

Caracterización *in silico* de la proteína Leishmanolisina (GP63) de *Leishmania braziliensis* y búsqueda de epitopes de utilidad para el desarrollo de vacunas.

In silico characterization of Leishmanolysin protein (GP63) from *Leishmania braziliensis* and epitope search for vaccine development.

Luis A. Rodríguez¹, Gustavo A. Sandoval¹

RESUMEN

Objetivo: Realizar la caracterización *in silico* de la Leishmanolisina GP63 de *Leishmania braziliensis* y búsqueda de epitopes de utilidad para el desarrollo de vacunas. **Material y métodos:** Se obtuvo la secuencia aminoacídica de la Leishmanolisina de *L. braziliensis* cepa MHOM/BR/75/M2904 (código de acceso CAM37262.1) a partir de la base de datos del GenBank. Dicha secuencia fue utilizada para determinar sus principales parámetros bioquímicos, dominios conservados entre proteínas cercanamente relacionadas, predicción de estructuras secundarias y modelamiento tridimensional, empleando las herramientas bioinformáticas ProtParam, Clustal Omega, PHD Secondary Structure Prediction, SWISS-MODEL, Phyre2 y Robetta, respectivamente. La visualización de modelos tridimensionales se realizó empleando PyMol y la comparación de estos mismos. **Resultados:** La Leishmanolisina GP63 de *L. braziliensis* es una proteína de 66.8 kDa con un pI de 5.68. Asimismo, presenta un 34.2% de láminas β , seguido por un 20.7% de α hélices y un 45.1% de loops internos, así como un dominio conservado en el sitio activo (HEEXXH) característico de la metaloproteasas dependientes de zinc. Finalmente, se encontraron tres clusters de péptidos correspondientes a epitopes potencialmente inmunogénicos distribuidos en la estructura tridimensional de la proteína. **Conclusión:** La Leishmanolisina GP63 de *L. braziliensis* es una metaloproteasa ácida de alto peso molecular cuyo motivo de unión al zinc contribuiría a su rol en la patogenicidad del parásito. Además, se constituye como un posible blanco para el desarrollo de vacunas y fármacos inhibidores.

Palabras Clave: *Leishmania braziliensis*, Leishmanolisina, metaloproteasa, HEEXXH.

ABSTRACT

Objective: To perform an *in silico* characterization of the Leishmanolysin GP63 from *Leishmania braziliensis* and epitope search for vaccine development. **Material and methods:** To carry out this analysis, the amino acid sequence of Leishmanolysin from *L. braziliensis* strain MHOM/BR/75/M290 was obtained from GenBank database (accession number CAM37262.1). This sequence was used to the determination of biochemical parameters, conserved domains, secondary structure prediction and tridimensional modelling, using ProtParam, Clustal Omega, PHD Secondary Structure Prediction SWISS-MODEL, Phyre2 and Robetta, respectively. Visualization of tridimensional models was performed using PyMol. **Results:** The Leishmanolysin GP63 from *L. braziliensis* is an acidic molecule with a molecular weight of 66.8 kDa and a pI of 5.68. It also shows a composition of 34.2% of β -sheet, 20.7% of α -helix and 45.1% of internal loops. This protein displays a conserved domain in the active site (HEEXXH) which is characteristic of zinc-dependent metalloproteases. Using sequence information, three clusters of peptides were predicted as potentially immunogenic epitopes which are distributed along the proteins structure. **Conclusion:** The Leishmanolysin GP63 from *L. braziliensis* is an acidic metalloprotease with a high molecular weight and its zinc-binding domain potentially contributes to its role in parasite infection. Furthermore, this protein will become a possible target for vaccine and inhibitor development.

Key words: *Leishmania braziliensis*, Leishmanolysin, metalloprotease, HEEXXH.

¹ Grupo de Estudio en Bioinformática Estructural. Facultad de Ciencias Biológicas. Universidad Nacional Mayor de San Marcos. Lima 01. Perú.

INTRODUCCIÓN

Los tripanosomatídeos son un grupo cosmopolita de flagelados que parasitan en conjunto tanto animales como plantas, los cuales se distinguen de otros protozoarios por la presencia de DNA kinetoplastidial, aislamiento de sus enzimas glicolíticas en un glicosoma, así como de una variación antigénica, trans-splicing, edición del RNA y una organización nuclear muy peculiar (1).

