to lead to comparatively more Treg cells (48). The most strongly T1D-protective HLA-DQA1*01:02-B1*06:02 heterodimer appears to protect through at least two mechanisms: 1) by binding strongly to diabetogenic epitopes, “stealing,” in fact, such peptides from T1D-susceptible HLA class II heterodimers, and 2) by inducing Treg cells able to block diabetogenic effector T cells (49,50). Such studies should be extended to heterodimers such as HLA-DQ7, shown to confer protection from T1D. Surprisingly, our results show that the HLA-DQ2 (*DQA1*05:01-B1*02:01*) haplotype, also associated with an increased risk of T1D, did not seem to predispose significantly to progression. The possible contribution of the HLA-DQ8 transdimer (*DQA1*05:01-B1*03:02*) may therefore need to be taken into account

(51). Furthermore, it should be noted that the HLA-DQ typing by oligonucleotide-specific clones and allele-specific single nucleotide polymorphisms in prior studies, missed the *DQB1*02:02* allele. This allele differs from *DQB1*02:01* by one residue, β 135Gly compared with Asp, respectively. In DPT-1 subjects, as well as in other populations, *DQB1*02:02* pairs in *cis* with *DQA1*02:01*, while *DQB1*02:01* pairs in *cis* with *DQA1*05:01* (52).

Progression associated motif VA at –18 β and β 57, discovered in the parenteral insulin arm, showed persistent associations in the entire cohort, while there was only a weak association in the oral insulin arm. It was therefore of interest that motif VA was found to be associated with elevated IAA levels, while motif AD was associated with reduced IAA levels. The earlier ad hoc results of potential preventive benefit from oral insulin among those with IAA levels exceeding 80 nU/mL (12,13) makes it interesting to extend the observed association of DQ motifs with IAA levels. Indeed, assessing allelic associations of DQ motifs with IAA positivity (<80 vs. \geq 80 nU/mL), we found that VA was positively (odds ratio 1.17; $P = 0.02$) and AD negatively (odds ratio 0.70; $P = 0.02$) associated with IAA levels (Supplementary Table 7). It cannot, therefore, be excluded that individuals with the VA motif may benefit from oral insulin.

A limitation of the study is that 41 (6%) of 711 subjects in the two arms of DPT-1, parenteral (12) and oral (13) insulin, were not available because of either lack of sample or poor DNA. Another limitation is the lack of studies that potentially could be used to validate the current findings. However, as both arms of DPT-1 did not prevent progression to clinical onset of diabetes, phase 2 or 3 clinical studies are not likely to materialize. The recent oral insulin study by TrialNet (53) had different inclusion criteria and protocol to validate observations in DPT-1.

The structural role of residue β 57 has been shown to be pivotal and multifaceted (28–31,54). It lies at the antigenic peptide's carboxy terminus and when occupied by an acidic Asp residue, its interaction with apposed α 76Arg, as well as with the carbonyl and amide groups of the bound antigenic peptide, adds to the stability of the complex (55).

In addition, residue $\beta 57$ is part of p9, the last of the five pockets that contribute to the binding of peptides to HLA-DQ. Because of its polymorphism, each different amino acid at this position leads to a different anchoring residue at p9 (42). The two alleles have a number of differing residues in the antigen-binding groove, ensuring different peptide-binding motifs, particularly at p4 and p9 (56) (Supplementary Table 6). Such differences ensure that the same autoantigenic epitopes would rarely fit into both molecules in the same binding register (32,41,56).

