

Teoría de unión al HLA clase II: teoría de probabilidad, combinatoria y entropía aplicadas a secuencias peptídicas

Javier Rodríguez

Universidad Nacional de Colombia, Departamento de Cirugía, Bogotá, Colombia.

BINDING TO CLASS II HLA THEORY: PROBABILITY, COMBINATORY AND ENTROPY APPLIED TO PEPTIDE SEQUENCES

Recibido: 15 Noviembre 2007

Aceptado: 29 Septiembre 2008

RESUMEN

La presentación de péptidos antigenicos por las moléculas de histocompatibilidad a los receptores T es fundamental en la especificidad y la variabilidad de la defensa por parte del sistema inmune, por lo tanto un entendimiento físico y matemático de este fenómeno es de vital importancia para el diseño de vacunas.

Se calcularon valores de probabilidad, combinatoria y entropía de 100 secuencias específicas de péptidos teóricos, promiscuos, sintéticos y naturales que se unen a moléculas de HLA clase II y 61 péptidos que no se unen. Estos valores se usaron para diferenciar los péptidos seleccionados y desarrollar una caracterización física y matemática, objetiva y reproducible para el fenómeno de unión.

Se desarrolló una predicción teórica del conjunto apropiado de secuencias que debe tener un péptido para que sea presentado por la molécula de HLA clase II, acertando en todos los casos.

La teoría desarrollada facilita el proceso de selección de los péptidos requeridos en el diseño de vacunas sintéticas o naturales.

PALABRAS CLAVE: Probabilidad/ Combinatoria/ Entropía/ Unión/ HLA clase II.

ABSTRACT

The antigenic peptide presentation by the histocompatibility molecules to the T receptors is fundamental in the specificity and variability of the host defense by the immune system, therefore a physical and mathematical understanding of this phenomenon is of vital importance for the design of vaccines.

Values of probability, combinatory and entropy from 100 specific sequences of theoretical, promiscuous, synthetic and natural peptides that bind to HLA class II molecules, as well as 61 peptides that do not bind were calculated. These values were used to differentiate the selected peptides and to develop an objective and reproducible physical and mathematical characterization of the binding phenomena.

A theoretical prediction of the appropriate group of sequences that a peptide should have so that it is presented by the HLA class II molecule was developed, hitting all the cases.

The developed theory facilitates the process of peptide selection required in the design of synthetic or natural vaccines.

KEY WORDS: Probability/ Combinatory/ Entropy/ Binding/ HLA class II.

INTRODUCCIÓN

La probabilidad es una medida predictiva de la posible ocurrencia de un evento, cuyo valor numérico se encuentra en el intervalo real $[0, 1]$. Kolmogorov definió la función de probabilidad como aquella que satisface tres axiomas elementales; la probabilidad de cualquier elemento del espacio muestral es mayor o igual a cero, la probabilidad de todo el espacio muestral es uno, y la probabilidad de la unión de eventos disyuntos es la suma de sus probabilidades⁽¹⁻⁴⁾. Esta teoría se usó para cuantificar las diferentes posibilidades de secuencias de péptidos teóricos y/o con capacidad o no de unión.

En el desarrollo de la Termodinámica y de la Mecánica Estadística, el concepto de entropía se ha ido reformulando y reinterpretando⁽⁵⁻⁸⁾. La formulación enunciada por Boltzmann está dada en términos del número de microestados y se aplica, entre otros, a los sistemas con una distribución equiprobable^(6,7). En este trabajo se estudia la aparición de aminoácidos en secuencias de péptidos que podrían unirse o no al HLA de clase II, asumiendo que cada uno de ellos tiene la misma probabilidad de conformar el péptido; por ello, en analogía con los microestados de la mecánica estadística, la aplicación de la formulación de Boltzmann del concepto de entropía lleva a una medida del sistema que caracteriza la unión peptídica al HLA de clase II.

En un sistema de muchos cuerpos se definen los conceptos de macroestado y microestado. El macroestado es una configuración general del sistema que se puede manifestar mediante múltiples microestados posibles. Por ejemplo, dado un número de partículas en un volumen fijo existen múltiples configuraciones con densidad uniforme, la densidad uniforme es una característica del sistema que se puede expresar a través de diferentes microestados, que serían los diferentes arreglos posibles de organización de las partículas en el volumen^(6,7). En la presente investigación se usa una analogía para representar las secuencias de aminoácidos en términos de micro y macroestados (ver Definiciones).

El antígeno leucocitario humano, HLA, es un complejo de proteínas especializadas cuya función es la de presentar asociados a células los antígenos de los microorganismos para su reconocimiento por las células T. Las moléculas de HLA clase I presentan péptidos a las células T CD8+, mientras que las moléculas de HLA clase II presentan péptidos a las células T CD4⁽⁹⁾. Cada molécula de HLA está anclada a la célula mediante dominios transmembrana y citoplasmáticos y posee una hendidura extracelular de unión al péptido seguida de un par de dominios tipo inmunoglobulina⁽¹⁰⁾.

En las moléculas de HLA clase II la hendidura de unión se presenta abierta en los dos extremos⁽¹¹⁻¹⁴⁾; debido a esto los péptidos presentados poseen un grado de variabilidad

en su tamaño, comprendido entre 10 y 24 residuos; un tamaño de 13 a 16 aminoácidos ha sido observado con mayor frecuencia⁽¹⁵⁻¹⁸⁾. Análisis de unión sugieren que únicamente una región central compuesta por nueve aminoácidos es esencial para la unión péptido-HLA clase II⁽¹⁹⁾, estudios de dicha región han relacionado una alta frecuencia de aparición de un grupo de aminoácidos en una posición específica, posiciones de anclaje, con los denominados motivos de unión⁽²⁰⁾, construidos a partir de grupos seleccionados, por ejemplo, de librerías M13 de péptidos presentados.

El objeto de estudio de la presente investigación es un sistema de 20 aminoácidos que se organizan en 9 espacios de ocupación en la región central de péptidos antigenicos; y el propósito es construir una teoría general basada en las leyes de probabilidad, combinatoria y entropía, capaz de predecir la unión de péptidos a moléculas de HLA clase II.

DEFINICIONES

Macroestados y microestados de las secuencias de aminoácidos (*Definición del autor*)

Macroestado: secuencias de péptidos elegidos para diferenciar matemática y físicamente la capacidad de unión al HLA clase II, y aquellos que tengan el valor predicho para poder unirse. Por lo que existen dos clases de macroestados, los de unión y los de no unión.

Microestado: cualquier secuencia específica de nueve aminoácidos.

Tipo de secuencia: microestados que presentan el mismo valor en su combinatoria y su evaluación con los criterios definidos. Definimos 7 tipos de secuencia iniciales. Dos de estos son representativos del macroestado de no unión, mientras que cinco son péptidos representativos del macroestado de unión.

Probabilidad Laplaciana

La probabilidad de un tipo de secuencia A se define como el cociente entre la cantidad de microestados N_A asociados a este tipo de secuencia y el total de posibles microestados $N^{(1,4)}$.

$$P(A) = \frac{\text{Microestados asociados al tipo de secuencia A}}{\text{Todos los microestados}} = \frac{N_A}{N}$$

Ecuación 1

Ley combinatoria para determinar la cantidad de microestados W

En la construcción del péptido, cada posición tiene asociada un número de posibles aminoácidos. La cantidad

TABLA I. Aminoácidos según motivos y grupos

Posición o grupo	4	6	7	9	Grupo cargados	Grupo 1	Grupo 2	Grupo 3
Residuos	M	A	L	L	D	Y	A	G
	L	G	Q	M	E	F	T	C
	A	S	M	A	H	W	Q	S
	V	T	N		K	M	N	P
		R			R	V		
		K				I		
Valor	4	6	4	3	5	7	9	4

de posibles microestados asociados a un tipo de secuencia se calcula multiplicando los valores asociados a cada posición, lo cual depende del número de posibles aminoácidos asociados a cada lugar; esto está determinado por el grupo al que pertenezca y sus repeticiones en el tipo de secuencia^(3, 4). Para las cuantificaciones específicas, ver tablas de criterios evaluadores y tablas de grupos de aminoácidos (Tablas I-IV).

Entropía

En un sistema cuyos microestados son equiprobables la entropía está dada por:

$$S = k \ln(W)$$

Ecuación 2

Donde k es igual a la constante de Boltzmann (1.38×10^{-23} J/k), W los posibles microestados y S el valor de la entropía^(6, 7). Se eligió la formulación más simple para mayor sencillez de la teoría.

MATERIAL Y MÉTODOS

El objeto de estudio de la presente investigación es el sistema de 20 aminoácidos clasificados según características químicas, motivos y grupos, que pueden ser organizados en 9 posibles lugares de ocupación en la región central del HLA clase II. Los eventos posibles de posible ocupación son estudiados con las probabilidades de aparición de cada uno de los 20 aminoácidos y sus posibles formas de combinación en la región central asociadas o no a la unión. Teóricamente es casi improbable esperar que se repita un mismo aminoácido 9 veces seguidas en la naturaleza; mientras que evaluar péptidos no repetitivos que se unan a diferentes alelos de HLA, como los péptidos promiscuos, son más probables, evaluando así existencia o inexistencia de secuencias de péptidos y que se unen o no a la molécula presentadora.

