Analysis of structure and function relationships of an autoantigenic peptide of insulin bound to H-2K^d that stimulates CD8 T cells in insulin-dependent diabetes mellitus

F. Susan Wong*†‡, Antonis K. Moustakas†§, Li Wen¶, George K. Papadopoulos§‖, and Charles A. Janeway, Jr.‖**

*Department of Pathology and Microbiology, School of Medical Sciences, University of Bristol, Bristol, BS8 1TD, United Kingdom; §Laboratory of Biochemistry and Biophysics, Faculty of Agricultural Technology, Technological Educational Institute of Epirus, GR47100 Arta, Greece; ¶Section of Endocrinology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520-8011; and **Howard Hughes Medical Institute, Section of Immunobiology, Yale University School of Medicine, 310 Cedar Street, New Haven, CT 06520-8011

Contributed by Charles A. Janeway, Jr., January 22, 2002

The recognition of MHC–peptide complexes by T cells is governed by structural considerations that are determined by the sequences of the individual components and their interaction with each other. We have studied the function of a highly diabetogenic CD8 T cell clone that is specific for insulin B15-23:H-2K^d. We have then related this to modeled MHC–peptide structures. The native peptide binds poorly to H-2K^d, because of the small glycine residue at peptide position p9 that is incapable of productive interactions with the hydrophobic residues of pocket F. In addition, electrostatic repulsions between the peptide glutamate residue at position 7 and 152D of the MHC molecule heavy chain contribute to the poor binding. However, B chain peptide 15-23 bound to K^d shows excellent T cell stimulation and the induction of CD8 cytotoxic T cells. Peptide substitution has also shown that p6G is likely to be a T cell antigen receptor interaction site. Our studies have shown that the predictions seen in the models correlate closely with the observed effects in functional assays and provide insight into how this peptide, which would not be predicted to stimulate these cells on H-2K^d binding studies alone, could activate such highly pathogenic T cells.

The identification of autoantigens in insulin-dependent diabetes became a major focus once this disease was recognized to have an autoimmune basis. A number of different autoantigens including insulin, glutamic acid decarboxylase, and the tyrosine phosphatase IA-2 had initially been suggested, based on the finding that specific auto-Abs were present in prediabetic or newly diabetic patients. Autoantigens have been identified for a number of pathogenic CD4 T cells in diabetes (1–4). In addition, the recent crystallization of the unusual MHC class II molecule I-A^{g7} in the nonobese diabetic (NOD) mouse model has given important insights into the binding of autoantigenic peptides recognized by CD4 T cells (5, 6).

Most recently, it has been recognized that CD8 T cells play a vital role in the early stages of the pathogenesis of diabetes. We have studied a CD8 T cell clone isolated from the islets of young, prediabetic NOD mice (7) that has been shown to be highly pathogenic in adoptive transfer. The MHC class I molecules in the NOD mouse are common to a number of mouse strains. However, other mice do not develop disease from the existence of potentially autoreactive MHC class I-restricted T cells. The autoantigen recognized by these highly pathogenic CD8 T cells is a peptide of the insulin B chain (amino acids 15–23, identical in human, mouse I, and II sequences), a sequence that is not in any way unique to the NOD mouse. Interestingly, it is part of the B9-23 sequence of the molecule that has been identified as the antigen for pathogenic CD4 T cells in the NOD mouse (1). In addition, it has just recently been shown to be an autoantigen for DQ8-restricted T cells in the peripheral blood of patients with newly diagnosed type-1 diabetes and prediabetic individuals (8).

Motifs have been established for the binding of peptides to various murine MHC class I molecules. The motif for H-2K^d would predict that there should be a tyrosine/phenylalanine at position 2 and a large hydrophobic residue at position 9 (refs. 9 and 10; see updated version at www.syfpeithi.de). A search through the preproinsulin molecule had not shown any peptide to conform to this motif. It was therefore a surprise that the antigenic peptide, identified by screening a cDNA library, should be a nonamer of insulin that binds very poorly to H-2K^d. Such poor binding would therefore make crystallization of this MHC–peptide complex difficult. To further examine structure–function relationships, we have made numerous amino acid substitutions within the peptide and then examined these relationships via MHC binding, T cell clonal proliferation, and T cell cytotoxicity, together with molecular modeling. The studies reported in this manuscript show interesting features in the interaction of the peptide with the MHC class I molecule H-2K^d and presumably with the cognate T cell antigen receptor (TCR) as well.

Materials and Methods

Insulin-Reactive CD8 T Cell Clone. This cloned T cell, derived from the islets of young NOD mice and designated G9C8, recognizes the peptide insulin B15-23 complex (7, 11). The clone is maintained in Click's media supplemented with 5% FCS and 5 units/ml of IL-2 (EL4 supernatant) and grown at 37°C in 5% CO_2. The cells are fed with irradiated NOD islets every 2 weeks.

^{51}Cr Release Cytotoxicity. P815 cells (1×10^6) were incubated with 0.1 μCi ^{51}Cr-sodium chromate in 100 μl for 2 h, washed, and resuspended at 10^4 cells per 50 μl. The cells were incubated at 25°C with different peptide concentrations in round-bottomed 96-well plates for 1 h. G9C8-cloned T cells were added to the plate in the effector target ratios of 15:1 and 5:1. The cell mixture was incubated for a further 4 h at 37°C. One hundred microliters of supernatant was assayed for ^{51}Cr release in a γ-counter. Specific lysis was determined as ((cytotoxic release-minimum release)/(maximum-minimum)) $\times$ 100.

