Systematic Analysis of the Combinatorial Nature of Epitopes Recognized by TCR Leads to Identification of Mimicry Epitopes for Glutamic Acid Decarboxylase 65-Specific TCRs[1]

Yasushi Uemura,* Satoru Senju,* Katsumi Maenaka,[§¶] Leo Kei Iwai,*[‖] Shinji Fujii,*[†] Hiroki Tabata,*[‡] Hirotake Tsukamoto,* Shinya Hirata,* Yu-Zhen Chen,* and Yasuharu Nishimura[2]*

Accumulating evidence indicates that recognition by TCRs is far more degenerate than formerly presumed. Cross-recognition of microbial Ags by autoreactive T cells is implicated in the development of autoimmunity, and elucidating the recognition nature of TCRs has great significance for revelation of the disease process. A major drawback of currently used means, including positional scanning synthetic combinatorial peptide libraries, to analyze diversity of epitopes recognized by certain TCRs is that the systematic detection of cross-recognized epitopes considering the combinatorial effect of amino acids within the epitope is difficult. We devised a novel method to resolve this issue and used it to analyze cross-recognition profiles of two glutamic acid decarboxylase 65-autoreactive CD4$^+$ T cell clones, established from type I diabetes patients. We generated a DNA-based randomized epitope library based on the original glutamic acid decarboxylase epitope using class II-associated invariant chain peptide-substituted invariant chains. The epitope library was composed of seven sublibraries, in which three successive residues within the epitope were randomized simultaneously. Analysis of agonistic epitopes indicates that recognition by both TCRs was significantly affected by combinations of amino acids in the antigenic peptide, although the degree of combinatorial effect differed between the two TCRs. Protein database searching based on the TCR recognition profile proved successful in identifying several microbial and self-protein-derived mimicry epitopes. Some of the identified mimicry epitopes were actually produced from recombinant microbial proteins by APCs to stimulate T cell clones. Our data demonstrate the importance of the combinatorial nature of amino acid residues of epitopes in molecular mimicry. ***The Journal of Immunology,*** **2003, 170: 947–960.**

Activation of autoreactive CD4$^+$ T cells is a crucial step in the development of T cell-mediated autoimmunity, and cross-reactivity between microbial and self Ags, a phenomenon known as molecular mimicry, is one of the mechanisms that account for the link between infection and autoimmunity (1–3). Thereby, identification of microbial Ags mimicking self Ags would provide insights into the disease process and new therapeutic or preventative strategies for autoimmune diseases. In the last decade, many studies have demonstrated degeneracy in Ag recognition by TCR, thus showing its high flexibility (2, 4–8). Degeneracy of TCR recognition can appear itself in various functional outcomes, depending on the affinity of the MHC/peptide ligand to TCR, which includes T cell responses ranging from full activation to strong antagonism (9–11).

In functional analysis using altered peptide ligands, it was reported that TCRs can discriminate not only between two peptides differing at a TCR contact residue but also between those differing at only a single MHC anchor residue, which does not significantly alter the binding affinity of a peptide to MHC (12, 13). In addition, it has been reported that certain TCR recognition is affected by amino acids adjacent or not adjacent to TCR contact residues or by amino acid combinations in antigenic peptides (14–17). These observations indicate that a substitution at a certain residue would induce conformational changes of peptides and may affect other residues. In contrast, a quantitative strategy using combinatorial peptide libraries with a positional scanning format (PS-SCLs)[3] and biometric score matrices dissected and predicted peptide mimicry ligands of a given cognate T cells (18, 19). However, the precise effect of successive combinations of residues in the antigenic peptide on recognition of certain TCR has heretofore not been clarified. In addition, there is no available technology that allows for systematic separation and identification of diverse T cell epitopes from a mixture of randomized peptide ligands.

Several groups, including ours, reported invariant chain (Ii)-based epitope-presenting vectors, in which class II-associated invariant chain peptide (CLIP, Ii_{89-101}) was replaced by antigenic

*Division of Immunogenetics, Department of Neuroscience and Immunology, Kumamoto University Graduate School of Medical Sciences and [†]First Department of Internal Medicine and [‡]Department of Pediatrics, Kumamoto University School of Medicine, Kumamoto, Japan; [§]Structural Biology Center, National Institute of Genetics, Mishima, Shizuoka, Japan; [¶]Biosystem Science, School of Advanced Science, Graduate University for Advanced Studies, Hayama, Kanagawa, Japan; and [‖]Laboratory of Immunology, Heart Institute (Incor), University of Sao Paulo Medical School, Sao Paulo, Brazil

Received for publication June 20, 2002. Accepted for publication November 13, 2002.

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

[1] This work was supported in part by Grants-in-Aid 11694294, 11557027, and 14370115 from the Ministry of Education, Science, Technology, Sports, and Culture, Japan, and a Research Grant for Intractable Diseases from Ministry of Health and Welfare, Japan. K.M. was supported in part by the year; 2000 Joint Research Project (Soken/K01-4) of Sokendai (The Graduate University for Advanced Studies) and a research grant from the Nakajima Foundation.

[2] Address correspondence and reprint requests to Dr. Yasuharu Nishimura, Division of Immunogenetics, Department of Neuroscience and Immunology, Kumamoto University Graduate School of Medical Sciences, Honjo 2-2-1, Kumamoto 860-0811, Japan. E-mail address: mxnishim@gpo.kumamoto-u.ac.jp

[3] Abbreviations used in this paper: PS-SCLs, positional scanning synthetic combinatorial peptide libraries; CLIP, class II-associated invariant chain peptide; GAD, glutamic acid decarboxylase; Ii, invariant chain; DC, dendritic cell; CDR, complementarity-determining region.

0022-1767/03/$02.00

peptides (20–23). Using this vector system, we reported a method to identify epitopes cross-recognized by autoreactive T cell clones from a library of randomized peptides (24).

In the current study, we used this vector to make libraries of the glutamic acid decarboxylase 65 epitope ($GAD65_{115–127}$)-based degenerate peptides, where three successive residues within the epitope were randomized. GAD65 is one of the important islet Ags implicated in autoimmunity of the NOD mouse and Type I diabetes in humans. We used two T cell clones established from Japanese patients with Type I diabetes and restricted by disease susceptible HLA-DR53 and they responded to GAD65 protein (25). The epitope ($GAD65_{115–127}$) used in this study was also reported to be immunodominant in studies using HLA-DR4 transgenic mice (26, 27). Epitopes stimulating GAD65-reactive $CD4^+$ T cell clones were isolated from the series of epitope libraries. Recognition properties of these TCRs were intensively analyzed, and combinatorial effects of amino acid residues within antigenic peptide on recognition by TCRs were investigated. The information obtained by screening of this epitope expression library included the combinations of amino acid residues with TCR agonism that could not be predicted using panels of single amino acid substituted peptide analogs. Data acquisition of stimulatory TCR ligands combined with a pattern match search allowed for identification of self or microbe-derived peptides cross-recognized by $CD4^+$ T cell clones autoreactive to GAD65.

Materials and Methods

T cell clones and T cell proliferation assay

Dodecamer peptide analogs with single-amino acid substitutions derived from $GAD65_{116–127}$ were purchased from Chiron Mimotopes (Clayton South, Victoria, Australia), and 13-mer peptides were synthesized as described (28). Two human $CD4^+$ T cell clones, SA32.5 and MK20.2, recognizing $GAD65_{116–127}$ (NILLQYVVKSFD) in the context of HLA-DR53 molecules (DRA*0101 + DRB4*0103) susceptible to type I diabetes were used throughout the study (25). T cells were fed weekly with 50 U/ml human rIL-2 and the irradiated DR53-matched allogenic PBMCs prepulsed with the $GAD65_{111–131}$ (LQDVMNILLQYVVKSFDRSTK) in RPMI 1640 supplemented with 10% heat-inactivated human plasma, 100 U/ml penicillin, 100 μg/ml streptomycin, and 2 mM L-glutamine. T cell proliferation assays were set up in 96-well flat-bottom culture plates (Falcon; BD Biosciences, San Jose, CA) with 3×10^4 T cells and irradiated (6000 cGy) DR53-positive 1.5×10^5 PBMCs or 1.0×10^4 dendritic cells (DCs) per well in the presence of peptides or recombinant proteins at various concentrations. After 48 h of culture, [^{3}H]thymidine was added (1 μCi/well); and after an additional 16 h, cells were harvested onto glass fiber filters, and radioactivity was counted on a beta scintillation counter (Wallac, Gaithersburg, MD).

In vitro generation of DCs

DCs were generated from $CD14^+$ monocytes purified by positive immunoselection from HLA-DR53-positive allogeneic PBMCs, using an anti-CD14 mAb coupled onto magnetic microbeads (CD14 microbeads; Miltenyi Biotec, Auburn, CA). The $CD14^+$ monocytes were cultured at 1×10^6 cells/ml in the presence of 100 ng/ml GM-CSF and 100 U/ml IL-4 (Ono Pharmaceutical, Osaka, Japan) in RPMI 1640 supplemented with 10% human plasma, 2 mM L-glutamine, and 100 μg/ml streptomycin. Cultures were fed on days 3 and 5 with fresh medium containing GM-CSF and IL-4. On day 5, DCs were treated with TNF-α (20 ng/ml). On day 7, the nonadherent cells were harvested and served as mature DCs. For the assessment of HLA class II-restricted presentation of recombinant microbial proteins, mature DCs were cultured with recombinant proteins for 14 h before use as APCs for proliferation assay.

Epitope expression library

The procedure for construction of an epitope-presenting library is described in our previous report (24). Briefly, oligonucleotide fragments encoding degenerate $GAD65_{115–127}$ were synthesized and purified using polyacrylamide gel (Genemed Synthesis, South San Francisco, CA). These oligonucleotide fragments were amplified by PCR with 5′-biotinylated primers, 5′-TCC CTC CTG GTG ACT CTG CTC CTC-3′ and 5′-ATT GTT ATC TGC TGT TCC GAC TTG-3′. The purified PCR products were digested with *Dra*I and *Sac*I, purified with streptavidin-agarose, and ligated to *Sma*I-*Sac*I-digested pCI, the CLIP-substituted epitope expression vector. The construct encodes Ii protein inserted with partially degenerate fusion peptides based on $GAD65_{115–127}$ (MNILLQYVVKSFD) instead of $Ii_{89–101}$ (SKMRMATPLLMQA) within the CLIP sequence. *Escherichia coli* (DH5α) was transformed with the ligation mixture, and the transformants were divided into 96-well culture blocks (Qiagen, Studio City, CA) to generate transformant pools of 30–50 clones and grown overnight in Luria-Bertani medium containing ampicillin (100 μg/ml). The amplified plasmid DNA was purified using a QIAprep 96 Turbo Miniprep system (Qiagen). The complexity of each degenerate-$GAD65_{115–127}$ expression sublibraries was $\sim 1.5 \times 10^4$–2.0×10^4.

Screening of epitope expression library by detecting IFN-γ production

Library DNA pools and expression vectors for HLA-DRA*0101 and DRB4*0103 were mixed with Transfectam reagent (Promega, Madison, WI) in serum-free DMEM. The DNA-Transfectam mixtures were then added to the COS-7 cells (1×10^4 cells/well) in 96-well flat-bottom culture plates and incubated for 90 min at 37°C. After removal of the transfection medium, the COS-7 cells were incubated overnight in DMEM supplemented with 10% FCS. After 24 h, cells were washed twice with DMEM, and T cells were added at a concentration of 3×10^4/well in RPMI 1640 supplemented with 10% heat-inactivated human plasma, 100 U/ml penicillin, 100 μg/ml streptomycin, and 2 mM L-glutamine. After 48 h of incubation, the supernatant was collected, and IFN-γ concentration was measured using a standard ELISA (Endogen, Woburn, MA). The library of DNA pools, for which a significant production of IFN-γ was detected was used to transform the bacteria to prepare a sublibrary of DNA pools consisting of $\sim$10 clones. Secondary screening was done as described above, using sublibraries. Single plasmid clones were obtained after three rounds of screening. DNA sequences of the purified plasmid clones were analyzed using BigDye Terminator Cycle Sequencing Ready Reaction Kits and the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). For construction of expression vectors encoding mimicry peptides or analogs of GAD65 epitope, both strands of oligonucleotide fragments encoding these peptides were synthesized (Espec Oligo Service, Tsukuba, Japan), annealed, and ligated to *Sma*I-*Sac*I-digested pCI. The amplified plasmid DNAs were purified for transfection. The agonistic activity was assessed in an IFN-γ secretion assay at various dilutions with pCI (wild Ii: irrelevant DNA).

Analysis of TCR V-(D)-J junctional regions of GAD65-autoreactive T cell clones

Total RNA was extracted from T cells using the TRIzol reagent (Life Technologies, Gaithersburg, MD), and first-strand cDNA was synthesized using Superscript RNase H^- reverse transcriptase (Life Technologies) and random hexamers. cDNA was subjected to PCR amplification for rearranged TCR-α with 29 5′-TCRAV family-specific oligonucleotides (Vα1–32) and a 3′-TCRAC (Cα) constant primer, and for rearranged TCR-β with 27 5′-TCRBV family-specific oligonucleotides (Vβ1–25) and a 3′-TCRBC (Cβ) constant primer (29). The amplified PCR products of the α-chain and of the β-chain were cloned into a plasmid vector, pGEM-T (Promega), and sequenced. The resulting sequences were analyzed using IMGT, the international ImMunoGeneTics database (http://imgt.cnusc.fr: 8104/).

