

Análisis bioinformático de la endolisina LysB4: Un agente antimicrobiano utilizado en el biocontrol de *Bacillus cereus* en alimentos

Bioinformatics analysis of LysB4 endolysine: An antimicrobial agent used in the biocontrol of *Bacillus cereus* in food

Renzo Punil¹ y Gustavo A. Sandoval¹

RESUMEN

Objetivo: Realizar el análisis bioinformático de la endolisina LysB4 empleada en el biocontrol de *Bacillus cereus* en la industria alimentaria. **Materiales y Métodos:** Se realizó la búsqueda de endolisinas bacterianas empleando la base de datos del GenBank. A partir de dicha información, se obtuvo la secuencia aminoacídica de la endolisina LysB4 (código de acceso AFF27501.1). 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, Prosite, PHD Secondary Structure Prediction, Cn3D, SWISS-MODEL, Phyre2 y Robetta, respectivamente. La visualización de modelos tridimensionales se realizó empleando PyMol. **Resultados:** La endolisina LysB4 posee 262 aminoácidos, un peso molecular de 27.8 kDa y un pI de 9.21, así como dominios conservados correspondientes a las superfamilias VanY y SH3, involucradas en su actividad catalítica, así como en la interacción con sustrato. Con respecto a su estructura secundaria, la endolisina LysB4 presenta 21.4% de hélices α , 28.6% de láminas β y 50% de loops internos. Además, de acuerdo con la predicción de su estructura tridimensional, la proteína en estudio reveló un plegamiento del tipo $\alpha+\beta$ con un core compuesto por láminas β antiparalelas rodeada de hélices α . Con respecto a su dominio de unión, esta región está conformada por láminas β antiparalelas. **Conclusión:** La endolisina LysB4 es una proteína de mediano peso molecular que muestra una arquitectura modular importante para su actividad lítica.

Palabras clave: LysB4, *Bacillus cereus*, tridimensional modelling.

ABSTRACT

Objective: To perform a bioinformatics analysis of endolysin B4 used for the biocontrol of *Bacillus cereus* in food industry. **Materials and Methods:** A intensive search of bacterial endolysins was performed using GenBank data base. Using this information, the aminoacid sequence of endolysin LysB4 was obtained (accession code: AFF27501.1), and used to determine its biochemical parameters, conserved domains between closely related proteins, prediction of secondary structures and tridimensional modelling, by means of different bioinformatics tools such as ProtParam, Prosite, PHD Secondary Structure Prediction, Cn3D, SWISS-MODEL, Phyre2 and Robett, respectively. The visualization of tridimensional models was performed in PyMol. **Results:** The endolysin B4 is formed by 262 aminoacids, and had a molecular weight of 27.8 kDa and a pI value of 9.21. It also formed by the conserved domains of VanY and SH3 superfamily proteins, which are involved in its catalytic activity and substrate interaction. Related to its secondary structure, endolysin LysB4 shows 21.4% of α -helix, 28.6% of β -sheet and 50% of internal loops. Furthermore, according to the tridimensional structure prediction, this protein reveals a folding of $\alpha+\beta$ type with a core formed by antiparallel β -sheets surrounded by several α -helix. On the other hand, the binding domain is exclusively formed by antiparallel β -sheets. **Conclusion:** Endolysine LysB4 is a medium MW protein with modular architecture important for its lytic activity.

Key words: LysB4, *Bacillus cereus*, tridimensional modelling.

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

INTRODUCCIÓN

Las bacterias causantes de enfermedades transmitidas por alimentos (ETA) constituyen una amenaza a la Salud Pública en todo el mundo. En los últimos años varias cepas patógenas de estas bacterias se han vuelto más agresivas y resistentes a los antibióticos [1]. Entre ellas se encuentra *Bacillus cereus* que es una bacteria gram-positiva en forma de varilla, anaerobia facultativa, móvil y formadora de esporas que puede causar intoxicación alimentaria mediante la producción de dos tipos de toxinas, una toxina emética y una toxina diarreica [2], cuya ingesta a partir de alimentos contaminados puede causar vómitos, náuseas y diarrea [3].

