



INFORME FINAL DE INVESTIGACIÓN

“CARACTERIZACIÓN BIOINFORMÁTICA DE LA PROTEASA ClpP DE *Streptococcus mutans* IMPLICADA EN LA FORMACIÓN DE CARIES DENTAL”

LIMA – PERÚ

2014

Resumen

Objetivo: Realizar el análisis bioinformático de la proteasa ClpP de *Streptococcus mutans* relacionada con su tolerancia al estrés químico durante la formación de caries dental humana.

Material y métodos: Para llevar a cabo este análisis se ha empleado la secuencia aminoacídica de la proteasa ClpP de *S. mutans* cepa UA159 obtenida a partir de la base de datos del GenBank (código de acceso Q8DST7.1), la cual fue utilizada para la determinación de sus parámetros bioquímicos, dominios conservados, predicción de estructuras secundarias, modelamiento por homología y la predicción del patrón de interacción entre proteínas empleando las herramientas bioinformáticas ProtParam, Prosite, PHD Secondary Structure Prediction, SWISS-MODEL y STRING respectivamente. La visualización de modelos tridimensionales se realizó empleando PyMol. **Resultados:** La proteasa ClpP de *S. mutans* es una molécula ácida y de mediano peso molecular y además presenta dos dominios altamente conservados relacionados con endopeptidasas bacterianas. A partir de la predicción de su estructura tridimensional, la proteasa ClpP presenta una conformación tetradecamérica similar a un anillo, por el cual pasarían las proteínas bacterianas que necesiten ser procesadas. De la red de interacción de proteínas, se tomó un score superior o igual 0.954, se determinó la relación filogenética entre la proteasa ClpP con clpX, clpE, clpC, groL y dnaK con identidades de homología de 62.5%, 62%, 55%, 63.4%, 79.6% respectivamente. **Conclusión:** La proteasa ClpP de *S. mutans* resulta importante para el proceso infectivo de esta bacteria lo cual permitirá generar una variedad de estrategias para la inhibición de la colonización de la cavidad bucal.

Palabras claves: *Streptococcus mutans*, *Streptococcus pneumoniae*, ClpP, endopeptidasa.

Abstract

Objective: To perform a bioinformatics analysis of the ClpP protease from *Streptococcus mutans*, which is related to its tolerance to chemical stress during formation of human dental caries. **Material and methods:** To carry out this analysis, the amino acid sequence of ClpP protease from *S. mutans* strain UA159 was obtained from GenBank database (accession number Q8DST7.1). This sequence was used to the determination of biochemical parameters, conserved domains, secondary structure prediction and homology modelling, by means of ProtParam, Prosite, PHD Secondary Structure Prediction and SWISS-MODEL bioinformatics tools, respectively. Visualization of tridimensional models was performed under PyMol. **Results:** ClpP protease from *S. mutans* is an acidic molecule with medium molecular weight and also shows two highly conserved domains related to other bacterial endopeptidases. Using predicted tridimensional structure, it was shown that ClpP protease display a tetradecameric conformation like a ring, which will be implied in the catalytic processing of bacterial proteins. **Conclusion:** ClpP protease from *S. mutans* is an important factor related to the infective process of *S. mutans*, which may allow the development of novel strategies to inhibit the colonization inside oral cavity.

Keywords: *Streptococcus mutans*, *Streptococcus pneumoniae*, ClpP, endopeptidase.

Introducción

La caries dental es considerada en la actualidad como una enfermedad ecológica y multifactorial en lugar de una enfermedad estrictamente microbiana [1], la cual está asociada a la pérdida gradual, progresiva y neta de los dientes, que afecta a diferentes grupos de personas, principalmente niños y adultos jóvenes. Además, esta condición se caracteriza principalmente por la destrucción de los tejidos dentarios como resultado de la desmineralización provocada por los ácidos generados por la placa bacteriana [1,2].

