

# Diferenciación matemática de péptidos de alta unión de MSP-1 mediante la aplicación de la teoría de conjuntos

J. Rodríguez

*Department of Immunology and Oncology, National Centre of Biotechnology/CSIC  
UAM-Campus de Cantoblanco, Madrid.*

## **MATHEMATICAL DIFFERENTIATION OF HIGH BINDING PEPTIDES FROM THE MEROZOITE SURFACE PROTEIN MSP-1 THROUGH THE APPLICATION OF SET THEORY**

*Recibido: 18 Septiembre 2007*

*Aceptado: 2 Mayo de 2008*

### RESUMEN

MSP-1 es una proteína de membrana de superficie del merozoíto de *Plasmodium falciparum* que interactúa con el receptor de membrana del eritrocito, y por lo tanto es una proteína implicada en el reconocimiento del merozoito al glóbulo rojo.

La teoría de conjuntos y datos experimentales fueron utilizados para realizar una caracterización matemática de 79 secuencias de la proteína MSP-1, mediante la definición de 5 reglas asociadas a la alta unión al receptor del glóbulo rojo.

Todos los péptidos de alta unión y de baja unión fueron caracterizados acertadamente según estudios experimentales.

Estudiando más proteínas implicadas en la invasión, ésta metodología podría ser útil para diferenciar péptidos de alta unión en la proteína MSP-1 y en otras proteínas de una forma objetiva y reproducible; el fenómeno de unión puede ser descrito a partir de características matemáticas.

**PALABRAS CLAVE:** Eritrocito / Plasmodium / MSP-1 / Péptidos.

### ABSTRACT

MSP-1 is a surface membrane protein of *Plasmodium falciparum* merozoites that interacts with the erythrocyte membrane receptor and therefore it is a protein involved in its recognition by the red globule.

The set theory and experimental data have been used to make a mathematical characterization of 79 different MSP-1 protein sequences, by means of the definition of five rules associated with high binding to the red globule receptor.

All the high binding and low binding peptides were correctly characterized according to experimental determinations.

Studying more proteins involved in the invasion, this methodology could be useful to detect peptides with high binding in the MSP-1 protein and in other proteins in a more objective and reproducible way. Thus, the binding phenomenon can be described starting from mathematical characteristics.

**KEY WORDS:** Erythrocyte / Plasmodium / MSP-1.

## INTRODUCCIÓN

La teoría de conjuntos permite diferenciar elementos particulares a partir de reglas generales al realizar operaciones como la unión, la intersección, la diferencia y la diferencia simétrica<sup>(1)</sup>. La teoría de conjuntos es una teoría matemática que se fundamenta en nociones como pertenencia, y cumple una serie de axiomas que definen las relaciones matemáticas entre elementos. Un conjunto es la reunión en un todo de objetos bien definidos y diferenciables entre sí, que se llaman elementos del mismo. Los elementos denotan relaciones de pertenencia a los conjuntos, de tal modo que si 'a' es un elemento del conjunto A, existe la relación de pertenencia  $a \in A$ , de lo contrario, si 'a' no es un elemento de A se denota  $a \notin A$ . De esta misma forma, entre los conjuntos se establecen relaciones que definen las operaciones posibles entre éstos.

Ahora bien, la malaria es una de las principales causas de morbimortalidad en el mundo entero<sup>(2)</sup>. Por sus características, la enfermedad se presenta principalmente en zonas tropicales, tomando allí la forma de epidemia, como ocurre en Colombia.

La reorientación e invaginación del merozoito durante la invasión al glóbulo rojo<sup>(3)</sup>, pueden ser aclaradas por interacciones receptor-ligando de alta y baja afinidad. Estas interacciones de membrana receptor-ligando son cruciales para la invasión de los glóbulos rojos por los merozoitos del *Plasmodium falciparum*. Es probable que la disección de estas interacciones conduzca al desarrollo de nuevas estrategias inmunoproliféricas<sup>(4)</sup>. Varias proteínas del merozoito del *Plasmodium falciparum* han mostrado interactuar con el glóbulo rojo, y estas incluyen: La proteína-1 de superficie del merozoito (MSP-1), el antígeno de unión del eritrocito (EBA-175kDa), la proteína de unión de la glicoforina (GBP-130), y la proteína-2 de superficie del merozoito (MSP-2), entre otras<sup>(5-7)</sup>.