Abbreviations: NOD, nonobese diabetic; TCR, T cell antigen receptor.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1l6q).

†F.S.W. and A.K.M. contributed equally to this work.

‡To whom reprint requests should be addressed. E-mail: Susan.Wong@bristol.ac.uk.

‖G.K.P. and C.A.J. contributed equally to this work.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

IMMUNOLOGY

H-2K^d Binding Assay. RMAS-K^d cells were kindly provided by A. Bothwell (Yale School of Medicine). The method used was modified from ref. 12. RMAS cells are deficient in transporter associated with antigen processing (TAP). MHC class I is stabilized on the surface of these cells by peptide. The RMAS-K^d cells were incubated at 30°C overnight. Peptide at various concentrations was added for 1 h at 30°C, and the cells were then incubated for a further hour at 37°C. After this, the cells were stained on ice with FITC-conjugated SF1-1.1 mAb that stains H-2K^d. Binding of peptide is determined by the degree of staining with SF1-1.1 and reported as mean channel fluorescence, as measured by flow cytometry.

Proliferation Assays. Ten thousand cloned T cells were incubated with 10^5 irradiated NOD spleen cells together with different concentrations of peptide for 72 h. One hundred microliters of supernatant was removed for cytokine assays, and 0.5 μCi-tritiated thymidine was added for a further 14-h incubation. The plates were harvested by using a Skatron (Sterling, VA) 96-well plate harvester, and thymidine uptake was measured in cpm, using a Wallac (Gaithersburg, MD) β-plate counter.

IFN-γ Assay. IFN-γ was measured by ELISA, using mAbs and recommended protocols (PharMingen).

Peptides. Peptides substituted with all of the amino acids at positions 5, 6, 7, and 8 were synthesized by Chiron. Alanine-substituted peptides at each position as well as peptides substituted with all of the amino acids at position 9 were synthesized at the W. M. Keck Biotechnology Facility of Yale University School of Medicine.

Modeling of H-2K^d-Insulin Peptide Complexes. Structural homology modeling of the H-2K^d molecule with the insulin B15-23 peptide was performed as described at an ambient pH value of 7.0 (13). In brief, the sequences of the H-2K^d molecule with the insulin B15-23 peptide were at first aligned to those of the human HLA-A2–Tax peptide complex (14). This human allele with a nonameric peptide was chosen over the slightly more homologous mouse allele H-2K^b for the following reasons. The binding motifs for H-2K^d and HLA-A2 are very similar [p2Y/F, p9V/L/I, vs. p2L, p9V/L/I, respectively (10)], whereas the motif for H-2K^b is different in the number and physicochemical character of the anchoring residues, as well as their relative position [p4Y, p6Y/F, p9L/M/I/V for nonamers (10, 15–17)]. Modeling based on an H-2K^b-nonamer peptide complex (Protein Data Bank ID code 2vab) yielded a peptide orientation that could not explain the cytotoxicity data regarding p6 and p8 (data not shown). In addition, key residues of the MHC molecule (97R, 99F) that determine the geometry of the backbone are identical in HLA-A2 and H-2K^d (Table 1). Second, a modeling of the mouse β2 microglobulin based on the structure of its human counterpart afforded us an internal check on how well our modeling procedure would approximate a real structure. Representations of all or part of the H-2K^d-insulin B15-23 complex were accomplished on a personal computer with the Micron Separations WEBLABVIEWER program (version 3.5). Unless otherwise stated, all representations of single antigenic residues in particular positions depict all heavy chain residues that have at least one atom at a distance of <5 Å from any atom of the particular peptide residues. The coordinates of the modeled H2-K^d-insulin B15-23 complex have been deposited in the Protein Data Bank.

Results

Diabetogenic-Cloned CD8 T Cells Are Stimulated by Insulin B15-23. The peptide B15-23 of insulin from human, pig, rat, or mouse has the sequence L Y L V C G E R G and binds very weakly to H-2K^d as shown in the RMAS-K^d assay (Fig. 1*A*). Despite this weak binding, the peptide is able to stimulate proliferation of cloned G9C8 CD8 T cells coupled with the production of IFN-γ (Fig. 1*B*), and cytotoxicity toward targets expressing H-2K^d (Fig. 1*C*). Our original experiments using overlapping peptides had shown that peptide pools containing the 9-mer B15-23 stimulated the cloned T cells to a greater extent than the 10-mer B15–24 (LYLVCGERGF) (11). In confirmation that the 9-mer rather than the 10-mer is the cognate peptide for the T cell clone, proliferation assays showed that the cells proliferated and produced 10-fold less IFN-γ in response to the 10-mer, and only at a high concentration of peptide (5 μg/ml), compared with the 9-mer (data not shown). The cytotoxic response was also lower, and 25 times more peptide was required to produce the same response as shown in Fig. 1*C*. Longer peptides can be cleaved by carboxypeptideases to shorter ones, but to our knowledge, no one has ever observed the reverse reaction.

Single alanine substitution at each position (p) in the wild-type peptide results in loss of T cell recognition for peptides substituted at p1, p2, p3, p4, p6, and very weak stimulation at p8 (Fig. 1*D*). There is nearly equivalent recognition for alanine substitutions at p5, p7, and p9 (see below).

Table 1. Variations of key residues around pocket B in three different MHC I heavy chains

Position	H-2K^b	HLA-A2	H-2K^d
7	Y	Y	Y
9	V	F	V
24	E	A	A
26	G	G	G
36	F	F	F
45	Y	M	F
59	Y	Y	Y
62	R	G	E
63	E	E	Q
64	T	T	T
67	A	V	A
74	F	H	F
97	V	R	R
99	S	Y	F

Data derived from ref. 31.