Northern blot analysis

Northern blot analysis was conducted as described (30). To prepare the probes, the MK20.2 cDNA was subjected to PCR amplification for 5′-TCRVα15- or 5′-TCRVα16-specific oligonucleotide and 3′-TCRJα region oligonucleotide primers, and each PCR product was TA cloned, digested, and gel purified. The two cDNA fragments of the *TCRAV* gene were labeled with [α-^{32}P]dCTP. After hybridization and quantitative analysis of signal intensities, probes were stripped, and a second hybridization was conducted using an β-actin probe. To assess cross-reactivity of the Vα-specific probes, TCRAV cDNA fragments were arrayed onto two copies of nylon membrane filters and hybridized using TCRAV cDNA probes.

Generation of recombinant proteins

Genomic DNA of *Legionella pneumophila* (strain Philadelphia-1) was kindly provided by Drs. T. Akaike and T. Akuta (Kumamoto University School of Medicine, Kumamoto, Japan). *Lactococcus lactis* (subsp. *lactis*) was provided by the Institute of Physical and Chemical Research (Wako, Japan). Bacterial genomic DNA were purified using DNeasy Tissue Kits (Qiagen). The *O*-succinylbenzoic acid-CoA ligase gene (coding for aa 101–201)/*L. lactis*, putative PTS system, lactose-specific component IIBC

gene (coding for aa 426–546)/*Streptococcus pyogenes* (ATCC 19615), putative dihydrolipoamide dehydrogenase gene (coding for aa 101–205)/*Neisseria meningitidis*, glutamine amidotransferase, class I gene (coding for aa 1–104)/*Streptococcus pneumoniae* (ATCC 49619), and the pilus assembly protein *PilB* gene (coding for aa 28–131)/*Legionella pneumophila* were PCR amplified and cloned into the plasmid vector (pGEM-T-easy vector system). Fusion proteins containing relatively small fragment (100–120 aa) of microbial proteins were generated, because larger recombinant proteins tend to become insoluble in bacteria and are difficult to purify. The inserted fragments were digested and ligated directionally into the prokaryotic expression vector pGEX-4T (Pharmacia, Peapack, NJ) to produce GST-fusion protein. The integrity of the constructs was confirmed by DNA sequencing. The procedure for protein induction and purification were described in our previous report (31). The purity and integrity of the fusion protein were confirmed by SDS-PAGE. The recombinant proteins were concentrated and separated from small peptide fragments with Centricon-30 (Amicon, MA), and the buffer was replaced with culture medium.

Results

A novel strategy to analyze combinatorial effects of amino acids on the antigenic peptide in exhibition of molecular mimicry

The goal of this study was to develop new and comprehensive methods for analysis of the combinatorial effects of residues in antigenic peptides on recognition by HLA class II-restricted TCRs. We also tried to clarify the significance of the effect of amino acid combination within antigenic peptide in recognition by TCR. By adopting the strategy, we sought to identify candidate mimicry epitopes for GAD65-autoreactive T cell clones established from type I diabetes patients.

Specificity of GAD65-reactive T cell clones SA32.5 and MK20.2 analyzed using single-amino acid residue-substituted peptide analogs

To investigate structural features of peptides cross-recognized by autoantigen-specific TCRs, we analyzed HLA-DR53-restricted SA32.5 and MK20.2 $CD4^+$ T cell clones reactive to GAD65 (peptide 116–127). These T cell clones were established from two independent type I diabetes patients in our previous study (25). At the beginning of this study, we verified that the two T cell clones expressed single TCRs, because it has been reported that a significant fraction of T cells in human peripheral blood expresses dual TCRs (32). The complementarity-determining region (CDR) sequences of the Ag contact sites were defined. As shown in Table I, TCR-α and TCR-β-chains of SA32.5 T cell clone revealed functional *TCRAV1S2* and *TCRBV9S1* gene rearrangements as well as an out of frame rearrangement of *AV27S1*. TCR α- and TCR β-chains of MK20.2 TCR revealed an in frame dual Vα rearrangement (*TCRAV15S1* and *TCRAV16S1* transcripts) and TCRBV3S1 rearrangement. In Northern blot analysis (Fig. 1*A*), T cell clone MK20.2 expressed the *TCRAV16S1* but not the *TCRAV15S1* gene transcript. Therefore, *TCRAV16S1* combined with the *TCRBV3S1* gene product mediates Ag recognition in MK20.2 TCR. These data indicated that reactivities of these T cell clones were determined by each single TCR. These T cell clones use distinct CDR3α and CDR3β sequences at the amino acid level and a distinct CDR3 length, whereas both T cell clones use the same Jβ rearrangement.

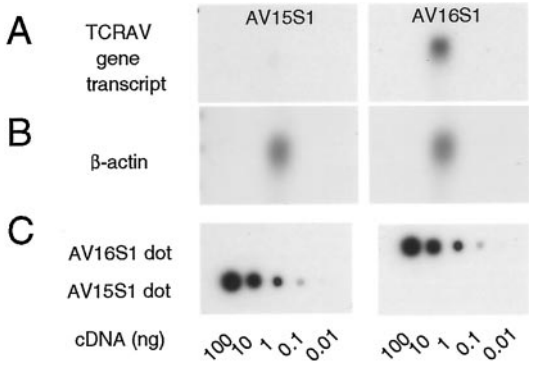


FIGURE 1. Identification of the TCR α-chain expressed in T cell clone MK20.2 by Northern blot analysis. *A*, Probes of TCRAV15S1 (*left lane*) and TCRAV16S1 (*right lane*) were hybridized in parallel with replica filters. *B*, Filters were probed with β-actin probe for control of amounts of RNA. *C*, Cross-hybridization analysis of TCRAV16S1 and TCRAV15S1 cDNA probes. To assess the cross-hybridization of each probe, cDNA dot hybridization was done using replica filters. Replicate dot blots on which the indicated amounts of TCRAV16S1 (*upper dots*) and TCRAV15S1 (*lower dots*) were manually spotted and hybridized with cDNA probes of TCRAV15S1 (*left*) or TCRAV16S1 (*right*). Cross-hybridization of each probe was not observed.

Next we analyzed these T cell clones to understand the scheme of Ag recognition properties, using conventional means, in which we examined proliferative responses of these T cell clones to the 35 peptide analogs carrying chemically conservative or nonconservative single-amino acid substitution in the native 12-mer epitope ($GAD65_{116-127}$). As shown in Fig. 2, several amino acids with different chemical properties (lysine, aspartic acid, asparagine, serine, and valine) were tested for position 117 (isoleucine in the native ligand). MK20.2 tolerated all of these substitutions. In contrast, SA32.5 did not tolerate the negatively charged aspartic acid. In the analysis using the same substitutions as position 117, these clones showed distinct specificity at position 118 (leucine in the native ligand) and at position 119 (leucine in the native ligand). With substitution at position 123 (valine in the native ligand), MK20.2 responded to hydrophobic residues (leucine, isoleucine, alanine, and methionine), in several orders of magnitude. In contrast, SA32.5 showed a significant response to only isoleucine among the tested substitutions at this position. SA32.5 tolerated

Table I. *TCR gene usage and TCR V-(D)-J junctional region sequences of α- and β-chains expressed in $GAD65_{116-127}$-autoreactive T cell clones SA32.5 and MK20.2*[a]

	TCRV	*FW*	*CDR3*	*FW*	*TCRJ*	CDR3 Length
SA32.5	***TCRAV1S2***	**CAV**	**SGQGAQKL**	**VFG**	***AJ 54*01***	8
	TCRAV27S1 (out of frame)	CAV	DSRVRNWSQZZADIWKRNNSECZT			
MK20.2	*TCRAV15S1*	CAD	SLLSPNSGSARQL	TFG	*AJ 22*01*	13
	TCRAV16S1	**CAA**	**WNNFNKF**	**YFG**	***AJ 21*01***	7
SA32.5 (*TCRBC2*)	***TCRBV9S1***	**CAS**	**SPTGQGAHTGEL**	**FFG**	***BJ 2-2*01***	12
MK20.2 (*TCRBC2*)	***TCRBV3S1***	CAS	SSTGVSPGEL	FFG	***BJ 2-2*01***	10

[a] Functional TCRs for each T cell clone are represented in boldface. The deduced amino acid sequence of the CDR3 loop is shown putatively supported by two framework branches (*FW*). Amino acids preceding "CA" of V region and those following the highly conserved "FG" of J region are not shown. Two T cell clones used distinct CDR3α and CDR3β sequences at the amino acid level and distinct CDR3 length. Both T cell clones used the same Jβ rearrangement.

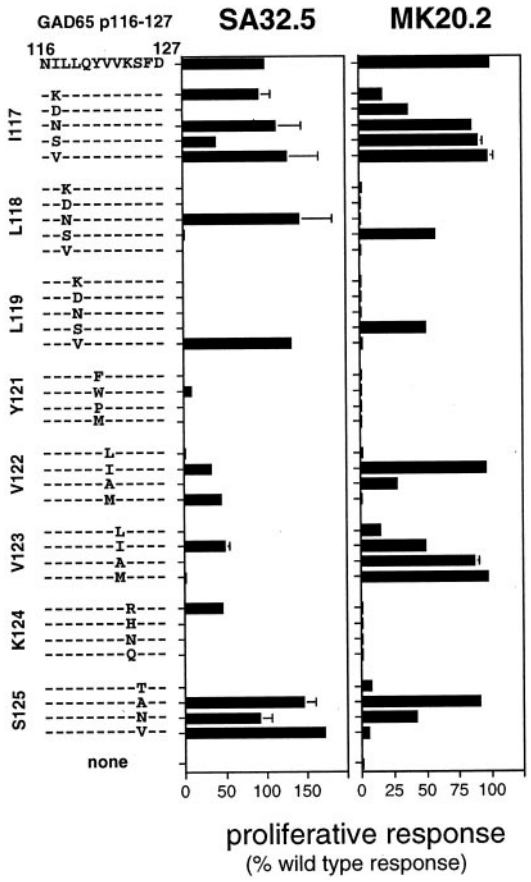


FIGURE 2. Proliferative response of $GAD65_{116-127}$-autoreactive human T cell clones to peptide analogs carrying a single-amino acid substitution in the $GAD65_{116-127}$ peptide. Data are given as percentages of wild-type response which were standardized by calculating the percentage to the response to wild-type peptide. The T cell responses to wild-type $GAD65_{116-127}$ were 27,021 cpm for SA32.5 and 23,146 cpm for MK20.2. Medium control response without peptide was <200 cpm. All data are expressed as the mean value of duplicate determinations ± SD.

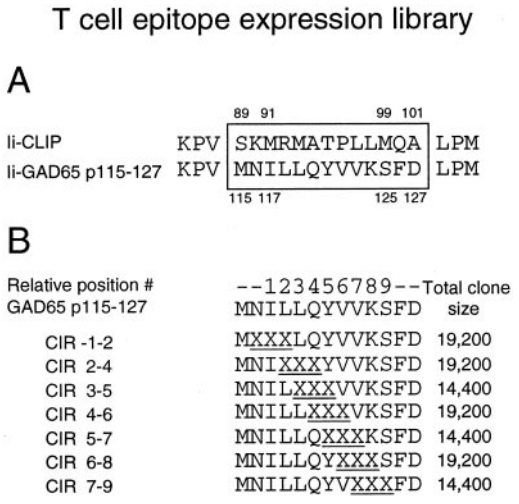


FIGURE 3. Design of a T cell epitope expression library using CLIP-substituted Ii genes. *A*, For construction of this library, the CLIP-encoding region (Ii_{89-101}) was genetically replaced with $GAD65_{115-127}$ to align I117 of $GAD65_{115-127}$ to M91 of Ii_{89-101}, the first DR-anchoring residue of the CLIP. *B*, The series of T cell epitope expression library was composed of seven sublibraries: CIR−1–2; CIR2–4; CIR3–5; CIR4–6; CIR5–7; CIR6–8; and CIR7–9. CIR−1–2 represents a sublibrary in which three successive residues from relative positions −1 to 2 on the $GAD65_{115-127}$ were replaced by randomized amino acids. Each sublibrary was aligned and is indicated in single-letter amino acid code. X indicates a random amino acid encoded by nucleotide triplets NNK, where N stands for any nucleotide and K stands for G or T. In each of three randomized codons, the third position was limited to G or T to minimize the appearance of stop codons.

tyrosine to tryptophan substitution at position 121 and lysine to arginine substitution at position 124, albeit with a significantly weaker response. In contrast, MK20.2 did not respond to these conservative substitutions at these positions. These data indicated that these T cell clones differ in responses against peptide analogs with a single-amino acid substitution. In particular, peptide analogs carrying replacement of Y121 and K124 with even conservative amino acids completely abrogated reactivity of MK20.2, suggesting that these residues are directly contacted by MK20.2 TCR.

