



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO



FACULTAD DE ODONTOLOGÍA

EXPRESIÓN RECOMBINANTE Y PURIFICACIÓN DE LA
PROTEÍNA CEMENTOGENINA (CMGN).

T E S I S

QUE PARA OBTENER EL TÍTULO DE

C I R U J A N A D E N T I S T A

P R E S E N T A:

ANDREA OLMOS BONILLA

TUTOR: Dr. ENRIQUE ROMO ARÉVALO

ASESOR: Dr. HIGINIO ARZATE



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Al Dr. Enrique Romo Arévalo, por creer en mi e inspirarme con su ejemplo, por su paciencia y por darme la oportunidad de crecer tanto profesional, como personalmente. Mi mayor reconocimiento y admiración a usted, le estaré por siempre agradecida.

Al Dr. Higinio Arzate, por abrirme las puertas de su laboratorio y permitirme realizar mi trabajo de titulación, por la paciencia y los consejos.

A la Dra. Lia Alioth Hoz Rodríguez, por ser una mujer admirable e inspirarme con su ejemplo, gracias por estar presente durante toda mi formación académica y haberme mostrado el mundo de la investigación. Que dicha encontrar mujeres empoderadas y tan inteligentes como tú.

Al Dr. Mikado A. Nidome Campos por su paciencia, la asesoría y consejos.

A mis revisores de tesis: Dra. Miryam Martínez Hernández, Dra. Claudia Patricia Pedraza Zamora y al Dr. Gonzalo Montoya Ayala, por su valioso tiempo, sus anotaciones y consejos.

Al Laboratorio de Biología Periodontal y Tejidos Mineralizados, por darme la oportunidad y proporcionarme las herramientas necesarias para realizar este trabajo.

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) IN203920.

A la Universidad Nacional Autónoma de México, la institución a la que debo mi formación académica.

DEDICATORIAS

A mi madre, por tu apoyo incondicional y amor con el que siempre me has criado. Estoy eternamente agradecida por tus consejos, la confianza y por animarte a ser mi primer paciente en la facultad. Te amo mamá, eres un ejemplo digno de seguir.

A mi padre, por tu paciencia y amor con el que siempre me has criado. Gracias por tu apoyo incondicional, por tu buen humor y por confiar en mí. Valoro todos tus consejos y enseñanzas, te amo y admiro.

A mi hermano, por ser mi compañero y cómplice de vida. Te amo y admiro.

A toda mi familia por su amor y consejos, pero principalmente por inspirarme con su ejemplo y pasión por sus trabajos.

A mi mejor amiga y hermana del alma, Rebeca. Gracias por compartir tu tiempo y vida conmigo, no imagino la universidad sin nuestra complicidad. Te amo y admiro.

Gracias Dios por la vida, la familia y la salud que me has dado.

ÍNDICE

Lista de abreviaturas	1
Resumen	2
Abstract	3
Introducción	4
Antecedentes	7
Planteamiento del problema	16
Justificación	16
Hipótesis	17
Objetivos	17
Metodología	18
Resultados	22
Discusión	38
Conclusiones	41
Referencias	42

Lista de abreviaturas

6x-His	Bandera de seis histidinas
ADN	Ácido desoxirribonucleico
BMP-2	Proteína morfogénica ósea 2
HACD1/CAP	Proteína de adhesión del cemento radicular
CEMP1	Proteína del cemento 1
Da	Dalton
DAB	Diaminobencidina
DTT	Ditiotreitol
FPLC	Cromatografía líquida de proteínas de alta velocidad
hrCMGN	Cementogenina recombinante humana
IPTG	Isopropil- β -D-1-tiogalactopiranosido
kDa	kilo Dalton
LB	Luria-Bertani
M	Molar
μ l	Microlitro
μ g	Microgramo
mg	Miligramo
mM	Milimolar
nm	Nanómetro
PBS	Amortiguador de fosfatos
PVDF	Fluoruro de polivinilideno
rpm	Revoluciones por minuto
SDS-PAGE	Electroforesis en gel de poliacrilamida dodecilsulfato de sodio
TGF- β	Factor de crecimiento transformante beta