La formación de la placa bacteriana se realiza de manera ordenada y la composición de la misma es variada, manteniéndose relativamente estable en el tiempo (homeostasis microbiana). El tejido dentario infectado presenta una especie predominante diferente a la de un tejido sano, sin embargo los patógenos putativos a menudo se encuentran en cantidades bajas en los tejidos sanos. En el tejido dentario infectado, se presencia un cambio proliferativo radical hacia el dominio de la comunidad por las especies acidogénicas tolerantes al ácido como *S. mutans* y algunas especies de lactobacilos [1].

Dentro de los principales agentes etiológicos causales se encuentra *Streptococcus mutans*, bacteria gram-positiva acidogénica y anaeróbica facultativa, cuya supervivencia y proliferación dentro de la cavidad bucal depende de diversos factores de virulencia, como su capacidad para metabolizar azúcares de los alimentos, principalmente sucrosa [3]; así como también su capacidad para formar biopelículas en la superficie de las piezas dentarias. Debido a que puede adherirse a diversas superficies de las mucosas, *S. mutans* también es considerada como causante de la endocarditis bacteriana [4,5].

A fin de contrarrestar los factores adversos que afecten la viabilidad de *S. mutans* en la cavidad bucal, esta bacteria expresa una gran variedad de proteínas entre las cuales destacan las proteasas ClpP [6]. Estas moléculas se encargan de proteger a las células bacterianas del daño producido por el acúmulo de proteínas que hayan sufrido una alteración en sus procesos normales de transcripción y traducción [7]. Debido a esta característica, estos componentes son conocidos como proteínas de choque térmico (del inglés *heat-shock proteins*), las cuales son

cruciales para el correcto plegamiento y estabilidad de proteínas reguladoras. A pesar de su gran significancia en el proceso patogénico de *S. mutans*, no se ha realizado un análisis de su estructura por lo que información adicional sobre este grupo de proteínas sería relevante para comprender el mecanismo de patogenicidad de *S. mutans* y así plantear alternativas para inhibir su proceso infectivo.

Por este motivo, en el presente trabajo se ha realizado un análisis bioinformático de la proteasa ClpP de *S. mutans* a partir de la información depositada en la base de datos del GenBank.

Materiales y métodos

Obtención de la secuencia aminoacídica de la proteasa ClpP

La secuencia primaria correspondiente a la proteasa ClpP de *Streptococcus mutans* cepa UA159 fue obtenida a través de la base de datos del GenBank (código de acceso Q8DST7.1) [8]. Dicha secuencia fue utilizada para el posterior análisis bioinformático y obtención de la información estructural de dicha proteína.

Determinación de parámetros bioquímicos y dominios conservados

Para determinar los principales parámetros bioquímicos de la proteasa ClpP se empleó la herramienta ProtParam del ExPASy (<http://web.expasy.org/protparam/>) [9], mientras que los dominios conservados presentes en dicha proteína fueron analizados mediante la herramienta Prosite del ExPASy (<http://prosite.expasy.org/>) [10].

Predicción de estructuras secundarias y alineamiento estructural

El análisis de predicción de estructuras secundarias presentes en la proteasa ClpP de *S. mutans* fue realizada empleando la herramienta PHD Secondary Structure Prediction Method (<http://npsa-prabi.ibcp.fr>) [11]. Por otro lado, el alineamiento estructural se realizó en la plataforma Cn3D (versión 4.3.1) utilizando como molde a la proteasa ClpP de *S. pneumoniae* (PDB ID: 1Y7O.1) [12].

Modelamiento por homología de la proteasa ClpP

Por otro lado, su modelo tridimensional se obtuvo empleando la herramienta SWISS-MODEL (<http://swissmodel.expasy.org>) [13], a partir de la estructura tridimensional de la proteasa ClpP de *S. pneumoniae* depositada en la base de datos del Protein Data Bank (PDB ID: 1Y7O.1) [13]. La visualización del modelo obtenido se realizó empleando el programa PyMol Molecular Graphics System (versión 1.3).