Modeling of H-2K^d-Insulin B15-23 and Its Amino Acid-Substituted Variants. The modeled structure of H-2K^d in complex with the insulin B15-23 peptide, based on the crystal structure of HLA-A2–Tax peptide, is very similar to the crystal structures of other mouse and human MHC class I molecules with bound peptides (14–18). The mouse β2 microglobulin structure, as modeled from its human counterpart, used as a criterion for the accuracy of the modeling process, exhibited very similar structure to its published crystal structure, in all respects. The rms values of the deviation in position for the α-carbons, the backbone atoms, and the heavy atoms are 1.33, 1.38, and 2.13 Å, respectively, indicating the accuracy of the structural determination by modeling.

Our modeling indicates that the insulin peptide fits well into the groove with the p2Y providing the primary anchor, and the p9G the secondary anchor, mainly through its free carboxylate end (Fig. 2*A*, Table 1). The peptide appears to have as TCR contact residues p1L (partial), p3L, p4V, p5C, p6G, and p8R (Fig. 2*A*). We note that p7E points into the groove and is not at all accessible for TCR contact. However, the assignment of the role of each residue in TCR recognition is provided by the combination of functional and structural studies as described later in this article.

The many MHC class I structures indicate depressions and "bumps" in the antigen-binding groove, yielding a fair idea of

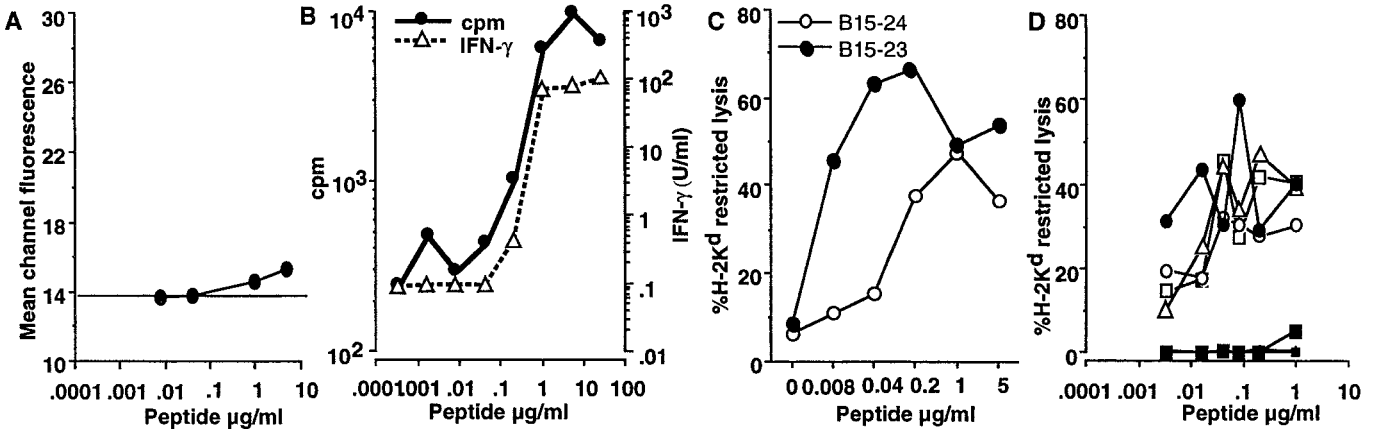

Fig. 1. H-2K^d binding of peptide is less sensitive than cytotoxicity or induction of IFN-γ secretion. (*A*) H-2K^d binding assay. Stabilization of H-2K^d by increasing concentrations of peptide is shown by the mean channel fluorescence of binding with anti-H-2K^d Ab SF1-1.1 on RMAS-K^d cells. The figure shows the mean staining by using anti-H-2K^d mAb in the presence of different concentrations of the insulin B chain peptide 15-23. The figure shows very poor binding as staining increases very little over background levels in the absence of peptide shown by the horizontal line. The scale has been chosen because it covers the range of the scale used for peptides that bind well. (*B*) Proliferation and IFN-γ production by the CD8 T cell clone G9C8 with increasing concentrations of peptide B15-23. The background level of proliferation on this occasion was 1,927 cpm and for IFN-γ production was 0.1 units/ml. (*C*) ^{51}Cr release cytotoxicity assay showing lysis of P815 targets by the CD8 T cell clone G9C8, with increasing concentrations of peptide B15-23 (closed circles) and peptide B15-24(open circles). (*D*) ^{51}Cr release cytotoxicity assay using peptides substituted at each position by alanine. Nomenclature used indicates the amino acid in the wild-type B15-23 peptide followed by the position of the amino acid, followed by the substituted amino acid, in this case, alanine (A). Wild-type peptide (closed circles), C5A (open squares), E7A (open triangles), and G9A (open circles) stimulate cytotoxicity. Substitutions L1A (*), Y2A (open diamond), L3A (©), V4A (X), G6A (+), and R8A (closed squares) abolish cytotoxicity.

