

# Identification and motif determination of naturally processed peptides from the murine Type I diabetes associated MHC class I allele H-2K<sup>d</sup>

James J Walters<sup>1,2</sup>, Anish Suri<sup>2</sup>, Emil R. Unanue<sup>2</sup> and Michael L. Gross<sup>1</sup>

1 - Center for Biomedical and Biorganic Mass Spectrometry, Washington University, St. Louis, MO 63130

2 - Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63130

## OVERVIEW

- Type I diabetes (T1D) in humans and in the non-obese diabetic (NOD) mouse is a result of a T cell-mediated autoimmune process directed against pancreatic islet  $\beta$ -cells.
- Most of the diabetogenic CD8<sup>+</sup> T cells in NOD mice are restricted to the class I MHC allele, H-2K<sup>d</sup>. The majority of literature on H-2K<sup>d</sup> bound peptides has been elucidated by peptide-H-2K<sup>d</sup> complex's ability to stimulate T cells *in vitro*, not naturally processed sequences of self-peptides presented on the cell surface.
- Using LC/LC-MS/MS, we provide here a spectrum of naturally processed peptides presented by H-2K<sup>d</sup>.
  - ➔ The majority of peptides are 8-10 amino acids in length, however, we identified a few longer peptides up to 18 amino acids.
  - ➔ Approximately 80% of the peptides have a motif defined by a Tyr or a Phe at P2 and a small hydrophobic residue, usually Ile/Leu or Val, at P5.
  - ➔ Mutational binding analysis confirmed this motif, as changing the P2 Tyr or Phe to an Ala abolished binding.
  - ➔ Interestingly, the longer peptides (>10mers) also bind to H-2K<sup>d</sup>, and this was affected by both the anchor residues and/or the peptide length.

## INTRODUCTION

- Type I diabetes (T1D) in humans and in the non-obese diabetic (NOD) mouse is a result of a T cell-mediated autoimmune process directed against pancreatic islet  $\beta$ -cells.
- Class I MHC molecules present peptides to CD8<sup>+</sup> T lymphocytes. Most class I variants present peptides 8-10 amino acids in length due to the closed ends of the binding groove (Figure 1).
- Most of the diabetogenic CD8<sup>+</sup> T cells in NOD mice are restricted to the class I MHC allele, H-2K<sup>d</sup>.
- The majority of literature on H-2K<sup>d</sup> bound peptides has been elucidated by peptide-H-2K<sup>d</sup> complex's ability to stimulate T cells *in vitro*<sup>1-4</sup>, not naturally processed sequences of self-peptides presented on the cell surface<sup>5</sup>.
- Past work to identify the sequence of self-peptides presented on the cell surface by other MHC alleles has offered insight into motif determination, defined allele differences, deciphered the determinant capture hypothesis, and offered potential to provide putative auto-immune peptides<sup>6</sup>.

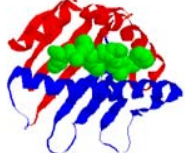


Figure 1. Space filling model of a typical class I MHC. The alpha 1 and alpha 2 domains form a binding groove with closed ends to interact with a peptide which is usually 8-10 amino acids long. Picture taken from www.immuno.path.cam.ac.uk