Este trabajo cuantificará fundamentalmente la región central de unión⁽¹⁹⁾. Inicialmente se define un espacio muestral,

TABLA II. Criterios evaluadores a partir de los cuales se calcula el número de microestados equivalentes para un tipo de secuencia específico

CRITERIO EVALUADOR
1 Si un péptido presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1 y en la segunda una Arginina (R) estas dos posiciones toman un valor de 3.
2 Si un péptido no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1 esta posición toma un valor de 20. Si además en la segunda posición no presenta una R y es diferente al aminoácido de la posición uno, esta posición toma un valor de 20.
3 Si la secuencia tiene un aminoácido cargado en la última posición y uno cargado en la posición 7 o en la 8 el número de microestados se multiplica por 13.
4 Si se cumple el primer criterio, el péptido presenta 5 repeticiones y en la posición cuatro presenta Metionina, Leucina, Valina o Alanina, M, L, V, A, el número de microestados se multiplica por 20.
5 Si un péptido no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición uno, en la segunda posición no presenta una R, y la secuencia tiene mas de 3 repeticiones el número de microestados se divide entre 20.
6 Si un péptido no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1, en la segunda posición no presenta una R, y la secuencia tiene menos de 4 repeticiones el número de microestados se multiplica por 10.
7 Si un péptido presenta una Asparagina, N, en la posición 1, en la segunda posición una R, presenta una L, en una de las posiciones restantes y no tiene repeticiones, la posición uno toma un valor de 6, la dos de 3 y el número de microestados se divide entre 8.
8 Si la secuencia tiene 4 de los aminoácidos hidrofóbicos característicos de unión de la posición 1 y 3 repeticiones el número de microestados se multiplica por 3.
9 Si la secuencia tiene más de 4 de los aminoácidos hidrofóbicos característicos de unión de la posición 1 y menos de 2 repeticiones el número de microestados se divide en 3.
10 Si la secuencia tiene más de 4 repeticiones el número de microestados se divide entre 10.

TABLA III. Aminoácidos según sus características químicas

Grupo	Aminoácidos	Valor
Cargados	E D	H R K
	Negativos	Positivos
Hidrofóbicos de alta unión	W F Y	I V
	Aromáticos	
Puentes de hidrógeno	S T C Y N Q	
Grupo 4	G P M L A	
		2 Negativos
		3 Positivos
		5
		6
		5

donde cada microestado es considerado como un posible evento. Se seleccionaron siete tipos de secuencias para el análisis: un péptido teórico que tiene el mismo aminoácido en todas sus posiciones; la región central del péptido de la Influenza Hemaglutinina, HA₃₀₉₋₃₁₇, que se une a la mayoría de los alelos DR^(21, 22); dos péptidos artificiales que contienen tres y cinco aminoácidos repetidos en las últimas posiciones⁽²³⁾; la región central de la cadena invariante del péptido asociada a la clase II, CLIP₉₁₋₉₉⁽²⁰⁾; la región central de un péptido de unión con tres lisinas en las posiciones 2, 5 y 7⁽²³⁾; y por último un péptido que contiene aminoácidos diferentes en todas las posiciones de la región central y no se une al HLA clase II⁽²³⁾. El criterio de selección de estos péptidos fue establecido con el fin de obtener un espectro completo de valores diferenciadores de probabilidad combinatoria y entropía de secuencias que se unen de las que no, para caracterizarlos matemática y físicamente. El péptido teórico se escogió para ser contrastado con las combinatorias de aminoácidos naturales y artificiales, los péptidos promiscuos se escogieron por su capacidad de unión a diferentes alelos, los péptidos artificiales y el péptido con tres lisinas, K, para caracterizar la importancia de las repeticiones de aminoácidos respecto a la unión, el último péptido que tiene todos los aminoácidos diferentes y no presenta unión, se escogió para compararlo con las secuencias de unión y con el péptido teórico donde todos los aminoácidos son repetidos, para diferenciar unión de no unión con repeticiones o no de aminoácidos.

Se calculó la probabilidad laplaciana del tipo de secuencia en el que todos los aminoácidos son iguales y se comparó con el valor obtenido para el tipo de secuencia en el cual las tres primeras posiciones están determinadas por un aminoácido, tomando como ejemplo la secuencia de la librería M13 de péptidos presentados⁽²¹⁾ donde las tres primeras posiciones son tres triptófanos (W).

Posteriormente se evaluó la probabilidad laplaciana de cada uno de los restantes tipos de secuencia, aplicando la ley combinatoria (ver definiciones), para calcular los microestados equivalentes a partir de la cantidad de

TABLA IV. Criterios evaluadores a partir de los cuales se calcula el número de microestados equivalentes para un tipo de secuencia específico

CRITERIO EVALUADOR

- 1 Si un péptido presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1 y en la posición cuatro presenta M, L, V, A y en la posición nueve presenta M, L, A y presenta A, G, S, T, R, K en la posición 6 o L, Q, M, N en la posición 7 y la secuencia tiene 3 de los aminoácidos hidrofóbicos característicos de unión de la posición 1 el número de microestados se divide en 2.
- 2 Si un péptido presenta Y, W, F, en la primera posición, R en la segunda, tiene 5 aminoácidos repetidos, y 8 aminoácidos hidrofóbicos, el número de microestados se multiplica por 5.
- 3 Si la secuencia tiene 3 aminoácidos consecutivos repetidos el número de microestados se divide entre 5.
- 4 Si la secuencia tiene 4 aminoácidos consecutivos cargados el número de microestados se divide entre 11.
- 5 Si la secuencia tiene 4 aminoácidos que hacen puentes de hidrógeno, tres o cuatro de los aminoácidos hidrofóbicos característicos de unión de la posición 1 y tres repeticiones el número de microestados se divide entre 12.
- 6 Si un péptido presenta Y, W, F, en la primera posición, R en la segunda, presenta una L y una N en una de las posiciones restantes, y tiene una repetición o ninguna, la posición uno toma un valor de 3, la dos de 3. Si tiene una repetición el número de microestados se divide entre 4, si no tiene repeticiones el número de microestados se divide entre 10.
- 7 Si un péptido presenta una N en la posición 1, en la segunda posición una R, presenta una L en una de las posiciones restantes, y no tiene repeticiones, la posición uno toma un valor de 6, la dos de 3 y el número de microestados se divide entre 8.
- 8 Si la secuencia tiene más de 4 de los aminoácidos hidrofóbicos característicos de unión de la posición 1 el número de microestados se divide en dos.
- 9 Si la secuencia tiene 5 aminoácidos hidrofóbicos y más de 2 aminoácidos que hacen puentes de hidrógeno o más de 2 aminoácidos cargados el número de microestados se divide en cuatro.
- 10 Si un péptido no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1 esta posición toma un valor de 13.

TABLA V. Pasos 1-2: Cálculos de probabilidad y número de microestados

Secuencia	Probabilidad del tipo de secuencia	Número de microestados
Péptido teórico: GGGGGGGGG*	$20/20^9 = 3.9 \times 10^{-11}$	20
CLIP ₉₁₋₉₉ : MRMATPLLM	$5184/20^9 = 1.01 \times 10^{-8}$	5184
Péptido artificial: YRAMAAAAA	$6480/20^9 = 1.26 \times 10^{-8}$	6480
Péptido artificial: YRAMSTAAA	$7776/20^9 = 1.51 \times 10^{-8}$	7776
YKKM E AVKL**	$60480/20^9 = 1.18 \times 10^{-7}$	60480
HA ₃₀₉₋₃₁₇ : YVKQNTLKL	$362880/20^9 = 7.08 \times 10^{-7}$	362880
Aminoácidos diferentes en todas las posiciones: GSVQYPITK	$604'800.000/20^9 = 1.1 \times 10^{-3}$	604'800.000

*Los valores para esta secuencia son calculados teóricamente, ver Anexo 1, y no se aplicaron las reglas del segundo paso

TABLA VI. Pasos 3-4: Cálculos de probabilidad y número de microestados

Secuencia	Probabilidad del tipo de secuencia	Número de microestados
CLIP ₉₁₋₉₉ : MRMATPLLM	$1080/20^9 = 2.10 \times 10^{-9}$	1080
Péptido artificial: YRAMAAAAA	$1500/20^9 = 2.92 \times 10^{-9}$	1500
Péptido artificial: YRAMSTAAA	$1800/20^9 = 3.51 \times 10^{-9}$	1800
YKKM E AVKL**	$1800/20^9 = 3.51 \times 10^{-9}$	1800
HA ₃₀₉₋₃₁₇ : YVKQNTLKL	$36000/20^9 = 7.03 \times 10^{-8}$	36000

aminoácidos que tendrían la posibilidad de ocupar cada posición. El número de aminoácidos asociados a cada posición depende del valor del grupo al que pertenezca y sus repeticiones en el tipo de secuencia. Para estos cálculos se realizaron cuatro pasos; en el primero se definieron 8 grupos de aminoácidos: los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1⁽²³⁾; aminoácidos cargados; los aminoácidos A, T, Q, N; un grupo con los aminoácidos G, C, S, P; y finalmente motivos reconocidos para posiciones 4, 6, 7 y 9 según resultados experimentales⁽²¹⁾ (Tabla I). El segundo paso consiste en evaluar diez criterios basados en las tres leyes teóricas y en observaciones experimentales (Tabla II). En el paso 3 se evalúan los péptidos que en el paso 1-2 presentaron el macroestado asociado a la unión, definiendo 4 grupos de aminoácidos: aminoácidos cargados, aminoácidos que hacen puentes de hidrógeno, 3 aminoácidos hidrofóbicos de baja unión⁽²³⁾ junto con la glicina y la prolina, y 5 aminoácidos hidrofóbicos característicos de alta unión en la posición 1⁽²³⁾ (Tabla III). El cuarto paso consiste en evaluar diez criterios basados en las tres leyes teóricas y en observaciones experimentales (Tabla IV).