Predicción de la interacción proteica

La red de interacción de proteínas se obtuvo usando el programa STRING [14]

Resultados

Del análisis de la secuencia primaria de la proteasa ClpP de *S. mutans* se obtuvieron un total de 196 aminoácidos correspondientes a una proteína de 21.6 kDa con un punto isoeléctrico de 5.47 y una proporción de 25 aminoácidos ácidos versus 22 aminoácidos básicos. Esto define a ClpP como una proteína ácida de bajo molecular. Por otro lado, se observó la presencia de dos regiones conservadas correspondientes al sitio activo de endopeptidasas bacterianas 88-TIVMGIAASMGY-99 y 110-RFMLPNAEYLHQP-123, donde se destaca la presencia de los aminoácidos Serina (S) e Histidina (H), respectivamente (Fig. 1).

Con respecto a su estructura secundaria, esta proteína es rica en hélices α (53.57%) y presenta una proporción menor de láminas β (16.33%). Del alineamiento estructural se obtuvo una homología del 85% a nivel de su secuencia de aminoácidos con la proteasa ClpP de *S. pneumoniae* (Fig. 2), mientras que el modelamiento por homología mostró que dicha proteína estructura tetradecamérica, compuesta de dos anillos heptaméricos (Fig. 3).

De la red de patrones de interacción proteica, con fines prácticos se tomó un score superior o igual 0.954 solo de las proteínas que tiene una estructura resuelta tridimensionalmente y se determinó la relación filogenética entre la proteasa ClpP con clpX, clpE, clpC, groL y dnaK con identidades de homología de 62.5%, 62%, 55%, 63.4%, 79.6% respectivamente; así como también la co-localización de ClpX y groL en tanto no hubo co-expresión de gens Dpr de *S. mutans* (Fig. 4).

ClpP *S. mutans*

1-MIPVVEQTSRGERSDIYSRLLKDRIIIMLTGPVEDNMANSIIAQLLFLD-50

*

Región 1

51-AQDNTKDIYLYINSPGGSVSAGLAIVDTMFIKSDVQ**TI VMGLAASMGT**I-100

*

*

*

*

*

*

Región 2

101-IASSGAKGKR**FMLPNAEYLIHQ**PMGGTGGGTQQSDMAIAAEQLLKT**RKKL**-150

*

151-EKILSDNSGKTIKQIHKDAERDYWMDAKETLKYGFIDEIMENNELK-196

Fig. 1. Estructura primaria del monómero de la proteasa ClpP de *S. mutans*. Se resaltan las regiones 1 y 2 altamente conservadas propias de endopeptidasas bacterianas. Además se muestra la ubicación de los aminoácidos no conservados (*).