Debido a la producción de toxinas y su resistencia a los antibióticos, la contaminación con *B. cereus* es una preocupación creciente en la industria alimentaria, y se necesita con urgencia métodos de control distintos al tratamiento antibiótico clásico. En los últimos años, se ha sugerido un enfoque basado en el empleo de endolisinas para controlar este tipo de patógenos transmitidos por los alimentos [4]. Las endolisinas son enzimas codificadas por los bacteriófagos y que se sintetizan al final de su ciclo de replicación. Como principal acción, hidrolizan el peptidoglicano de la pared celular bacteriana, lo cual permite la liberación de la progenie del fago de la célula huésped [6]. Además, existen otras ventajas del empleo de las endolisinas como agentes de control biológico, las cuales incluyen: (i) baja probabilidad de desarrollar resistencia bacteriana; (ii) actividad lítica específica de especie sin afectar a otras bacterias; y (iii) una alta actividad enzimática que permite lisis a las células bacterianas dentro de minutos o incluso segundos [7]. La mayoría de las endolisinas conocidas tienen una arquitectura modular que consiste de un dominio catalítico N-terminal y un dominio de unión C-terminal conectados por un péptido enlazador corto o linker, el dominio de unión sirve para reconocer y unirse específicamente a la superficie celular de la bacteria huésped independiente de la actividad lítica del dominio catalítico [8].

La endolisina LysB4 tiene una capacidad lítica muy eficaz, no solo en su célula huésped natural que es *B. cereus*; sino también, en muchas otras bacterias gram-positivas y gram-negativas, es estable en un amplio rango de valores de temperatura y pH; volviéndose así en un poderoso antimicrobiano [9]. Por esta razón, en el presente trabajo se ha

caracterizado a nivel *in silico* la endolisina LysB4 importante para el biocontrol de *Bacillus cereus* en la industria alimentaria. De esta forma, al conocer la estructura 3D podremos comprender mejor la eficiencia de su actividad catalítica y el amplio rango de huéspedes que presenta su dominio de unión.

MATERIALES Y METODOS

Obtención de la secuencia aminoacídica de la endolisina LysB4

La secuencia primaria de la endolisina fue obtenida a través de la base de datos del GenBank (código de acceso AFF27501.1). Esta secuencia proviene de un estudio donde caracterizan esta endolisina a nivel experimental [9]. Dicha secuencia aminoacídica fue utilizada para los ensayos de predicción de estructura tridimensional.

Determinación de parámetros bioquímicos y dominios conservados

Para determinar los principales parámetros bioquímicos de la endolisina LysB4 se empleó la herramienta ProtParam del ExPASy (<http://web.expasy.org/protparam/>) [10]. Los dominios conservados en la endolisina fueron analizados mediante la herramienta Prosite del ExPASy (<http://prosite.expasy.org/>) [11]. La búsqueda y obtención de secuencias relacionadas fue realizada empleando la herramienta BLAST del NCBI [12]. Las secuencias fueron alineadas utilizando la herramienta Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) [13].

Predicción de estructuras secundarias y alineamiento estructural

El análisis de predicción de estructuras secundarias presentes en la endolisina LysB4 fue realizada empleando la herramienta PHD Secondary Structure Prediction Method (<http://npsa-prabi.ibcp.fr/>) [14]. Además, el alineamiento estructural se realizó en la plataforma Cn3D (versión 4.3.1) utilizando como molde diversas proteínas estructuralmente similares [15] con estructuras tridimensionales reportadas y depositadas en el PDB, filtrando diversos parámetros como un mayor Score, menor E-value y una mayor resolución ($< 2.0 \text{ \AA}$).