Resumen

El conocimiento e identificación de proteínas cemento específicas ha sido clave para la generación de conocimiento sobre la biología y estructura del cemento radicular. Adicionalmente, la clonación y expresión recombinante de las proteínas cemento específicas ha permitido desarrollar terapias experimentales encaminadas a la regeneración de tejidos mineralizados como son el hueso y cemento radicular. El objetivo del presente estudio fue realizar la expresión recombinante, la purificación e identificación de la proteína Cementogenina (CMGN) como un posible nuevo marcador del cemento radicular humano. Nuestra metodología consistió en la expresión de CMGN mediante el sistema recombinante *E. coli Rosetta* transformado con el plásmido pET-28a (+) /CMGN que contiene un promotor inducible con isopropil- β -D-1-tiogalactopiranosido (IPTG), para su selección cuenta con un gen de resistencia a kanamicina y para su purificación e inmunodetección cuenta con una bandera de seis histidinas (6x-His). El cultivo de las células transformadas se realizó en medio LB (Luria-Bertani) adicionado con 100 μ g/ml de kanamicina y la inducción con IPTG a una concentración final de 0.5 mM. La purificación se realizó con tres pasos de cromatografía líquida de alta velocidad (FPLC): para la fase de captura (primer paso) se empleó una cromatografía de afinidad a iones metálicos con una columna de Níquel (HisPrep FF/16), la purificación intermedia (segundo paso) se realizó con una cromatografía de afinidad a hidroxapatita (HA ultrogel) y para el pulido final (tercer paso) se realizó una ultrafiltración/diálisis y cromatografía de afinidad a hidroxapatita. La identificación de la proteína se hizo mediante inmunodetección por Western blot con el anticuerpo α -His. Nuestros resultados indican que con el protocolo de purificación propuesto se obtiene un rendimiento final 1.7mg de hrCMGN por cada litro de cultivo. El presente trabajo establece la primera etapa en la caracterización de la proteína CMGN y abre la posibilidad de investigar su expresión nativa en el cemento radicular.

Abstract

The identification of specific cementum proteins has been the key for the generation of knowledge about biology and structure of root cementum. Additionally, the recombinant expression of specific cementum proteins has allowed the development of experimental therapies aimed to the regeneration of mineralized tissues such as bone and root cementum. The objective of the present study was to carry out the recombinant expression, purification and identification of the cementogenin protein (CMGN) as a possible new marker of human root cementum. Our methodology consisted in the expression of CMGN using the recombinant *E. coli Rosetta* system transformed with the pET-28a(+)/CMGN plasmid that contains an inducible promoter with isopropyl- β -D-1-thiogalactopyranoside (IPTG), for its selection. It has a kanamycin resistance gene and for its purification and immunodetection it contains a six-histidine flag (6x-His). The culture of the transformed cells was carried out in LB medium (Luria-Bertani) added with 100 μ g/ml of kanamycin and induction with IPTG at a final concentration of 0.5 mM. The purification was carried out with three steps of fast protein liquid chromatography (FPLC): for the capture phase (first step) a metal ion affinity chromatography with a Nickel column (HisPrep FF/16) was used, the intermediate purification (second step) was performed with hydroxyapatite affinity chromatography (HA ultrogel) and for final polishing (third step) ultrafiltration/dialysis and hydroxyapatite affinity chromatography were performed. The identification of the protein was made by Western blot immunodetection with an anti-his antibody. Our results indicate that with this purification protocol, a final yield of 1.7mg of hrCMGN is obtained for each liter of culture. Conclusion: this work establishes the first stage in the characterization of the CMGN protein and opens the possibility of investigating its native expression in root cementum.

INTRODUCCIÓN

Una proteína recombinante es aquella que es producida en un organismo que ha sido sujeto a modificación genética y que, por lo general, es diferente a aquel en que la proteína se produce de forma nativa. La secuencia codificante de la proteína de interés se inserta a una secuencia de ADN heterólogo (plásmido), con el propósito de producir la proteína de interés en un modelo de expresión seleccionado. El templado de ADN que contiene la secuencia de interés y que deseamos aislar, puede ser obtenido mediante procesos enzimáticos para después ser ligado al vector apropiado y posteriormente insertarlo en el modelo de clonación. Para la clonación en bacterias, uno de los vectores más utilizados son los plásmidos (elementos genéticos extracromosomales que contienen uno o más genes de resistencia antibiótica), estos vectores son capaces de replicarse junto con el ADN de la bacteria. El plásmido debe contener al menos un promotor y una región de iniciación para dirigir la expresión de la secuencia codificante, un marcador seleccionable, y elementos replicables. Adicionalmente, el vector debe contener otros elementos genéticos para facilitar la detección, purificación, o solubilización de la proteína, como pueden ser secuencias codificantes para banderas de aminoácidos.¹⁻³