Construction of the T cell epitope expression library using CLIP-substituted Ii genes

We constructed a set of T cell epitope expression libraries by using an Ii-based epitope presenting plasmid vector, pCI, in which the CLIP region of Ii was substituted with MHC class II-restricted epitopes (Fig. 3*A*). In the libraries, sequences of peptides are derived from the $GAD65_{115-127}$, and three successive residues within the sequence were totally randomized, the theoretical maximum complexity of each library being 20^3 (Fig. 3*B*). We defined the I117 of $GAD65_{115-127}$ as the relative position 1 in this study. As shown in Fig. 3*B*, the library set (series CIR) is composed of seven sublibraries: CIR−1–2; CIR2–4; CIR3–5; CIR4–6; CIR5–7; CIR6–8; and CIR7–9. (CIR−1–2 represents a library in which three successive residues from a relative position −1 to 2 in the T cell epitope $GAD65_{115-127}$ are replaced by randomized amino acids.) The positions of inserted randomized amino acids are serially overlapped between individual sublibraries covering the core epitope $GAD65_{116-125}$. Individual libraries contained ~14,400–19,200 DNA clones and were divided into subpools composed of 30–50 clones. Summation of all the complexity of peptides is estimated to be at least 120,000 species.

Two $GAD65_{116-127}$-autoreactive and HLA-DR53-restricted TCRs responded differently to T cell epitope expression libraries: cross-reactivity scanning

The epitope recognition by the two $GAD65_{116-127}$-reactive TCRs was further investigated using the T cell epitope expression system to identify the diverse peptide ligands of these T cell clones. The two T cell clones were examined in parallel with their production of IFN-γ in response to a set of T cell epitope expression libraries. Fig. 4 shows the frequency of pools stimulating T cell clones in the epitope libraries. The vector DNA encoding native $GAD65_{115-127}$ mixed with a 100× excess amount of the wild-type Ii gene was assigned an arbitrary activity of 1 U, leading to determination of the relative activity of each separated pool. Thus, these TCRs responded to pools of each epitope library at different frequencies. MK20.2 responded to many of the pools of CIR−1–2 (99.7%) and CIR2–4 (66.0%) libraries with >0.5 U of response. In contrast, this TCR responded to a few of the pools in relatively C terminus-randomized libraries (CIR3–5 to CIR7–9 libraries) (<7%). These results suggest that the highest specificity of MK20.2 TCR exists at the relatively C-terminal side of $GAD65_{115-127}$. In contrast, SA32.5 showed a broader response profile against several libraries spanning the epitope functional core. SA32.5 TCR showed a number of strong responses with library CIR6−8 (73.4%). These results collectively indicate that the spectrum of fine specificity in TCR recognition was clearly different between these two TCRs.

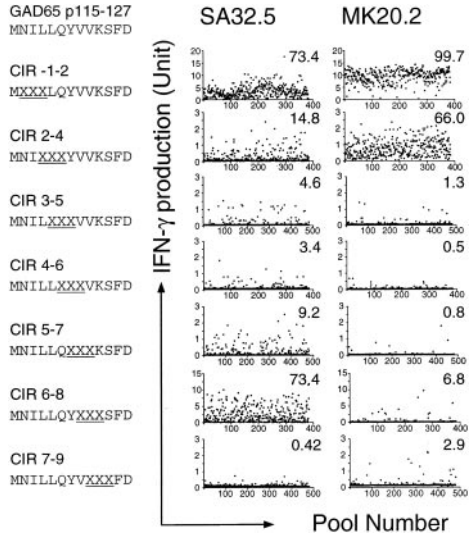


FIGURE 4. Response of $GAD65_{116-127}$-autoreactive TCRs to pools of a peptide expression library. Each pool was expected to contain ~30–50 distinct peptide ligands. IFN-γ production for each pool is expressed in terms of units. The response to pCIG (relevant DNA; $GAD65_{115-127}$) 100-fold diluted with pCI (irrelevant DNA; CLIP) was defined to be a relative value 1 U IFN-γ production ranging from 232 to 506 pg/ml. Percentages of positive pools for individual sublibraries are indicated. Positive pools that stimulated a significant IFN-γ production (relative value >0.5) were assigned. The negative control culture was stimulated by pCI, and IFN-γ production stimulated by pCI ranged between 5.75 and 36.3 pg/ml. The experiment was repeated twice with reproducible results. Randomized residues are underlined.

Isolation of diverse stimulatory peptides recognized by TCRs

The T cell epitope expression library was used to identify a series of agonistic peptide ligands for these TCRs. We screened 2,796 pools (total clone size was estimated to be 120,000); and after 3 rounds of screening, possible stimulatory peptides with higher antigenicity for SA32.5 (85 ligands) and MK20.2 (63 ligands) TCRs, respectively, were identified. The wild-type $GAD65_{115-127}$ sequence was isolated from several libraries but not from library CIR−1–2, CIR3–5, and CIR6−8. For some peptide sequences identified, the same sequences were isolated more than twice; and in ~70% of such cases, peptides were encoded by distinct nucleotide sequences. The sequences of stimulatory peptides and their stimulatory activity are shown in Fig. 5. MK20.2 TCR responded to most of the pools (99.7%) of library CIR−1–2 in the first round of screening (Fig. 4). Therefore, agonistic peptide ligands for MK20.2 TCR were not isolated from this library, and sequences isolated from library CIR−1–2 as stimulatory ligands for SA32.5 TCR were tested for their capacity to stimulate MK20.2 TCR (Fig. 5*A*). With regard to library CIR2–4 for SA32.5 TCR (14.8%) and MK20.2 TCR (66.0%) and the library CIR6−8 for SA32.5 TCR (73.4%), stimulatory ligands were not exhaustively isolated because of the high frequency of positive pools.

Analysis of stimulatory ligands from library CIR−1–2 (MXXXLQYVVKSFD library)

The panel of stimulatory ligands isolated from library CIR−1–2 (Fig. 5*A*) shows that both SA32.5 and MK20.2 tolerated phenylalanine, methionine, leucine, isoleucine, valine, and cysteine at relative position 1 (isoleucine in the native ligand). These amino acids were compatible with the HLA-DR primary anchor residue that was restricted by the Val/Gly dimorphism at DRβ86 where valine is used in the HLA-DR53 molecule (33, 34). For SA32.5 TCR, relative position 2 showed a strong preference for asparagine and leucine. MK20.2 TCR tolerated most of the agonistic sequences identified with SA32.5 TCR with different stimulatory capacities. However, MK20.2 TCR did not respond to peptides that incorporated asparagine at relative position 2.

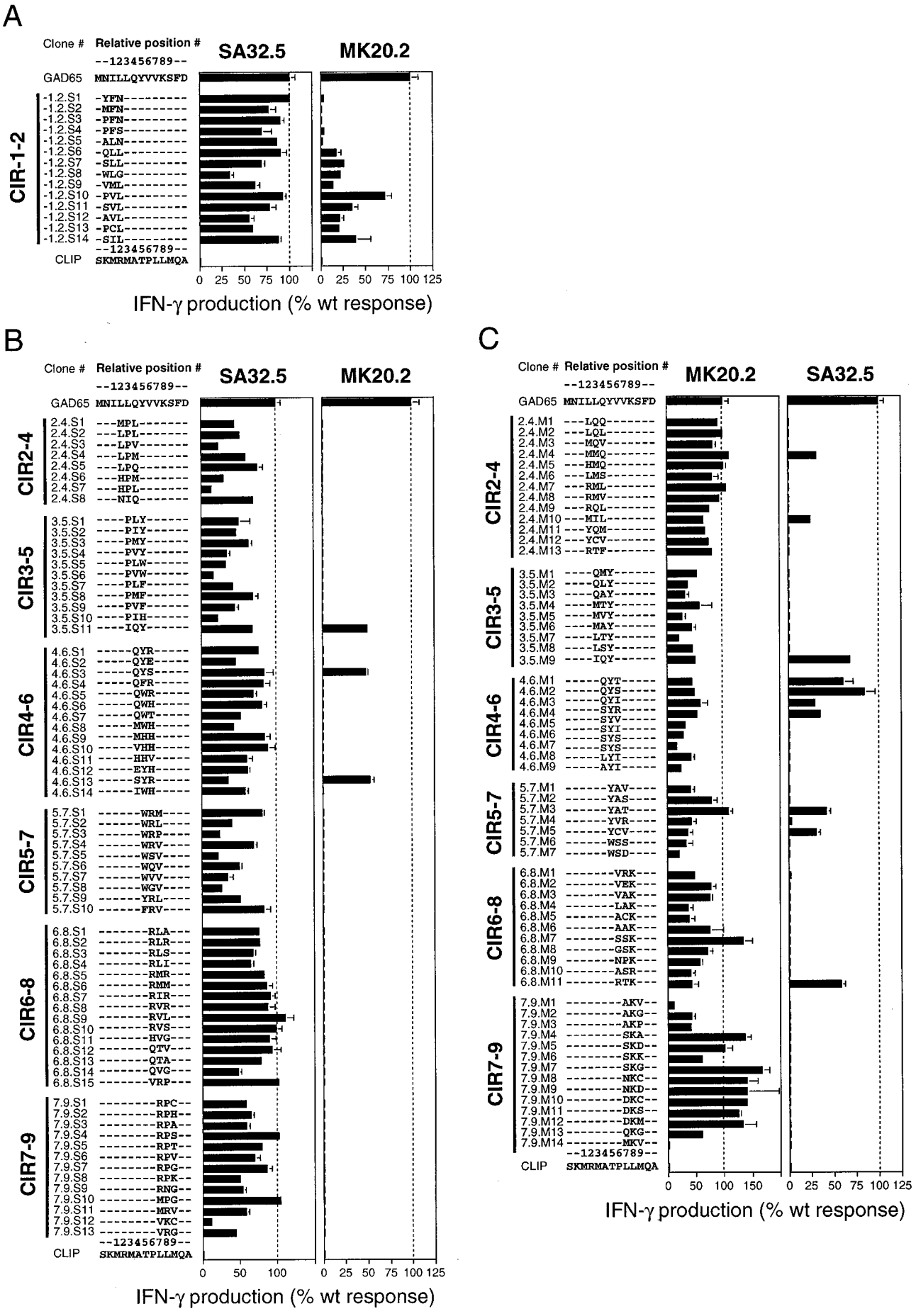
Analysis of stimulatory ligands for SA32.5 TCR: CIR2–4 to CIR7–9 libraries

SA32.5 TCR had a preference for leucine and asparagine at relative position 2 in screening of CIR−1–2 (Fig. 5*A*) and CIR2–4 (Fig. 5*B*), and methionine and histidine were also tolerated in CIR2–4. This TCR showed a strong preference for proline with an imino ring at relative position 3 in the screening of two separate libraries, CIR2–4 and CIR3–5. At this position, isoleucine sharing a chemically conservative side chain with native leucine was also tolerated. At relative position 4, leucine, methionine, valine, and glutamine were tolerated almost equally in CIR2–4 and CIR3–5 libraries. In CIR3–5, isoleucine was also tolerated; however, glutamine at relative position 4, the same amino acid as in the native ligand, was the most frequently observed (7 of 14) in CIR4−6. At this position, histidine, glutamic acid, and serine were also tolerated in CIR4−6. At relative position 5, located at the center of the epitope, this TCR did not tolerate chemically conservative single amino acid substitutions except for tryptophan in the analysis peptide analogs with single-amino acid residue substitution (Fig. 2). However, ligands containing tyrosine, phenylalanine, and tryptophan, which share an aromatic ring, and histidine with an imidazole ring at this position were isolated from the two libraries, CIR3–5 and CIR4−6. Neighboring residues may compensate for conformational mimicking.

Tryptophan at relative position 5 was the most frequently observed amino acid (8 of 10) in stimulatory ligands isolated from CIR5–7. At relative position 6, this TCR has a preference for positively charged arginine and histidine in CIR4−6, CIR5–7, and CIR6−8. However, glutamic acid, serine, threonine, and valine were also tolerated in CIR4−6. In addition, serine, glutamine, valine, and glycine were also tolerated in CIR5–7. Similarly, glutamine and valine were also tolerated in CIR6−8. At relative position 7, this TCR showed a preference for hydrophobic residues (methionine, leucine, valine, and proline) in CIR5–7 and (leucine, methionine, valine, and isoleucine) in CIR6−8. This preference in chemical character is not significantly changed in these two libraries. However, there is a drastic change in the preference in CIR7–9 at this position in which the 3-mer randomized portion was moved to the C-terminal side only by one amino acid from CIR6−8. This TCR has a preference for positively charged arginine at relative position 7. In addition, threonine and arginine in CIR6−8 and methionine and valine in CIR7–9 were also permitted at this position. At relative position 8, several amino acids with different chemical properties (arginine, proline, serine, isoleucine, methionine, leucine, valine, alanine, and glycine) were tolerated almost equally in CIR6−8. On the contrary, similar drastic changes in preference observed at position 7 were also observed at position 8 in CIR7–9, and this TCR has a strong preference for proline at this position. Several amino acids with different chemical properties were tolerated at relative position 9.