Modelamiento tridimensional de la endolisina LysB4

El modelo tridimensional del dominio catalítico de la endolisina se obtuvo empleando la herramienta SWISS-MODEL (<http://swissmodel.expasy.org>) [16] a partir de la información depositada en la base de datos del Protein Data Bank (PDB). Por otro lado, se empleó la herramienta Phyre2

(<http://www.sbg.bio.ic.ac.uk/phyre2/>) [17], así como el servidor Robetta (<http://rosetta.bakerlab.org/>) [18], 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).

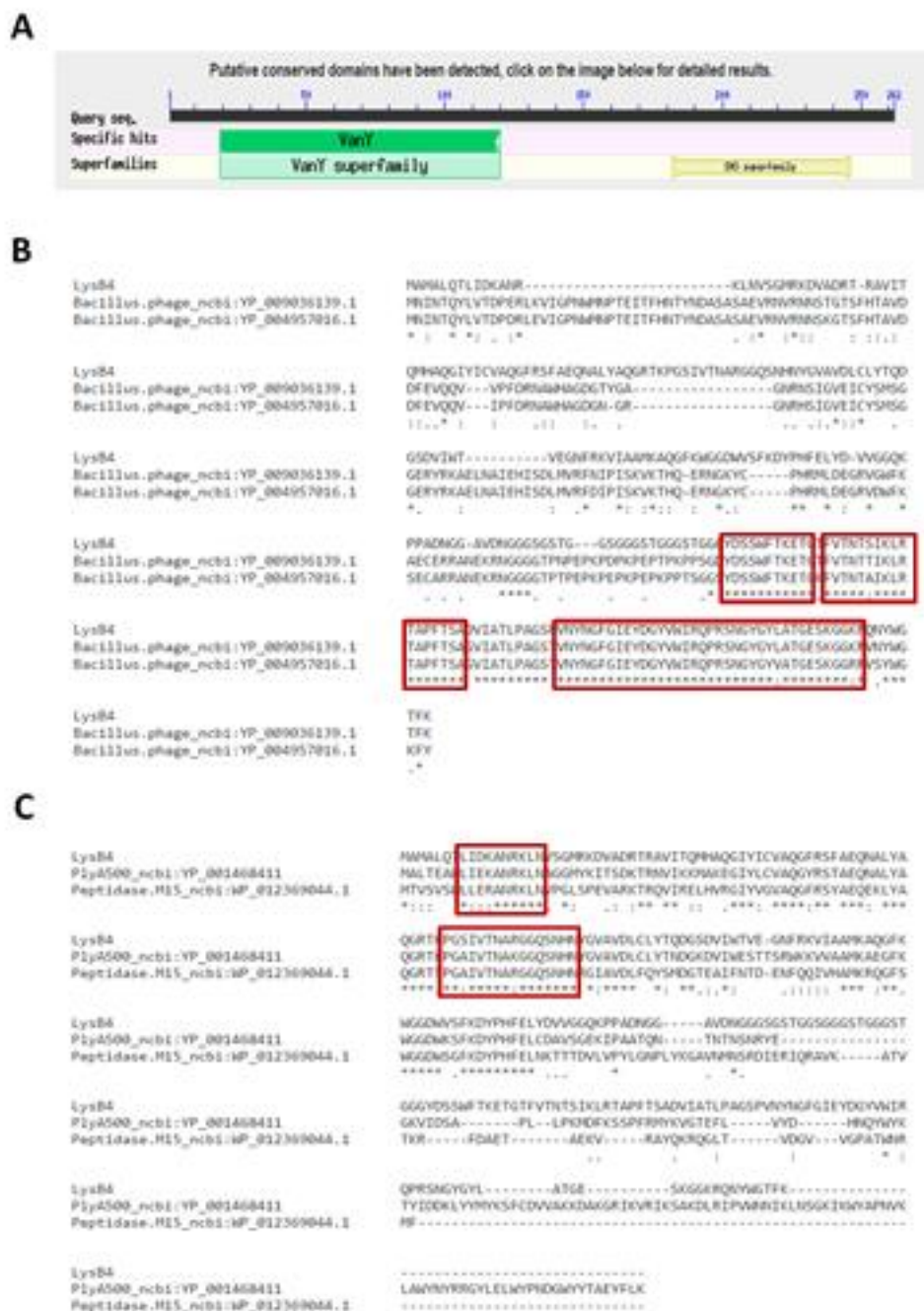


Fig. 1. (A) Dominios obtenidos con la herramienta BLAST. (B) Alineamiento múltiple mostrando las tres regiones conservadas correspondientes al dominio de unión. (C) Alineamiento múltiple mostrando las dos regiones conservadas correspondientes al dominio catalítico. Se resaltan los aminoácidos conservados (*) a lo largo de la secuencia.