Uno de los desafíos al trabajar con proteínas es trabajar con un modelo fácilmente escalable y una relación costo-beneficio favorable, la mayoría de los esfuerzos en este contexto están dirigidos a desarrollar y optimizar los sistemas de expresión recombinante. La expresión recombinante de proteínas de humano mediante el uso de plásmidos y bacterias es una de las posibilidades que ofrece la biotecnología, y que posibilita obtener proteínas humanas con fines terapéuticos o para estudiar a fondo procesos biológicos y profundizar en el conocimiento de estructuras de proteínas de reciente caracterización. Esta técnica es de gran valor porque las bacterias tienen tasas de división rápidas y su costo de mantenimiento es relativamente bajo, de esta forma se pueden obtener en poco tiempo muchas copias del gen humano inserto en el ADN bacteriano, y producir grandes cantidades de proteínas recombinantes.⁴

Los avances en la tecnología recombinante han permitido el estudio y caracterización de las proteínas cemento específicas: la proteína del cemento 1 CEMP1⁵ y la proteína de adhesión del cemento radicular HACD1/CAP⁶; se ha demostrado que ambas proteínas promueven el proceso de regeneración ósea y derivado de su caracterización bioquímica y estructural, se han diseñado una variedad de péptidos osteogénicos con la capacidad de ejercer un efecto similar a la proteína nativa o recombinante.⁷

El cemento radicular está compuesto por una matriz inorgánica (mineral) formada por cristales de hidroxiapatita y una matriz orgánica compuesta principalmente por colágena tipo I (90%), colágena tipo III (5%) y en menor proporción colágena tipo XII; proteínas no colágenas como sialoproteína ósea, osteopontina, osteocalcina y algunos factores de crecimiento como el Factor de Crecimiento Transformador - Beta (TGF β) y la Proteína Morfogénica Ósea (BMP-2).⁸ Además de su rol indispensable como tejido de anclaje al hueso alveolar circundante, el cemento radicular tiene funciones adaptativas y de reparación. Aunque el cemento es muy similar al tejido óseo, existen diferencias en cuanto a vascularización, componentes celulares y tasa recambio. Mientras que la mayoría de sus proteínas extracelulares son constituyentes típicos de la matriz de los tejidos mineralizados a base de colágena, otras son específicas del cemento CEMP1 y HACD1/CAP.⁹ La proteína HACD1/CAP promueve la migración celular para las células del ligamento periodontal y del hueso alveolar, promueve la adhesión celular mediada por las integrinas $\alpha\beta 1$,¹⁰ incrementa la actividad específica de la fosfatasa alcalina y promueve la condrogénesis y mineralización en células mesenquimatosas (induciendo la diferenciación celular).¹¹ Por otra parte, la proteína CEMP1 participa en el depósito y formación de cristales de hidroxiapatita, por lo que tiene un rol fundamental en la diferenciación de cementoblastos y en la mineralización de la matriz extracelular del cemento.^{12, 13} En cultivos tridimensionales ha demostrado estimular la proliferación de células del ligamento periodontal e inducir la formación de estructuras semejantes a tejidos mineralizados como hueso, cemento radicular y material similar al cartílago.¹⁴ Con un enfoque terapéutico, CEMP1 ha demostrado ser capaz de regenerar defectos óseos de tamaño crítico en ratas Wistar. Al colocar a CEMP1 en una matriz de gelatina y colocarla en defectos de 9mm en calvarias de rata a las 16 semanas se observó mediante cortes

histológicos e imágenes de microtomografía, que CEMP1 logró regenerar 97% del defecto óseo.⁷

Es por esto, que la investigación e identificación de nuevos marcadores del cemento radicular ayudará a mejorar la comprensión de la biología de dicho tejido y eventualmente generar terapias encaminadas al desarrollo de terapias exitosas en la regeneración de los tejidos periodontales dañados, así como para otros tejidos mineralizados.¹⁵