Epitope expression library revealed the importance of linear combinations of residues in recognition by SA32.5 TCR

By observing the preference of residues in peptides in recognition by SA32.5 TCR, preferable residues at each position in the N-terminal side were relatively similar in each separated library. However, in the peptide C-terminal side (Fig. 5*B*, CIR5–7,


A
Clone #
Relative position #
--123456789--
SA32.5
MK20.2
GAD65 MNILLQYVVKSFD
CIR-1-2
-1.2.S1 -YFN---------
-1.2.S2 -MFN---------
-1.2.S3 -PFN---------
-1.2.S4 -PFS---------
-1.2.S5 -ALN---------
-1.2.S6 -QLL---------
-1.2.S7 -SLL---------
-1.2.S8 -WLG---------
-1.2.S9 -VML---------
-1.2.S10 -PVL---------
-1.2.S11 -SVL---------
-1.2.S12 -AVL---------
-1.2.S13 -PCL---------
-1.2.S14 -SIL---------
--123456789--
CLIP SKMRMATPLLMQA
0 25 50 75 100
0 25 50 75 100 125
IFN-γ production (% wt response)
B
Clone #
Relative position #
--123456789--
SA32.5
MK20.2
GAD65 MNILLQYVVKSFD
CIR2-4
2.4.S1 ---MPL-------
2.4.S2 ---LPL-------
2.4.S3 ---LPV-------
2.4.S4 ---LPM-------
2.4.S5 ---LPQ-------
2.4.S6 ---HPM-------
2.4.S7 ---HPL-------
2.4.S8 ---NIQ-------
CIR3-5
3.5.S1 ----PLY------
3.5.S2 ----PIY------
3.5.S3 ----PMY------
3.5.S4 ----PVY------
3.5.S5 ----PLW------
3.5.S6 ----PVW------
3.5.S7 ----PLF------
3.5.S8 ----PMF------
3.5.S9 ----PVF------
3.5.S10 ----PIH------
3.5.S11 ----IQY------
CIR4-6
4.6.S1 -----QYR-----
4.6.S2 -----QYE-----
4.6.S3 -----QYS-----
4.6.S4 -----QFR-----
4.6.S5 -----QWR-----
4.6.S6 -----QWH-----
4.6.S7 -----QWT-----
4.6.S8 -----MWH-----
4.6.S9 -----MHH-----
4.6.S10 -----VHH-----
4.6.S11 -----HHV-----
4.6.S12 -----EYH-----
4.6.S13 -----SYR-----
4.6.S14 -----IWH-----
CIR5-7
5.7.S1 ------WRM----
5.7.S2 ------WRL----
5.7.S3 ------WRP----
5.7.S4 ------WRV----
5.7.S5 ------WSV----
5.7.S6 ------WQV----
5.7.S7 ------WVV----
5.7.S8 ------WGV----
5.7.S9 ------YRL----
5.7.S10 ------FRV----
CIR6-8
6.8.S1 -------RLA---
6.8.S2 -------RLR---
6.8.S3 -------RLS---
6.8.S4 -------RLI---
6.8.S5 -------RMR---
6.8.S6 -------RMM---
6.8.S7 -------RIR---
6.8.S8 -------RVR---
6.8.S9 -------RVL---
6.8.S10 -------RVS---
6.8.S11 -------HVG---
6.8.S12 -------QTV---
6.8.S13 -------QTA---
6.8.S14 -------QVG---
6.8.S15 -------VRP---
CIR7-9
7.9.S1 --------RPC--
7.9.S2 --------RPH--
7.9.S3 --------RPA--
7.9.S4 --------RPS--
7.9.S5 --------RPT--
7.9.S6 --------RPV--
7.9.S7 --------RPG--
7.9.S8 --------RPK--
7.9.S9 --------RNG--
7.9.S10 --------MPG--
7.9.S11 --------MRV--
7.9.S12 --------VKC--
7.9.S13 --------VRG--
--123456789--
CLIP SKMRMATPLLMQA
0 50 100
0 25 50 75 100 125
IFN-γ production (% wt response)
C
Clone #
Relative position #
--123456789--
MK20.2
SA32.5
GAD65 MNILLQYVVKSFD
CIR2-4
2.4.M1 ---LQQ-------
2.4.M2 ---LQL-------
2.4.M3 ---MQV-------
2.4.M4 ---MMQ-------
2.4.M5 ---HMQ-------
2.4.M6 ---LMS-------
2.4.M7 ---RML-------
2.4.M8 ---RMV-------
2.4.M9 ---RQL-------
2.4.M10 ---MIL-------
2.4.M11 ---YQM-------
2.4.M12 ---YCV-------
2.4.M13 ---RTF-------
CIR3-5
3.5.M1 ----QMY------
3.5.M2 ----QLY------
3.5.M3 ----QAY------
3.5.M4 ----MTY------
3.5.M5 ----MVY------
3.5.M6 ----MAY------
3.5.M7 ----LTY------
3.5.M8 ----LSY------
3.5.M9 ----IQY------
CIR4-6
4.6.M1 -----QYT-----
4.6.M2 -----QYS-----
4.6.M3 -----QYI-----
4.6.M4 -----SYR-----
4.6.M5 -----SYV-----
4.6.M6 -----SYI-----
4.6.M7 -----SYS-----
4.6.M8 -----LYI-----
4.6.M9 -----AYI-----
CIR5-7
5.7.M1 ------YAV----
5.7.M2 ------YAS----
5.7.M3 ------YAT----
5.7.M4 ------YVR----
5.7.M5 ------YCV----
5.7.M6 ------WSS----
5.7.M7 ------WSD----
CIR6-8
6.8.M1 -------VRK---
6.8.M2 -------VEK---
6.8.M3 -------VAK---
6.8.M4 -------LAK---
6.8.M5 -------ACK---
6.8.M6 -------AAK---
6.8.M7 -------SSK---
6.8.M8 -------GSK---
6.8.M9 -------NPK---
6.8.M10 -------ASR---
6.8.M11 -------RTK---
CIR7-9
7.9.M1 --------AKV--
7.9.M2 --------AKG--
7.9.M3 --------AKP--
7.9.M4 --------SKA--
7.9.M5 --------SKD--
7.9.M6 --------SKK--
7.9.M7 --------SKG--
7.9.M8 --------NKC--
7.9.M9 --------NKD--
7.9.M10 --------DKC--
7.9.M11 --------DKS--
7.9.M12 --------DKM--
7.9.M13 --------QKG--
7.9.M14 --------MKV--
--123456789--
CLIP SKMRMATPLLMQA
0 50 100 150
0 25 50 75 100 125
IFN-γ production (% wt response)

CIR6–8, and CIR7–9), there is a different preference at each position depending on which successive residues were randomized. It is conceivable that this phenomenon is affected by fixed residues flanking the randomized regions. Amino acids at relative position 6 in library CIR6–8 are not fixed, so that this position tolerates several amino acids with different chemical properties. In the isolated ligands from CIR6–8, if positively charged amino acids (arginine and histidine) located at relative position 6, hydrophobic amino acids (leucine, methionine, isoleucine, and valine) were followed at relative position 7. Subsequently, arginine, serine, or hydrophobic and aliphatic amino acids were followed at relative position 8 (see clones 6.8.S1–S11 in Fig. 5*B*). At the same time, it shapes the most dominant motif in CIR6–8, and 11 of 15 agonistic clones (73%) were positive for this motif. Also if glutamine located at relative position 6, neutral or hydrophobic amino acids were followed at relative position 7, and aliphatic amino acids were followed at relative position 8 successively (6.8.S12–S14). In the case of valine located at relative position 6, arginine at relative position 7 and proline at relative position 8 were followed successively (6.8.S15 and 7.9.S1–S8). Valine at relative position 6 was fixed in the CIR7–9 library, so that it preferentially made up the most dominant V-R-P (relative position 6–8) motif (62%) and excluded other motifs. If methionine was located at relative position 7, proline or arginine followed at relative position 8 (7.9.S10 and -S11). In case of valine at relative position 7, positively charged arginine or lysine followed at relative position 8 (7.9.S12 and -S13). These observations indicate that fixed residues sometimes limit adjacent residues and exclude other recognition motifs. In this way, peptide libraries with three successive randomized residues and their overlapping system yielded different information related to TCR recognition motifs depending on which successive ones are randomized.

In summary, SA32.5 exhibited a preference for proline at relative position 3 of peptides in screening of two separate libraries; CIR2–4 and CIR3–5. This residue does not correspond to the native sequence. In addition, this TCR tolerated phenylalanine, tryptophan, and histidine with similar structural side chains at relative position 5 when combined with some successive residues different from native ones. In the C-terminal side of peptide, three overlapping libraries, CIR5–7, CIR6–8, and CIR7–9, presented distinct recognition profiles. In this case, linear combinations of residues formed by the chemical properties of each were significantly influenced. With these promiscuous interactions, all of the TCR contacts may be altered by substitutions of other residues.

Analysis of stimulatory ligands for MK20.2 TCR: CIR2–4 to CIR7–9 libraries

MK20.2 did not respond to most of the agonistic sequences among many peptides identified with SA32.5 (Fig. 5*B*). At relative position 2, MK20.2 TCR has a preference for methionine, leucine, and tyrosine with hydrophobic side chains and for histidine and arginine with positively charged side chains in CIR2.4 (Fig. 5*C*). It seems that this position tolerates bulky side chains, because there is no relationship among chemical properties of the amino acid side chains identified. At relative position 3, this TCR indicates a strong preference for methionine and glutamine in two overlapped libraries, CIR2–4 and CIR3–5, that do not correspond to the native sequence. In addition, isoleucine, cysteine, and threonine in CIR2–4 and leucine and isoleucine in CIR3–5 are also tolerated. At relative position 4, this TCR has a preference for leucine, methionine, valine, glutamine, and serine in two overlapping libraries (CIR2–4 and CIR3–5). In addition, phenylalanine in CIR2–4 and alanine and threonine in CIR3–5 were also tolerated. In contrast, with the screening of CIR4–6, this TCR has a preference for glutamine and serine at this position, and leucine and alanine were also tolerated at this position. At relative position 5, this TCR showed almost exclusive specificity for tyrosine in three consecutive libraries, CIR3–5, CIR4–6, and CIR5–7, which is the same amino acid as the native sequence. In addition, when serine is at relative position 6, tryptophan was tolerated at position 5. At relative position 6, several amino acids with different chemical properties were tolerated. In CIR4–6, MK20.2 TCR tolerated serine, threonine, arginine, valine, and isoleucine. In CIR5–7 and CIR6–8, small amino acids such as alanine, valine, glycine, and serine were preferred. At relative position 7, valine, serine, threonine, arginine, and aspartic acid in CIR5–7 and valine, cysteine, serine, and alanine, in CIR6–8 were permitted. However, this TCR has a preference for alanine, serine, asparagine, and aspartic acid, at this position in CIR7–9. At relative position 8, MK20.2 TCR was highly specific to lysine in two overlapping libraries, CIR6–8 and CIR7–9, as in the native $GAD65_{115–127}$ sequence. Several amino acids with different chemical properties were tolerated at relative position 9. SA32.5 responded to a limited fraction of the agonistic sequences identified with MK20.2.

In summary, MK20.2 TCR was specific for tyrosine at relative position 5 and lysine at relative position 8 of the antigenic peptide, as judged by screening data from most of the separate libraries. However, several combinations of other residues were observed depending on which successive residues were randomized.

Combinatorial effects of multiple residues in exhibition of antigenicity

Several investigators reported that single amino acid modifications in a given peptide exert positive or negative effects when combined in one peptide species containing multiple substitutions (16, 17, 35). These phenomena were also observed in a series of agonistic sequences we identified using the CLIP-substituted Ii library, as shown in Fig. 6. As shown in Fig. 6, *A* and *B*, SA32.5 shows decreased or totally abrogated T cell reactivity against Y5W (standing for a peptide analog having Y to W substitution at the relative position 5), Y5F, Y5H, V7R, V7 M, K8P, and S9T. However, when these residues in each position were combined with other adjacent residues, these peptides had a completely restored T cell reactivity. For instance, when QWH (4.6.S6 clone in Fig. 5; substituted residues are underlined), QFR (4.6.S4), MHH (4.6.S9), at relative positions 4–6, and RPT (7.9.S5), and MPG (7.9.S10), at relative positions 7–9, were introduced into the peptide in combination, SA32.5 exhibited marked IFN-γ production and proliferation. Although the MK20.2 T cell clone exhibited strong specificity for lysine at relative position 8 (Fig. 5*C*), combinatorial

FIGURE 5. Identification of diverse cross-reactive epitopes recognized by GAD65-autoreactive T cell clones using a T cell epitope expression cloning strategy. *A,* Agonistic sequences isolated with SA32.5 TCR from library CIR−1–2 and their capacities to stimulate IFN-γ production by SA32.5 and MK20.2. *B,* Agonistic sequences isolated with SA32.5 TCR from CIR2–4, CIR3–5, CIR4–6, CIR5–7, CIR6–8, and CIR7–9 libraries and responses of SA32.5 and MK20.2. *C,* Agonistic sequences isolated with MK20.2 TCR from CIR2–4, CIR3–5, CIR4–6, CIR5–7, CIR6–8, and CIR7–9 libraries and the response of SA32.5 and MK20.2. The sequences of CLIP-substituted peptides and relative stimulatory activity of IFN-γ production measured in culture with a mixture of relevant: irrelevant DNA ratios at 1:50 are summarized. Data are given as percent wild-type (wt) response. All data are expressed as the mean value of duplicate determinations ± SD.