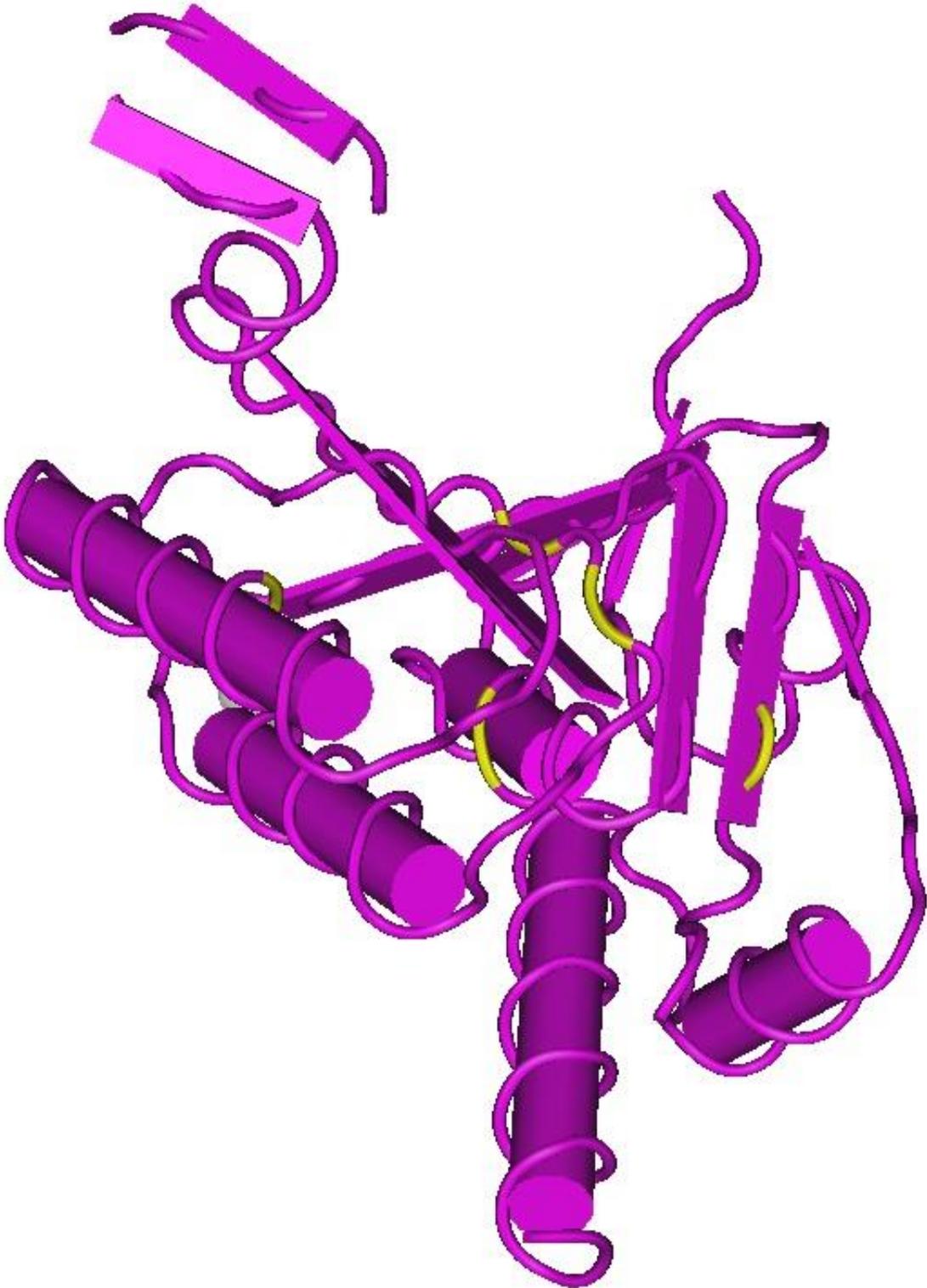


Fig. 2. Alineamiento estructural de las proteasas ClpP de *S. mutans* y *S. pneumoniae*. Se observan las porciones que corresponde a aminoácidos no conservados (amarillo).