Los pasos tres y cuatro, aplicados a los macroestados de unión resultados de la aplicación de los pasos 1 y 2, se usaron para refinar los cálculos.

El procedimiento de cálculo para los pasos uno y tres es asignar a cada posición un número según el aminoácido encontrado en cada lugar dependiendo del

conjunto en el que se encuentre. Se realiza el mismo procedimiento con cada aminoácido que aparece consecutivamente, considerando además la disminución de sus posibilidades en la medida que van apareciendo aminoácidos de un mismo grupo y las repeticiones de un mismo aminoácido. Posteriormente, para calcular la probabilidad laplaciana de un tipo de secuencia (ver definiciones), se multiplican dichos valores hallando el número de microestados asociados al tipo de secuencia (Tablas V y VI). La metodología desarrollada es exemplificada para cuatro secuencias específicas (Anexo 1).

Todas las posibles combinaciones de aminoácidos generan secuencias específicas que pueden ser asociadas al fenómeno de unión, o no. Si se considera que en cada lugar de ocupación cada aminoácido tiene la misma probabilidad de aparición, y que basados en datos de la biología molecular damos valores diferentes para aminoácidos característicos de la unión en ciertas posiciones, y se evalúa cada secuencia específica como un microestado (ver definiciones), es posible caracterizar matemáticamente con la ley de entropía todas las posibles asociaciones de aminoácidos y relacionarlas con unión al HLA o no.

Partiendo del número de microestados posibles se realizaron los cálculos de entropía para cada tipo de secuencia aplicando la ley de Boltzmann (Ecuación 2), buscando caracterizaciones matemáticas estrictas para cualquier cadena nonamérica (Tablas VII y VIII).

TABLA VII. Pasos 1 y 2: Cálculos de entropía

Secuencia	Entropía del tipo de secuencia
Péptido teórico: GGGGGGGGG*	4,13x10 ⁻²³
CLIP91-99: MRMATPLLM	1,18 x10 ⁻²²
Péptido artificial: YRAMAAAAA	1,21x10 ⁻²²
Péptido artificial: YRAMSTAAA	1,23 x10 ⁻²²
YKKMEEAVKL**	1,51 x10 ⁻²²
HA309-317: YVKQNTLKL	1,76 x10 ⁻²²
Aminoácidos diferentes en todas las posiciones: GSVQYPITK	2,79x10 ⁻²²

*Los valores para esta secuencia son calculados teóricamente (Ver Anexo 1) y no se aplicaron las reglas del segundo paso.

Predicción teórica

A partir del rango de microestados y de entropía asociado a la unión, se realiza la predicción teórica del número de microestados equivalentes y la configuración particular para los péptidos de unión a la molécula de HLA clase II, a partir de la ecuación de Boltzmann (Ecuación 2). Despejando W, que es el número de microestados asociados a cada tipo de secuencia, física y matemáticamente se deduce cuales son las secuencias de aminoácidos que se unen (Ecuación 3).

$$W = e^{S/k}$$

Ecuación 3

Posteriormente, a partir de la ecuación 2, se despejó el valor de S/k (Ecuación 4).

$$\ln(W) = S/k$$

Ecuación 4

Con la proporción S/k se predicen cuales son los valores de los péptidos asociados al macroestado unión y cuales son los asociados al macroestado de no unión (Tablas IX y X).

TABLA IX. Predicción teórica del número de microestados W asociados a un valor particular de entropía S el valor de la proporción S/k asociado, y su porcentaje. Donde k es igual a la constante de Boltzmann con un valor de 1.38x10⁻²³ (J/k), paso 1-2

S	W	Proporción S/k*	Porcentaje	Predicción
4,13411x10 ⁻²³	20	3	31,7%	NO UNIÓN
8,48489 x10 ⁻²³	468	6		
1,10488 x10 ⁻²²	3000	8,01		
1,18036 x10 ⁻²²	5184	8,55	15,8%	UNIÓN
1,60614 x10 ⁻²²	113400	11,6		
1,76665 x10 ⁻²²	362880	12,8		
2,1152 x10 ⁻²²	4536000	15,3	52,5%	NO UNIÓN
2,1946 x10 ⁻²²	8064000	15,9		
3,427 x10 ⁻²²	60949324800	24,8		
3,72 x10 ⁻²²	5,12E+11	27		

TABLA VIII. Pasos 3 y 4: Cálculos de entropía

Secuencia	Entropía del tipo de secuencia
CLIP ₉₁₋₉₉ : MRMATPLLM	9,63 x10 ⁻²³
Péptido artificial: YRAMAAAAA	1,00 x10 ⁻²²
Péptido artificial: YRAMSTAAA	1,03 x10 ⁻²²
YKKMEEAVKL**	1,03 x10 ⁻²²
HA ₃₀₉₋₃₁₇ : YVKQNTLKL	1,44 x10 ⁻²²

A continuación se calcularon los valores de probabilidad, microestados y entropía para 12 regiones centrales de péptidos probados experimentalmente de unión⁽²³⁾; todas las cadenas nonámeras sobreapiladas de 12 secuencias de 13 aminoácidos de largo de no unión, probadas experimentalmente⁽²³⁾ (Tabla XI); y 83 péptidos de la página Web: www.immuneepitope.org⁽²⁴⁾ (Anexo 2). Para estos últimos los cálculos se efectuaron a todas las posibles cadenas nonámeras sobreapiladas (Tabla XII). Fueron utilizados un total de 161 péptidos, 100 de unión y 61 de no unión, contrastando los resultados con la predicción teórica.

Finalmente, por medio de un algoritmo se generó una secuencia matemática de 500 aminoácidos, encontrando todas las cadenas nonámeras sobreapiladas para calcular sus valores de probabilidad, combinatoria, y entropía (ver Anexo 3 y Tabla XIII), y se estudiaron sus implicaciones.

RESULTADOS

El espacio muestral obtenido contiene 20⁹ posibles microestados, péptidos nonaméricos, asociados a los 20 aminoácidos esenciales; de estos existen 20 microestados que cumplen la característica de tener el mismo aminoácido en todas sus posiciones, por lo tanto su probabilidad es

TABLA X. Predicción teórica del número de microestados W asociados a un valor particular de entropía S y el valor de la proporción S/k asociado, y su porcentaje. Donde k es igual a la constante de Boltzmann con un valor de 1.38×10^{-23} (J/k), paso 3-4

S	W	Proporción S/k*	Porcentaje	Predicción
$4,13411 \times 10^{-23}$	20	3	25,9%	NO UNIÓN
$8,48489 \times 10^{-23}$	468	6,15		
$9,3873 \times 10^{-23}$	900	6,8		
$9,63891 \times 10^{-23}$	1080	6,98	13%	UNIÓN
$1,28165 \times 10^{-22}$	10800	9,29		
$1,4478 \times 10^{-22}$	36000	10,5		
$2,1152 \times 10^{-22}$	4536000	15,3	61,1%	NO UNIÓN
$2,1946 \times 10^{-22}$	8064000	15,9		
$3,427 \times 10^{-22}$	60949324800	24,8		
$3,72 \times 10^{-22}$	5,12E+11	27		

TABLA XI. Cálculos realizados para regiones centrales de péptidos probados experimentalmente⁽²³⁾. Resultados escogidos de 24 secuencias. W1, P1, S1, son los valores de microestados, probabilidad y entropía para el paso 1-2. W2, P2, S2, son los valores de microestados, probabilidad y entropía para el paso 3-4.