MSP-1 es una proteína polimórfica de superficie de membrana<sup>(7)</sup>, que contiene secuencias variables, semiconservadas y constantes. Durante su maduración, la proteína es sometida a una serie de rupturas proteolíticas. La primera, independiente de calcio, rompe la proteína en cuatro fragmentos: un fragmento terminal N de 83 kDa, dos fragmentos centrales de 30 y 38 kDa, y un fragmento terminal C de 42 kDa<sup>(8,9)</sup>. Los cuatro fragmentos proteolíticos se autoensamblan en una asociación compleja no covalente en la membrana del merozoito, la cual está involucrada directamente en el proceso invasivo<sup>(10)</sup>. Después de tal ruptura, e inmediatamente antes de la invasión, el fragmento de 42 kDa es sometido a una proteólisis dependiente de calcio por una serin proteasa presente en la membrana del merozoito<sup>(9)</sup>, resultando en un fragmento terminal N de 33 kDa

y un fragmento terminal C de 19 kDa. Este último permanece unido a la membrana del merozoito y es el único fragmento de proteína llevado por el parásito dentro del glóbulo rojo<sup>(3)</sup>.

La MSP-1 completa<sup>(11)</sup>, sus fragmentos de proteína recombinantes<sup>(12)</sup> y los péptidos sintéticos basados en la MSP-1<sup>(13)</sup> inducen inmunidad protectora parcial o completa en simios. Los péptidos también inducen inmunidad parcial en humanos infectados experimentalmente<sup>(14)</sup>; anticuerpos monoclonales y policlonales dirigidos contra la MSP-1 inhiben la invasión del glóbulo rojo por merozoitos *in vitro*<sup>(15)</sup>. Los péptidos sintéticos contra la malaria, SPf66, contienen una secuencia conservada del fragmento de 83 kDa<sup>(14)</sup>.

En un trabajo previo, donde se estudiaron los aminoácidos de las secuencias de la MSP-1 involucrados en la unión e invasión a las células rojas de sangre en humanos<sup>(4)</sup>; se escogieron 79 péptidos no sobrelapados con un tamaño de 20 aminoácidos, que fueron sintetizados químicamente usando la técnica de síntesis de fase sólida para múltiples péptidos<sup>(16,17)</sup>, encontrando que 9 de ellos presentan alta unión al receptor del glóbulo rojo.

Los estudios de la unión de péptidos a las moléculas de HLA de clase I y II han encontrado una alta aparición de aminoácidos específicos en posiciones específicas, los denominados motivos de unión<sup>(18)</sup> que han sido el fundamento para el desarrollo de métodos predictivos basados en secuencias<sup>(19)</sup>, en este trabajo se realiza una analogía de dichos motivos cuantificando aminoácidos específicos a través de reglas que definen conjuntos. Desde la perspectiva de la teoría de conjuntos en un trabajo previo, Rodríguez desarrolló un trabajo de péptidos de unión al HLA de clase II, encontrando un comportamiento caracterizable a partir de reglas halladas experimentalmente en péptidos nonaméricos específicos, logrando diferenciar péptidos de unión y no unión mediante operaciones matemáticas<sup>(20)</sup>.

El propósito de la presente investigación es construir una metodología capaz de caracterizar matemáticamente la unión de péptidos del MSP-1 al receptor del glóbulo rojo, partiendo de resultados experimentales<sup>(4)</sup> y de un análisis matemático por medio de la teoría de conjuntos.