ANTECEDENTES

Periodonto

El periodonto es una estructura dinámica sustentada por un arreglo vascular único, un sistema linfático y elementos nerviosos altamente especializados. Puede definirse de manera simple como “los tejidos que revisten y brindan soporte a los dientes”. Y consta de los siguientes tejidos: hueso alveolar, cemento radicular, ligamento periodontal y encía. Todos estos tejidos se desarrollan y funcionan como una unidad.¹⁶ La mayoría de los tejidos periodontales se desarrollan a lo largo de la formación radicular del diente y su posterior erupción, y tienen su origen en el folículo dental que se deriva de la cresta neural.¹⁷ Una vez formado el tubo neural por invaginación del ectodermo suprayacente, las células neuroepiteliales pluripotentes, migran desde la región media dorsal del tubo neural para invadir los arcos branquiales en desarrollo. Al salir del tubo neural, las células de la cresta neural pierden su naturaleza epitelial y asumen un fenotipo mesenquimatoso capaz de una migración celular dirigida. Las células ectomesenquimales de la cresta neural interactúan de forma recíproca e inductiva con el epitelio oral temprano para formar primordios dentales. Las células derivadas de la cresta neural dan lugar a los condrocitos, osteoblastos, fibroblastos del ligamento periodontal, cementoblastos y odontoblastos. La diferenciación final del fenotipo está regulada por la interacción de las células ectomesenquimales con los factores extrínsecos, como son los factores de crecimiento. Así, podemos determinar que, durante el desarrollo dental, la papila dental da origen a odontoblastos y a la pulpa dental, mientras que el folículo dental da origen al cemento, el ligamento periodontal y el hueso alveolar.¹⁷

Debido a su particular composición, el periodonto es un conjunto de tejidos interesantes de estudiar, así como su reparación y capacidad regenerativa. La estructura y función del periodonto se ven severamente comprometidas cuando sus componentes se ven afectados principalmente por enfermedades inflamatorias crónicas.¹⁸

La enfermedad periodontal es una enfermedad importante a nivel mundial, en la que los tejidos que conforman el periodonto resultan afectados, ocasionando inflamación crónica y pérdida del hueso alveolar. Una vez que el daño ocurre, los tejidos duros y blandos del periodonto no pueden ser restaurados a su forma original. Actualmente se aplican protocolos clínicos como injertos óseos, regeneración tisular guiada y factores de crecimiento, sin embargo, sus indicaciones y pronóstico son limitados.¹⁹ En este sentido, se ha determinado que el tejido clave para lograr la regeneración de los tejidos que conforman el periodonto es el cemento radicular, un tejido del que hasta ahora solo se han identificado dos proteínas cemento específicas: la proteína de adhesión del cemento radicular (HACD1/CAP) y la proteína del cemento 1 (CEMP1), proteínas que promueven la adhesión, migración y diferenciación celular, incrementan la actividad condrogénica, mineralizante, participan de manera activa y llevan a cabo un papel fundamental en la homeostasis del cemento radicular.²⁰

Cemento radicular

El cemento radicular es un tejido mineralizado que se encuentra sobre la dentina radicular, desde la unión cemento esmalte hasta el ápice del diente, en cuya formación participan los cementoblastos. Este tejido provee anclaje entre la raíz dental y el hueso alveolar, mediante el ligamento periodontal.²¹ Los dientes humanos están compuestos por 3 variedades principales de cemento:

1. Cemento acelular afibrilar, se encuentra en menor proporción y recubre principalmente la unión cemento esmalte.
2. Cemento acelular con fibras extrínsecas, se encuentra principalmente en el tercio cervical.
3. Cemento celular con fibras intrínsecas, tiene la capacidad de producir estratos de cemento acelular con fibras extrínsecas, formando el cemento celular mixto estratificado y se encuentra principalmente en las porciones apicales y en las furcas.