PREDICCIÓN DE UNIÓN		PREDICCIÓN DE NO UNIÓN	
YDRADAGWL:	YRRLYNGLK:	LRDFWGNPL:	ASLVITDLT:
W1: 40320.0	W1: 15552.0	W1: 34560	W1: 2.016E7
P1: 7.87E-8	P1: 3.03E-8	P1: 6.75E-08	P1: 3,93E-05
S1: 1.46E-22	S1: 1.33E-22	S1: 1.44E-22	S1: 2.32E-22
W2: 3600.0	W2: 3600.0	W2: 43200	W2: 62400
P2: 7.03E-9	P2: 7.03E-9	P2: 8,43E-08	P2: 1,21E-07
S2: 1.13E-22	S2: 1.13E-22	S2: 1.47E-22	S2: 1.52E-22
YKKMEAVKL:	YAFLHATDL:	LSDNYTLDH:	DTVVVALDL:
W1: 60480.0	W1: 241920.0	W1: 2358720	W1: 3360
P1: 1.18E-7	P1: 4.72E-7	P1: 4,60E-06	P1: 6,56E-09
S1: 1.51E-22	S1: 1.71E-22	S1: 2.02E-22	S1: 1.12E-22
W2: 1800.0	W2: 14400.0	W2: 10800	W2: 1560
P2: 3.51E-9	P2: 2.81E-8	P2: 2,10E-08	P2: 3,04E-09
S2: 1.03E-22	S2: 1.32E-22	S2: 1.28E-22	S2: 1.01E-22
FTGGIGNKL:		GSVQYPITK:	
W1: 90720.0		W1: 6.048E8	
P1: 1.77E-7		P1: 1,18 E-03	
S1: 1.57E-22		S1: 2.79E-22	
W2: 36000.0		W2: 1123200	
P2: 7.03E-8		P2: 2,19E-06	
S2: 1.44E-22		S2: 1.92E-22	

$20/20^9 = 3.9 \times 10^{-11}$; mientras que existen 20^6 microestados que cumplen la característica de tener, en las tres primeras posiciones, un aminoácido definido –por ejemplo triptófanos–, y en las posiciones restantes cualquier aminoácido esencial, lo cual lleva a una probabilidad de $20^6/20^9 = 1.25 \times 10^{-4}$. El cociente entre las probabilidades calculadas es $3,12 \times 10^{-7}$, lo cual muestra una diferencia entre el péptido presentado

y el péptido teórico de no unión; en órdenes de magnitud hay 10 millones de microestados con las tres primeras posiciones fijas por cada microestado con todos los aminoácidos iguales.

El cálculo de las probabilidades y el número de microestados asociados a los tipos de secuencia seleccionados para el paso 1 y su posterior evaluación con los criterios definidos (paso

TABLA XII. Cálculos realizados para regiones centrales predichas de unión para 10 péptidos de la página Web: www.immunepitope.org (24). Resultados escogidos de 83 secuencias. W1, P1, S1, son los valores de microestados, probabilidad y entropía para el paso 1-2. W2, P2, S2, son los valores de microestados, probabilidad y entropía para el paso 3-4.

PÉPTIDOS Y CADENAS ASOCIADAS AL MACROESTADO UNIÓN			
1. FNNIPSRYNL	3. SEFAYGSFV	5. DTGFYTLH	8. FAYGSFVRT:
YDKMLPLDD:	RTVSLPGADE:	VIKSDLV	FAYGSFVRT:
FNNIPSRYN:	FAYGSFVRT:	LHVIKSDLV:	W-1: 302400.0
W-1: 226800.0	W-1: 302400.0	W-1: 75600.0	P-1: 5.90E-7
P-1: 4.42E-7	P-1: 5.90E-7	P-1: 1.47E-7	S-1: 1.74E-22
S-1: 1.70E-22	S-1: 1.74E-22	S-1: 1.55E-22	W-2: 36000.0
W-2: 36000.0	W-2: 36000.0	W-2: 3600.0	P-2: 7.03E-8
P-2: 7.03E-8	P-2: 7.03E-8	P-2: 7.03E-9	S-2: 1.44E-22
S-2: 1.44E-22	S-2: 1.44E-22	S-2: 1.13E-22	9. LQLSNDNR
IPSRYNLYD:	YGSFVRTVS:	6. PLNTSYRS	TLTLLSV:
W-1: 120960.0	W-1: 136080.0	GENLNLS:	LQLSNDNR:
P-1: 2.36E-7	P-1: 2.65E-7	YRSGENNL:	W-1: 282240.0
S-1: 1.61E-22	S-1: 1.63E-22	W-1: 15552.0	P-1: 5.51E-7
W-2: 18000.0	W-2: 27000.0	P-1: 3.03E-8	S-1: 1.73E-22
P-2: 3.51E-8	P-2: 5.27E-8	S-1: 1.33E-22	W-2: 10800.0
S-2: 1.35E-22	S-2: 1.40E-22	W-2: 18000.0	P-2: 2.10E-8
YNLYDKMLP:	4. SWYKGER	P-2: 3.51E-8	S-2: 1.28E-22
W-1: 181440.0	VDGNRQII:	S-2: 1.35E-22	LSNDNRTLT:
P-1: 3.54E-7	WYKGERVDG:	7. RHNWVNH	W-1: 60480.0
S-1: 1.67E-22	W-1: 201600.0	AVPLAMKLI:	P-1: 1.18E-7
W-2: 5400.0	P-1: 3.93E-7	WVNHAVPLA:	S-1: 1.51E-22
P-2: 1.05E-8	S-1: 1.68E-22	W-1: 302400.0	W-2: 3600.0
S-2: 1.18E-22	W-2: 7200.0	P-1: 5.90E-7	P-2: 7.03E-9
LYDKMLPLD:	P-2: 1.40E-8	S-1: 1.74E-22	S-2: 1.13E-22
W-1: 16800.0	S-2: 1.22E-22	W-2: 21600.0	NRTLTLLSV:
P-1: 3.28E-8	YKGGERVDGN:	P-2: 4.21E-8	W-1: 268800.0
S-1: 1.34E-22	W-1: 181440.0	S-2: 1.37E-22	P-1: 5.25E-7
W-2: 1080.0	P-1: 3.54E-7	VNHAVPLAM:	S-1: 1.72E-22
P-2: 2.10E-9	S-1: 1.67E-22	W-1: 60480.0	W-2: 1625.0
S-2: 9.63E-23	W-2: 7200.0	P-1: 1.18E-7	P-2: 3.17E-9
2. LQLSNGN	P-2: 1.40E-8	S-1: 1.51E-22	S-2: 1.02E-22
RTLTLFNV:	S-2: 1.22E-22	W-2: 10800.0	10. LRDNIQG
LSNGNRTLT:		P-2: 2.10E-8	ITKPAIRR:
W-1: 36288.0		S-2: 1.28E-22	IQGITKPAI:
P-1: 7.08E-8			W-1: 254016.0
S-1: 1.44E-22			P-1: 4.96E-7
W-2: 7200.0			S-1: 1.71E-22
P-2: 1.40E-8			W-2: 27000.0
S-2: 1.22E-22			P-2: 5.27E-8
			S-2: 1.40E-22

2), se muestran en la Tabla V y la Figura 1. Los cálculos de probabilidad, número de microestados, para el paso 3 y su posterior evaluación con los criterios definidos en el paso 4, se muestran en la Tabla VI. Los valores de probabilidad para todos los tipos de secuencia en el paso 1-2, oscilaron entre 3.9×10^{-11} y 1.1×10^{-3} , la probabilidad calculada para el CLIP y el HA es de 1.01×10^{-8} y 7.08×10^{-7} , respectivamente; la región limitada por estos valores de probabilidad está dentro de un

rango característico de unión. Los valores de probabilidad en el paso 3-4, oscilaron entre 2.10×10^{-9} y 7.03×10^{-8} .

El cálculo del valor de entropía para cada tipo de secuencia aplicando el paso 1 y su posterior evaluación con los criterios definidos (paso 2) se muestran en la Tabla VII y la Figura 2. Los valores de entropía para el paso 3 y su posterior evaluación con los criterios definidos (paso 4), se muestran en la Tabla VIII.

TABLA XIII. Péptidos seleccionados de la totalidad de secuencias predichas de unión para la proteína teórica generada

PÉPTIDOS	
1.	WPQCRNNRF
2.	VFIVFMEPQ
3.	FIVFMEPQS
4.	IQTIVRTCF
5.	FKLAESRSM
6.	LAESRSMFF
7.	MFFLKHTVD
8.	FFLKHTVDT
9.	FLKHTVDTF
10.	WRAMDMILCP

Se encontró que en el paso 1-2 todos los posibles péptidos de unión están incluidos en los tipos de secuencia que presentan valores de entropía comprendida en el rango $[1,18 \times 10^{-22}, 1,76 \times 10^{-22}]$, mientras que los péptidos de no unión están incluidos en los tipos de secuencia que presentan valores de entropía mayores que $1,76 \times 10^{-22}$ o menores que $1,18 \times 10^{-22}$. En el paso 3-4 los tipos de secuencia asociados a la unión tienen valores de entropía comprendida en el rango $[9,63 \times 10^{-23}, 1,44 \times 10^{-22}]$.