which positions might also constitute secondary anchors (Fig. 2 *A* and *B*). The orientation of several heavy chain residues near the amino terminus of the peptide allow for anchoring of the charged amino-terminal group into pocket A, and for favorable interactions of many polar residues with it (Fig. 2*A*). In addition, several hydrophobic residues make a large pocket B where an aromatic residue (Y, F) fits well. Furthermore, there are four depressions as viewed from above, in the manner that a T cell receptor would approach the H-2K^d-insulin B15-23 complex. The first one is at p3 [pocket D in the nomenclature of Fremont *et al.* (15)], the second between p5 and p6 (pocket C), the third around p7 (pocket E), and the fourth around p9 (pocket F). Depending on the peptide, the orientation of certain residues at specific positions may vary (14). Here p3L fits into pocket C, and p7E into pocket E, whereas p9G barely enters the spacious pocket F (Fig. 4, which is published as supporting information on the PNAS web site, www.pnas.org). The listing of peptides that are documented to bind to H-2K^d contains several different residues at p3-p7, with very few aromatics (10). Indeed, only pocket E is large enough to accommodate such residues (see below). The backbone of the insulin peptide, as viewed from the level of the β-sheet floor, curves upward, because of the 97R-99F combination that is right in the middle of this floor, as well as the 74F/116F pair that points into the groove at the level of the p7/p8 backbone residue (Fig. 4). This is a general finding in many MHC class I alleles, including HLA-A2 and −Aw68, which have bulky and inflexible residues (R/F/Y) at these positions (14, 18–20). H-2K^b has the smaller valine and serine residues at positions 97 and 99, respectively, but maintains F/Y at 74 and 116, thus this rise in the backbone is not as steep (15–17). The peptide backbone curves into pocket F (p9), as is typical of all peptides bound to MHC class I molecules (Fig. 4). Consequently, the TCR contact residues at the carboxyl-terminal part of the peptide are p6G and p8R, situated at the same level above the β-sheet floor and next to each other (Fig. 4). This arrangement shapes the selection of cognate TCRs, as no substitutions are tolerated at these two positions, as shown by cytotoxicity assays (Fig. 1*D* and below).

Anchor Residues at p2 and p9. The fitting of the p2Y residue in pocket B is particularly tight (Fig. 2*B*). A close inspection of this arrangement indicates that two key substitutions from the HLA-A2 sequence permit an aromatic residue at this pocket: 9F → V and 67V → A. The residues at these positions make contact with the p2Y of the peptide, so that the presence of a smaller V and A makes room for the aromatic residue in pocket B. This space, however, is not sufficient to accommodate a W, which has not been observed at this position, either in H-2K^d epitope or in antigenic peptide extraction and sequencing studies (10). The proximal aromatic residues 7Y, 36F, 45F, 59Y, and 99F interact with p2Y/F via aromatic–aromatic interactions (21).

Anchoring at p9 is enhanced by the formation of hydrogen bonds by residues 143T and 146K with the carboxylate end of the peptide (Fig. 2*A*). In our model, residues 80T and 84Y are not close enough to form hydrogen bonds to this carboxylate end but might do so by means of interposed water molecules, as in the case of HLA-A2–influenza virus matrix peptide with 80T (14). In addition, the wild-type p9G is too small to fill the p9 pocket. It is thus no surprise that the binding of the wild-type insulin B chain peptide 15-23 to H-2K^d is very weak (Fig. 1*A*). By contrast, peptide variants with p9A/V/L/I show much better binding (Fig. 3*A*, and data not shown), and the modeled structures reveal substantial van der Waals interactions of the p9 side chain with residues lining the p9 pocket (Fig. 5 *A* and *B*, which is published as supporting information on the PNAS web site, and data not shown). The binding by several variants at p9 does not lead to any displacement of the peptide backbone of the H-2K^d molecule (data not shown), and therefore, the TCR recognition of the new complexes is equal to or greater than that of the wild-type peptide (Fig. 4). The variant with p9Y does not show significant binding (Fig. 3*A*), as this residue does not fit well into the p9 pocket (data not shown). However, even this weak binding is in the correct orientation for the cognate T cell to recognize the complex and effect an equivalent cytotoxic reaction toward H-2K^d target cells bearing this peptide variant (Fig. 3*B*).

We have also tested and modeled the binding and T cell proliferation of residue substitutions at p5-p8 (see supporting text and Figs. 6 and 7, which are published on the PNAS web