## METHODS

- **Peptide Isolation and Identification**
  - ➔ N6-1, an insulinoima, and NOD.C3, an APC line generated from NOD splenocytes were grown in large numbers (5-10 x10<sup>6</sup> cells).
  - ➔ H-2K<sup>d</sup>-peptide complexes were affinity purified using a SF1.1.1 monoclonal antibody conjugated to Sepharose beads.
  - ➔ Peptides were acid-eluted and separated by off-line strong-cation exchange chromatography.
  - ➔ Ten eluted fractions were then analyzed by on-line capillary HPLC nanoESI-MS and MS/MS using an LCO-Deca quadrupole ion-trap mass spectrometer.
  - ➔ Product ion spectra were submitted to database searching using SEQUEST.
  - ➔ All spectra were manually inspected to confirm identification.
- **Binding Analyses**
  - ➔ The cell line T-2K<sup>d</sup> was utilized which:
    - has an intrinsic defect in peptide loading onto the class I molecules.
    - has very few stable class I molecules on the cell surface.
    - expresses H-2K<sup>d</sup> on it's cell surface which is stabilized and upregulated by the binding of a high affinity peptide.
  - ➔ Synthetic peptides corresponding to identified naturally processed sequences (or mutated sequences) were mixed with T-2K<sup>d</sup> cells at varying concentrations.
  - ➔ The cells were then stained with a biotinylated anti-H-2K<sup>d</sup> (SF1.1.1) monoclonal primary antibody.
  - ➔ Binding of peptide was determined with Phyco-erythrin (PE)-labeled streptavidin and reported by mean fluorescence intensity as measured by FACS

## RESULTS and DISCUSSION

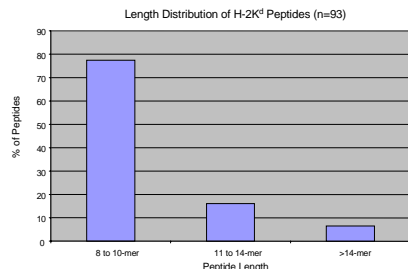


Figure 2. Of the 93 peptides identified by LC-MS/MS, approximately 21% contained greater than 10 amino acids, which is the approximate length of the MHC class I binding groove. This would indicate that the longer peptides may form a loop structure at the center of the peptide<sup>7</sup>, or the peptide could protrude from either the carboxy- or amino terminal<sup>8</sup>. Both of the models would allow the binding dependent amino acids to maintain interaction with the class I molecule.

## Product-ion spectrum of the peptide GFGDLKTPAGLQVLND

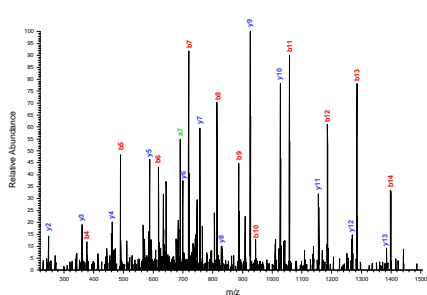


Figure 3. Sequel analysis of the MS/MS spectrum of the doubly charged parent ion of m/z 823.0. The Xcorr score is 5.16 and the  $\Delta$ score is 0.30. Note the presence of a substantial b ion series and y ion series. The protein source was identified as eukaryotic translation elongation factor beta. This shows the ability of H-2K<sup>d</sup> to bind peptides of greater than 10 amino acids.

## Amino acid distribution of the non-anchor residues of H-2K<sup>d</sup>

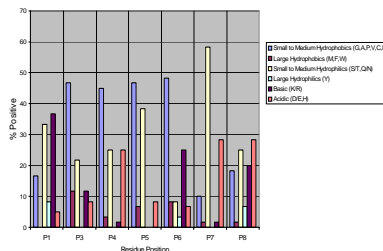


Figure 4. Amino acid distribution of non-anchor residues on naturally processed peptides they did not identify by LC-MS/MS. There is not any consensus at any of these positions, suggesting that they do not play a major role in peptide binding to the MHC.

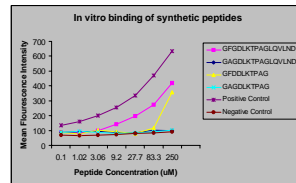


Figure 5. *In vitro* binding of select synthetic peptides. The **red** peptide binds with the greatest affinity. The **yellow** binds, however with weaker affinity indicating that the extension of the carboxy terminus contributes to the interaction with the MHC molecule. The other two peptides with the F at P2 changed to A, lose the ability to bind regardless of the length of the peptide. This suggests that the peptide is bound to the expected register with F at P2 and a small hydrophobic residue at P5.