A partir de los cálculos realizados se observa que el valor mínimo de entropía corresponde a las secuencias que tienen un único aminoácido, mientras que la máxima entropía corresponde a secuencias con todos los aminoácidos diferentes; siendo estos extremos los casos en que no se presenta unión por parte del péptido, mientras que los tipos de secuencia asociados a unión tienen valores intermedios de entropía.

Para el tipo de secuencia asociado al CLIP, se encontró que posee 5184 microestados equivalentes y que la proporción S/k es $1,18 \times 10^{-22} / 1,38 \times 10^{-23} = 8,55$, mientras que para el tipo de secuencia asociado al HA que posee 326880 microestados equivalentes, la proporción S/k es $1,76 \times 10^{-22} / 1,38 \times 10^{-23} = 12,8$ (Tablas IX y X).

La metodología desarrollada en este trabajo permite predecir cadenas nonámeras de péptidos que se unen al HLA clase II en la región central para cualquier péptido sintetizado o natural, a partir de las características matemáticas establecidas para la diferenciación de péptidos que se unen de los que no. De los 161 péptidos probados, la predicción de macroestados de unión y no unión acertó en un 100%. A partir de la relación S/k se deduce matemáticamente el porcentaje de péptidos asociados a la unión con respecto a la totalidad de los tipos de secuencia posibles, encontrando valores de 15.8% para el paso 1-2 y 13% para el paso 3-4 (Tablas IX y X). Para la secuencia teórica se encontraron 492 péptidos nonaméricos sobreapilados, al calcular sus valores

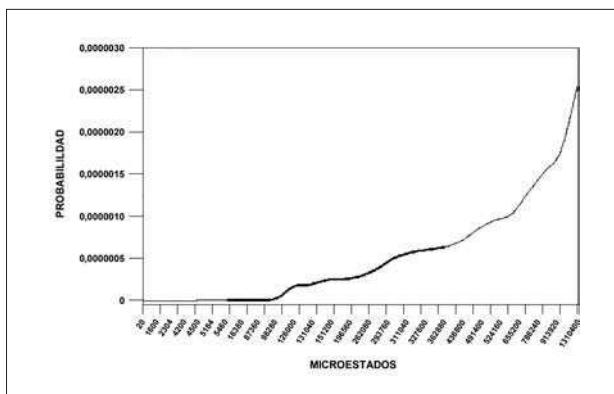


Figura 1. Logaritmo natural de la probabilidad de aparición de las diferentes secuencias, paso 1-2. El logaritmo se evaluó para hacer las magnitudes comparables.

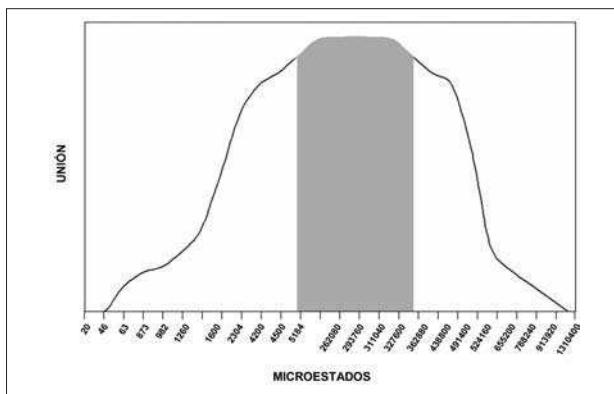


Figura 2. Unión asociada a los microestados para el paso 1-2.

de probabilidad, combinatoria, y entropía se encontraron 85 regiones centrales asociadas al macroestado de unión y 407 al de no unión (Tabla XIII, Anexo 3).

DISCUSIÓN

Este es el primer trabajo donde se realiza una teoría de predicción de unión de péptidos presentados por moléculas de HLA clase II basada en leyes de probabilidad, de combinatoria y de entropía. Se encontraron valores numéricos de probabilidad y de combinatoria que caracterizan matemáticamente cualquier posible microestado. Con la relación S/k y la ecuación más simple de entropía se predijeron cuáles son los valores que corresponden al macroestado de unión.

Los valores de probabilidad y de combinatoria permitieron diferenciar y caracterizar matemáticamente los péptidos que se unen de los que no. Los valores de entropía calculados

se encuentran dentro de valores finitos para cualquier cadena nonamérica, similar a la generalización de la entropía desarrollada por Bekenstein⁽²⁵⁾, donde no hay valores infinitos para la entropía. Además, estos valores no definen una flecha unidireccional en el fluir del tiempo, pues no evalúan la evolución del sistema. A partir de la ley de entropía se obtiene la relación S/k , la cual posee proporciones de una ley física de la naturaleza que depende de la forma en que los microestados W , tienen un orden de combinación, permitiendo predecir físicamente los péptidos nonámeros de unión y no unión. Las predicciones de entropía comprueban cómo péptidos construidos sintéticamente son presentados por la molécula de HLA clase II, evidenciando que esta teoría puede generar teóricamente y predecir cualquier posible experimento diseñado para unión. La relación S/k como el valor predicho, en un rango finito, comparado con el rango total de valores posibles, demuestra matemáticamente la alta especificidad del fenómeno de unión en la naturaleza comparada con todas las posibilidades matemáticas. Desde esta nueva perspectiva teórica comparada con el trabajo experimental de la inmunología y basados en las predicciones de los péptidos posibles de unión de la proteína teórica desarrollados, es muy posible predecir péptidos que se unan a la molécula de HLA clase II que jamás hayan sido probados experimentalmente, similarmente a la mecánica Newtoniana a partir de la cual se predijo la existencia de un planeta⁽²⁶⁾ que no había sido visto previamente, lo que elevaría la inmunología a nivel de ciencia fundamental en el caso de la presentación por las moléculas de HLA clase II.

Se calculó el número de microestados para cada tipo de secuencia y la entropía asociada a este número, partiendo de observaciones experimentales que permitieron desarrollar generalizaciones. Desde esta perspectiva teórica se definieron 10 criterios evaluadores para el paso 2 y 10 criterios para el paso 4 que permitieron caracterizar la especificidad del macroestado de unión. Los cinco primeros para el paso 2 y los dos primeros del paso 4 definieron la especificidad para las 24 regiones centrales estudiadas⁽²³⁾ y los siete macroestados seleccionados, mientras que los demás criterios permitieron refinar los cálculos, esto enmarcado desde las leyes teóricas aplicadas para calcular microestados particulares.

En las Tablas VI y VII se observa cómo la ley de entropía, que depende de W y es creciente, caracterizó los aminoácidos que se unen al HLA de clase II y los diferenció de los que no; lo cual se corrobora. Por ejemplo, al CLIP se asocian 1080 microestados y tiene una entropía de $1,18 \times 10^{-22}$, lo cual corresponde a la predicción de unión, mientras que el péptido GSVQYPITK con 604 800.000 microestados tiene una entropía

de $2,79 \times 10^{-22}$ y corresponde a la predicción de no unión. Al deducir la ecuación 3 mediante un proceso algebraico, que conserva la relación funcional de sus variables, es posible predecir el número de microestados que tendría un péptido que se una al HLA clase II. De la misma manera, la ecuación de la entropía en este contexto estudia la manera en la que conjuntos de aminoácidos se asocian en un orden, análogo a un sistema de partículas en el estudio de la materia, y el resultado general caracteriza la unión o la no unión de esos conjuntos de aminoácidos. En conjunto, la teoría desarrollada lleva a la predicción de la existencia o inexistencia de secuencias de aminoácidos, como en la secuencia teórica de un mismo aminoácido en las nueve posiciones, y la diferenciación con secuencias de aminoácidos diferentes evaluando de aquellos que se pegan de los que no.

Trabajos anteriores en predicción de unión de péptidos han utilizado Artificial Neural Networks, ANN, programación lineal, o vectores de máquina⁽²⁷⁻³⁰⁾; estos métodos buscan patrones de unión caracterizando las posiciones y sus respectivos motivos. Su generalización es limitada, pues depende de la muestra usada. La teoría desarrollada en este trabajo confirmó los resultados experimentales, en términos de su representación numérica y de las combinaciones de aminoácidos apropiadas. A diferencia de los trabajos mencionados, el presente es de naturaleza teórica e independiente de la población de estudio. Ahora bien, el método experimental de péptidos sobreapilados, al que Berzofsky definió como "método de fuerza bruta"⁽³¹⁾, consiste en la síntesis de péptidos, pertenecientes a la secuencia de una misma proteína que se sobreapan. La capacidad de dichos péptidos para estimular respuestas en los linfocitos T, tanto proliferativas como citotóxicas, es probada luego *in vitro*. Aunque este método es completo, es costoso e implica mucho tiempo y trabajo. La teoría planteada en este trabajo permite evaluar todas las secuencias nonámeras sobreapiladas para cualquier secuencia proteica, simplificando el método empleado pues se dejan a un lado los procesos experimentales de ensayo y error por predicciones numéricas rigurosas.