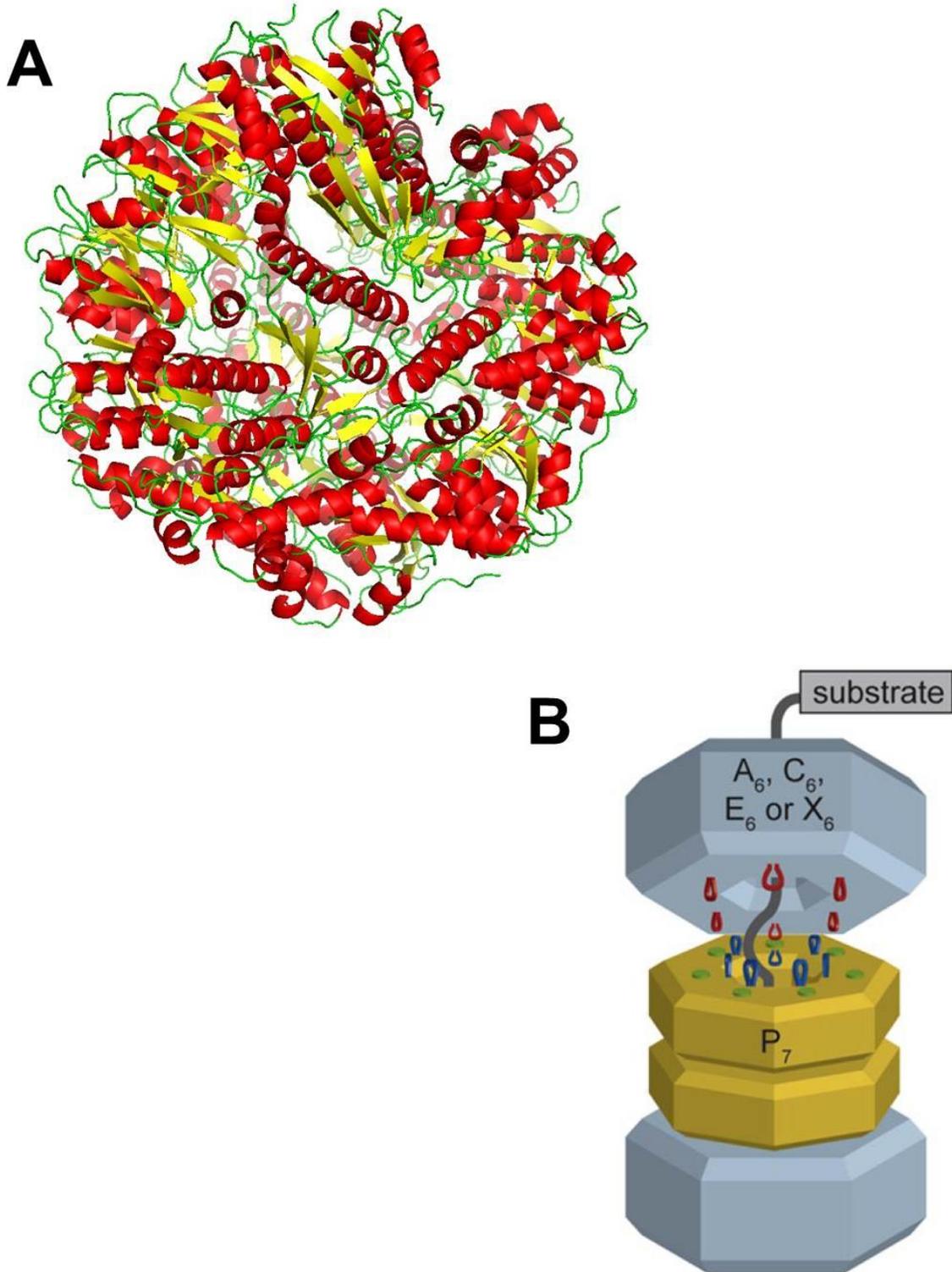


Fig. 3. (A) Modelo tridimensional de la estructura tetradecamérica de la proteasa ClpP de *S. mutans*. (B) Arquitectura de las proteasas-chaperonas Clp. Se destaca la ubicación de la proteasa ClpP (naranja) como parte fundamental del anillo [7].