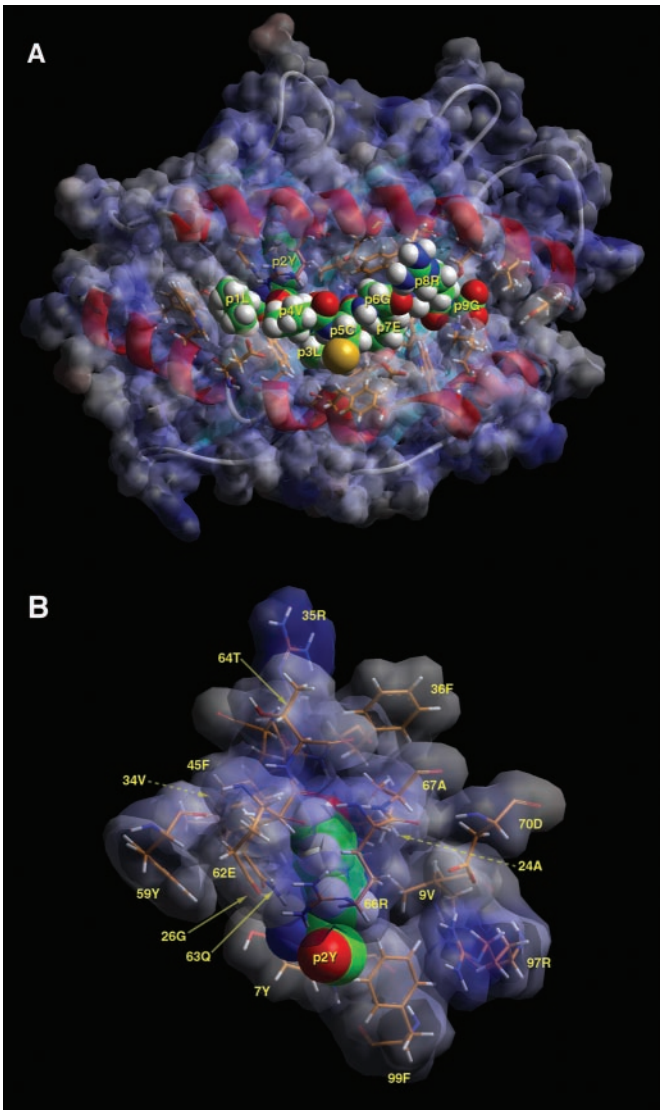

Fig. 2. The modeled structure of H-2K^d binding the insulin B chain peptide 15-23. (*A*) View of the modeled structure of the $\alpha1\alpha2$ region of the H-2K^d molecule in complex with the murine insulin B15-23 peptide, as seen from above (TCR view). Several ways of depiction are shown simultaneously to appreciate how the peptide fits into the groove. The antigenic peptide is in space-filling form with its carbon atoms shown in green, nitrogen in blue, oxygen in red, hydrogen in white, and sulfur in orange. The $\alpha1\alpha2$ domain of the molecule is depicted according to its secondary structure in different regions: α-helix in red, β-pleated sheet in turquoise, and random coil in gray. The solvent-accessible surface of the $\alpha1\alpha2$ domain is shown in gray with colorings according to the electrostatic surface potential (blue for positive, red for negative, and intermediate hues for neutral). Depressions in the surface within the antigen-binding groove (not shown) are potential pockets. The surface of the heavy chain is made transparent so that peptide residue p2Y, which is buried in pocket B (see also Fig. 4), as well as the heavy chain residues making contact with the insulin peptide can be seen, albeit in a lighter color. These residues are shown with their carbon atoms in orange. (*B*) View of pocket B (position p2) containing the insulin B16Y residue and surrounding heavy chain amino acids, in the complex of H-2K^d and insulin B15-23, as seen from above (TCR view). The p2Y residue, nearly completely buried in this pocket, is in space-filling form, whereas the residues forming the p2 pocket are in stick form with van der Waals surface representation. Unlike HLA-A2, pocket B in H-2K^d can accommodate an aromatic residue because of F9V and V67A mutations. In both cases, the smaller residues are found in H-2K^d, hence the specific difference in the p2 motif. The several aromatic residues from the heavy chain are favorably situated to interact with p2Y/F. Color, surface electrostatic, and transparency conventions are as in *A*.

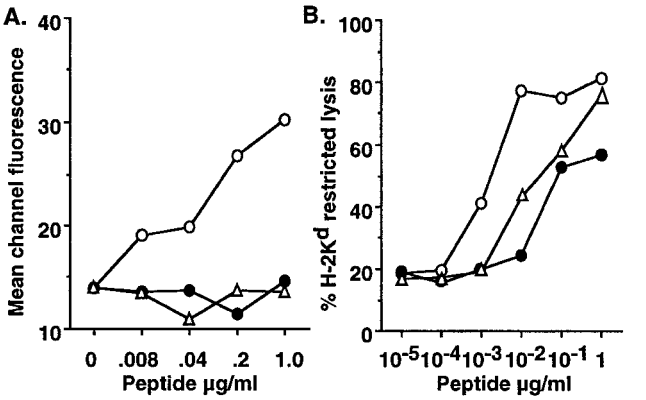

Fig. 3. Binding to H-2K^d as determined by surface stabilization of H-2K^d in RMAS-K^d cells, compared with ^{51}Cr release cytotoxicity assays. (*A*) H-2K^d binding assay for native peptide and mutants at position 9. Stabilization of H-2K^d by increasing concentrations of peptide is shown by the mean channel fluorescence of binding with anti-K^d Ab SF1-1.1. The wild-type peptide containing glycine at position 9 (G9G), shown as closed circles, binds very poorly to the MHC, and increasing concentrations of peptide do not increase the staining with the Ab. A similar result is found with the substitution using tyrosine (G9Y), shown as open triangles. By contrast, substitution of the glycine by valine (G9V), shown as open circles, improves the binding to the H-2K^d molecule. (*B*) ^{51}Cr release cytotoxicity assay for same mutants as in *A*, showing that the mutants that bind better to the H-2K^d molecule also stimulate improved cytotoxicity by the G9C8 T cell clone, the best binding found with G9V (open circles) also stimulating cytotoxicity at a lower concentration of peptide. However, even the poorly binding wild-type peptide (closed circles) or the tyrosine-substituted peptide, G9Y (open triangles), will still stimulate cytotoxicity at levels as low as 0.01 μg/ml.

site). Those at p5 can have a profound effect on the ability of the CD8 T cell clone to recognize insulin B15-23, whereas those at p7 show only small effects. By contrast, not even conservative substitutions are tolerated at p6 and p8.