Entre los miembros de esta familia destacan los géneros *Leishmania* y *Trypanosoma*, los cuales son responsables de muchas enfermedades en humanos, principalmente de áreas tropicales y subtropicales. El avance de la enfermedad depende del estado inmunológico del huésped y de la cepa, donde las formas infectantes o promastigotes pueden sobrevivir en los macrófagos de los individuos afectados y causar diferentes tipos de lesiones, entre ellas lesiones cutáneas ulcerosas (2).

Durante el proceso infeccioso, existen diversos tipos de proteínas que intervienen asegurando su efectividad. Un grupo de éstas induce la activación de Cdc42 para reorganizar la red de actinas alrededor de los parásitos internalizados, permitiendo la regulación del citoesqueleto de las células afectadas (3). Otras le permiten al parásito la evasión de la respuesta inmune innata desde los momentos tempranos de la infección, bloqueando la inducción y regulación de factores cruciales como la producción de óxido nítrico, factor de necrosis tumoral alfa (TNF α), interleuquina 12 (IL-12), especies reactivas de oxígeno, entre otros. Asimismo puede influenciar en las vías de señalización intracelular, sobre todo a nivel de la fosforilación y activación de varias quinasas de las familias JAK y MAPK, y afectando así a los factores de transcripción críticos para la activación completa de mediadores inflamatorios (STAT1 α , NF- κ B, AP-1) (2).

Entre las principales etapas del proceso infeccioso se encuentra la internalización del parásito, en la cual destacan diversas moléculas como el lipofosfoglicano (LPG) y la glicoproteína GP63 anclada a glicofosfatidilinositol (GPI). Esta proteína, denominada Leishmanolisina o también MSP (Major Surface Protease), es una metaloproteasa muy abundante (0.5-1% de la proteína celular total) y se distribuye en toda la superficie del promastigote, incluyendo el flagelo y el bolsillo flagelar, siendo crucial en su ingreso al macrófago y en la activación del complemento (3). En estudios previos se ha determinado la estructura tridimensional de GP63 de *Leishmania*

major (causante de Leishmaniasis en el viejo mundo), donde se ha encontrado un motivo HEXXHXXGXXH y un pro-péptido N-terminal que se remueve durante la maduración de la proteína (4, 5); sin embargo no existen estudios similares sobre la Leishmanolisina de *L. braziliensis*, el cual es considerado el principal agente causal de Leishmaniasis cutánea en el neotrópico (5, 6).

Por este motivo, en el presente trabajo se ha realizado la caracterización *in silico* de la Leishmanolisina GP63 de *Leishmania braziliensis* y la búsqueda de epítopes de utilidad para el desarrollo de vacunas.

MATERIALES Y MÉTODOS

Obtención de la secuencia aminoacídica de la Leishmanolisina GP63 y determinación de parámetros bioquímicos

La secuencia primaria de la Leishmanolisina de *L. braziliensis* cepa MHOM/BR/75/M2904 fue obtenida a través de la base de datos del GenBank (código de acceso CAM37262.1). Dicha secuencia fue utilizada para los ensayos de predicción de estructura tridimensional. Para determinar los principales parámetros bioquímicos de la enzima en estudio, se empleó la herramienta ProtParam del ExpASY (<http://web.expasy.org/protparam/>) (7).

Alineamiento múltiple de secuencias aminoacídicas

Para los ensayos de alineamiento múltiple se obtuvieron las secuencias aminoacídicas de *L. mexicana* (código de acceso: XP_003872884.1), *L. major* (código UniProtKB: P08148), *L. panamensis* (códigos de acceso: AAB97895.1, XP_010696764.1) y *L. guyanensis* (código de acceso: Q00689.1). Dichas secuencias fueron alineadas y editadas utilizando la herramienta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) (8)

Predicción de estructuras secundarias y alineamiento estructural

El análisis de predicción de estructuras secundarias presentes en la Leishmanolisina GP63 fue realizada empleando la herramienta PHD Secondary Structure Prediction Method (<http://npsa-prabi.ibcp.fr>) (9). Además, el alineamiento estructural se realizó en la plataforma

Cn3D (versión 4.3.1) utilizando como molde diversas proteínas estructuralmente similares.