Trabajos experimentales evidenciaron pérdida de la energía de unión debido a la disminución del tamaño del péptido asociado a la molécula de HLA clase II, el cual generalmente varía entre 13 y 16 aminoácidos⁽¹⁵⁻¹⁸⁾. Una compensación a esta energía perdida es evidenciada a través de las interacciones de la región central, y puede ser provista por la adición de aminoácidos en las posiciones de anclaje⁽²³⁾. Desde estos resultados experimentales sería posible diseñar antígenos sintéticos que se presenten a las moléculas de HLA clase II con las longitudes predichas por esta teoría, o determinar el lugar de la región central de la secuencia

antigénica sin importar su longitud. Desde este punto de vista propongo predecir los posibles péptidos nonaméricos de unión naturales o sintéticos; hacer la contrastación experimental de unión; y posiblemente introducir aminoácidos que cambien los valores de los microestados, buscando el diseño de péptidos promiscuos, lo importante del tercer y cuarto paso, de acuerdo a las evaluaciones matemáticas. Finalmente, hacer las pruebas experimentales de respuesta y memoria inmunológica y así poder desarrollar vacunas a partir de esta metodología física y matemática.

Este trabajo muestra el desarrollo de una teoría a partir de tres leyes aplicadas, evidenciando órdenes físicos y matemáticos subyacentes a la presentación antigénica. La teoría desarrollada logra simplificar este fenómeno tan complejo, partiendo de solo siete péptidos con los que se construye la generalización que obvia la especificidad alélica; pues el fenómeno de unión es caracterizado a partir de representaciones numéricas de tipos de secuencias y no de aminoácidos específicos. Como en las teorías fundamentales actuales de la física⁽³²⁾ no hay causas sino órdenes físicos y matemáticos acausales para la descripción y comprensión del fenómeno estudiado. Por ser una teoría no requiere de análisis estadísticos ni del uso de grandes cantidades de datos experimentales.

FUTURAS APLICACIONES

Se contrastará esta teoría con las bases de datos experimentales desarrolladas y se comparará con los modelos predictivos actuales. Se compararán predicciones de secuencias antigénicas u oncocénicas de proteínas con péptidos que se conoce son presentados por moléculas de HLA clase II para la corroboración de esta teoría.

CONCLUSIONES

1. Las probabilidades y las combinatorias para todos los microestados posibles se encuentran dentro de un rango finito caracterizando de forma objetiva y reproducible la unión.

2. Los valores de entropía para todos los microestados posibles se encuentran dentro de un rango finito y predicen todos los valores asociados a el macroestado de unión y al de no unión. La evidencia matemática de la alta especificidad del fenómeno de unión en la naturaleza está representada por el valor predicho que debe tener el sistema para que se una, el cual está dado por la relación S/k .
3. Este estudio permitió encontrar órdenes matemáticos subyacentes al fenómeno de unión del HLA clase II superando los procesos experimentales de ensayo y error. La predicción física y matemática realizada en la presente investigación facilita el proceso de selección de los péptidos requeridos en el desarrollo de vacunas; optimizando el tiempo y recursos económicos.
4. La predicción teórica de secuencias de péptidos naturales no evidenciados experimentalmente puede ser posible.

AGRADECIMIENTOS

A todos los que me ayudaron y a todos los que me rechazaron, por que sin todos no hubiera sido posible llegar. Infinitas gracias a todos y a todo. Infinitas gracias a todos los que están y estuvieron en el Grupo Insight.

Dedicado a Sir. Isaac Newton, MAESTRO del camino.

CONFLICTO DE INTERÉS

El autor declara no tener ningún conflicto de interés financiero.

CORRESPONDENCIA:

Javier Rodríguez, M.D.
Director del Grupo Insight.
Universidad Nacional de Colombia,
Departamento de Cirugía.
Carrera 79b número 51-16 Sur
Bogotá, Colombia
Phone: (057-1) 4527541
E-mail: grupoinsight2025@yahoo.es

ANEXO UNO. CÁLCULOS DE MICROESTADOS Y ENTROPIA

1. Tipo de secuencia: GGGGGGGGG

G	G	G	G	G	G	G	G	G	= 20 microestados asociados
20	1	1	1	1	1	1	1	1	

Este tipo de secuencia tiene la característica de tener en todas sus posiciones el mismo aminoácido, por esto a la primera posición se le asigna un valor de 20 pues puede aparecer cualquiera de los 20 aminoácidos esenciales, las siguientes posiciones tienen un valor de 1 pues se repite el mismo aminoácido de la primera posición.

El valor de entropía asociado a este tipo de secuencia tiene un valor de:

$$S = k \ln(20) = 4,13 \times 10^{-23}$$

2. Tipo de secuencia: YKKMEAVKL

Paso uno:

Y	K	K	M	E	A	V	K	L	= 60480 microestados equivalentes
7	5	1	4	4	6	6	1	3	

Este tipo de secuencia está asociado a un péptido de unión (22), la primera posición Y tiene un valor de 7 pues pertenece al grupo 1 (Tabla I) en la posición dos a K se le otorga un valor de 5, dado que está dentro del grupo de cargados, la posición 3 tiene un valor de uno pues el aminoácido ya había aparecido, a M le corresponde 4 porque pertenece al grupo de motivos en esa posición, a E le corresponde un valor de 4 porque está en el grupo de aminoácidos cargados y el valor disminuye pues en la posición dos apareció un aminoácido del mismo grupo. A tiene valor 6, pues pertenece al grupo de motivos en esa posición; a V le corresponde un valor de 6 porque está en el grupo 1 y el valor disminuye pues en la posición uno apareció un aminoácido del mismo grupo. La posición 8 tiene un valor de 1, pues el aminoácido ya había aparecido; finalmente a L le corresponde 3 pues pertenece al grupo de motivos en esa posición.

Paso dos:

Para este tipo de secuencia no se cumple ningún criterio, por lo tanto se mantiene el valor calculado en el primer paso. Posteriormente se calcula el valor de entropía asociado al tipo de secuencia por medio de la ecuación 2:

$$S = k \ln(60480) = 1,51 \times 10^{-22}$$

Dado que el tipo de secuencia asociado está dentro de los rangos de unión, se realiza el tercer paso:

Paso tres:

Y	K	K	M	E	A	V	K	L	= 7200 microestados equivalentes
5	3	1	5	2	4	4	1	3	

Este tipo de secuencia está asociado a un péptido de unión (22). A la primera posición Y se le otorga un valor de 5, pues pertenece al grupo de 5 aminoácidos hidrofóbicos de alta unión (Tabla III). En la posición dos a K se le otorga un valor de 3, dado que está dentro del grupo de cargados positivos; a la posición 3 se le otorga un valor de uno, pues el aminoácido ya había aparecido en la posición 2. A la posición 4 (M) se le da un valor de 5, pues pertenece a los aminoácidos del grupo 4. En la posición cinco (E) se le otorga un valor de 2, ya que está dentro del grupo de cargados negativos. A A le corresponde 4 porque pertenece al grupo 4, y el valor decrece debido a que en la posición 4 apareció un aminoácido del mismo grupo. A V le corresponde 4 porque pertenece al grupo de hidrofóbicos de alta unión y el valor decrece por la aparición de otro aminoácido del mismo grupo en la posición 1. A la posición 8 (K) se le otorga un valor de uno, pues el aminoácido ya había aparecido en la posición 2. Finalmente, en la última posición 9 se le da un valor de tres, porque pertenece al grupo 4 y el valor decrece debido a que en la posición 4 y 6 aparecieron aminoácidos del mismo grupo.

Posteriormente se calcula el valor de entropía asociado al tipo de secuencia por medio de la ecuación 2:

$$S = k \ln(7200) = 1,22 \times 10^{-22}$$

Paso cuatro:

Para este tipo de secuencia no se cumple ningún criterio, por lo tanto se mantiene el valor calculado en el tercer paso.

3. Tipo de secuencia: MRMATPLLM

Paso uno:

M	R	M	A	T	P	L	L	M = 20160 microestados equivalentes
7	5	1	4	9	4	4	1	1

Este tipo de secuencia está asociado al péptido CLIP. La primera posición M tiene un valor de 7 pues pertenece al grupo 1 (Tabla I); en la posición dos a R se le otorga un valor de 5, dado que está dentro del grupo de cargados; a M le corresponde un valor de 1, pues el aminoácido ya había aparecido. La A siguiente tiene un valor de 4 porque pertenece al grupo de motivos en esa posición; el siguiente aminoácido tiene un valor de 9 pues pertenece al grupo 2 y no hay motivo para esa posición; la P tiene un valor de 4 pues pertenece al grupo 3 y no es motivo para esa posición. La posición 7 tiene un valor de 4, pues pertenece al grupo de motivos para esa posición; las dos últimas posiciones tienen un valor de 1, pues estos aminoácidos ya habían aparecido.