Discussion

Relatively little is known of the autoantigens that CD8$^+$ T cells respond to in type 1 diabetes. In humans, cytotoxic CD8$^+$ T cells responsive to glutamic acid decarboxylase have been studied (22, 23). The insulin B15-23 peptide is the first epitope identified for an MHC class I allele in type 1 diabetes in which the source of the peptide is mainly in pancreatic β cells, and the cognate G9C8-cloned T cells reactive to the peptide have been shown to be highly pathogenic when transferred into NOD mice. Our previous studies have shown that CD8 T cells recognizing this peptide are present in the very early phases of insulitis, long before clinical diabetes becomes manifest (11). It is therefore important to understand the structure–function relationships of the participating molecules in the context of the disease. Interestingly, the B15-23 peptide binds very poorly to the MHC molecule in question, H-2K^d, and relatively high concentrations of peptide (0.1–1 μg/ml) are required to induce the cells to become cytolytic or to proliferate. Although both the B15-23 nonamer and the B15-24 decamer were able to induce cytotoxicity, it was unlikely that the decamer was the cognate peptide for the clone, because the stimulation occurred at 100-fold lower concentration with the nonamer in both cytotoxicity and production of IFN-γ. It is possible that the decamer stimulates less well because the nonamer binds in register and the presence of the F residue in the decamer interferes with TCR recognition. Indeed, such peptides that extend by one or more residues out of the carboxyl-terminal pocket of an MHC class I molecule have been identified by pool sequencing and x-ray crystallography (10, 24, 25). However, the universal mode of TCR binding to MHC I–peptide complexes would be considerably disturbed by the

presence, at the carboxyl terminus of the insulin peptide, of the bulky and inflexible phenylalanine residue protruding out of the p9 pocket. This amino acid protrusion would certainly cause rearrangement of nearby H-2K^d heavy chain residues and perhaps instability of the complex, as observed in the HLA-A2–calreticulin decamer peptide complex [added p10Gly (24)]; thus, such a complex would most probably not be recognized by the cognate TCR. Alternatively, the decamer preparations could have smaller amounts of the nonamer as a breakdown product, and this minute amount is in fact stimulating to the T cell clone. Last, but much less likely, both the nonamer and the decamer could stimulate recognition by the TCR if both p9G and p10F are able to act as anchor residues to the MHC, but it would be difficult to explain our peptide substitution data if this were the case. The two peptide–MHC I complexes herein would have distinctly different TCR recognition surfaces, making their nearly equivalent recognition by a single TCR species highly unlikely.

Homology modeling is a well tested method of approximating the three-dimensional structure of proteins, as shown in a number of retrospective studies (26, 27). In essence, the base and the modeled protein must share at least a 40% identity in amino acid residues and a similarity of biochemical function. In the case at hand, we have deliberately chosen a human MHC class I protein (HLA-A2) that shares considerable identity (74.2% in the antigen binding $\alpha 1\alpha 2$ domain) with the modeled protein, H-2K^d. In addition, the murine protein has a nearly identical peptide-binding motif and peptide length and backbone orientation as its human counterpart, HLA-A2, because identical or very similar residues in the β-sheet floor and the two α-helices determine the shape and selectivity of the antigenic peptide (14, 18–20). Of all of the different MHC class I proteins whose crystal structure has been determined, HLA-A2 is the only one that has such a close similarity to H-2K^d in the peptide binding motif and consequent identity in the residues lining the antigen-binding groove (Table 1 and ref. 28). Further discussion is found in the supporting text.

The structural modeling performed of the 34 different variants at p5, p6, p7, p8, and p9 (data shown in supporting text) provides us with a very good framework from which to explain the binding data obtained with nearly all of these variants. The residues on the two α-helices and the β-sheet floor are so versatile that they can reorient themselves to accommodate most substitutions. However, T cell recognition is quite sensitive to most substitutions including a number that is conservative or semiconservative. For example, no replacement of p6G is tolerated by the cognate TCR, whereas several residues (Q, S, V, I, L, and A) can replace the glycine at the p9 (pocket F) anchor. At p7 all variants bind to H-2K^d better than the wild-type peptide, yet no cytotoxic effect is evident for several of these, apparently because of the widely accommodating pocket E on the one hand and the retention of a similar conformation of the resultant H-2K^d–peptide complexes only in the few variants that stimulate cytotoxicity. The latter include conservative D and nonconservative Q, M, G, and A and can substitute for p7E that points downwards into the groove (pocket E). These substitutions can induce cytotoxicity at lower peptide doses than that of the wild-type peptide. The conformation of the region around residues p6–p8 is crucial, as not even conservative substitutions are tolerated at p6 or p8.

The poor binding of the peptide to the MHC has important consequences for the selection and peripheral activation of T cells responsive to this peptide. In both humans and mice, it has been shown that there are cells that express autoantigens in the thymus (29–33), although it is likely that the levels of insulin there will be low. It is possible that the poor binding may not allow expression of sufficient H-2K^d–insulin B chain 15-23 peptide complexes to drive negative selection, allowing escape of these potentially autoreactive T cells, and thus predisposing to autoimmunity. However, in the endocrine pancreas, much higher levels of insulin B15-23 peptide would presumably be available to stimulate these autoreactive T cells. A transgenic mouse that expressed proinsulin in the thymus was shown to develop tolerance and did not develop diabetes (34). Although CD4 T cells reactive to proinsulin were suggested to be affected in this model, it is equally possible that proinsulin-reactive CD8 T cells may also have been deleted. Although not shown in the NOD mouse, in humans, *IDDM2* is the second gene locus found to have major effects on susceptibility to disease and is located in the insulin gene (*INS*) Variable Number of Tandem Repeats (*VNTR*) regulatory region (35, 36). There are polymorphisms at this locus and, in Caucasoid populations, class I INS VNTR alleles encoding shorter repeat sequences predispose to diabetes. In these individuals, lower levels of INS mRNA and protein are expressed in the thymus and higher levels are found in the pancreas. In contrast, individuals with class III INS VNTR alleles encoding many more repeat sequences, and who have higher levels of INS mRNA and protein in the thymus and lower levels in the pancreas, are dominantly protected (29, 33). It was suggested that the higher levels of INS mRNA may indicate greater levels of expression and enhance tolerance to preproinsulin in the thymus (29, 33). Our finding suggests that thymocytes recognizing a poorly binding peptide may escape from negative selection and cause disease. This escape from the thymus could indicate that they can be activated peripherally in individuals who develop diabetes.

