

## ABSTRACT

The absence of mutations in the desmuslin gene did not impact its function. Nevertheless, the single-nucleotide polymorphisms mapped in this study are highly disequilibrated and can be used for disambiguation studies of this region of chromosome 15q26.3.

## INTRODUCTION

MHC molecules are responsible for presenting antigens to T cells, which are small peptides with amino acid sequence restrictions that affect the cellular immune response. To understand the immune response, it is essential to know the regulations for identifying peptides that can be provided by specific MHC alleles. As peptide receptors, major histocompatibility complex (MHC) molecules can attach to various peaks of these molecules, but each MHC allele exhibits a preference for sets of notably conserved sequences with unique features. Crystallographic studies of MHC-peptide complexes and sequence analysis of peptides eluted from these complexes reveal that MHC genetic polymorphism is responsible for the differences in peptide binding between MHC molecules. Our argument is that the selection of peptides by a particular MHC allele can be expressed as influenced by the MHC sequence, rather than the conventional method where the allele itself is dependent on the candidate peptide. By using amino acids that occupy the different positions in the MHC molecule as arguments for this function instead of referenced to an allele, it is possible to extend the generalizations from available elution data to include cases where no empirical data is yet found. In the past decade, there has been a surge of data on peptide binding affinity towards the same MHC molecule (much more abundant than predicted). Are the data derived from crystallographic and chemical data solely relevant to understanding the T cell-specific peptide requirements of the different MHC alleles? The challenge is to find a suitable theoretical framework for investigating and analyzing the sequences of these MHC-peptide complexes. Geometry provides the theoretical basis for this, as we use it to describe vector amino acid sequences in a metric space. This means that two sequenced amino acids are at approximately the same distance from each other. Our analytical tools can manipulate these transformations to study the relationships among MHC-peptide complexes. By utilizing the duality principle in geometry and dual spaces, we can shift from measuring the distances between amino acid sequences and the sequence properties of MHC-peptide complexes to measuring their correlation. The data in this context demonstrates that the sequence of an MHC molecule affects the range of sequence restrictions for the peptide bound to the same structure. This is demonstrated through proximity studies of specific MHC molecules and their sequence-binding proxies. In this research we focus on the class I human leucocyte antigen (HLA) molecules and nonameric peptides bound to them. Where  $S$  is the space for these complexes, the algebraic and geometric structure of MHC-peptide complex has been determined. In this space, the MHC-peptide complex is characterized by a point that corresponds to its vector amino acid sequence properties. The co-ordinates of this point are determined by these same properties, as well as other peptide and MHC (specifically polymorphic positions in the respective MHC domains). Peptide binding data facilitates the population of this space with points. These points are represented by matrix  $M$  of size  $n$ , where  $a$  is the space's

dimensionality, and each row contains the vector representing each point. In this study, 2535 nonameric peptides that bind alleles of the HLA-A1 family, as well as other alleles from different regions of Mexico, were identified in the MHCPEP database. A property of the type MHC-peptide complex is represented by each point in this space. The matrix  $M'$  is transposed to a vector of  $M$ , with each point represented by  $n$  points in  $S'$ . Let  $x$  be defined as "[ ]" having amino acid D in peptide position 3',  $y$  as [ ],  $Y$  in penicillin antibiotics 3-4" and  $z$  as (in space  $S$ ) 2). Defining the MHC-peptide complexes  $a, b, c, d$  (points in space  $S$ ) as: [HLA-A\*0101-IADMGHLKY], hLA-1-0102-STEPVNILY], HLA-2202-2042-TYSAGIVQI], and Hla-DOHOH2+DVPTAAV can help us to define these complex points. Leaving aside other points and dimensions, space  $S$  for MHC-peptide complex is defined as:  $a(1,0,1)$ ,  $b(0.1,1)$ ,  $c(0.0;0)$ , where each point is identified by its irrational lengths, and the values of these length are used to determine whether he or she has '1' or not ('0'). The matrix can represent space  $S$ , with each row containing a point in the MHC-peptide complex located at its three coordinates. Additionally, the transposed space of sequence properties is defined as  $S' = x, 1,0, 0, 10,00 y, z, 11,0$  and  $0$ , and it can be represented by the matrix with four coordinate values. The transposed space of space  $S'$  is the matrix of "space" that is shown as  $s'$ , which is known as the corresponding inverse of peptide  $i$ . This means that, in this example, if we take our subject in the first sentence, then it is our predicate who takes our object in second sentence; and of course, dualism: as Ramsey points out there is no sense in distinguishing 'individual' from having 'quality' or 'particular' which has an equal proportion, since the two types are always symmetrical. Translating the point  $S$  into  $S'$  by transposing it allows for measurement of distance between points and sequence properties of MHC-peptide complexes, specifically focusing on distance (and indeed time) between these points. The degree to which the MHC molecule and its sequence characteristics are similar or different could provide insight into how peptides bind to MH molecules and how MHA alleles dictate their sequence requirements for binding candidates. Our data unit is the "MHC-peptider complex," which includes two types of categorical variables. Firstly, there are two types of variables:  $Paa$ , which is found at position  $Pp$  in the peptide-these are called 'p' and the set of all those amino acids in it-Hydro acidhyde ('haa') at location  $Hp$  within the MHC molecule- these are known as essentially different variables from each other and are named correspondingly with  $Haa$ . To clarify, a partition of space  $S'$  is used to separate the class of peptide sequence properties ( $P$ ) from the group of MHC sequence characteristics ( $H$ ), resulting in  $K = P, H, S$  means  $P - H$  and  $P + H = 0$ . Bipolarity is a characteristic of MHC sequence properties, where the amino acid threonine is present in position 9 of the peptide, and only polymorphic positions are considered in the alpha 1 and 2 domains. Hence, nonamers possess 180 sequence properties, one for each of the 20 amino acids at nine positions, while MHC alleles have 207 sequence characteristics for every possible amino acid at 73 polymorphic positions in alpha-1 and alpha-2 domains. At every polymorphic site, only a few amino acids are present; for example, G, Q, and R are found in multiple alleles at position 65, while A and D are absent in positions 90 and C. The Cartesian product  $HP$  into  $R$  (the line of real numbers) provides the sequence properties of MHC molecules and peptides bound to them by determining the distance between each pair  $(h, p)$ . A distance measure is used to define the identity of a function that takes two inputs: an HLA sequence property and peptide property. In  $S'$ -based calculations, the separation between two properties-points or vectors-can be determined using different approaches. The closest alternative is the Ochiai similarity index, which is defined as:  $a = hp + (h) + ph$ ;  $b: [h] + P (h/p)$ ;  $c: [h], 2(r), 3 (d), 4 (k) * 0$ . To put it