Paso dos:

Para este tipo de secuencia se cumple el criterio número 1, pues en la posición uno tenemos M un aminoácido hidrofóbico de unión, y además en la segunda posición aparece una R; entonces estas dos posiciones toman un valor de 3:

M	R	M	A	T	P	L	L	M = 5184 microestados equivalentes
3	3	1	4	9	4	4	1	1

Posteriormente se calcula el valor de entropía asociado al tipo de secuencia por medio de la ecuación 2:

$$S = k \ln(5184) = 1,18 \times 10^{-22}$$

Paso tres:

M	R	M	A	T	P	L	L	M = 2160 microestados equivalentes
5	3	1	4	6	3	2	1	1

A la primera posición M se le otorga un valor de 5, pues pertenece al grupo 4 (Tabla III); en la posición dos a R se le otorga un valor de 3, dado que está dentro del grupo de cargados positivos. Luego aparece nuevamente M, y se le otorga un valor de 1; a la posición 4 le corresponde 4 ya que pertenece al grupo 4 y el valor decrece debido a que en la posición uno apareció un aminoácido del mismo grupo; la T siguiente tiene un valor de 6 pues pertenece al grupo que hacen puentes de hidrógeno. La P siguiente tiene un valor de 3, pues pertenece al grupo 4 y el valor decrece; la siguiente posición tiene un valor de 2 pues pertenece al grupo 4 y el valor decrece; las dos últimas posiciones tienen un valor de 1, pues están repetidos.

Paso cuatro:

Para este tipo de secuencia se cumple el criterio número 1, pues en la posición uno, tres, siete, ocho y nueve tenemos aminoácidos hidrofóbicos de unión, es decir, la secuencia tiene un total de cinco aminoácidos hidrofóbicos de unión, entonces el número de microestados se divide entre dos.

M	R	M	A	T	P	L	L	M = 2160/2= 1080 microestados equivalentes
5	3	1	4	6	3	2	1	1

Posteriormente se calcula el valor de entropía asociado al tipo de secuencia por medio de la ecuación 2:

$$S = k \ln(1080) = 9,63 \times 10^{-23}$$

4. Tipo de secuencia: GSVQYPITK

Paso uno:

G	S	V	Q	Y	P	I	T	K = 1.8144x10 ⁶ microestados asociados
5	3	7	9	6	2	5	8	5

Este tipo de secuencia tiene la característica de tener en todas sus posiciones diferentes aminoácidos. La primera posición tiene un valor de 4, pues pertenece a un aminoácido del grupo tres (Tabla I); la siguiente posición toma un valor de tres, pues pertenece a este mismo grupo y el valor decrece, y la siguiente posición toma un valor de 7 pues pertenece al grupo 1. La siguiente posición toma un valor de 9 pues pertenece al grupo 2; la posición 5 toma un valor de 6 pues pertenece al grupo 1 y el valor decrece debido a la aparición de un aminoácido del mismo grupo en la posición 3. La siguiente posición toma un valor de 2, pues pertenece a este mismo grupo y el valor decrece; la posición 7 toma un valor de 5, pues pertenece al grupo 1 y el valor decrece; la posición 8 toma un valor de 8, pues pertenece al grupo 2 y el valor decrece por la anterior aparición en la posición 4 de un aminoácido del mismo grupo. Finalmente, la posición 9 toma un valor de 5, pues es un aminoácido cargado.

Paso dos:

Para este tipo de secuencia se cumple el criterio dos, pues no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1, entonces esta posición toma un valor de 20. Además, en la segunda posición no presenta una R, entonces también toma un valor de 20; por otra parte se cumple el criterio 8, pues no presenta uno de los 7 aminoácidos hidrofóbicos característicos de unión en la posición 1, en la segunda posición no presenta una R, y la secuencia tiene menos de 4 repeticiones, entonces el número de microestados se multiplica por 10:

G	S	V	Q	Y	P	I	T	K = $6.048 \times 10^7 \times 10 =$ microestados asociados
20	20	7	9	6	2	5	8	5

El valor de entropía asociado a este tipo de secuencia tiene un valor de:

$$S = k \ln(6.048 \times 10^8) = 2,79 \times 10^{-22}$$

ANEXO DOS. PÉPTIDOS UTILIZADOS DE LA PÁGINA WEB: WWW.IMMUNEEPITOPE.ORG⁽²⁴⁾

- | | | |
|-------------------------|--------------------------|----------------------------|
| 1. DFEFEQMFTDAMG | 29. FNNIPSRYNLYDKMLELDD | 57. QFSLWRRPVVTAHIE |
| 2. IIGYVIGTQQATPGP | 30. DPIELNATLSAVA | 58. RHNVVNHAAPLAMKLI |
| 3. GREIIYPNASLIIQN | 31. PRYISLIPVNVVAD | 59. EKVPVSEVMGTTLAEMSTPEAT |
| 4. LQLSNGNRTLTLFNV | 32. PLNTSYRSGENLNLS | 60. IYRRRLMKQDFSVPLPHS |
| 5. TLTLFNVTRNDTASY | 33. YTNQNINISMERNLMKHGFH | 61. KRCCLLHLAVIGALLAVGATKV |
| 6. TVSAELPKPSISSNN | 34. AYGSFVRTVSL | 62. PGPVTAQVVLQAAIPLTSCGSS |
| 7. KVKILPKDRWTQHTTG | 35. LTNQNINIDQEFNLMKHGFH | 63. QNILLSNAPLGPQFP |
| 8. YSWRINGIPQQHTQV | 36. FAYGSFVRT | 64. HQAISPRTLNSPAIF |
| 9. LQLSNDNRTLTLSSV | 37. QRLLLTTASLLTFWNP | 65. KILEPFRKYTAFTIP |
| 10. QEIDPLSNSYIIPVNSN | 38. ITVSASGTSPGLSAG | 66. FAKIKNTHTNGVRL |
| 11. LSYNYIPVN | 39. ADLVPTATLLDTY | 67. GFLGFLATAGSAMGA |
| 12. SSVFNVVNNSIGLIM | 40. FAYGSFVRTVSLPGADE | 68. HGIRPVVSTQLLL |
| 13. VVFPASFFIKLPIILA | 41. AYGSFVRTVSLPGADE | 69. ILQRLSATLQRIREV |
| 14. FATCFLIPLTSQFFLP | 42. TFWNPPTAKLTIES | 70. IRPVVSTQLLL |
| 15. RVYQEPQVSPQRAET | 43. EFAYGSFVRTVSLPGADE | 71. ITTEQEIQFQQSKNS |
| 16. IVAVHVASFIEAEV | 44. ALFNRLDDLG | 72. PKISFEPPIP HYCAPAGFAI |
| 17. SAMGAASLTATQSR | 45. YGSFVRTVSLPGADE | 73. VVWGIKQLQARVLAVERYLKD |
| 18. THGIRPVVSTQLLL | 46. EFAYGSFVRTVSL | 74. ENPVVHFFANIVTPRTP |
| 19. TIVLMAVHCMNFKRR | 47. VLLLHVNLQPHLFGY | 75. IPQEWEKPAITVKVLPA |
| 20. VILLRIVIYIVQMLA | 48. FDLEMLGDVESPS | 76. KSKYKLATSLVLAGLL |
| 21. VQCTHGIRPVVSTQLL | 49. NNSIVKSITVSASGT | 77. KTVTAMDVYALKRQ |
| 22. VSTVQCTHGIRPVVSTQL | 50. TVGIMIGVLVGLI | 78. LTSQFFLPALPVFTWL |
| 23. WDFISTPPLVRLVFN | 51. SWYKGERVDGNRQII | 79. LWWSTMYLTHHYFVDL |
| 24. WLSTYAVRITWYSKN | 52. KVKILPRDRWTQHTTG | 80. VKNVIGPFMKAVCVE |
| 25. WRRDNRRGLRMAKQN | 53. TISPSYTYYRPGVNL | 81. SSIIFGAFPSLHSGCC |
| 26. LIRILQRALFMHFRG | 54. YLWWVNNQSLPVSPR | 82. LAAIIIFLFGPPTALRS |
| 27. FNNIPSRYNLYDKMLPLDD | 55. DTGFYTLHVIKSDLV | 83. THHYFVDLIGGAMLSL |
| 28. NNIPSRYNLYDKMLDDL | 56. LRDNIQGITKPAIRR | |