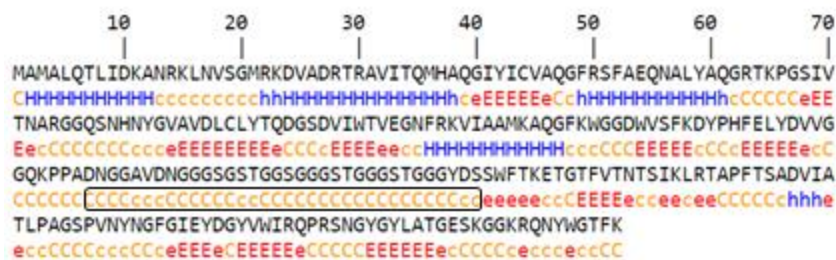


Fig. 2. Predicción de estructuras secundarias de la endolisina LysB4. Se observan las regiones correspondientes a las hélices α (Hh), láminas β (Ee) y loops internos (Cc) distribuidos a lo largo de la secuencia de aminoácidos de la proteína. La porción encerrada pertenece a la región linker.

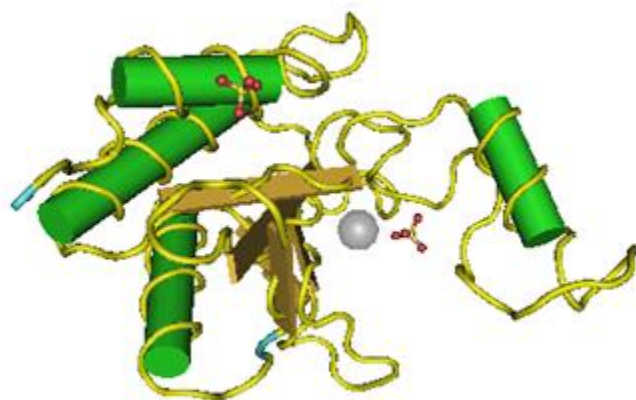


Fig. 3. Alineamiento estructural de la endolisina LysB4 con el dominio catalítico de la endolisina Ply500. La estructura molde presenta 4 hélices α , 5 láminas β y un catión divalente (Zn^{+2}) ubicado en los aminoácidos H80, D87 y H133. Las líneas de color amarillo corresponden a las regiones conservadas predichas por el Cn3D. Estas dos regiones conservadas comprenden casi la totalidad de la proteína.