Este tejido está compuesto por fibras de Sharpey calcificadas, colágena, glicosaminoglicanos, proteoglicanos y una matriz inorgánica de hidroxiapatita, toda esta composición le otorga su más conocida función: brindar anclaje. Sin

embargo, en 1942, Gottlieb²² postuló que la deposición continua de capas de cemento, parecían ser de gran importancia al ser una barrera contra la disminución del nivel epitelial, lo que fue confirmado por sus observaciones en pacientes con hipofosfatemia, cuyos dientes estaban desprovistos de cemento. Estos pacientes adquirieron enfermedad periodontal severa, sugiriendo que las anomalías en la deposición del cemento pueden dar como resultado un periodonto susceptible a las invasiones microbianas y, por ende, una destrucción de los tejidos.²³

Proteínas cemento específicas

Proteína del cemento 1 (CEMP1)

Esta proteína se aisló por primera vez de células derivadas de un cementoblastoma humano, es expresada por cementoblastos, subpoblaciones celulares del Ligamento Periodontal (LP) y células troncales de origen mesenquimal adyacentes al LP.²⁰ El gen que codifica para esta proteína está localizado en el cromosoma 16 en el brazo corto región 13.3. La secuencia primaria se dedujo de una cadena de ADN complementario (de origen humano) que mostró un 98% de homología con una secuencia predicha encontrada en el *Pan troglodytes* (Chimpancé). Es una proteína conformada por 247 aminoácidos, de peso molecular teórico de 25.9 kDa y punto isoeléctrico de 9.73. Es una proteína altamente desordenada, con un 62.5% de random coil, 28.6% de hélice alfa y 9.9% de hojas beta, además posee sitios de glicosilación en los aminoácidos 20 y 25 que inducen la unión de iones de calcio (Ca^{2+}) en su superficie, y de esta manera participa en la regulación de los cristales de hidroxiapatita. Adicionalmente, se ha demostrado que se encuentra implicada en la regulación de la osificación endocondral, remodelación y regeneración ósea. CEMP-1 tiene alta afinidad a la hidroxiapatita, incluso sin las modificaciones post traduccionales, induciendo la formación de cristales. Todas estas propiedades indican que CEMP-1 juega un rol importante durante el proceso de biomineralización y es necesaria para la síntesis de cristales.¹⁵

En 2007, Carmona puso a prueba CEMP-1, induciendo células no mineralizantes (fibroblastos) mediante la sobreexpresión de dicha proteína. Estas células

mostraron aumento en la proliferación, formación de nódulos mineralizados, aumento en la actividad de fosfatasa alcalina y expresión de *novo* de osteocalcina, osteopontina y sialoproteína; todas las anteriores, moléculas asociadas con la formación de hueso y cemento. Este estudio demostró de manera contundente la habilidad que tiene CEMP1 de cambiar el fenotipo de una célula no mineralizante a un fenotipo mineralizante.²⁵ Otro estudio en un modelo canino con necrosis pulpar, demostró la capacidad de CEMP1 para reclutar células troncales de origen mesenquimal provenientes del LP, lo que promovió la proliferación y mineralización de estas células.^{26, 27} Un estudio demostró la función de CEMP1 en la formación de cristales de fosfato de calcio utilizando el sistema libre de células por contradifusión en gel de silicato. Donde se observó la formación de esférulas desde el día 7, manteniéndose hasta el último periodo de tiempo observado (60 días), donde la relación Ca/P de los cristales inducidos por CEMP1 indica que la fase mineral que se formó fue fosfato octacálcico, el cual es el estado de transición hacia hidroxapatita.²⁸

Todos estos datos indican que CEMP1 es una proteína única que posee múltiples propiedades como la mineralización, proliferación, diferenciación y es modulador celular, además tiene potencial terapéutico para el tratamiento de defectos óseos y la regeneración de otros tejidos mineralizados.

Proteína de adhesión del cemento radicular (HACD1/CAP)

La Hidroxiacil coenzima A dehidratasa 1 (HACD1) / Proteína de Adhesión del Cemento (CAP), está constituida por 140 aminoácidos, tiene un peso molecular teórico de 14.9 kDa y un punto isoeléctrico de 7.82. El gen que codifica para la proteína HACD1/CAP se localiza en el cromosoma 10 región p13-14. Es considerada una isoforma de la proteína HACD1 relacionada con el metabolismo de ácidos grasos, diferenciándose entre ellas por los últimos 15 aminoácidos del extremo carboxilo de la secuencia.¹⁵ Su estructura secundaria está conformada por α -hélice (43.2%), láminas β (8.9%), giros β (2%) y sin estructura (45.9%).²⁸⁻²⁹ Debido a su alto porcentaje sin estructura, HACD1/CAP es considerada como una proteína intrínsecamente desordenada, este tipo de proteínas se caracterizan por ser multifuncionales y por tener diferentes propiedades de unión