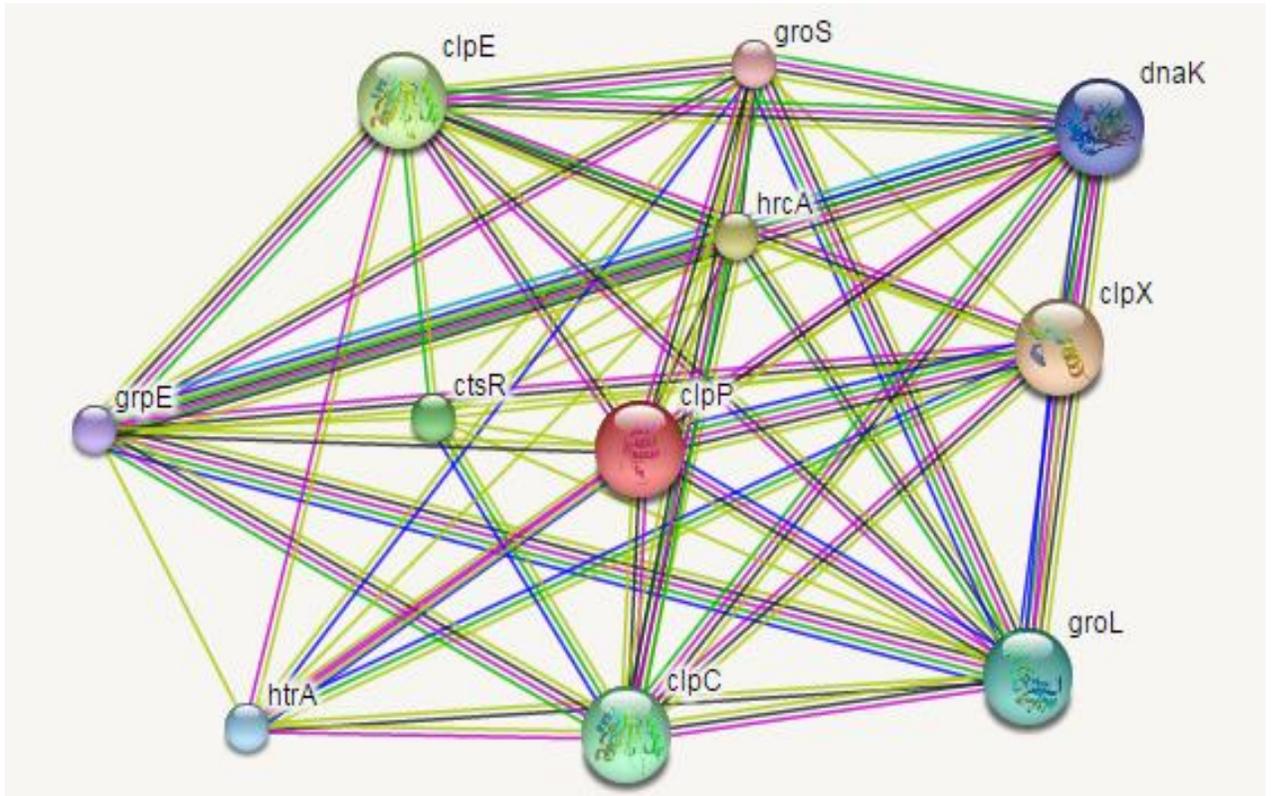


Fig. 4. Red de patrón de interacción de la proteína Dpr de *S. mutans* con las proteínas clpX, clpE, clpC, groL y dnaK con identidades de homología de 62.5%, 62%, 55%, 63.4%, 79.6% respectivamente.

Discusión

En el presente artículo se ha realizado el análisis bioinformático de la proteasa ClpP de *Streptococcus mutans*. Las proteasas del tipo ClpP corresponden a un grupo de complejos enzimáticos dependientes de ATP altamente difundidos en bacterias [7], las cuales están compuestas de múltiples subunidades capaces de formar ensamblajes proteicos lo suficientemente grandes como para procesar sustratos proteicos [12,15]. De los resultados obtenidos se ha encontrado que la proteasa ClpP de *S. mutans* es una proteína ácida de mediano peso molecular y que logra formar estructuras homotetradecaméricas. Así mismo, estas subunidades están dispuestas en dos anillos heptaméricos que forman un poro central a través de la cual ingresan las proteínas para ser procesadas. En esta región resulta importante la presencia del sitio catalítico compuesto por los aminoácidos Serina (S) e Histidina (H).

Con respecto a la estructura tridimensional de las proteasas ClpP, hasta la fecha existen diversos reportes que mencionan una alta similitud de estas proteínas con los sistemas de chaperonas Hsp70 y DnaK presentes en *Escherichia coli*, así como con el complejo Hsp110-Hsc70 presente en *Saccharomyces cerevisiae* [15], lo cual demuestra la importancia de esta proteasa en el mantenimiento de la integridad y viabilidad de *Streptococcus mutans*. Dentro de las posibles estrategias para disminuir el proceso infectivo de *S. mutans* estaría el empleo de inhibidores competitivos y no competitivos de endopeptidasas del tipo serinoproteasas, los cuales pueden encontrar en distintas fuentes naturales como vegetales, los cuales son producidos como parte de una respuesta de defensa natural y que podrían utilizarse en el campo clínico para el tratamiento de la caries dental [16].