### Definiciones

**Conjunto:** selección de elementos que cumplen una regla  $P(x)$ , definida previamente. El conjunto A se define como:

$$A = \{x \mid x \text{ cumple con la regla } P(x)\}$$

**Subconjunto:** Si A es subconjunto de B, A está contenido en B:

$$(A \subset B) \Leftrightarrow (\forall x \in A \Rightarrow x \in B)$$

**ANEXO I.** Para ejemplificar la metodología realizada se tomarán como ejemplo 3 secuencias de las 79 estudiadas, de éstas 2 son de alta unión y 1 de no unión según resultados experimentales (3) (Tablas I, II, III, IV y V)

**TABLA I.** Péptidos usados como ejemplo de la metodología

Número	Péptido	Alta unión
1513	GYSLFQKEKMVLNEGTSFTA	Sí
1577	FKVLSKLEGKLDNLEKK	Sí
1527	LEKRIDLTKKNENIKELYDK	No

**TABLA II.** Aplicación de la Regla 1 a las secuencias ejemplo

Péptido	Frecuencia de aparición	Sumatoria	Pertenencia al Conjunto I
GYSLFQKEKMVLNEGTSFTA	2E	2	∈
FKVLSKLEGKLDNLEKK	2E + 1D	3	∈
LEKRIDLTKKNENIKELYDK	3E + 2D + 1R	6	∈

**TABLA III.** Aplicación de la Regla 2 a las secuencias ejemplo

Péptido	Posición 1	Posición 11	Posición 15	Pertenencia al Conjunto II
GYSLFQKEKMVLNEGTSFTA	G	V	G	∈
FKVLSKLEGKLDNLEKK	F	L	L	∈
LEKRIDLTKKNENIKELYDK	L	N	K	∈

**TABLA IV.** Aplicación de la Regla 3 a las secuencias ejemplo

Péptido	Pertenencia al Conjunto II
GYSLFQKEKMVLNEGTSFTA	∈
FKVLSKLEGKLDNLEKK	∈
LEKRIDLTKKNENIKELYDK	∈

**TABLA V.** Aplicación de la Regla 4 a las secuencias ejemplo

Péptido	Suma frecuencia de aparición	Sumatoria	Pertenencia al Conjunto IV
GYSLFQKEKMVLNEGTSFTA	2K + 2E	4	∈
FKVLSKLEGKLDNLEKK	6K + 2E + 1D	9	∈
LEKRIDLTKKNENIKELYDK	5K + 3E + 2D + 1R	11	∉

**Intersección:** La intersección de los conjuntos A y B, se define como:

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\}$$

**Unión:** La unión de los conjuntos A y B, se define como:

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\}$$

**MATERIAL Y MÉTODO**

Se estudiaron 79 secuencias de péptidos no sobrelapados de la proteína MSP-1 de *Plasmodium falciparum*, a partir de la distribución de los aminoácidos que las conforman y de las características químicas de éstos. Posteriormente se establecieron cinco reglas que definen la

**TABLA I.** Número de péptidos que pertenecen ( $\in$ ) y no pertenecen ( $\notin$ ) a cada conjunto

Conjunto	$\in$	$\notin$
I	54	25
II	43	36
III	24	55
IV	43	36

pertenencia o no de un péptido a cuatro conjuntos diferentes.

La determinación de estas reglas se hizo partiendo de observaciones experimentales de los péptidos, con las propiedades químicas involucradas en la unión de los aminoácidos. Las reglas 1 y 5 cuantifican la aparición de aminoácidos cargados positivos y negativos en el péptido, la regla 2 realiza una analogía con los motivos de unión<sup>(18,19)</sup>, cuantificando aminoácidos específicos en posiciones específicas, finalmente las reglas 3 y 4 cuantifican la forma en que se repiten aminoácidos específicos, para con ellas encontrar diferencias concretas de los péptidos que se unen y los que no para sintetizarlas y así poder diferenciar los péptidos de alta unión de la MSP-1

Las reglas para la determinación de los conjuntos son:

- Regla 1. Valores de 2, 3, 4, 6 ó 7 en la suma de la frecuencia de aparición de los aminoácidos D, E, H y R en la secuencia.
- Regla 2. Presencia de los aminoácidos K, N, G, A, V, L, F, M, P, R, E, T ó I en las posiciones 1, 11 y 15 o en las posiciones 2, 3, 6, 7, 8, 10, 11, 16, 19, y 20.
- Regla 3. Presencia de la cadena KK, TT, SS, NN-ENE, ó KXX en el péptido. Donde NN-ENE quiere decir que además de la cadena NN hay una cadena ENE separada en la secuencia y KXX se define por la presencia de dos lisinas (K) separadas por Y ó E.
- Regla 4. Presencia de la cadena LL, EE, MM, VV, ó VNV en el péptido.
- Regla 5. Valores de 4, 5, 6, 7 ó 9 en la suma de la frecuencia de aparición de los aminoácidos K, D, E, H y R en la secuencia.

A partir de las reglas anteriores se definieron los siguientes conjuntos:

- I = {x | x cumplen la primera regla}
- II = {x | x cumplen la segunda regla}
- III = {x | x cumplen la tercera regla y no cumplen la cuarta regla}
- IV = {x | x cumplen la quinta regla}

Donde  $x \in S$ . Y S es el conjunto de todas las posibles cadenas peptídicas de 20 aminoácidos de MSP-1.

**TABLA II.** Número de péptidos que pertenecen ( $\in$ ) y no pertenecen ( $\notin$ ) a cada conjunto

Conjunto	$\in$	$\notin$
I	54	25
I $\cap$ II	33	46
(I $\cap$ II) $\cap$ III	11	68
((I $\cap$ II) $\cap$ III) $\cap$ IV)	9	70
((I $\cap$ II) $\cap$ III) $\cap$ IV): conjunto unión		

Posteriormente, estas reglas fueron evaluadas en los 79 péptidos no sobrelapados de la proteína MSP-1, con un tamaño de 20 aminoácidos cada uno, tomados de un estudio experimental ((4), ver Anexo 1). Finalmente se realizan las intersecciones de los cuatro conjuntos, definida ésta como el conjunto (Unión): Unión =  $\{(I \cap II) \cap III) \cap IV\}$ . Por tratarse de una metodología matemática no requiere de análisis estadísticos ni el uso de una gran cantidad de datos experimentales.

**RESULTADOS**

De los 79 péptidos evaluados, 54 pertenecen al conjunto I, 43 pertenecen al conjunto II, 24 pertenecen al conjunto III, y 43 pertenecen al conjunto IV (Tabla I).

De los 79 péptidos, 33 pertenecen al conjunto I  $\cap$  II, 11 pertenecen al conjunto (I  $\cap$  II)  $\cap$  III, y 9 pertenecen al conjunto ((I  $\cap$  II)  $\cap$  III)  $\cap$  IV), (Tabla II), estos 9 péptidos corresponden a los péptidos de alta unión comprobados experimentalmente: 1513, 1522, 1577, 1582, 1585, 1589, 1590, 1591, 1598, siendo este conjunto el que define los péptidos de alta unión.

**DISCUSIÓN**

Este es el primer trabajo en el que, a partir de la teoría de conjuntos, de observaciones experimentales y de reglas diferenciadoras, se caracteriza el fenómeno de alta unión de la proteína de membrana MSP-1 al receptor de glóbulo rojo, encontrando órdenes matemáticos que diferencian los péptidos que presentan alta unión, de los péptidos que no, de forma objetiva y reproducible.

Por medio de observaciones experimentales, se establecieron 5 reglas que definen la pertenencia a 4 conjuntos de péptidos. Estas reglas se basan en patrones matemáticos que representan las propiedades de las cadenas aminoácidas componentes de los péptidos, evidenciando desde de la teoría de conjuntos órdenes matemáticos

acausales asociados a los péptidos de alta unión, que permiten diferenciarlos de los péptidos de baja unión. Las operaciones entre los conjuntos establecidos por las cuatro reglas describen un orden físico y matemático acausal que determina los péptidos de alta unión al glóbulo rojo.