ANEXO TRES. SECUENCIA TEÓRICA GENERADA, Y PÉPTIDOS DE ESTA SECUENCIA PREDICHOES COMO DE UNIÓN

DQKQELCKWPQCRNNRFHTSTEDCHPDFHVIVFMEPQSHIQTIVRTCFKLAESRSRSMFFLKHTVDTFRKLFVLMVFHQGNGLV-	HAMCGVWHWCLHDDEKMALCIRMCKGKGCQKNRKISGSVKREVCQPNFMRGKSCATETIWLKQREITVPNRHSIMLEFIAHTLN-		
WEFTAGPHVCDVVCQHIDMMQTHDMKALSDMETKFTIWECHQPALADANSSESSGCNDFHVHLECRQKQPRGCMVVHQRIEN-	HIPHQCWRWIKGIPQMWEASHLGHMGKNIWNQWHDKAQMPSLGARTRGHNLEGMPWRAMDMLCPEMEHAPALFIDWAAG-		
5.FKLAESRSM	TAIKWCPDEWCDGSTSIMCREQNCVCSQNTMSTVFTTRHRVM SHALEFNEMMAKFDHKTRPDTANARCVGMVVMLFTPWKC-		
6.LAESRSMFF	DARCRSPNQVVIVSCLVHVGGFGHINWEKMHGIPVSTLMRD PAGARDITHQMHTEVHIVKEDTEKGSKWCWLRAIQCMTNLVQ		
7.MFFLKHTVD			
8.FFLKHTVDT			
9.FLKHTVDTF			
10.LKHTVDTFR			
11.VDTFRKLFV			
12.LFVLMVFHQ			
13.FVLMVFHQG			
14.VFHQGNGLV			
15.FHQGNGLVH			
16.LVHAMCGVW			
17.VHAMCGVWH			
18.MCGVWHWCL			
19.VWHWCLHD			
20.MALCIRMC			
21.LCIRMCGKG			
22.IRMCGKGQ			
23.MCGKGQKN			
1.WPQCRNNRF	24.ISGSVKREV	47.WQWHDKAQM	70.MVVMLLFTP
2.VFIVFMEPQ	25.VKREVCQPN	48.SRLGARTRG	71.VVMLLFTPW
3.FIVFMEPQS	26.MLEFIAHTL	49.LGARTRGHN	72.VMLLFTPWK
4.IQTIVRTCF	27.VCDVVCQHI	50.LEGMPWRAM	73.WKCDARCR
5.FKLAESRSM	28.VVCQHIDMM	51.MPWRAMDM	74.LSPNQVVIV
6.LAESRSMFF	29.VCQHIDMMQ	52.WRAMDMLC	75.VVIVSCLVH
7.MFFLKHTVD	30.IDMMQTHDM	53.MLCPEMEHA	76.VIVSCLVHV
8.FFLKHTVDT	31.MMQTHDMKA	54.LCPEMEHAP	77.IVSCLVHVG
9.FLKHTVDTF	32.LSDMETKFT	55.MEHAPALFI	78.VSCLVHVGG
10.LKHTVDTFR	33.METKFTIWE	56.FIDWAAGTA	79.VGGFGHINW
11.VDTFRKLFV	34.FHQPALADA	57.IDWAAGTAI	80.INWEKMHGI
12.LFVLMVFHQ	35.LADANSESS	58.WAAGTAIKW	81.MHTEVEHIV
13.FVLMVFHQG	36.VHLECRQKQ	59.IKWCPEDEWC	82.VEHIVKEDT
14.VFHQGNGLV	37.LECRQKQPR	60.MCREQNCVC	83.IVKEDTEKG
15.FHQGNGLVH	38.CRQKQPRGC	61.VCSQNTMST	84.VKEDTEKGS
16.LVHAMCGVW	39.PRGCMVVHV	62.VFTTRHRVM	85.WCWLRAIQC
17.VHAMCGVWH	40.VVHVQRIEN	63.FTTRHRVMS	
18.MCGVWHWCL	41.QRIENHIPH	64.WALEFNEMM	
19.VWHWCLHD	42.IENHIPHQC	65.LEFNEMMAK	
20.MALCIRMC	43.WIKGIPQM	66.MMAKFHDHK	
21.LCIRMCGKG	44.MWEASHLGH	67.FDHKTRPDT	
22.IRMCGKGQ	45.WEASHLGHM	68.TRPDTANAR	
23.MCGKGQKN	46.IFNWQWHDK	69.ARCSVGMVVM	

BIBLIOGRAFÍA

1. Laplace P. *Ensayo filosófico sobre las probabilidades*. Barcelona: Altaya 1995.
2. Feynman RP, Leighton RB, Sands M. *Probabilidad*. En: Feynman RP, Leighton RB, Sands M. *Física*. Vol. 1. Wilmington: Addison-Wesley Iberoamericana, S. A. 1964: pp. 6-1, 6-16.
3. Mood A, Graybill F, Boes D. *Introduction to the theory of statistics*. 3rd. Ed. Singapore Mc. Graw-Hill. 1974.
4. Blanco L. *Probabilidad, notas de clase*. Universidad Nacional de Colombia. Departamento de Matemáticas y Estadística: 1996.
5. Feynman RP, Leighton RB, Sands M. *Leyes de la Termodinámica*. En: Feynman RP, Leighton RB, Sands M. *Física*. Vol. 1. Wilmington: Addison-Wesley Iberoamericana, S. A. 1964. pp. 44-1, 44-19.
6. Matváev A. *Física molecular*. Moscú: MIR, 1987.
7. Tolman R. *Principles of statistical mechanics*. New York: Dover Publications. 1979.
8. Claude S. *The mathematical theory of communication*. Chicago: University of Illinois Press, 1980.
9. Wang J-H, Reinherz E. Structural basis of T-cell recognition of peptides bound to MHC molecules. *Mol Immunol* 2001; 38:1039-1049.
10. Abbas A, Lichtman A, Pober J. *Inmunología celular y molecular*. 4^a Ed. McGraw Hill, 2002.
11. Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol* 1995; 13:587-622.
12. Fremont DH, Hendrickson WA, Marrack P, Kappler J. Structures of an MHC class II molecule with covalently bound single peptides. *Science* 1996; 272:1001-1004.
13. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, Wiley DC. 3-Dimensional structure of the human Class-II histocompatibility antigen HLADR1. *Nature (Lond.)* 1993; 364:33-39.
14. Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, Wiley DC. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature (Lond.)* 1994; 368:215-221.
15. Hunt DF, Michel H, Dickinson TA, Shabanowitz J, Cox AL, Sakaguchi K, Appella E. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science* 1992; 256:1817-1820.
16. Chicz RM, Urban RG, Gorga JC, Vignali DAA, Lane WS, Strominger JL. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med* 1993; 178:27-47.
17. Rudensky AY, Preston-Hurlburt P, Hong S-C, Barlow A, Janeway CA. Sequence analysis of peptides bound to MHC class II molecules. *Nature* 1991; 353:622-627.
18. Chicz RM, Urban RG, Lane WS, Gorga JC, Stern LJ, Vignali DAA, Strominger JL. Predominant naturally processed peptides bound to HLA DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* 1992; 358:764-768.
19. Rammennsee HG, Friede T, Stevanovic S. MHC ligands and peptide motifs: First listing. *Immunogenetics* 1995; 41:178-228.
20. Sinigaglia F, Hammer J. Motifs and Supermotifs for MHC Class II Binding Peptides. *J. Exp Med* 1995; 181:449-451.
21. Hammer J, Valsasnini P, Tolba K, Bolin D, Higelin J, Takacs B, Sinigaglia F. Promiscuous and Allele-Specific Anchors in HLA-DR-Binding Peptides. *Cell* 1993; 74:197-203.
22. Busch J, Strang G, Howland K, Rothbard J. Degenerate binding of immunogenic peptides to HLA-DR proteins on B cell surfaces. *Int Immunol* 1990; 2:443-448.
23. Hammer J, Belunis C, Bolin D, Papadopoulou J, Walskyt R, Higelin J, et al. High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. *Immunology* 1994; 91: 4456-4460.
24. Peters B, Sidney J, Bourne P, Bui HH, Buus S, Doh G, et al. The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol* 2005 Mar; 3:e91
25. Bekenstein J. *La información en el Universo holográfico*. Investigación y Ciencia: Octubre 2003; 36-43.
26. Feynman R, Leighton R, Sands M. *Física*. Vol 1: Mecánica, radiación y calor. pp 7-8. Wilmington: Addison-Wesley Iberoamericana. 1987.
27. Nielsen M, Lundsgaard C, Worning P, Sylvester S, Lamberth K, Buus S, et al. Improved prediction of MHC class I and II epitopes using a novel Gibbs sampling approach. *Bioinformatics* 2004; 20:1388-1397.
28. Southwood S, Sidney J, Kondo A, del Guercio MF, Appella E, Hoffman S, et al. Several Common HLA-DR Types Share Largely Overlapping Peptide Binding Repertoires. *J Immunol* 1998; 160:3363-3373.
29. Brusic V, Rudy G, Honeyman M, Hammer J, Harrison L. Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network. *Bioinformatics* 1998; 14:121-130.
30. Dönnies P, Eofsson A. Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics* 2002; 3:25.
31. Meister GE, Roberts CGP, Berzofsky JA, De Groot AS. Two novel T cell epitope prediction algorithms based on MHC-binding motifs; comparison of predicted and published epitopes from *Mycobacterium tuberculosis* and HIV protein sequences. *Vaccine*, 1995; 13:581-591.
32. Fernandez Rañada A. *Orden y Caos. Introducción*. Barcelona: Prensa Científica, 1990.