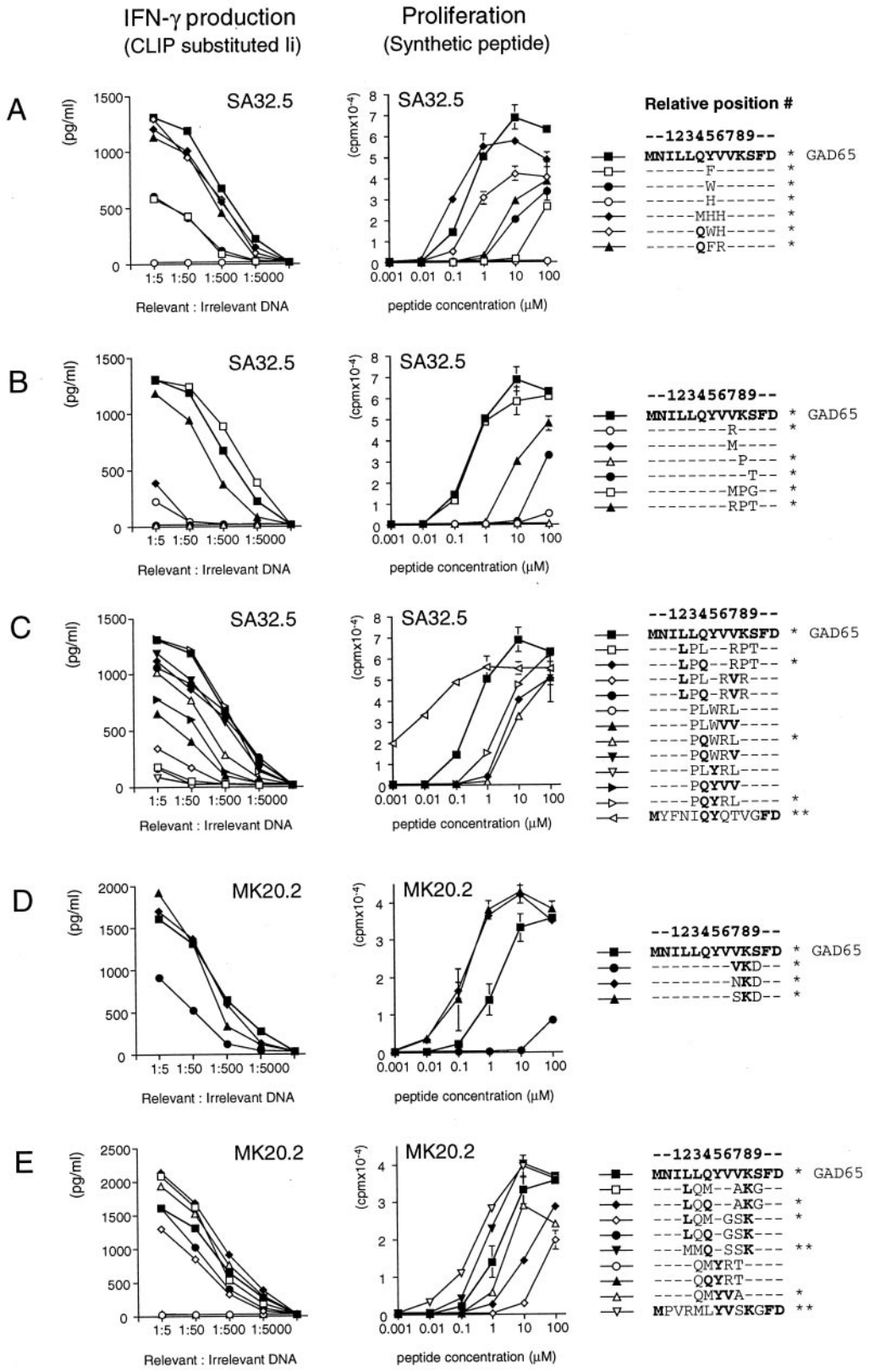


FIGURE 6. Combinatorial effects on TCR-recognition of multiple residues in antigenic peptide. *Left panels*, IFN-γ production stimulated by the CLIP substituted Ii chain epitope expression system in $GAD65_{116-127}$-specific T cell clones. All data are expressed as the mean value of duplicate determinations. *Right panels*, Proliferative responses of $GAD65_{116-127}$-specific T cell clones to the synthetic peptides at indicated concentrations. Medium control response without peptide were <200 cpm. All data are expressed as the mean value of triplicate determinations ± SD. Sequences of CLIP-substituted peptides encoded for by mutated Ii genes and synthetic peptides were indicated. Residues identical with $GAD65_{115-127}$ are represented in boldface. *, Peptides tested in both transfection and synthetic peptide assay; **, peptides tested only in proliferation assay to synthetic peptides.

effects were also observed for this TCR (Fig. 6*D*). For example, substitution of a single residue S9D decreased T cell reactivity but in combination with asparagine (7.9.M9) or serine (7.9.M5) at relative position 7 in combination exhibited an increased response. These data indicate that this strategy allows one to identify linear combinations of residues in antigenic peptide triggering TCRs.

To determine whether combinations of two stimulatory sequences consisting of three successive amino acids would affect TCR recognition, we constructed CLIP-substituted Ii expression vectors that encode peptides that incorporated two identified stimulatory sequences. These peptides did not always exhibit strong agonism for these TCRs (Fig. 6, *C* and *E*). It seems that substitution of multiple residues have a substantial effect on the recognition of SA32.5 TCR. It is conceivable that the overall linear combination of residues in antigenic peptide significantly affect recognition by SA32.5 TCR.

We asked whether the combinatorial effects we found in the analysis using the CLIP-substituted vector system were reproducible when epitopes were added as synthetic peptides. We synthesized 13-mer peptides and tested then for proliferative response of T cell clones to the peptides at various doses. As shown in Fig. 6, the results of proliferation assay using synthetic peptides were almost in parallel with those obtained in experiments using epitope presenting vector and quantification of IFN-γ produced by the T cell clones (Fig. 6). Therefore, we could verify the combinatorial effect also in the experiments using synthetic peptides. Thereafter, 13-mer peptides incorporating three successive amino acids which have been proven to stimulate T cell response. MYFNIQYQTVGFD for SA32.5 and MPVRMLYVSKGFD for MK20.2, were synthesized. As shown in the right panels of Fig. 6, *C* and *E*, these peptides with no resemblance to $GAD65_{115–127}$ activated T cell clones, respectively, at lower concentrations than did the native sequence, thereby indicating that simultaneous multiple residue modification deduced from the results obtained using epitope expression cloning strategy could generate superagonists.

These data strongly suggest that linear combinations of residues on antigenic peptides affect recognition by TCRs. Therefore, amino acid combinatorial effects should be considered when searching for cross-reactive epitopes. We also suggest that T cell epitope expression cloning can provide the strategy for efficient identification of optimal sequences.

Identification of microbial and self mimics for $GAD65_{115–127}$-autoreactive TCRs

The results obtained from T cell epitope expression library were used to establish recognition motifs for SA32.5 and MK20.2 TCRs. As shown in Fig. 7, amino acids with similarities in chemical characteristics and with successive stimulatory combination with residues obtained from library screening that may induce mimicry were considered to provide search criteria for peptide mimics of $GAD65_{115–127}$. As for the SA32.5 TCR, motifs consisting of amino acids leucine, isoleucine, and proline at relative position 3, and tyrosine, tryptophan, and phenylalanine at relative position 5 were considered because these residues were critical for recognition by SA32.5 TCR (Fig. 5). With regard to the peptide C-terminal side (relative positions 6–8), three kinds of search criteria were considered (Fig. 7*A*), because a successive combination of residues in the antigenic peptide may significantly affect the recognition by SA32.5 TCR. As for MK20.2 TCR, amino acids methionine, glutamine, isoleucine, and leucine at relative position 3 with strong preference were considered. Amino acids tyrosine at relative position 5 and lysine at relative position 8 were fixed on the basis of exquisite specificity of MK20.2 TCR (Fig. 7*B*).

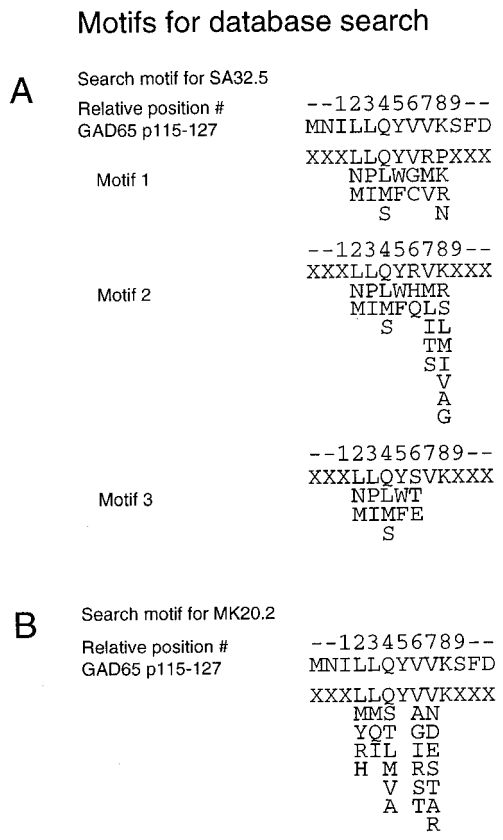


FIGURE 7. Summary of motifs recognized by TCRs for a database search of peptide mimics. The search motifs were considered by amino acid preference in TCR recognition. With regard to SA32.5 TCR, three kinds of search motifs were considered depending on specific recognition properties. Search motifs are given in single-letter amino acid code; X indicates all amino acids.

To identify potential mimicry epitopes from natural proteins, a pattern match search for microbial and self proteins was conducted using the SWISS-PROT (104,559 protein entries) database, the TrEMBL (560,376 protein entries) database, and the ScanProsite program (http://www.expasy.ch/tools/scnpsit2.html; January 2002). Among the candidates conforming to the criteria, we selected 47 sequences with successive stimulatory combinations recognized by TCRs. In addition, with respect to the selection of microbial peptide, proteins derived from possible infectious pathogens were considered and listed (Table II). To determine whether these candidate peptides would activate T cell clones, we constructed CLIP-substituted Ii expression vectors encoding these peptide candidates. These candidates were examined on their potential to stimulate production of IFN-γ by T cell clones using COS-7 cells expressing the CLIP-substituted Ii and HLA-DR53. Among the 47 candidates selected, SA32.5 responded to 7 candidates, *L. lactis*, *S. pyogenes*, *N. meningitidis*, *Chlamydia pneumoniae*, *Homo sapiens*/claudin-17, tafazzin, and tafazzin-like protein-derived peptides. T cell clone MK20.2 responded to five candidates, *Legionella pneumophila*, *S. pneumoniae*, *Staphylococcus aureus*, *Rickettsia prowazekii*, and human herpesvirus 6-derived peptides.

On the basis of the data obtained, we synthesized 12 candidate peptides and tested their capacity to stimulate T cell clones (Fig. 8). These peptides stimulated each T cell clone respectively and with different activity. Peptides derived from *N. meningitidis*; putative dihydrolipoamide dehydrogenase (147–159) and self protein tafazzin (207–219) activated SA32.5 at concentrations of <1 nM

Table II. *Candidate selection for mimicry peptides*[a]