Modelamiento tridimensional de la Leishmanolisina GP63 de L. braziliensis

El modelo tridimensional de la enzima en estudio se obtuvo empleando la herramienta SWISS-MODEL (<http://swissmodel.expasy.org>) (10), a partir de la información depositada en la base de datos del Protein Data Bank (PDB). Por otro lado, se empleó la herramienta Phyre2 Protein Homology/analogy Recognition Engine (<http://www.sbg.bio.ic.ac.uk/phyre2/>) (11), así como el servidor Robetta Full Chain Protein Structure Prediction (<http://rosetta.bakerlab.org/>) (12), para la predicción de estructuras mediante modelamiento por *threading* y *ab initio*, respectivamente. La visualización de los modelos obtenidos se realizó empleando el programa PyMol Molecular Graphics System (versión 1.3).

Predicción de epítopes con potencial inmunogénico presentes en Leishmanolisina GP63 de L. braziliensis

A partir de la secuencia aminoacídica obtenida previamente, se realizó la predicción de péptidos inmunogénicos en el servidor MHCpred 2.0 (<http://www.ddg-pharmfac.net/mhcpred>) (13). Para este fin se utilizaron los alelos HLA-A*0201 y HLA-DRBQ*0101, eligiéndose los 5 resultados de mayor confianza de predicción. Los péptidos elegidos fueron alineados estructuralmente utilizando el modelo tridimensional generado en este trabajo utilizando el servidor de predicción de epítopes The Pepitope Server (<http://pepitope.tau.ac.il/>) (14), determinándose su ubicación en la proteína en estudio.



Fig. 1: Alineamiento múltiple de diversas proteínas Leishmanolisinas empleando Clustal Omega. Se resalta el sitio activo y la expansión del motivo catalítico, así como los aminoácidos conservados (*) a lo largo de la secuencia primaria de las proteínas

RESULTADOS

La Leishmanolisina GP63 de *L. braziliensis* es una proteína de 66.8 kDa y presenta un pI de 5.68. Del total de residuos de aminoácidos, 48 residuos están cargados negativamente (ácido aspártico y ácido glutámico) y 36 están cargados positivamente (arginina y lisina). A partir de alineamiento múltiple realizado se observan regiones conservadas a lo largo de la proteína en estudio, las cuales se corresponden en parte con α hélices y con láminas β , y/o sus loops circundantes. Asimismo se puede observar el inicio de la proteína madura, que es común para todas las secuencias (Fig. 1).

El modelo tridimensional obtenido por homología concuerda con las tres regiones reportadas previamente para la Leishmanolisina de *L. major*, presentando un dominio N-terminal, un dominio central y un dominio C-terminal bien definidos (Fig. 2).

En dicha representación se resalta la posición del sitio activo, observándose que dos de las histidinas forman parte de un α hélice de manera similar a lo encontrado en la GP63 de *L. major*. Además, la tercera histidina del sitio activo se encuentra en un loop a la misma distancia en la estructura y el modelo obtenido por homología. Luego del análisis de predicción de estructuras secundarias, se observa un 34.2% de láminas β , seguido por un 20.7% de α hélices y un 45.1% de loops internos. Esto fue corroborado empleando los servidores Phyre 2 y Robetta, identificando la misma distribución en el molde 1LML de GP3 de *L. major*. Finalmente, el mapeo de los epítopes inmunogénicos seleccionados permitió determinar su distribución en cluster a nivel de la superficie de la proteína en estudio (Fig. 3).