In the NOD mouse, there is evidence that CD8 T cells are involved both in the initiation (37–40) as well as in the final effector phases of disease (41). Less is known about the early stages of the human disease. The association of disease with the MHC class I alleles is less striking than that with MHC class II alleles in humans. However, there is evidence that the MHC class I region is important in determining susceptibility to disease, and that this is not simply because of the linkage disequilibrium with MHC class II molecules (42, 43). In studies where patients with newly diagnosed diabetes have undergone pancreatic biopsy or in postmortem when patients have died within a year after onset of disease, CD8 T cells were seen to be predominant in islets (44–46). The early attempts at hemipancreatic isografting without immunosuppression between identical twins, in which the diabetic twin was transplanted with one half of a pancreas from the nondiabetic identical twin, showed that disease recurrence was very rapid, and that within 6 weeks, abundant CD8 T cells were present in the transplanted pancreas in 3 of 4 cases (47). Hence, CD8 T cells and their interaction with MHC class I molecule–peptide complexes are likely to play an important role in the human disease. Recent studies in the NOD mouse have indicated that it is of great importance to take into account CD8 T cell priming in the use of oral or nasal immunotherapy (48) and failure to do this may lead to unexpected effects. It is clear that further study of this important interaction will be crucial to development of future preventive immunotherapy for diabetes.

We thank Demetrios Kyrkas and Liliana Marquesini for technical assistance, Jennifer Boucher-Reid for secretarial help, and Irene Visintin for assistance with the manuscript. This work was funded by National Institutes of Health/Juvenile Diabetes Research Foundation (JDRF) Grant PO1-DK53015. The work of A.K.M. and G.K.P. was supported in part by a grant from the Technological Educational Institute of Epirus Research Committee. F.S.W. previously held a Career Development Award from the JDRF and is currently a Wellcome Trust Senior Fellow in Clinical Science. L.W. has a Career Development Award from the JDRF. C.A.J. is an Investigator of the Howard Hughes Medical Institute.