	Mimicry Candidates for T Cell Clone SA32.5			Mimicry Candidates for T Cell Clone MK20.2		
	No.	Sequence	Source	No.	Sequence	Source
		Relative position no. —123456789—			Relative position no. —123456789—	
		MNILLQYVVKSFD	*H. sapiens*/GAD65$_{115–127}$		**MNILLQYVVKSFD**	*H. sapiens*/GAD65$_{115–127}$
Motif 1	Y519	KNLLIQFVRNHYG	*Haemophilus influenzae*	M307	**KMSLLQYIVKNNKI**	*Legionella pneumophila*
	Y520	DKNLIQFVVKGSN	*Borrelia burgdorferi*	M308	KTSLQQYGSKDQL	*Mycoplasma pneumoniae*
	Y521	TRLNPLWVRPVGS	Hepatitis G virus	M309	**DDIMMSYAAKGFV**	*S. pneumoniae*
Motif 2	Y531	LLSLPLYHVSGQG	*H. influenzae*	M310	**PYVLLAYIVKSVG**	*Staphylococcus aureus*
	Y532	LIVNPLYRLKGYG	*Bacillus subtilis*	M311	DAILQMYVSKFLF	*Mycoplasma capricolum*
	Y533	LATLLLWHVVGAT	*Mycobacterium avium*	M312	DTVLQMYGEKHAG	*S. typhimurium*
	Y534	LGLMPLYHVVGFF	*Pseudomonas* sp.	M313	SIILMLYATKFFE	*Mycoplasma pulmonis*
	Y535	LLSLPLFHVSGQG	*Salmonella typhimurium*	M314	RKLLQAYGAKLVL	*Listeria innocua*
	Y536	**LVILPMFHVSGLS**	*L. lactis*	M315	**ELVLMSYRDKLVL**	*Rickettsia prowazekii*
	Y537	YNIMLQYRVKVES	*Vibrio cholerae*	M316	KYLRLVYGNKILS	*Prevotella ruminicola*
	Y538	AVLLPLYRLRQYA	*Bacillus halodurans*	M317	AQNRMSYSNKDYD	*Listeria monocytogenes*
	Y539	VQFNPQWQLALVA	*Pasteurella multocida*	M318	GESMMAYAVKGHR	*P. aeruginosa*
	Y540	**REILPQYQLVILA**	*S. pyogenes*	M319	RDVMIAYATKAHV	*S. pneumoniae*
	Y541	KKNLLQWQTSADS	*Clostridium acetobutylicum*	M320	IAVRMAYSSKTPT	*Bacillus subtilis*
	Y542	QEFLISYRLKIVD	*Helicobacter pylori*	M321	TPMRLSYIEKKKG	*S. typhimurium*
	Y543	**PVILPQWQSLGNR**	*N. meningitidis*	M322	VNVMQTYTVKPGT	*Staphylococcus aureus*
	Y544	PKVLPQYQSLQNW	*Chlamydia pneumoniae*	M323	**KELLQSYVSKNNN**	Human herpesvirus 6
	Y545	**QAMNILYQTVQAF**	*C. pneumoniae*	M324	MRNMLQYVSKNLD	Orf virus
	Y546	WLMNPLFRLISKA	*H. influenzae*	M325	IADLQQYRNKLET	Human rotavirus
	Y547	SFVLPLFRVAALL	*Pseudomonas aeruginosa*	M326	GEVRQAYGAKGFS	*H. sapiens*/Glypican-6 precursor 31–43
	Y548	RPINPLFHILVET	*S. pneumoniae*			
	Y549	ILMLLQFRVLDRR	*Brucella melitensis*			
	Y550	YLINPMFRIIANT	*B. melitensis*			
	Y551	**TTLLPQWRVSAFV**	*H. sapiens*/claudin-17 24–36			
	Y552	**PIILPLWHVGMND**	*H. sapiens*/tafazzin 207–219			
	Y553	**PIILPLWHVGEPG**	*H. sapiens*/tafazzin-like protein 138–150			
Motif 3	Y554	QYVMLQFTVKERP	*Treponema paraluiscuniculi*			

[a] Of the peptide mimicry candidates we obtained from a pattern match search, sequences with stimulatory combinations of three successive amino acids for recognition by TCRs were selected. Candidates for mimicry peptide sequences are given in a single-letter amino acid code aligned with GAD65$_{115–127}$. Sequences stimulating significant IFN-γ production (>200 pg/ml) determined using the CLIP-substituted epitope expression system are represented in boldface.

despite a limited sequence homology to the native GAD65 sequence (Fig. 8, *A* and *B*). Peptide mimics derived from *Legionella pneumophila*, pilus assembly protein PilB (39–51) and *Staphylococcus aureus*, and hypothetical protein SAV0107 (227–239) activated MK20.2 at concentrations of <100 nM (Fig. 8*C*). In total, seven agonistic peptides derived from microbial or self protein (21.6% of the peptide mimic candidates conforming to the search criteria) were identified for SA32.5, and five agonistic peptides derived from microbial protein (25% of the peptide mimic candidates) were identified for MK20.2.

Cross-reactivity of T cell clones SA32.5 and MK20.2 to naturally processed microbial Ags

To investigate whether these microbial peptide mimics identified by our library scanning can be naturally processed and presented, five recombinant proteins carrying mimicry epitopes were generated as GST-fusion proteins and tested for their capacity to induce proliferative responses of these T cell clones. As APCs, we used DCs generated from CD14$^+$ monocytes. As shown in Fig. 9, T cell clone SA32.5 responded to fusion proteins of putative dihydrolipoamide dehydrogenase (101–205) derived from *N. meningitidis* and *O*-succinylbenzoic acid-CoA ligase (101–201) derived from *L. lactis.* T cell clone MK20.2 responded to a fusion protein of glutamine amidotransferase, class I (1–104) derived from *S. pneumoniae.* However, fusion proteins derived from *S. pyogenes* and *Legionella pneumophila* did not activate either T cell clone. Furthermore, these responses were markedly inhibited by HLA-DR mAb L243 (data not shown). Cross-reactivity of the SA32.5 T cell clone to fusion proteins carrying mimicry epitopes for MK20.2 were not observed, and vice versa. These data suggest that some of these microbial peptide mimics identified by our strategy can be naturally processed in DCs and presented to GAD65-specific CD4$^+$ T cell clones.

Discussion

In this study, we developed a novel strategy to analyze the combinatorial effects of residues in the antigenic peptides on recognition by TCRs. It allows for systematic separation and identification of diverse T cell epitopes from a mixture of randomized peptides. We showed that the combinatorial effects can be classified by analyzing the epitope sequences with agonistic properties. We searched protein databases with the defined search criteria incorporating combinatorial effect, and we identified mimicry epitopes of microbial origin, which stimulate GAD65-autoreactive T cell clones established from type I diabetes patients.

Because of the importance of CD4$^+$ T cells in autoimmunity, much effort has been directed toward identifying cross-reactive epitopes of microbial Ags recognized by autoreactive CD4$^+$ T cells. Thus far, cross-reactive epitopes have been predicted and identified by primary sequence homology, the data obtained from single-residue-substituted peptide analogs, or PS-SCLs (2, 7, 8, 36, 37). The majority of these approaches using synthetic peptides have been fundamentally based on the concept that Ag recognition surface of TCRs is relatively flat and that each amino acid on each

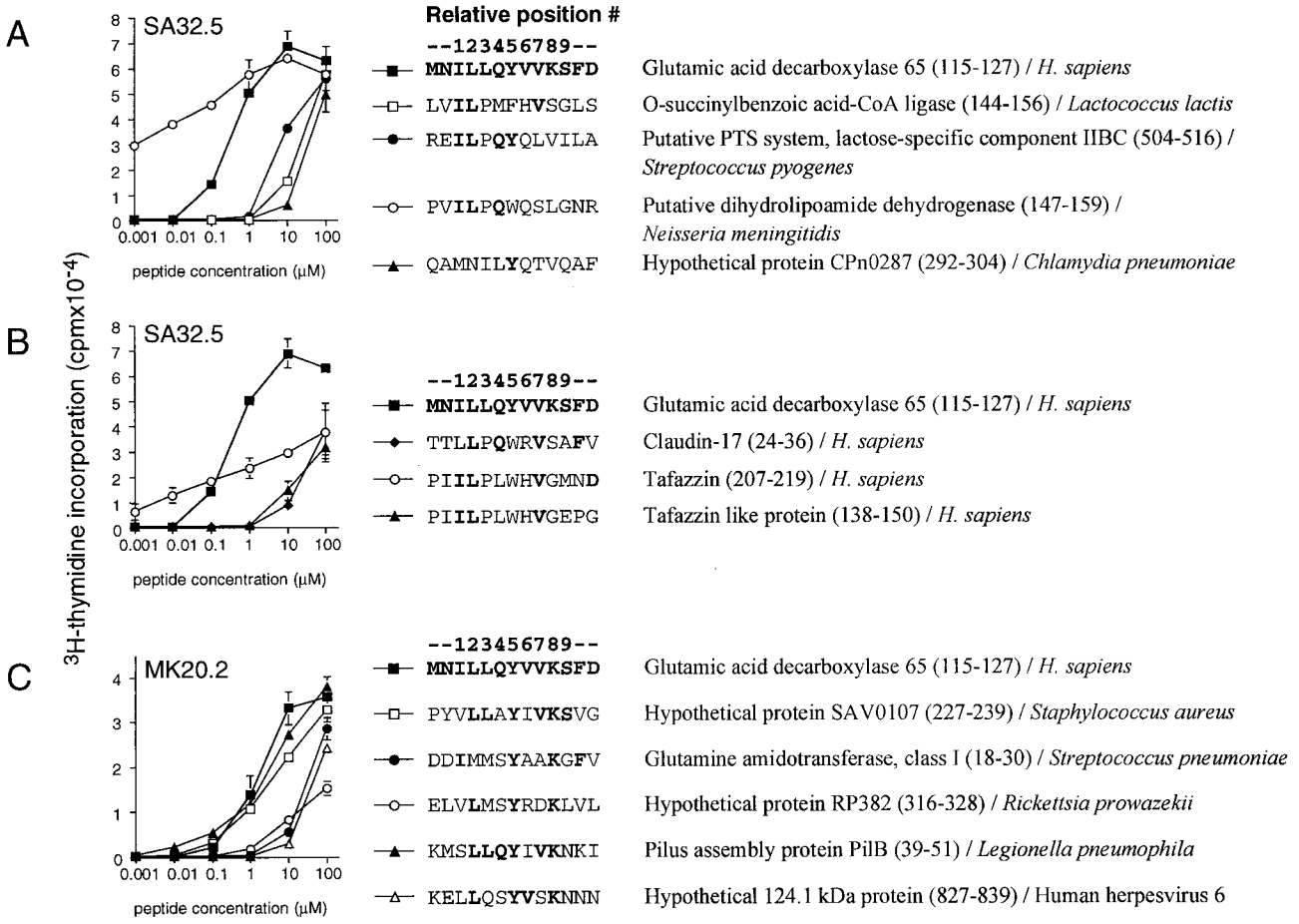


FIGURE 8. Proliferative response of $GAD65_{116-127}$-specific T cell clones to the mimicry peptides in comparison with $GAD65_{115-127}$. *A*, Response of SA32.5 to microbial mimicry peptides. *B*, Response of SA32.5 to self mimicry peptides. *C*, Response of MK20.2 to microbial mimicry peptides. Sequences of mimicry peptides are indicated, and residues identical with $GAD65_{115-127}$ are represented in boldface. Medium control responses without peptide were <200 cpm. All data are expressed as the mean value of triplicate determinations ± SD.

position of the peptide independently contributes to recognition by TCR (7, 38). On the basis of this assumption, mimicry epitopes have been searched in protein databases using pattern match searching with information on suitable amino acids for each position of the peptide. The use of PS-SCLs for analyses of T cell recognition has yielded the many useful concepts. However, some T cell clones did not respond to PS-SCLs (39) or responded to PS-SCLs ambiguously (40). In addition, artificial peptides composed of the optimal residue for each position selected based on analysis with PS-SCLs do not necessarily show agonistic activity, as we previously described (39). The major drawback of this method is that it identifies only the relative importance of each amino acid for each position of the peptide, and one could not directly analyze the agonistic activity induced by combinations of residues in antigenic peptides.

To address these problems, we used a totally different approach. The novelty of this system is that it enables one to directly identify sequences of T cell epitopes from plasmid-based epitope libraries composed of thousands of randomized sequences. Based on the accumulated information on sequences of agonistic peptides, one can determine the combinations of residues with agonistic properties. We used two different GAD65-specific T cell clones, SA32.5 and MK20.2, expressing distinct TCR but recognizing the same epitope with the same restriction element (25). At first, we verified that two T cell clones expressed single TCR, then we identified epitopes agonistic to each of these two TCRs from the epitope libraries and compared their sequences.

It is clear that TCRs of SA32.5 and MK20.2 represent distinct cross-reactivity and different recognition profiles (Fig. 5, *B* and *C*). It was observed that SA32.5 TCR tolerates structurally related amino acids at position 5 only when combined with specific adjacent residues (Figs. 5*B* and 6*B*). The patterns of the recognition profiles significantly differ depending on where randomized residues had been inserted. These findings were observed mainly on the peptide C-terminal side (Figs. 5*B* and 6*B*). Also, overall combinations of residues in the antigenic peptide affect the recognition by SA32.5 TCR more significantly than that by MK20.2 TCR (Figs. 5*B* and 6, *A–C*). Notably SA32.5 TCR permits the exchange of residues at all of the positions, as reported (41). In contrast, MK20.2 TCR permits specifically tyrosine and lysine for relative positions 5 and 8, respectively, even when residues around them are randomly exchanged (Fig. 5*C*). These residues are also conserved in the original $GAD65_{115-127}$. However, even in the case of this TCR, significant amino acid combinatorial effects were observed in some positions (Figs. 5*B* and 6*D*).