The identification of the polymorphism at -18β is novel and unexpected. The localization in the signal peptide may involve translocation of the β -chain, pairing with the α -chain and its expression on the cell surface, but this remains to be investigated. Signal peptides, marking proteins for membrane insertion or secretion, are mostly hydrophobic with interspersed basic residues. We are not aware of how variations in one residue, as -18β Ala/Val of HLA-DQB1, could affect the process of biosynthesis and, perhaps, the function of the mature HLA-DQ protein (Supplementary Fig. 5). Mutations in the signal peptide of the human preprorenin gene are linked to severe renotubular kidney disease (57). Likewise, mutations at two positions of the human preproinsulin signal peptide result in mutant insulin gene-induced diabetes of youth (58). These mutations in the signal peptides are nonconservative and affect the expression of the respective mature proteins. There is also a single nucleotide substitution at codon 17 (Thr \rightarrow Ala) of the signal peptide of CTLA-4 (position -19 in our numbering scheme) that along with two other polymorphisms in the same gene (one in the 5' upstream region, the other in the 3' downstream region) have been associated with T1D susceptibility (termed *IDDM3*) (59) as shown in Supplementary Fig. 5. The structure of bovine prolactin signal peptide in complex with the ribosome and endoplasmic reticulum channel protein Sec61 (through which the nascent polypeptide chain proceeds) show that -18 Val of this signal peptide interacts with residues from the channel protein (47). The role of residue -18 Val compared with Ala in the signal peptide of HLA-DQB is unknown and needs to be examined through the entire process of biosynthesis of membrane/secreted

proteins anchoring the synthesizing machinery to the endoplasmic reticulum SRP receptor through the signal peptide and SRP and proceeding to the Sec61 channel protein. It is intriguing that decamer and 11-mer sequences within the signal peptide (e.g., the 6–15 SLRIPGGLRA/V and 8–21 RIPGGLRA/VATV for DQ7 and DQ8, respectively) may serve as binding peptides for HLA-A2 (anchors as indicated by bold-face), the most common of HLA-A alleles in Caucasians and perhaps also other alleles. Future research should explore the possibility that the signal peptide of MHC II molecules serve a similar stabilizing function as peptides of HLA class I signal peptides for HLA-E molecules (60). It cannot be excluded that the -18β homozygous states may affect intracellular and membrane HLA-DQ heterodimer levels in either the resting or the cytokine/pathogen-stimulated state. The subjects of this study, being T1D autoantibody positive, are certainly in the latter state that may enhance or decrease HLA-DQ expression. Decreased HLA-DQ cell surface mean fluorescence intensity with an increasing number of islet autoantibodies was observed in CD16⁺ cells, CD14⁺CD16⁻ classical monocytes, CD4⁺ T cells, and CD8⁺ T cells in 10–15-year-olds at genetic risk for T1D (61). Critical amino acid residues in the -18β , $\beta 57$ pair may explain the reduction in HLA-DQ cell surface expression.

In conclusion, we report that HLA-DQ motifs of -18β and $\beta 57$ were associated with the rate of progression to T1D in DPT-1 subjects and that these associations were paralleled by higher or lower autoantibody levels. By using genotypes of HLA-DQ motifs, we effectively identified subjects with either more rapid or slower progression to T1D. Our approach to identify predicted amino acid residues rather than alleles also allows rare risk haplotypes to be included in the analyses of risk. This may improve the design and stratification strategies of future prevention trials beyond what is possible by examining individual alleles or genotypes. Identification of these two specific residues provides specific experimental targets to investigate the functional role of $\beta 57$ and the possible regulatory role of -18β . Such a role might include influencing cell membrane protein levels or in any other manner the autoimmune

response associated with specific HLA-DQ molecules. Finally, further studies are needed to determine whether DQ motifs influence the preventive potential by either parenteral or oral insulin.

Acknowledgments. The authors thank the National Institute of Diabetes and Digestive and Kidney Diseases Central Repository for approving their request to access DNA samples and for expeditiously delivering these to their laboratory. The authors also thank Dr. Ramanujan S. Hegde of the Medical Research Council Laboratory of Molecular Biology, Cambridge, U.K., for help in guiding G.K.P. to the relevant signal peptide-Sec61 structure.