Este trabajo permite, mediante reglas específicas que se sintetizan con la teoría de conjuntos, encontrar un orden en la unión de los péptidos de alta unión al receptor del glóbulo rojo en el proceso de invasión a éste, a diferencia del trabajo de unión al HLA clase II basado también en la teoría de conjuntos<sup>(20)</sup> donde se diferencian péptidos de unión a la molécula presentadora en el proceso antigénico para la respuesta inmune.

Algunos trabajos de predicción de unión de péptidos usan algoritmos evolucionarios junto con redes neuronales artificiales partiendo de grandes bases de datos para predecir qué péptidos se unirán, basados en resultados experimentales<sup>(21-25)</sup>. Otro tipo de desarrollo para predecir unión de péptidos es un SVM (Support Vector Machine), que no utiliza una gran cantidad de datos, sin embargo sus predicciones no obtienen altos porcentajes de correlación experimental<sup>(24)</sup>. Este es un trabajo teórico práctico que desarrolla una metodología que evidencia órdenes matemáticos objetivos y reproducibles, determinados por medio de la teoría de conjuntos, cuyas correlaciones se corroboran experimentalmente. No se necesitan análisis estadísticos, pues los péptidos son caracterizados a través de reglas simples, establecidas desde la teoría de conjuntos.

### Limitaciones del estudio

La metodología presentada para la determinación de péptidos de alta unión ha sido corroborada hasta ahora únicamente para caracterizar la unión al glóbulo rojo en esta proteína específica. Para su generalización sería útil estudiar reglas similares a éstas, para moléculas importantes en la unión ligando-receptor de malaria y de otras enfermedades como tuberculosis, VIH y cáncer, entre otras.

### AGRADECIMIENTOS

Al Grupo Insight por obligarme a publicar éste trabajo, que estaba hecho hace muchos años.

Al doctor Manuel Elkin Patarroyo por apoyar mi internado especial y mi rural especial en los años 1996 y 1997 en el Instituto de Inmunología, los cuales dediqué a profundizar en la esencia teórica acausal del Universo de la física moderna (Mecánica estadística, Mecánica cuántica y Teoría del Caos). DEDICACIÓN: A mis hijos. A Alejandro Cáce-

res, "científico de cafetería", quien fue mi Maestro de Física Cuántica y quien me dio formación para entender los fenómenos acausalmente y confiar en este resultado acausal en esta rama del conocimiento tan causalista y determinista, como la interacción ligando-receptor de la Biología Molecular.

### CONFLICTO DE INTERÉS

El autor declara no tener conflicto de interés.

#### CORRESPONDENCE TO:

Javier Rodríguez, M.D.  
Director del Grupo Insight.  
Carrera 79b número 51-16 sur  
Bogotá, Colombia  
Phone (057-1) 4527541  
E-mail: grupoinight2025@yahoo.es

### BIBLIOGRAFÍA

1. Hrbacek KJ. Introduction to set theory. 3rd Ed, Marcel Dekker, Inc. New York. 1999; pp 1-33.
2. World malaria report 2005. Geneva, World Health Organization, United Nations Children's Fund, 2005.
3. Hadley TJ, Klotz F,W, Miller LH. Invasion of erythrocytes by malaria parasites: A cellular and molecular overview. *Annu Rev Microbiol* 1986; 40:451-77.
4. Urquiza M, Rodríguez LE, Suarez JE, Guzmán F, Ocampo M, Curtidor H, et al. Identification of Plasmodium falciparum MSP-1 peptides able to bind to human red blood cells. 1996. Blackwell Science Ltd. *Parasite Immunol*, 1996; 18: 515-526.
5. Kochan J, Perkins ME, Ravetch JV. A tandemly repeated sequence determines the binding domain for an erythrocyte receptor-binding protein of Plasmodium falciparum. *Cell* 1985;44:689-696.
6. Camus D, Hadley TJ. A Plasmodium falciparum antigen that binds to host erythrocytes and merozoites. *Science* 1985;230:553-556.
7. Perkins ME, Rocco LJ. Sialic acid-dependent binding of Plasmodium falciparum merozoite surface antigen, Pf200, to human erythrocytes. *J Immunol* 1988;141:3190-3196.
8. Lyon JA, Haynes JD, Diggs CL, Chulay JD, Haidaris CG, Pratt-Rossiter J. Monoclonal antibody characterization of the 195-kilodalton major surface glycoprotein of Plasmodium falciparum malaria schizonts and merozoites: Identification of additional processed products and a serotype-restricted repetitive epitope. *J Immunol* 1987;138: 895-901.
9. Blackman MJ, Holder AA. Secondary processing of the Plasmodium falciparum Merozoite surface Protein (MSP-1) by a calcium-dependent membrane-bound serine protease shedding of MSP33