As shown in Table III, MK20.2 had the same Vβ3.1 usage as did HA1.7, and it was reported that Vβ3.1 was predominantly expressed in $HA_{306-318}$-specific TCRs which preferred P8K of $HA_{306-318}$ (42). Molecular modeling (Fig. 10*A*) of the fit of MK20.2 TCR to the recent crystal structure of HA1.7 TCR-HA peptide/HLA-DR1 complex (42) was done. As shown in Fig. 10, the model predicted that the D28β and E30β of the CDR1 loop of TCRVβ3.1 made a charged interaction with lysine at position 8 of

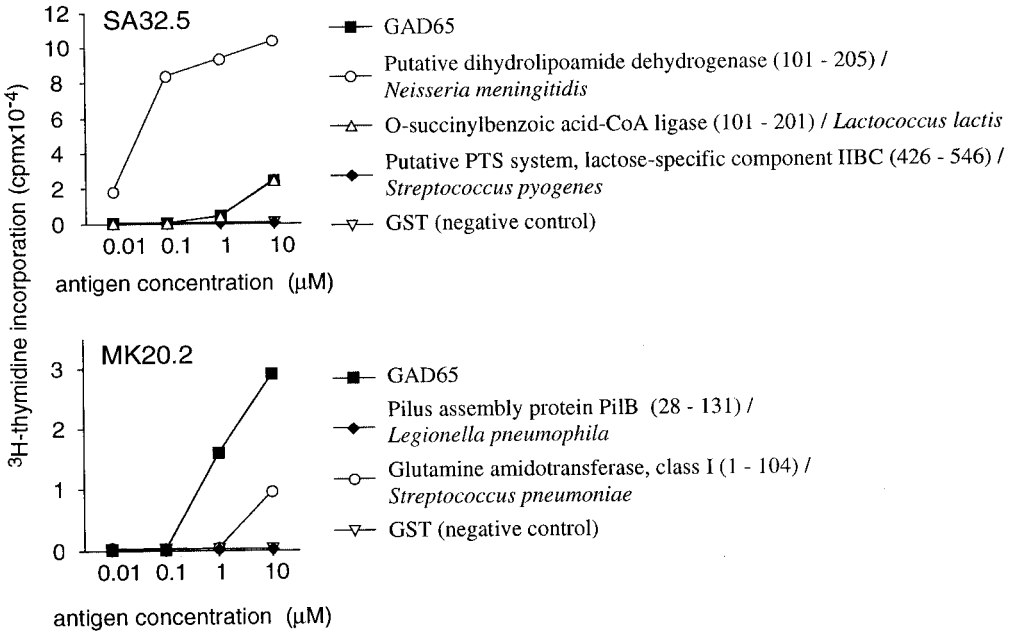


FIGURE 9. Proliferative response of $GAD65_{116-127}$-specific T cell clones to recombinant proteins carrying the mimicry peptides in comparison with GAD65 protein. *A,* Response of SA32.5 to microbial proteins: putative dihydrolipoamide dehydrogenase (101–205)/*N. meningitidis*, *O*-succinylbenzoic acid-CoA ligase (101–201)/*L. lactis,* and putative PTS system, lactose-specific component IIBC (426–546)/*S. pyogenes. B,* Response of MK20.2 to microbial proteins: glutamine amidotransferase, class I (1–104)/*S. pneumoniae* and pilus assembly protein PilB (28–131)/*Legionella pneumophila*. Medium control responses without recombinant protein were <150 cpm. Data are expressed as the mean value of duplicate determinations ± SD.

HLA-DR53-bound $GAD65_{115-127}$ peptide (Fig. 10*B*), thus implying the exclusive specificity for lysine at relative position 8. In contrast with BV3S1, BV9S1 of SA32.5 TCR uses the small amino acid glycine in 28β and the negatively charged amino acid aspartic acid in 30β (Table III), which would weaken electrostatic interactions toward lysine at position 8. This observation seems consistent with the finding that SA32.5 TCR did not have a strong preference for lysine at position 8 (Fig. 5*B*).

The HLA-DR53 binding motif has not yet been determined, and only a few peptides with binding affinity to HLA-DR53 molecule have been reported. However, the data obtained in our study led us to speculate on the peptide-binding properties to HLA-DR53 molecules. In a recent study on HLA-DR53-binding peptides, the molecule did not show any amino acid preference in P1 position (45). In fact, although HLA-DR53 possess pocket-1 consisting of valine at DRβ86, which commonly accommodates aliphatic residues, it permits lysine and tyrosine (45). These results are consistent with our finding that relative position 1 permits several residues with different chemical properties (Fig. 1). However, the peptide sequences isolated from the epitope expression library indicate that relative position 1 preferred aliphatic residues. A more comprehensive scan may yield the information of this issue. In celiac disease, the disease-susceptible HLA-DR53 molecule selectively binds a large number of gliadin-derived peptides that are extremely rich in glutamine and proline (46). Accordingly, it was suggested that one or more peptide binding pockets of HLA-DR53 molecules preferentially engaged the side chain of glutamine or proline. In the TCR recognition motif of the T cell clone SA32.5, proline at relative position 3 and glutamine at relative positions 4 and 6 are the most important residues for TCR agonism (Fig. 5*B*). Although the two TCRs represented a distinct specificity at relative positions 3, 5, 7, or 8 of the peptides, SA32.5 TCR and MK20.2 TCR represented similar preferences for certain amino acids at some positions. For example, at relative position 1, both TCR tolerate phenylalanine, methionine, leucine, isoleucine, valine, and cysteine; at relative position 4 they tolerate leucine, methionine, glutamine,

Table III. *Comparison of CDR sequences of TCRβ chains of $GAD65_{115-127}$-specific MK20.2, SA32.5, and $HA_{306-318}$-specific HA1.7*[a]

T Cell Clone	Vβ usage	CDR1β	CDR2β	CDR3β
		28 30	48 51 54 56	96 99
HA1.7	**TCRBV3S1**	ECVQD M**D**H**E**NMFW	IYFSYDVKMKEKGD	CASS**S**TGLP--YGYTF
		K(P8)	DR67α K DR39α K	K(P8)
MK20.2	**TCRBV3S1**	ECVQD M**D**H**E**NMFW	IYFSYDVKMKEKGD	CASS**S**TGVS--PGELF
SA32.5	TCRBV9S1	KCEQN LGHDTMYW	MFSYNNKELIINET	CASSPTGQGAHTGELF

[a] As determined in a crystallographic analysis (42), interactions between residues that provide TCR-DR binding sites between α-helical structure of the DRα chain and the CDR2 region of HA1.7 TCR are lined with dotted doublet lines. Electrostatic interactions between P8K of HA (peptides 306–318) and the CDR region of HA1.7 are indicated by solid doublet lines. Predicted interactions between residues of CDR region of MK20.2 or SA32.5 TCR and those of the $GAD65_{115-127}$)/DR53 complex are indicated and lined.

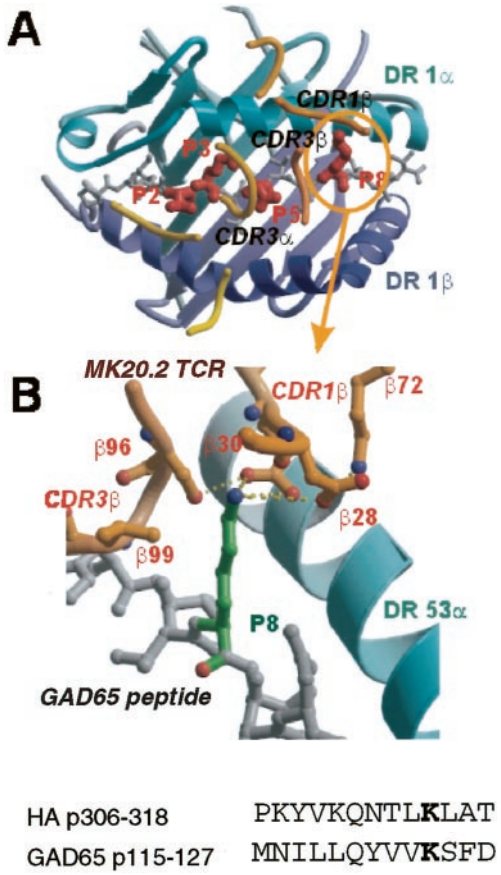


FIGURE 10. Structure of HA1.7 TCR-HA/DR1 complex (*A*) and molecular modeling of MK20.2 TCR-GAD65/DR53 complex (*B*). *A,* Structure of the HA1.7 TCR-HA/DR1 complex (42) and relative orientation of the CDR loops of HA1.7 TCR on top of HA/DR1 complex. CDR loops of TCR Vα and Vβ chains are displayed in tubes and colored yellow and orange, respectively. $HA_{306-318}$ peptide is shown in a ball-and-stick model; its TCR binding residues are red and the others are gray. DR α-chain is cyan, and the β-chain is blue. *B,* Molecular modeling of predicted electrostatic interactions between P8 lysine of $GAD65_{115-127}$ and acidic residues of MK20.2 TCR. Here the homology model of MK20.2 TCR-GAD65/DR53 complex was based on the above HA1.7 TCR-HA/DR1 complex. The residues on CDR1 and three of MK20.2 TCR Vβ chain are orange. The GAD65 peptide is shown in a ball-and-stick model, and P8K is green. The DR53 α-chain is displayed in cyan. Electrostatically active atoms are red and blue (negatively and positively charged, respectively). Numbers represent the position of relevant side chain residues. Hydrogen bonding and electrostatic interactions are dotted yellow lines. This figure was produced by BOBSCRIPT (43) and Raster 3D (44). The peptide sequences of $HA_{306-308}$ and $GAD65_{115-127}$ are given in single-letter amino acid code.

and serine; and at the relative position 6 neutral or positively charged residues and small aliphatic amino acids are allowed (Fig. 5). These residues may be HLA-DR53 anchor residues; this remains to be elucidated. It is conceivable that findings observed in this study also reflect peptide-binding characteristics in which anchor combinations are important for HLA binding, as described (47). Not only requisition for direct TCR contact but also constraints of HLA-DR53 binding could explain the combinatorial effect observed on the peptide C-terminal side, i.e. a positively charged amino acid is required at positions 122–124 in combination with a bulky amino acid at position 121 for the recognition by SA32.5 TCR.

Several groups of investigators have reported epidemiologic correlations between infection and type I diabetes (48, 49). Among the cross-reactive ligands identified in this study, epitopes derived from *N. meningitidis*, *S. pneumoniae*, and *L. lactis* can actually be produced from proteins by APCs to stimulate the T cells (Fig. 9). *N. meningitidis* is a Gram-negative and pathogenic bacterium. The outcome of meningococcal infection ranges from asymptomatic carriage to meningitis and fulminant meningococcemia in children and young people (50). *S. pneumoniae* is a Gram-positive and pathogenic bacterium, which causes bacterial sepsis, pneumonia, meningitis, and otitis media in young children (51). *N. meningitidis* and *S. pneumoniae*, which are also human respiratory commensals, in common can spread in the bloodstream under an immunocompromised state, etc. It is conceivable that these bacteria normally residing as harmless commensals may spread and participate in priming of autoreactive T cells and increase memory pools in the periphery. Also, asymptomatic carriages may predispose genetically susceptible children and initiate the autoimmune process. *L. lactis* is a nonpathogenic Gram-positive bacterium, used to produce fermented foods and dairy goods. When taken orally with food, this bacterium is metabolically active in all compartments of the intestinal tract (52). It is presumed to be taken up by M cells in intestinal lymphoid organs and to sensitize T cells. Although its relationship to autoimmune process has yet to be reported, it is intriguing in consideration of a recent report that short term breast-feeding and the early introduction of cow's milk-based infant formula can predispose genetically susceptible children to type I diabetes (53).

In conclusion, we established a novel system to analyze the combinatorial effects of residues in the antigenic peptide on recognition by HLA class II-restricted TCR. The degree of combinatorial effects differs depending on part of the antigenic peptide and structure of TCRs, even when the epitope and restriction molecule are identical. Importantly, the defined TCR recognition motif incorporating combinatorial effects proved useful in identifying mimicry epitopes of autoreactive TCRs. The findings demonstrate the importance of combinatorial effects of each amino acid residue in the antigenic peptide on TCR recognition and propose a new direction for examining cross-reactive epitopes of TCR in investigating autoimmunity.

Acknowledgments

We thank Dr. S. Matsushita (Saitama Medical School, Saitama, Japan) and Dr. N. Sato (Niigata University School of Dentistry, Niigata, Japan) for helpful suggestions and Drs. T. Akaike and T. Akuta (Kumamoto University School of Medicine) for the *Legionella pneumophila* gene, We thank Kirin Brewery (Tokyo, Japan) for human GM-CSF and M. Ohara (Fukuoka, Japan) for helpful comments.