Funding. The study was supported by National Institute of Diabetes and Digestive and Kidney Diseases grant 1R56 DK117276. The molecular simulation studies by G.K.P., G.B.P., and A.K.M. and reported here were supported by grants from the Technological Educational Institute of Epirus Research Committee, a Biotech4 grant from the European Union (contract BIO4 CT95 0263), a supplementary grant from the Hellenic Secretariat of Research and Technology, and a grant from the European Union's Third EU Regional Development Framework for Greece (EPEAEK II scheme, Program Archimedes).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. L.P.Z., J.S., G.K.P., D.E.G., and Å.L. researched and analyzed the data and wrote the manuscript. J.S. and A.P. were investigators who conducted DPT-1 trial. R.W., C.-W.P., W.C.N., and D.E.G. performed the NGTS and researched data. G.P.B., J.A.N., and A.K.M. performed molecular simulations and graphical representations of select HLA-DQ molecules and contributed to the Conclusions section of the manuscript. All authors reviewed and commented on the manuscript with respect to structural implications and discussions. L.P.Z. and Å.L. are guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Pociot F, Lernmark Å. Genetic risk factors for type 1 diabetes. *Lancet* 2016;387:2331–2339
- Tuomilehto J. The emerging global epidemic of type 1 diabetes. *Curr Diab Rep* 2013;13:795–804
- Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T. Concordance for islet autoimmunity among monozygotic twins. *N Engl J Med* 2008;359:2849–2850
- Atkinson MA, Eisenbarth GS, Michels AW. Type 1 diabetes. *Lancet* 2014;383:69–82
- Concannon P, Rich SS, Nepom GT. Genetics of type 1A diabetes. *N Engl J Med* 2009;360:1646–1654
- Robertson CC, Inshaw JR, Onengut-Gumuscu S, et al.; Type 1 Diabetes Genetics Consortium. Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. *Nat Genet* 2021;53:962–971