1. Daniel, D. & Wegmann, D. R. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 956–960.
2. Kaufman, D. L., Clare-Salzler, M., Tian, J., Forsthuber, T., Ting, G. S., Robinson, P., Atkinson, M. A., Sercarz, E. E., Tobin, A. J. & Lehmann, P. V. (1993) *Nature (London)* **366,** 69–72.
3. Tisch, R., Yang, X. D., Singer, S. M., Liblau, R. S., Fugger, L. & McDevitt, H. O. (1993) *Nature (London)* **366,** 72–75.
4. Zekzer, D., Wong, F. S., Ayalon, O., Millet, I., Altieri, M., Shintani, S., Solimena, M. & Sherwin, R. S. (1998) *J. Clin. Invest.* **101,** 68–73.
5. Corper, A. L., Stratmann, T., Apostolopoulos, V., Scott, C. A., Garcia, K. C., Kang, A. S., Wilson, I. A. & Teyton, L. (2000) *Science* **288,** 505–511.
6. Latek, R. R., Suri, A., Petzold, S. J., Nelson, C. A., Kanagawa, O., Unanue, E. R. & Fremont, D. H. (2000) *Immunity* **12,** 699–710.
7. Wong, F. S., Visintin, I., Wen, L., Flavell, R. A. & Janeway, C. J. (1996) *J. Exp. Med.* **183,** 67–76.
8. Alleva, D. G., Crowe, P. D., Jin, L., Kwok, W. W., Ling, N., Gottschalk, M., Conlon, P. J., Gottlieb, P. A., Putnam, A. L. & Gaur, A. (2001) *J. Clin. Invest.* **107,** 173–180.
9. Rammensee, H.-G., Falk, K. & Rotzschke, O. (1993) *Annu. Rev. Immunol.* **11,** 213–244.
10. Rammensee, H.-G., Friede, T. & Stevanoviic, S. (1995) *Immunogenetics* **41,** 178–228.
11. Wong, F. S., Karttunen, J., Dumont, C., Wen, L., Visintin, I., Pilip, I. M., Shastri, N., Pamer, E. G. & Janeway, C. A. (1999) *Nat. Med.* **5,** 1026–1031.
12. Deng, Y., Yewdell, J. W., Eisenlohr, L. C. & Bennink, J. R. (1997) *J. Immunol.* **158,** 1507–1515.
13. Moustakas, A. K., Routsias, J. & Papadopoulos, G. K. (2000) *Diabetologia* **43,** 609–624.
14. Madden, D. R., Garboczi, D. N. & Wiley, D. C. (1993) *Cell* **75,** 693–708.
15. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257,** 919–927.
16. Fremont, D. H., Stura, E. A., Matsumura, M., Peterson, P. A. & Wilson, I. A. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 2479–2483.
17. Matsumura, M., Fremont, D. H., Peterson, P. A. & Wilson, I. A. (1992) *Science* **257,** 927–934.
18. Bouvier, M. & Wiley, D. C. (1994) *Science* **265,** 398–402.
19. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. (1992) *Cell* **70,** 1035–1048.
20. Silver, M. L., Guo, H. C., Strominger, J. L. & Wiley, D. C. (1992) *Nature (London)* **360,** 367–369.
21. Burley, S. K. & Petsko, G. A. (1985) *Science* **229,** 23–28.
22. Ou, D., Jonsen, L. A., Metzger, D. L. & Tingle, A. J. (1999) *Hum. Immunol.* **60,** 652–664.
23. Panina-Bordignon, P., Lang, R., van Endert, P. M., Benazzi, E., Felix, A. M., Pastore, R. M., Spinas, G. A. & Sinigaglia, F. (1995) *J. Exp. Med.* **181,** 1923–1927.
24. Collins, E. J., Garboczi, D. N. & Wiley, D. C. (1994) *Nature (London)* **371,** 626–629.
25. Rotzschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, G. & Rammensee, H. G. (1990) *Nature (London)* **348,** 252–254.
26. Johnson, M. S., Srinivasan, N., Sowdhamini, R. & Blundell, T. L. (1994) *Crit. Rev. Biochem. Mol. Biol.* **29,** 1–68.
27. Martin, A. C., MacArthur, M. W. & Thornton, J. M. (1997) *Proteins*, Suppl. 1, 14–28.
28. Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. (1987) *Sequences of Proteins of Immunological Interest* (U.S. Dept. Health Hum. Serv., Bethesda).
29. Pugliese, A., Zeller, M., Fernandez, A., Jr., Zalcberg, L. J., Bartlett, R. J., Ricordi, C., Pietropaolo, M., Eisenbarth, G. S., Bennett, S. T. & Patel, D. D. (1997) *Nat. Genet.* **15,** 293–297.
30. Pugliese, A., Brown, D., Garza, D., Murchison, D., Zeller, M., Redondo, M., Diez, J., Eisenbarth, G. S., Patel, D. D. & Ricordi, C. (2001) *J. Clin. Invest.* **107,** 555–564.
31. Smith, K. M., Olson, D. C., Hirose, R. & Hanahan, D. (1997) *Int. Immunol.* **9,** 1355–1365.
32. Sospedra, M., Ferrer-Francesch, X., Dominguez, O., Juan, M., Foz-Sala, M. & Pujol-Borrell, R. (1998) *J. Immunol.* **161,** 5918–5929.
33. Vafiadis, P., Bennett, S. T., Todd, J. A., Nadeau, J., Grabs, R., Goodyer, C. G., Wickramasinghe, S., Colle, E. & Polychronakos, C. (1997) *Nat. Genet.* **15,** 289–292.
34. French, M. B., Allison, J., Cram, D. S., Thomas, H. E., Dempsey-Collier, M., Silva, A., Georgiou, H. M., Kay, T. W., Harrison, L. C. & Lew, A. M. (1997) *Diabetes* **46,** 34–39.
35. Bennett, S. T., Lucassen, A. M., Gough, S. C., Powell, E. E., Undlien, D. E., Pritchard, L. E., Merriman, M. E., Kawaguchi, Y., Dronsfield, M. J., Pociot, F., *et al.* (1995) *Nat. Genet.* **9,** 284–292.
36. Kennedy, G. C., German, M. S. & Rutter, W. J. (1995) *Nat. Genet.* **9,** 293–298.
37. Serreze, D. V., Leiter, E. H., Christianson, G. J., Greiner, D. & Roopenian, D. C. (1994) *Diabetes* **43,** 505–509.
38. Wicker, L. S., Leiter, E. H., Todd, J. A., Renjilian, R. J., Peterson, E., Fischer, P. A., Podolin, P. L., Zijlstra, M., Jaenisch, R. & Peterson, L. B. (1994) *Diabetes* **43,** 500–504.
39. Serreze, D. V., Chapman, H. D., Varnum, D. S., Gerling, I., Leiter, E. H. & Shultz, L. D. (1997) *J. Immunol.* **158,** 3978–3986.
40. Wang, B., Gonzalez, A., Benoist, C. & Mathis, D. (1996) *Eur. J. Immunol.* **26,** 1762–1769.
41. Kay, T. W., Parker, J. L., Stephens, L. A., Thomas, H. E. & Allison, J. (1996) *J. Immunol.* **157,** 3688–3693.
42. Fennessy, M., Metcalfe, K., Hitman, G. A., Niven, M., Biro, P. A., Tuomilehto, J. & Tuomilehto-Wolf, E. (1994) *Diabetologia* **37,** 937–944.
43. Langholz, B., Tuomilehto-Wolf, E., Thomas, D., Pitkaniemi, J. & Tuomilehto, J. (1995) *Genet. Epidemiol.* **12,** 441–453.
44. Itoh, N., Hanafusa, T., Miyazaki, A., Miyagawa, J., Yamagata, K., Yamamoto, K., Waguri, M., Imagawa, A., Tamura, S., Inada, M., *et al.* (1993) *J. Clin. Invest.* **92,** 2313–2322.
45. Moriwaki, M., Itoh, N., Miyagawa, J., Yamamoto, K., Imagawa, A., Yamagata, K., Iwahashi, H., Nakajima, H., Namba, M., Nagata, S., *et al.* (1999) *Diabetologia* **42,** 1332–1340.
46. Foulis, A. K., Liddle, C. N., Farquharson, M. A., Richmond, J. A. & Weir, R. S. (1986) *Diabetologia* **29,** 267–274.
47. Sibley, R. K., Sutherland, D. E., Goetz, F. & Michael, A. F. (1985) *Lab. Invest.* **53,** 132–144.
48. Hanninen, A., Braakhuis, A., Heath, W. R. & Harrison, L. C. (2001) *Diabetes* **50,** 771–775.