References

1. Oldstone, M. B. 1998. Molecular mimicry and immune-mediated diseases. *FASEB J. 12:1255.*
2. Wucherpfennig, K. W., and J. L. Strominger. 1995. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell 80:695.*
3. Panoutsakopoulou, V., M. E. Sanchirico, K. M. Huster, M. Jansson, F. Granucci, D. J. Shim, K. W. Wucherpfennig, and H. Cantor. 2001. Analysis of the relationship between viral infection and autoimmune disease. *Immunity 15:137.*
4. Loftus, C., E. Huseby, P. Gopaul, C. Beeson, and J. Goverman. 1999. Highly cross-reactive T cell responses to myelin basic protein epitopes reveal a nonpredictable form of TCR degeneracy. *J. Immunol. 162:6451.*
5. Nanda, N. K., K. K. Arzoo, H. M. Geysen, A. Sette, and E. E. Sercarz. 1995. Recognition of multiple peptide cores by a single T cell receptor. *J. Exp. Med. 182:531.*
6. Quaratino, S., C. J. Thorpe, P. J. Travers, and M. Londei. 1995. Similar antigenic surfaces, rather than sequence homology, dictate T-cell epitope molecular mimicry. *Proc. Natl. Acad. Sci. USA 92:10398.*
7. Hemmer, B., B. T. Fleckenstein, M. Vergelli, G. Jung, H. McFarland, R. Martin, and K. H. Wiesmuller. 1997. Identification of high potency microbial and self ligands for a human autoreactive class II-restricted T cell clone. *J. Exp. Med. 185:1651.*

8. Hausmann, S., M. Martin, L. Gauthier, and K. W. Wucherpfennig. 1999. Structural features of autoreactive TCR that determine the degree of degeneracy in peptide recognition. *J. Immunol. 162:338.*
9. De Magistris, M. T., J. Alexander, M. Coggeshall, A. Altman, F. C. Gaeta, H. M. Grey, and A. Sette. 1992. Antigen analog-major histocompatibility complexes act as antagonists of the T cell receptor. *Cell 68:625.*
10. Combadiere, B., M. Freedman, L. Chen, E. W. Shores, P. Love, and M. J. Lenardo. 1996. Qualitative and quantitative contributions of the T cell receptor ζ chain to mature T cell apoptosis. *J. Exp. Med. 183:2109.*
11. Sloan-Lancaster, J., B. D. Evavold, and P. M. Allen. 1993. Induction of T-cell anergy by altered T-cell-receptor ligand on live antigen-presenting cells. *Nature 363:156.*
12. Kersh, G. J., and P. M. Allen. 1996. Structural basis for T cell recognition of altered peptide ligands: a single T cell receptor can productively recognize a large continuum of related ligands. *J. Exp. Med. 184:1259.*
13. Kersh, G. J., M. J. Miley, C. A. Nelson, A. Grakoui, S. Horvath, D. L. Donermeyer, J. Kappler, P. M. Allen, and D. H. Fremont. 2001. Structural and functional consequences of altering a peptide MHC anchor residue. *J. Immunol. 166:3345.*
14. Ausubel, L. J., C. K. Kwan, A. Sette, V. Kuchroo, and D. A. Hafler. 1996. Complementary mutations in an antigenic peptide allow for crossreactivity of autoreactive T-cell clones. *Proc. Natl. Acad. Sci. USA 93:15317.*
15. Leggatt, G. R., A. Hosmalin, C. D. Pendleton, A. Kumar, S. Hoffman, and J. A. Berzofsky. 1998. The importance of pairwise interactions between peptide residues in the delineation of TCR specificity. *J. Immunol. 161:4728.*
16. Hemmer, B., C. Pinilla, B. Gran, M. Vergelli, N. Ling, P. Conlon, H. F. McFarland, R. Houghten, and R. Martin. 2000. Contribution of individual amino acids within MHC molecule or antigenic peptide to TCR ligand potency. *J. Immunol. 164:861.*
17. Wilson, D. B., C. Pinilla, D. H. Wilson, K. Schroder, C. Boggiano, V. Judkowski, J. Kaye, B. Hemmer, R. Martin, and R. A. Houghten. 1999. Immunogenicity. I. Use of peptide libraries to identify epitopes that activate clonotypic $CD4^+$ T cells and induce T cell responses to native peptide ligands. *J. Immunol. 163:6424.*
18. Hemmer, B., B. Gran, Y. Zhao, A. Marques, J. Pascal, A. Tzou, T. Kondo, I. Cortese, B. Bielekova, S. E. Straus, et al. 1999. Identification of candidate T-cell epitopes and molecular mimics in chronic Lyme disease. *Nat. Med. 5:1375.*
19. Zhao, Y., B. Gran, C. Pinilla, S. Markovic-Plese, B. Hemmer, A. Tzou, L. W. Whitney, W. E. Biddison, R. Martin, and R. Simon. 2001. Combinatorial peptide libraries and biometric score matrices permit the quantitative analysis of specific and degenerate interactions between clonotypic TCR and MHC peptide ligands. *J. Immunol. 167:2130.*
20. van Bergen, J., S. P. Schoenberger, F. Verreck, R. Amons, R. Offringa, and F. Koning. 1997. Efficient loading of HLA-DR with a T helper epitope by genetic exchange of CLIP. *Proc. Natl. Acad. Sci. USA 94:7499.*
21. Malcherek, G., C. Wirblich, N. Willcox, H. G. Rammensee, J. Trowsdale, and A. Melms. 1998. MHC class II-associated invariant chain peptide replacement by T cell epitopes: engineered invariant chain as a vehicle for directed and enhanced MHC class II antigen processing and presentation. *Eur. J. Immunol. 28:1524.*
22. Fujii, S., S. Senju, Y. Z. Chen, M. Ando, S. Matsushita, and Y. Nishimura. 1998. The CLIP-substituted invariant chain efficiently targets an antigenic peptide to HLA class II pathway in L cells. *Hum. Immunol. 59:607.*
23. Senju, S., S. Hirata, H. Matsuyosi, M. Masuda, Y. Uemura, K. Araki, K. Yamamura, and Y. Nishimura. Generation and genetic modification of dendritic cells derived from mouse embryonic stem cells. *Blood. In press.*
24. Fujii, S., Y. Uemura, L. K. Iwai, M. Ando, S. Senju, and Y. Nishimura. 2001. Establishment of an expression cloning system for $CD4^+$ T cell epitopes. *Biochem. Biophys. Res. Commun. 284:1140.*
25. Tabata, H., T. Kanai, H. Yoshizumi, S. Nishiyama, S. Fujimoto, I. Matsuda, M. Yasukawa, S. Matsushita, and Y. Nishimura. 1998. Characterization of self-glutamic acid decarboxylase 65-reactive $CD4^+$ T-cell clones established from Japanese patients with insulin-dependent diabetes mellitus. *Hum. Immunol. 59:549.*
26. Wicker, L. S., S. L. Chen, G. T. Nepom, J. F. Elliott, D. C. Freed, A. Bansal, S. Zheng, A. Herman, A. Lernmark, D. M. Zaller, et al. 1996. Naturally processed T cell epitopes from human glutamic acid decarboxylase identified using mice transgenic for the type 1 diabetes-associated human MHC class II allele, DRB1*0401. *J. Clin. Invest. 98:2597.*
27. Patel, S. D., A. P. Cope, M. Congia, T. T. Chen, E. Kim, L. Fugger, D. Wherrett, and G. Sonderstrup-McDevitt. 1997. Identification of immunodominant T cell epitopes of human glutamic acid decarboxylase 65 by using HLA-DR(α*0101,β1*0401) transgenic mice. *Proc. Natl. Acad. Sci. USA 94:8082.*
28. Chen, Y. Z., S. Matsushita, and Y. Nishimura. 1996. Response of a human T cell clone to a large panel of altered peptide ligands carrying single residue substitutions in an antigenic peptide: characterization and frequencies of TCR agonism and TCR antagonism with or without partial activation. *J. Immunol. 157:3783.*
29. Misko, I. S., S. M. Cross, R. Khanna, S. L. Elliott, C. Schmidt, S. J. Pye, and S. L. Silins. 1999. Crossreactive recognition of viral, self, and bacterial peptide ligands by human class I-restricted cytotoxic T lymphocyte clonotypes: implications for molecular mimicry in autoimmune disease. *Proc. Natl. Acad. Sci. USA 96:2279.*
30. Masuda, M., S. Senju, S. Fujii Si, Y. Terasaki, M. Takeya, S. Hashimoto Si, K. Matsushima, E. Yumoto, and Y. Nishimura. 2002. Identification and immunocytochemical analysis of DCNP1, a dendritic cell-associated nuclear protein. *Biochem. Biophys. Res. Commun. 290:1022.*
31. Fujita, H., S. Senju, H. Yokomizo, H. Saya, M. Ogawa, S. Matsushita, and Y. Nishimura. 1998. Evidence that HLA class II-restricted human $CD4^+$ T cells specific to p53 self peptides respond to p53 proteins of both wild and mutant forms. *Eur. J. Immunol. 28:305.*
32. Padovan, E., G. Casorati, P. Dellabona, S. Meyer, M. Brockhaus, and A. Lanzavecchia. 1993. Expression of two T cell receptor α chains: dual receptor T cells. *Science 262:422.*
33. Busch, R., C. M. Hill, J. D. Hayball, J. R. Lamb, and J. B. Rothbard. 1991. Effect of natural polymorphism at residue 86 of the HLA-DR β chain on peptide binding. *J. Immunol. 147:1292.*
34. Rammensee, H. G., T. Friede, and S. Stevanoviic. 1995. MHC ligands and peptide motifs: first listing. *Immunogenetics 41:178.*
35. Vergelli, M., B. Hemmer, M. Kalbus, A. B. Vogt, N. Ling, P. Conlon, J. E. Coligan, H. McFarland, and R. Martin. 1997. Modifications of peptide ligands enhancing T cell responsiveness imply large numbers of stimulatory ligands for autoreactive T cells. *J. Immunol. 158:3746.*
36. Grogan, J. L., A. Kramer, A. Nogai, L. Dong, M. Ohde, J. Schneider-Mergener, and T. Kamradt. 1999. Cross-reactivity of myelin basic protein-specific T cells with multiple microbial peptides: experimental autoimmune encephalomyelitis induction in TCR transgenic mice. *J. Immunol. 163:3764.*
37. Maier, B., M. Molinger, A. P. Cope, L. Fugger, J. Schneider-Mergener, G. Sonderstrup, T. Kamradt, and A. Kramer. 2000. Multiple cross-reactive self-ligands for *Borrelia burgdorferi*-specific HLA-DR4-restricted T cells. *Eur. J. Immunol. 30:448.*
38. Hiemstra, H. S., P. A. van Veelen, N. C. Schloot, A. Geluk, K. E. van Meijgaarden, S. J. Willemen, J. A. Leunissen, W. E. Benckhuijsen, R. Amons, R. R. de Vries, et al. 1998. Definition of natural T cell antigens with mimicry epitopes obtained from dedicated synthetic peptide libraries. *J. Immunol. 161:4078.*
39. Tanaka, Y., H. Ohyama, M. Ogawa, Y. Nishimura, and S. Matsushita. 1999. Identification of peptide superagonists for a self-K-*ras*-reactive $CD4^+$ T cell clone using combinatorial peptide libraries and mass spectrometry. *J. Immunol. 162:7155.*
40. Judkowski, V., C. Pinilla, K. Schroder, L. Tucker, N. Sarvetnick, and D. B. Wilson. 2001. Identification of MHC class II-restricted peptide ligands, including a glutamic acid decarboxylase 65 sequence, that stimulate diabetogenic T cells from transgenic BDC2.5 nonobese diabetic mice. *J. Immunol. 166:908.*
41. Hemmer, B., M. Vergelli, B. Gran, N. Ling, P. Conlon, C. Pinilla, R. Houghten, H. F. McFarland, and R. Martin. 1998. Predictable TCR antigen recognition based on peptide scans leads to the identification of agonist ligands with no sequence homology. *J. Immunol. 160:3631.*
42. Hennecke, J., A. Carfi, and D. C. Wiley. 2000. Structure of a covalently stabilized complex of a human $\alpha\beta$ T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J. 19:5611.*
43. Esnouf, R. M. 1997. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model 15:132.*
44. Merritt, E. A., and M. E. P. Murphy. 1994. Raster3D version-2.0: a program for photorealistic molecular graphics. *Acta Cryst. D50:869.*
45. Texier, C., S. Pouvelle-Moratille, M. Busson, D. Charron, A. Menez, and B. Maillere. 2001. Complementarity and redundancy of the binding specificity of HLA-DRB1, -DRB3, -DRB4 and -DRB5 molecules. *Eur. J. Immunol. 31:1837.*
46. Clot, F., C. Gianfrani, M. C. Babron, F. Bouguerra, S. Southwood, M. F. Kagnoff, R. Troncone, S. Percopo, J. F. Eliaou, F. Clerget-Darpoux, A. Sette, and L. Greco. 1999. HLA-DR53 molecules are associated with susceptibility to celiac disease and selectively bind gliadin-derived peptides. *Immunogenetics 49:800.*
47. Geluk, A., K. E. van Meijgaarden, S. Southwood, C. Oseroff, J. W. Drijfhout, R. R. de Vries, T. H. Ottenhoff, and A. Sette. 1994. HLA-DR3 molecules can bind peptides carrying two alternative specific submotifs. *J. Immunol. 152:5742.*
48. Baum, H., V. Brusic, K. Choudhuri, P. Cunningham, D. Vergani, and M. Peakman. 1995. MHC molecular mimicry in diabetes. *Nat. Med. 1:388.*
49. Akerblom, H. K., and M. Knip. 1998. Putative environmental factors in Type 1 diabetes. *Diabetes Metab. Rev. 14:31.*
50. Taha, M. K., A. E. Deghmane, A. Antignac, M. L. Zarantonelli, M. Larribe, and J. M. Alonso. 2002. The duality of virulence and transmissibility in *Neisseria meningitidis. Trends Microbiol. 10:376.*
51. Giebink, G. S. 2001. The prevention of pneumococcal disease in children. *N. Engl. J. Med. 345:1177.*
52. Steidler, L., W. Hans, L. Schotte, S. Neirynck, F. Obermeier, W. Falk, W. Fiers, and E. Remaut. 2000. Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science 289:1352.*
53. Kimpimaki, T., M. Erkkola, S. Korhonen, A. Kupila, S. M. Virtanen, J. Ilonen, O. Simell, and M. Knip. 2001. Short-term exclusive breastfeeding predisposes young children with increased genetic risk of type I diabetes to progressive β-cell autoimmunity. *Diabetologia 44:63.*