7. Herold KC, Bundy BN, Long SA, et al.; Type 1 Diabetes TrialNet Study Group. An anti-CD3 antibody, teplizumab, in relatives at risk for type 1 diabetes. *N Engl J Med* 2019;381:603–613
8. Insel RA, Dunne JL, Atkinson MA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care* 2015;38:1964–1974
9. Ilonen J, Hammäs A, Laine AP, et al. Patterns of β -cell autoantibody appearance and genetic associations during the first years of life. *Diabetes* 2013;62:3636–3640
10. Krischer JP, Lynch KF, Schatz DA, et al.; TEDDY Study Group. The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study. *Diabetologia* 2015;58:980–987
11. Krischer JP, Lynch KF, Lernmark Å, et al.; TEDDY Study Group. Genetic and Environmental interactions modify the risk of diabetes-related autoimmunity by 6 years of age: the TEDDY study. *Diabetes Care* 2017;40:1194–1202
12. Diabetes Prevention Trial–Type 1 Diabetes Study Group. Effects of insulin in relatives of patients with type 1 diabetes mellitus. *N Engl J Med* 2002;346:1685–1691
13. Skyler JS, Krischer JP, Wolfsdorf J, et al. Effects of oral insulin in relatives of patients with type 1 diabetes: the Diabetes Prevention Trial–Type 1. *Diabetes Care* 2005;28:1068–1076
14. Katsarou A, Gudbjörnsdóttir S, Rawshani A, et al. Type 1 diabetes mellitus. *Nat Rev Dis Primers* 2017;3:17016
15. Krischer JP, Liu X, Vehik K, et al.; TEDDY Study Group. Predicting islet cell autoimmunity and type 1 diabetes: an 8-year TEDDY study progress report. *Diabetes Care* 2019;42:1051–1060
16. Ziegler AG, Rewers M, Simell O, et al. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* 2013;309:2473–2479
17. Beyerlein A, Bonifacio E, Vehik K, et al.; TEDDY Study Group. Progression from islet autoimmunity to clinical type 1 diabetes is influenced by genetic factors: results from the prospective TEDDY study. *J Med Genet* 2019;56:602–605
18. Vehik K, Cuthbertson D, Ruhlig H, Schatz DA, Peakman M; DPT-1 and TrialNet Study Groups. Long-term outcome of individuals treated with oral insulin: Diabetes Prevention Trial–Type 1 (DPT-1) oral insulin trial. *Diabetes Care* 2011;34:1585–1590
19. Dayan CM, Korah M, Tatovic D, Bundy BN, Herold KC. Changing the landscape for type 1 diabetes: the first step to prevention. *Lancet* 2019;394:1286–1296
20. Butty V, Campbell C, Mathis D, Benoist C; DPT-1 Study Group. Impact of diabetes susceptibility loci on progression from pre-diabetes to diabetes in at-risk individuals of the Diabetes Prevention Trial–Type 1 (DPT-1). *Diabetes* 2008;57:2348–2359
21. Redondo MJ, Babu S, Zeidler A, et al.; Diabetes Prevention Trial Type 1 Study Group. Specific human leukocyte antigen DQ influence on expression of antiislet autoantibodies and progression to type 1 diabetes. *J Clin Endocrinol Metab* 2006;91:1705–1713
22. Pugliese A, Gianani R, Moromisato R, et al. HLA-DQB1*0602 is associated with dominant protection from diabetes even among islet cell antibody-positive first-degree relatives of patients with IDDM. *Diabetes* 1995;44:608–613
23. Pugliese A, Boulware D, Yu L, et al.; Type 1 Diabetes TrialNet Study Group. HLA-DRB1*15:01-DQA1*01:02-DQB1*06:02 haplotype protects autoantibody-positive relatives from type 1 diabetes throughout the stages of disease progression. *Diabetes* 2016;65:1109–1119
24. Graham J, Hagopian WA, Kockum I, et al.; Diabetes Incidence in Sweden Study Group; Swedish Childhood Diabetes Study Group. Genetic effects on age-dependent onset and islet cell autoantibody markers in type 1 diabetes. *Diabetes* 2002;51:1346–1355
25. Diabetes Prevention Trial–Type 1 Diabetes Study Group. Effects of insulin in relatives of patients with type 1 diabetes mellitus. *N Engl J Med* 2002;346:1685–1691
26. Nelson WC, Pyo CW, Vogan D, et al. An integrated genotyping approach for HLA and other complex genetic systems. *Hum Immunol* 2015;76:928–938
27. Dallmann-Sauer M, Fava VM, Gzara C, et al. The complex pattern of genetic associations of leprosy with HLA class I and class II alleles can be reduced to four amino acid positions. *PLoS Pathog* 2020;16:e1008818
28. Fremont DH, Monnaie D, Nelson CA, Hendrickson WA, Unanue ER. Crystal structure of I-Ak in complex with a dominant epitope of lysozyme. *Immunity* 1998;8:305–317
29. Bondinas GP, Moustakas AK, Papadopoulos GK. The spectrum of HLA-DQ and HLA-DR alleles, 2006: a listing correlating sequence and structure with function. *Immunogenetics* 2007;59:539–553
30. Brown JH, Jardetzky TS, Gorga JC, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 1993;364:33–39
31. Stern LJ, Brown JH, Jardetzky TS, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994;368:215–221
32. Zhao LP, Papadopoulos GK, Kwok WW, et al. Next-generation HLA sequence analysis uncovers seven HLA-DQ amino acid residues and six motifs resistant to childhood type 1 diabetes. *Diabetes* 2020;69:2523–2535
33. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: Wiley, 2002
34. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22
35. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;42:121–130
36. Zhao LP, Prentice RL, Self SG. Multivariate mean parameter estimation by using a partly exponential model. *J R Stat Soc B* 1992;54:805–811
37. Zhao LPPRL. *Use of a quadratic exponential model to generate estimating equations for means, variances, and covariances*. Oxford, Clarendon Press, 1991
38. Storey JD. The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann Stat* 2003;31:2013–2035
39. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 2003;100:9440–9445
40. Moustakas AK, Papadopoulos GK, Friede T, et al. The antigenic motif of the type-1-diabetes resistant HLA-DQA1*0501/B1*0301 allele compared to the neutral allele DQA1*0301/B1*0301 [abstract 333]. *Diabetologia* 2005;48(Suppl. 1):A125
41. Ettinger RA, Kwok WW. A peptide binding motif for HLA-DQA1*0102/DQB1*0602, the class II MHC molecule associated with dominant protection in insulin-dependent diabetes mellitus. *J Immunol* 1998;160:2365–2373
42. Lee KH, Wucherpfennig KW, Wiley DC. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat Immunol* 2001;2:501–507
43. Stepniak D, Wiesner M, de Ru AH, et al. Large-scale characterization of natural ligands explains the unique gluten-binding properties of HLA-DQ2. *J Immunol* 2008;180:3268–3278
44. Beaty JS, West KA, Nepom GT. Functional effects of a natural polymorphism in the transcriptional regulatory sequence of HLA-DQB1. *Mol Cell Biol* 1995;15:4771–4782
45. Beaty JS, Sukiennicki TL, Nepom GT. Allelic variation in transcription modulates MHC class II expression and function. *Microbes Infect* 1999;1:919–927
46. Reichstetter S, Kwok WW, Kochik S, Koelle DM, Beaty JS, Nepom GT. MHC-peptide ligand interactions establish a functional threshold for antigen-specific T cell recognition. *Hum Immunol* 1999;60:608–618
47. Voorhees RM, Hegde RS. Structure of the Sec61 channel opened by a signal sequence. *Science* 2016;351:88–91
48. Li MO, Rudensky AY. T cell receptor signalling in the control of regulatory T cell differentiation and function. *Nat Rev Immunol* 2016;16:220–233
49. Wen X, Yang J, James E, Chow IT, Reijonen H, Kwok WW. Increased islet antigen-specific regulatory and effector CD4⁺ T cells in healthy individuals with the type 1 diabetes-protective haplotype. *Sci Immunol* 2020;5:eaax8767
50. Petsiou A, Paschou SA, Vartholomatos G, et al. A modified flow cytometry method for objective estimation of human CD4⁺ regulatory T cells (CD4⁺ Tregs) in peripheral blood, via CD4/CD25/CD45RO/FoxP3 labeling. *Cytometry B Clin Cytom* 2020;98:259–269
51. Chow IT, Gates TJ, Papadopoulos GK, et al. Discriminative T cell recognition of cross-reactive islet-antigens is associated with HLA-DQ8 transdimer-mediated autoimmune diabetes. *Sci Adv* 2019;5:eaaw9336
52. Zhao LP, Papadopoulos GK, Kwok WW, et al. Motifs of three HLA-DQ amino acid residues (α 44, β 57, β 135) capture full association with the risk of type 1 diabetes in DQ2 and DQ8 children. *Diabetes* 2020;69:1573–1587
53. Krischer JP, Schatz DA, Bundy B, Skyler JS; Writing Committee for the Type 1 Diabetes TrialNet Oral Insulin Study Group. Effect of oral insulin on prevention of diabetes in relatives of patients with type 1 diabetes: a randomized clinical trial. *JAMA* 2017;318:1891–1902
54. Painter CA, Stern LJ. Conformational variation in structures of classical and non-classical MHCII proteins and functional implications. *Immunol Rev* 2012;250:144–157
55. Siebold C, Hansen BE, Wyer JR, et al. Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers strong susceptibility to narcolepsy. *Proc Natl Acad Sci U S A* 2004;101:1999–2004