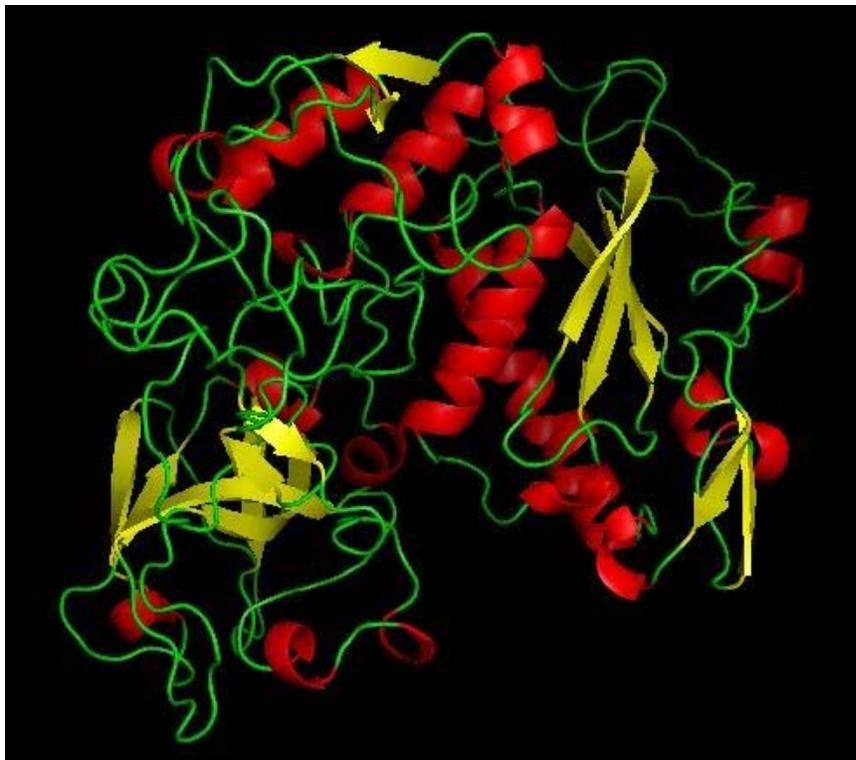


Fig. 2: Modelo tridimensional de la enzima Leishmanolisina GP63 de *L. braziliensis* empleando modelamiento por homología. Las α hélices se muestran de color rojo mientras que las láminas β se muestran de color amarillo

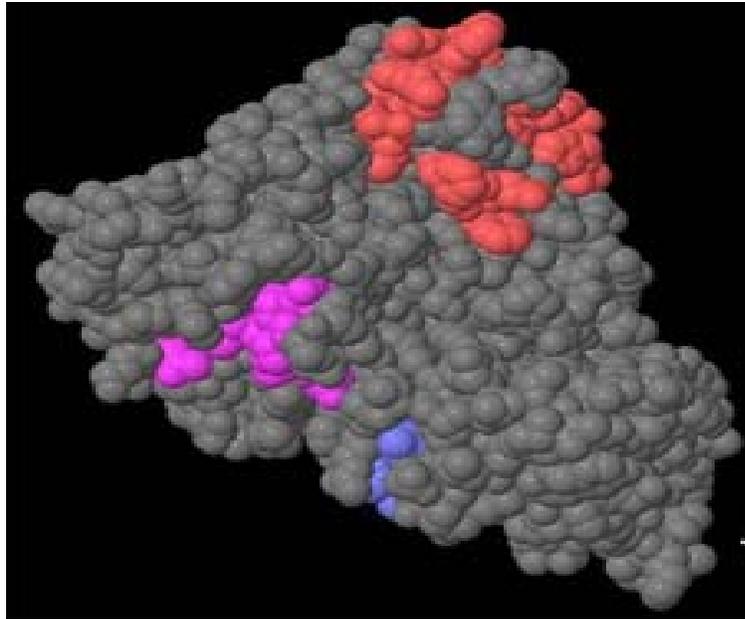


Fig. 3: Distribución de epítopes inmunogénicos potencial de la enzima Leishmanolisina GP63 de *L. braziliensis*. El cluster 1 corresponde a los péptidos FLGGFSEFL, LVTYLIPQA y FLDYCPFIV (color naranja), el cluster 2 corresponde al péptido YLIPQALQL (color azul) y el cluster 3 corresponde al péptido QVFSDDHPA (color rosado).

DISCUSIÓN

En el presente estudio se ha realizado la caracterización *in silico* de la proteína Leishmanolisina GP63 de *L. braziliensis*. Se observó que presenta porcentajes de identidad relativamente altos con las provenientes de otras especies del viejo y nuevo mundo. Asimismo, la identidad entre la secuencia problema y la secuencia obtenida a partir de la estructura cristalográfica es del 70%, lo cual permite deducir la importancia de esta proteína para el proceso infeccioso de *L. braziliensis*. Van der Auwera et al. (15) realizaron un secuenciamiento multilocus con fin de obtener una filogenia confiable de las especies de *Leishmania* a partir de lo cual obtuvieron dos clusters, uno que albergaba especies como *L. mexicana* y *L. tropica* y otro que albergaba especies como *L. peruviana*, *L. braziliensis* y *L. guyanensis*. Las regiones conservadas se corresponden principalmente con residuos formadores de estructuras secundarias y los loops que serían importantes para la estabilización de la proteína (16). Con respecto al dominio conservado en el sitio activo (HEEXH) se ha encontrado que es bastante similar al de proteasas dependientes de zinc como la astacina o la colagenasa, donde la presencia de una molécula de agua sostenida