(asociados a la biomineralización como la nucleación y/o control del crecimiento de los cristales de hidroxiapatita).³⁰⁻³³ Se ha demostrado que HACD1/CAP se une a hidroxiapatita, y con mayor fuerza al cemento que a la superficie dentinaria. Se une también con fibronectina, pero es 150 veces más fuerte su unión a hidroxiapatita. Estas observaciones sugieren que HACD1/CAP tiene un papel importante en la formación del cemento.¹⁵ Olson y colaboradores,³⁴ demostraron que HACD1/CAP promueve la unión de fibroblastos, células endoteliales y algunas células musculares. Tiene, también, la capacidad de dirigir la migración celular desde el hueso alveolar, permitiéndole jugar un rol importante durante el reclutamiento de células específicas en el proceso de regeneración, así como mediar sus funciones biológicas.¹⁵ HACD1/CAP podría ser la vía natural por la que los cementoblastos progenitores son atraídos a la superficie radicular durante la homeostasis y la regeneración, pues se ha reportado que HACD1/CAP posee capacidad de anclar células progenitoras del Ligamento Periodontal y esto se encuentra directamente relacionado con la expresión de fosfatasa alcalina y la formación de tejido mineralizado.^{35, 36} En un estudio se puso a prueba la formación *in vitro* y la regeneración *in vivo* que induce HACD1/CAP, los resultados indicaron que esta proteína es capaz de ensamblar nanoesferas de 7.0nm - 27nm de diámetro, mientras que los resultados funcionales demostraron que promueve la nucleación de cristales de hidroxiapatita *in vitro*, y los estudios *in vivo* practicados en defectos críticos de rata mostraron que HACD1/CAP generó 73% y 87% de formación ósea a las 4 y 8 semanas respectivamente.⁶

Aunque originalmente HACD1/CAP se identificó como una proteína del cemento, ha demostrado ser muy efectiva induciendo la reparación ósea y regeneración, por lo tanto, esta proteína tiene un gran potencial para su aplicación en la bioingeniería de tejidos y regeneración tisular.

Esto es importante en la regeneración periodontal (donde las células que se desean reclutar son los fibroblastos y cementoblastos, pero no se desean células epiteliales) pues HACD1/CAP permite determinar qué células son requeridas y las dirige durante su diferenciación para lograr la homeostasis o regeneración

(que involucra la formación de cemento y tejido conectivo de anclaje). Estas acciones pueden lograrse mediante la producción de péptidos derivados de HACD1/CAP y CEMP-1, como se ha demostrado en los estudios previamente citados.

El avance en la investigación de la biología del cemento radicular y en particular, la investigación, identificación y expresión recombinante de marcadores específicos (CEMP1 y HACD1/CAP), ha hecho posible conocer con mayor fidelidad la biología molecular de dicho tejido. Sin embargo, continuar con esta investigación es esencial, para determinar si existen otras proteínas cemento específicas que no han sido identificadas.

Proceso para identificar una proteína nueva y su depósito en la base de datos HUGO Gene Nomenclature Committee

La base de datos Human Gene Nomenclature Committee (HGNC) es una división del comité internacional del Human Genome Organization (HUGO) que se encuentra involucrado en varios aspectos del campo de la genética. La división encargada de registrar nuevas proteínas es UnitProtKB/Swiss-Prot, al registrar su base de datos una nueva proteína el HUGO registrará símbolo estandarizado. El HGNC es el responsable de aprobar símbolos y nombres únicos para locus de origen humano, incluyendo genes codificantes, genes de ARN no codificante y pseudogenes, todo esto para evitar ambigüedades en la comunicación científica. Los símbolos aprobados se almacenan en la base de datos HGNC www.genenames.org, un recurso en línea que contiene nomenclatura de genes, grupos de genes y recursos asociados que incluyen enlaces externos a información específica acerca de genómica, proteómica y fenotipo.