56. Kwok WW, Domeier ML, Raymond FC, Byers P, Nepom GT. Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. *J Immunol* 1996;156:2171–2177
57. Živná M, Kidd K, Zaidan M, et al. An international cohort study of autosomal dominant tubulointerstitial kidney disease due to REN mutations identifies distinct clinical subtypes. *Kidney Int* 2020;98:1589–1604
58. Guo H, Xiong Y, Witkowski P, et al. Inefficient translocation of preproinsulin contributes to pancreatic β cell failure and late-onset diabetes. *J Biol Chem* 2014;289:16290–16302
59. Qu HQ, Bradfield JP, Grant SF, Hakonarson H, Polychronakos C; Type 1 Diabetes Genetics Consortium. Remapping the type 1 diabetes association of the CTLA4 locus. *Genes Immun* 2009;10(Suppl. 1):S27–S32
60. Pietra G, Romagnani C, Moretta L, Mingari MC. HLA-E and HLA-E-bound peptides: recognition by subsets of NK and T cells. *Curr Pharm Des* 2009;15:3336–3344
61. Andersson Svärd A, Maziarz M, Ramelius A, Lundgren M, Lernmark Å; DiPiS Study Group. Decreased HLA-DQ expression on peripheral blood cells in children with varying number of beta cell autoantibodies. *J Transl Autoimmun* 2020;3:100052