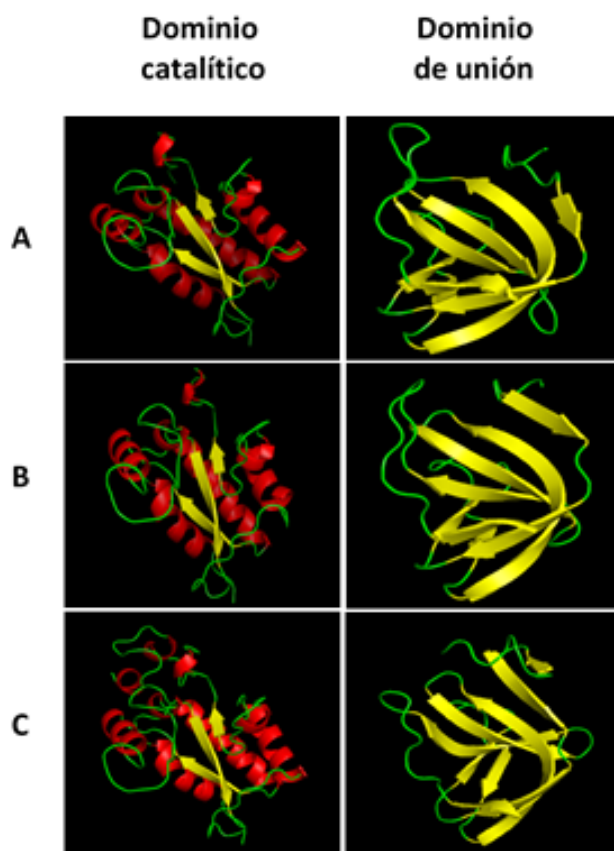


Fig. 4. Modelamiento tridimensional del dominio catalítico y el dominio de unión de la endolisina LysB4 empleando modelamiento por homología (A), por *threading* (B) y *ab initio* (C). Las hélices α se muestran de color rojo mientras que las láminas β se muestran de color amarillo.

RESULTADOS

La endolisina LysB4 presenta 262 aminoácidos con un peso molecular de 27.8 kDa y un pI de 9.21. Mediante búsqueda en BLAST, se encontraron dos dominios pertenecientes a las superfamilias VanY, correspondiente a la enzima D-alanil-D-alanina carboxipeptidasa, así como a la superfamilia SH3 cuyos dominios permiten la interacción de proteínas. A partir de secuencias con similitudes entre 94% y 88%, se puede realizar el alineamiento múltiple, observándose tres regiones conservadas que corresponden tanto al dominio de unión como al dominio catalítico (Fig. 1).

En relación al análisis de estructuras secundarias, se observa que la endolisina LysB4 presenta 21.4% de hélices α , 28.6% de láminas β y 50% de loops internos en su estructura (Fig. 2).

Con respecto a la predicción de la estructura tridimensional, el molde más idóneo para el alineamiento estructural de la región catalítica correspondió al dominio catalíticamente activo de la endolisina Ply500 del bacteriófago A500, el cual infecta a *Listeria monocytogenes* (PDB ID: 2VO9) (Fig. 3) [19]. En el modelamiento por homología del dominio catalítico se observó un plegamiento del tipo $\alpha+\beta$ con un core compuesto por láminas β antiparalelas rodeada de hélices α , lo cual fue confirmado en los modelamientos por *threading* y *ab initio*. En el modelamiento tanto por homología como por *threading*, el dominio catalítico comprende los primeros 146 aminoácidos. En el caso del modelamiento del dominio de unión, éste también se realizó por homología, *threading* y *ab initio*; mostrando en su estructura tridimensional solo láminas β antiparalelas (Fig. 4). El dominio de unión representa aproximadamente los últimos 90 aminoácidos.

DISCUSIÓN

En el presente artículo se han determinado las características estructurales *in silico* de la endolisina LysB4 y sus dominios, mediante herramientas bioinformática. Esta proteína tiene un alto potencial como un agente antimicrobiano eficaz en el control de patógenos transmitidos por alimentos, además de *Bacillus cereus* también ha demostrado un alto potencial lítico en bacterias como *Listeria monocytogenes*, *Bacillus subtilis* y algunas

gram-negativas, previo tratamiento con EDTA. LysB4 se caracterizó experimentalmente como una endopeptidasa que muestra una temperatura y pH óptimo de 50 °C y 8.5, respectivamente [9].