- as a noncovalently associated complex with other fragments of the MSP-1. *Mol Biochem Parasitol* 1992;50:307-316.
10. McBride JS, Heidrich HG. Fragments of the polymorphic Mr 185000 glycoprotein from the surface of isolated *Plasmodium falciparum* Merozoites from an antigenic complex. *Mol Biochem Parasitol* 1987;23:71-84.
  11. Siddiqui WA, Tam LQ, Kramer KJ, Hui GS, Case SE, Yamaga KM, et al. Merozoite surface coat precursor protein completely protects Aotus monkeys against *Plasmodium falciparum* malaria. *Proc Natl Acad Sci USA* 1987; USA 84: 3014-3018.
  12. Etlinger HM, Casper P, Matile H, Schoenfeld HJ, Stueber D, Takacs B. Ability of recombinant or native proteins to protect monkeys against heterologous challenge with *Plasmodium falciparum*. *Infect Immunity* 1991;59:3498-3503.
  13. Patarroyo ME, Romero P, Torres ML, Clavijo P, Moreno A, Martínez A, et al. Induction of protective immunity against experimental infection with malaria using synthetic peptides. *Nature* 1987; 328: 629-632.
  14. Patarroyo ME, Amador R, Clavijo P, Moreno A, Guzman F, Romero P, et al. A synthetic vaccine protects human against challenge with asexual blood stages of *Plasmodium falciparum* malaria. *Nature* 1988;332:156-159.
  15. Hui GSN, Siddiqui WA. Serum from Pf195 protected Aotus monkeys inhibit *Plasmodium falciparum* growth in vitro. *Experimental Parasitol* 1987;64:519-522.
  16. Merrifield RB. Solid phase peptide synthesis I. The synthesis of a tetrapeptide. *J Amer Chem Soc* 1963;85:2149-2154.
  17. Houghten RA. General method for the rapid solid phase synthesis of large numbers of peptides: Specificity of antigen antibody interaction at level of individual aminoacids. *Proc Natl Acad Sci USA* 1982;82:5131-5135.
  18. Hammer J, Takacs B, Sinigaglia F. Identification of a motif for HLA-DR1 binding peptides using M13 Display Libraries. *J Exp Med* 1992;4:1007-1013.
  19. Zhao B, Sakharkar KR, Lim ChS, Kanguane P, Sakharkar MK. MHC-Peptide binding prediction for epitopes based vaccine design. *Int J Integrative Biol* 2007;2:127-140.
  20. Rodríguez Velásquez J. Teoría de conjuntos aplicada a la caracterización matemática de unión de péptidos al HLA clase II. *Rev Cienc Salud* 2008;1:9-15.
  21. Reche PA, Glutting JP, Zhang H, Reinherz EL. Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. *Immunogenetics* 2004;56:405-419.
  22. 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;6:581-591
  23. Nielsen M, Lundegaard C, Worning P, Sylvester C, Lamberth K, Buus S, et al. Improved prediction of MHC class I and II epitopes using a novel Gibbs sampling approach. *Bioinformatics Advance* 2004;20:1388-1397.
  24. Dönnes P, Elofsson A. Prediction of MHC class I binding peptides, using SVMHC *BMC Bioinformatics* 2002;3:3-25.
  25. Brusica 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;2:121-130.