entre el zinc el oxígeno del ácido glutámico es crucial para la acción catalítica de esta enzima (4). Finalmente, en cuanto al análisis de epítopes inmunogénicos, se eligió el HLA-A (MHC-I: proteínas citosólicas) y el HLA-DR (MHC-II: proteínas del medio externo) por la capacidad de los promastigotes de internalizarse y ser degradados dentro, como por la posible degradación externa del parásito. Al haberse usado solo péptidos con alta confianza para la predicción de inmunogenicidad, destaca el péptido YLIPQALQL, afín a ambas clases de MHC. Además, en casos del HLA-A, se encuentran péptidos que forman cluster a nivel de la estructura tridimensional de la proteína modelada.

Conclusiones:

-La Leishmanolisina GP63 de *L. braziliensis* es una metaloproteasa ácida de alto peso molecular y que presenta un alto grado de identidad con sus homólogas en especies cercanas. La conservación de su motivo de unión a zinc contribuiría con su rol en la patogenicidad de los promastigotes del parásito. Además, se constituye como un posible blanco para el desarrollo de vacunas y fármacos inhibidores.

AUTOR DE CORRESPONDENCIA

Gustavo A. Sandoval, MSc.
Facultad de Ciencias Biológicas
Universidad Nacional Mayor de San Marcos
Av. Venezuela s/n - Ciudad Universitaria, Lima
- Perú.
Teléfono: +51-1-6197000 Anexo 1558.
E-mail: gsandovalp@unmsm.edu.pe

REFERENCIAS BIBLIOGRAFICAS

1. Santos AL, Branquinha MH, D'Avila-Levy CM. The ubiquitous gp63-like metalloprotease from lower trypanosomatids: in the search for a function. *An Acad Bras Cienc.* 2006;78(4):687-714.
2. Olivier M, Atayde VD, Isnard A, Hassani K, Shio MT. Leishmania virulence factors: focus on the metalloprotease GP63. *Microbes Infect.* 2012;14(15):1377-89.
3. Hallé M, Gomez MA, Stuble M, Shimizu H, McMaster WR, Olivier M, et al. The Leishmania surface protease GP63 cleaves multiple intracellular proteins and actively participates in p38 mitogen-activated protein kinase inactivation. *J Biol Chem.* 2009;284(11):6893-908
4. Schlagenhauf E, Etges R, Metcalf P. The crystal structure of the Leishmania major surface proteinase leishmanolysin (gp63). *Structure.* 1998;6(8):1035-46.
5. Yao C, Donelson JE, Wilson ME. The major surface protease (MSP or GP63) of Leishmania sp. Biosynthesis, regulation of expression, and function. *Mol Biochem Parasitol.* 2003;132(1):1-16.
6. Mears ER, Modabber F, Don R, Johnson GE. A review: The current in vivo models for the discovery and utility of new anti-leishmanial drugs targeting cutaneous Leishmaniasis. *PLoS Negl Trop Dis.* 2015;9(9):e0003889.
7. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy Server. In: Walker JM, editor. *The Proteomics Protocols Handbook.* Humana Press; 2005. p. 571-607.
8. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011 Oct 11;7:539.
9. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol.* 1993 Jul 20;232(2):584-99.
10. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014 Jul;42(Web Server issue):W252-8.
11. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10(6):845-58.
12. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 2004;32(Web Server issue):W526-31.
13. Guan P, Doytchinova IA, Zygori C, Flower DR. MHCpred: A server for quantitative prediction of peptide-MHC binding. *Nucleic Acids Res.* 2003;31(13):3621-4.
14. Mayrose I, Shlomi T, Rubinstein ND, Gershoni JM, Ruppin E, Sharan R, et al. Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. *Nucleic Acids Res.* 2007;35(1):69-78.
15. Van der Auwera G, Ravel C, Verweij JJ, Bart A, Schönian G, Felger I. Evaluation of four single-locus markers for Leishmania species discrimination by sequencing. *J Clin Microbiol.* 2014;52(4):1098-104.
16. Sádlová J, Volf P, Victor K, Dujardin JC, Votýpka J. Virulent and attenuated lines of Leishmania major: DNA karyotypes and differences in metalloproteinase GP63. *Folia Parasitol (Praha).* 2006;53(2):81-90.

Recibido: 28/09/2015
Aceptado: 18/12/2015