Las regiones conservadas obtenidas pertenecen a los 2 dominios de la endolisina y en el alineamiento estructural del dominio catalítico se encontró una similitud del 71% con la bien caracterizada región catalítica de la endolisina Ply500. Este dominio presenta láminas β antiparalelas rodeada de hélices α que es característico de algunas enzimas, mientras que el dominio de unión presenta solo láminas β . El dominio catalítico posee 3 aminoácidos que soportan al cofactor Zn^{+2} , el cual es necesario para que este dominio sea activo, dicho aminoácidos son histidina (H80), glutamina (D87) e histidina (H133), los cuales se conservan en las endolisinas Ply500 y Ply118 [19]. La predicción de las estructuras secundarias se conserva marcadamente en los modelos predichos por homología, *threading* y *ab initio* para ambos dominios. La importancia de la estructura tridimensional de esta proteína radica en sus dominios independientes y en su alto rango de bacterias sensibles a ésta. Ambos dominios se pueden cambiar sin limitaciones funcionales graves, lo cual abre nuevas perspectivas para poder lograr diferentes funciones útiles a través de la ingeniería de proteína [20].

CONCLUSIONES

La endolisina LysB4 es una proteína de mediano peso molecular que muestra una arquitectura modular compuesta de dos dominios independientes y conectados por un linker. La conservación de su sitio de unión al Zn^{+2} es importante para su actividad lítica. Se requieren estudios complementarios para la predicción de la interacción entre el dominio de unión de la endolisina y el sitio de unión de la pared celular bacteriana para poder comprender la especificidad que presenta.

AUTOR DE CORRESPONDENCIA

Dr. Gustavo A. Sandoval
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 BIBLIOGRÁFICAS

1. De W Blackburn, Clive, and Peter J. McClure, eds. *Foodborne pathogens: hazards, risk analysis and control*. Elsevier, 2009.
2. Arnesen LPS, Fagerlund A, Granum PE. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev* 2008;32:579-606.
3. Bottone EJ. *Bacillus cereus*, a volatile human pathogen. *Clin Microbiol* 2010. *Rev* 23:382-98.
4. O'Flaherty S, Ross RP, Coffey A. Bacteriophage and their lysins for elimination of infectious bacteria. *FEMS Microbiol Rev*; 2009;33:801-19.
5. Young, R. Y., Wang, N., & Roof, W. D. Phages will out: strategies of host cell lysis. *Trends in microbiology*; 2000;8(3),120-8.
6. Schuch R., Nelson D., Fischetti VA. A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature* 2002;418:884-9.
7. Fischetti, V. A. Bacteriophage lysins as effective antibacterials. *Current opinion in microbiology*; 2008;11(5),393-400.
8. Loessner, M. J., Kramer, K., Ebel, F., & Scherer, S. C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Molecular microbiology*; 2002;44(2):335-49.
9. Son, Bokyoung, et al. Characterization of LysB4, an endolysin from the *Bacillus cereus*-infecting bacteriophage B4. *BMC microbiology* 12.1 (2012):33.
10. 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.
11. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.* 2006 Jul 1;34(Web Server issue):W362-5.
12. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389-402
13. 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.
14. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol.* 1993;232(2):584-99.
15. Wang Y, Geer LY, Chappey C, Kans JA, Bryant SH. Cn3D: sequence and Structure views for Entrez. *Trends Biochem Sci.* 2000;25(6):300-2.
16. 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.
17. 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.
18. 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.
19. Korndörfer IP, Kanitz A, Danzer J, Zimmer M, Loessner MJ, Skerra A. Structural analysis of the L-alanyl-D-glutamate endopeptidase domain of *Listeria* bacteriophage endolysin Ply500 reveals a new member of the LAS peptidase family. *Acta Crystallographica Section D: Biological Crystallography.* 2008;64(6), 644-50.
20. Dong Q, Wang J, Yang H, Wei C, Yu J, Zhang Y, et al. Construction of a chimeric lysin Ply187N-V12C with extended lytic activity against staphylococci and streptococci. *Microbial biotechnology*; 2015;8(2),210-20.

Recibido: 16/09/2016

Aceptado: 25/11/2016