La predicción de patrones de redes de interacción permite visualizar y analizar de manera precisa la relación filogenética, co-expresión de genes y proteínas relacionadas en algunas u otras especies. En nuestro resultado cuando se observa el Gene Neighborhood, los genes ortólogos en diferentes organismos se ubican repetidamente y se localizan cercanamente entre ellas con una homología de más del 50% entre la proteína ClpC-ClpX, es decir estas proteínas se encuentran altamente conservadas. No se observó una co-expresión de los genes de las proteínas ClpC, pero sí de proteínas ortólogas como ClpC y ClpX en *Escherichia coli* K12

relacionadas con ClpC de *S. mutans*. Hubo presencia de co-localización de las proteínas ClpC y ClpX altamente conservadas en la gran mayoría de organismos del árbol filogenético. Se observó una gran cantidad de reportes experimentales de las proteínas clpX, clpE, clpC, groL y dnaK con un alto grado de score. Además existe una semejanza funcional entre estas cuatro proteínas, ya que todas por diferentes mecanismos actúan en el procesamiento proteolítico de las proteínas con plegamiento incorrecto [17].

Conclusiones

La proteasa ClpP de *Streptococcus mutans* es una proteína ácida de mediano peso molecular implicada en la degradación dependiente de ATP de proteínas a nivel intracelular lo cual permitiría la sobrevivencia de la bacteria frente a diversos tipos de estrés en la cavidad bucal.

Literatura citada

1. Marsh PD. Dental plaque as a biofilm and a microbial community - implications for health and disease. *BMC Oral Health*. 2006;6 Suppl 1:S14.
2. Ding Y, Wang W, Fan M, Tong Z, Kuang R, Jiang W, et al. Antimicrobial and anti-biofilm effect of Bac8c on major bacteria associated with dental caries and *Streptococcus mutans* biofilms. *Peptides*. 2014;52:61-7.
3. Kuramitsu HK, Wang BY. Virulence properties of cariogenic bacteria. *BMC Oral Health*. 2006 Jun 15;6 Suppl 1:S11.
4. Gould K, Ramirez-Ronda CH, Holmes RK, Sanford JP. Adherence of bacteria to heart valves in vitro. *J Clin Invest*. 1975 Dec;56(6):1364-70.
5. Nomura R, Nakano K, Nemoto H, Fujita K, Inagaki S, Takahashi T, et al. Isolation and characterization of *Streptococcus mutans* in heart valve and dental plaque specimens from a patient with infective endocarditis. *J Med Microbiol*. 2006 Aug;55(Pt 8):1135-40.
6. Kajfasz JK, Martinez AR, Rivera-Ramos I, Abranches J, Koo H, Quivey RG Jr, Lemos JA. Role of Clp proteins in expression of virulence properties of *Streptococcus mutans*. *J Bacteriol*. 2009 Apr;191(7):2060-8.
7. Kress W, Maglica Z, Weber-Ban E. Clp chaperone-proteases: structure and function. *Res Microbiol*. 2009 Nov;160(9):618-28.
8. Ajdić D, McShan WM, McLaughlin RE, Savić G, Chang J, Carson MB, et al. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc Natl Acad Sci U S A*. 2002 Oct 29;99(22):14434-9.
9. 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.
10. 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.
11. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol*. 1993 Jul 20;232(2):584-99.

12. Gribun A, Kimber MS, Ching R, Sprangers R, Fiebig KM, Houry WA. The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *J Biol Chem*. 2005 Apr 22;280(16):16185-96.
13. 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.
14. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, Simonovic, N., Roth, A., Santos, A., Tsafou, KP., Kuhn, M. Bork, P., Jensen, LJ., von Mering, C. (2014). STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, gku1003.
15. Doyle SM, Genest O, Wickner S. Protein rescue from aggregates by powerful molecular chaperone machines. *Nat Rev Mol Cell Biol*. 2013 Oct;14(10):617-29.
16. Jamal F, Pandey PK, Singh D, Khan MY. Serine protease inhibitors in plants: nature’s arsenal crafted for insect predators. *Phytochemistry Reviews*. 2013 Mar;12(1):1-34.
17. Tong, Z., Huang, L., Ling, J., Mao, X., Ning, Y., & Deng, D. (2014). Effects of intracanal irrigant MTAD Combined with nisin at sub-minimum inhibitory concentration levels on *Enterococcus faecalis* growth and the expression of pathogenic genes. *PloS one*, 9(3), e90235.