## Amino Acid Sequence Binding (nM)

Amino Acid Sequence	EXPT.1	EXPT.2
AMKRHGLDN (Neg control)	NB <sup>a</sup>	NB
KYNKANWFL (Pos control)	30.0	30.0
NYGPMKGGSFGG	30.0	27.0
NAGPMKGGSFGG	NB	NB
NYGPMKGGG	30.0	17.0
NAGPMKGGG	NB	NB
GFGDLKTPAGLQVLND	30.0	25.0
GAGDLKTPAGLQVLND	NB	NB
CGFGDLKTPAG	70.0	40.0
GAGDLKTPAG	NB	NB
IFIKPGADLSTGHDEL	170.0	250.0
IAIKPGADLSTGHDEL	NB	NB
IFIKPGADLS	~1000.0	250.0
IAIKPGADLS	NB	NB
KYGEVTNLL	8.0	4.0
KAGEVTNLL	NB	NB
SYLEMGGHDI	0.5	0.7
SALEMGGHDI	-500.0	150.0
KYLTKVDYL	23.0	5.5
KALTKVDYL	-1000.0	NB
KYHSANVL	32.0	25.0
KAHSANVL	NB	45.0
VYNASNELL	15.0	8.0
KYKASENAI	8.0	7.0
KYMEDVTQI	2.0	1.0
KYKDIYTEL	9.0	5.0

Table 1. Synthetic peptide binding analysis to H-2K<sup>d</sup>, presented as the peptide concentration necessary for a 3X increase of H-2K<sup>d</sup> expression on the cell surface. Peptides shown in blue were identified by LC-MS/MS. Other variants of these peptides were tested to determine if peptide length played a role in binding, and/or if a mutation at the putative P2 site would abolish binding. Changing P2 to an A, greatly decreases or completely prevents binding to H-2K<sup>d</sup>. Interestingly, shortening the peptide had either no effect on binding or only slightly decreased binding.

<sup>a</sup>NB = a non-binder

## CONCLUSIONS

- The majority (~80%) of the naturally processed peptides presented by the Type I diabetes associated MHC class I allele, H-2K<sup>d</sup>, display a consensus motif with a Tyr or a Phe at P2 and a small hydrophobic residue at P5, and are 8-10 amino acids in length. This corresponds to previous published reports<sup>9</sup>.
- Interestingly, approximately 20% of the peptides identified were longer than 10 amino acids. This was confirmed by synthetic peptide binding analyses. This suggests that H-2K<sup>d</sup> could bind longer peptides *in vivo* and play an undetermined role in the onset of Type I diabetes.
- When the longer peptides are mutated by carboxy-terminal truncation, binding is only marginally inhibited if at all. Additionally, altering the P2 site from a Tyr/Phe to an Ala always abolishes binding. Together, these results indicate that the longer peptides are binding in the same register with the shorter peptides, perhaps by forming internal loop structures and/or with carboxy-terminal overhangs.

## FUTURE DIRECTIONS

- Further characterize the interaction of peptides greater than 10 amino acids by *in vitro* determination of peptide off-rates.
- Define the presence or absence of *in vivo* CD8<sup>+</sup> T-cells which are restricted to longer peptides (>10 amino acids) presented by H-2K<sup>d</sup> in the NOD mouse after immunization.
- Analyze and correlate X-ray crystal data of H-2K<sup>d</sup> with peptides which are 8-10 amino acids in length and those which are longer.
- Explore the mechanism by which longer peptides interact with H-2K<sup>d</sup> by using H/D exchange and mass spectrometry.

## REFERENCES

- Wong et al., *PNAS*, 99 (8) 5551, 2002.
- Jiang et al., *J Gen Virol*, 83, 429, 2002.
- Fan et al., *J Immunol*, 164, 1689, 2000.
- Falk et al., *Nature*, 351, 290, 1991.
- Suri et al., *PNAS*, 100 (9) 5330, 2003.
- Fremont et al., *Science*, 257, 919, 1992.
- Stylianou et al., *Eur J Immunol*, 30, 3099, 2000.

## ACKNOWLEDGMENTS

This research supported by NIH National Center for Research Resources Grant P41RR00954.