Requisitos generales:

- Cada símbolo debe ser único.
- El símbolo debe ser una representación corta o abreviada, del nombre descriptivo del gen.
- Los símbolos deben contener letras en latín y números arábigos.
- Los símbolos no contendrán puntuación

- Los símbolos no deben contener “G” para gen.
- Los símbolos no deben referenciar la especie, por ejemplo, “H/h” en caso de humanos.

Requisitos específicos para enzimas y proteínas.⁴²

Aplican todas las reglas generales, y se incluyen las siguientes: los nombres de los genes que codifican para enzimas deben basarse en las recomendaciones del Comité de Nomenclatura de la Unión Internacional de Biología Molecular y Bioquímica, por ejemplo: *FPGS* “folil poliglutamato sintetasa”. Los nombres de los genes codificantes para proteínas plasmáticas, hemoglobinas y proteínas especializadas se basan en nombres estándar y aquellos recomendados por sus respectivos comités, por ejemplo, *HBA1* “hemoglobina, alfa 1”.

Estatus del símbolo:

Cada símbolo registrado puede tener uno de los siguientes estatus.

- Pendiente: símbolos en progreso.
- Aprobado: los símbolos oficiales están disponibles públicamente.
- Reservado: símbolos oficiales que no están disponibles públicamente, estos se publicarán en un periodo de seis meses.
- Símbolos retirados: símbolos previos que ahora tienen símbolos diferentes aprobados.
- Registro retirado: para símbolos que se ha demostrado que no existen.

En caso de no contar con datos sobre la función de la proteína, se debe basar en las siguientes reglas:

1. Basado en dominios estructurales conocidos en el gen.
2. Basado en genes homólogos con el genoma humano.
3. Basado en genes homólogos de otras especies.
4. Basado en la presencia de un marco de lectura

Siempre que sea posible, los genes relacionados se nombran utilizando un símbolo de raíz común para permitir la agrupación, generalmente en función de la homología de secuencia, la función compartida o la pertenencia a complejos de proteínas.

Proteína Cementogenina (CMGN)

En el presente trabajo se hace la propuesta de la expresión recombinante y purificación de la proteína cementogenina (CMGN), como la primera etapa para establecer si esta proteína es un posible nuevo marcador del cemento radicular humano. El marco de lectura de CMGN consta de 729 pb y da por resultado una proteína de 243 aminoácidos. La propuesta se encuentra en su etapa experimental y se ha generado un sistema de expresión recombinante procarionte, para posteriormente realizar la caracterización bioquímica y determinar su función biológica e inmunolocalización en líneas celulares y tejidos humanos.

Secuencia de la región codificante para CMGN:

```
1   ATGCAAACATTTGGGGAACAACCAGGGGCCAGTGCATGGCTGCCACAGT
51  GGCTCAGAGGACGGCAAGGACCATGGCATGCAAACATTTGGGGAACAACC
101 AGGGGCCAGTGCATGGCTGCCACAGTGGCTCAGAGGACGGCAAGGACCG
151 ATGGATGGACGACAGCCCCGAGAGCCAGCTGAGTTTGGAGACGGGGCTGT
201 GTGGCCATCCCGGCTGCCCGGGGCTCTTGCTCACCTCGTCACCTCTG
251 GCATCTCCCCGGCAGCACTTAGCAGCCCAGTGAGAGAGCGGTCAAGGCAG
301 GCGTGTGGATGGATCCAGGGCCCTCTCTCAAGGGTTCTGACAGATGAGGG
351 TCGAAGAAGGGCCAGGCTGTCCAGTCTCTGGGCCAGAGAGGATGGGCAA
401 GTCAAGGAACAGGCAGGGTCAGCAGCCGGGGTGCAGGGAGAGGCAGGTGT
451 ATGCTGGAAGGTCAGGCCCCACGAGGGCCCTTGGGGGCATCAAGTGGGAC
501 AGGGCCCCGGCGCACCTCAGTGCCCTGGGCAGGCCTCGGGCCAGTGATG
551 TGGTCACTCCCTCGGGAAATGGCGCGCCGGGCCCGCCGGGCCCGCCGGGC
601 TTTGCGGGCCCGCCGGGCGCTGTTGGGACCGGACCACCCACCCTGGGCCA
651 CACCCATCCCGCCTCAGGCCTGGATTTTCTGTTGCCGCAGATGTTTGCGT
701 GTCAGAGCTCCCGCAGATGTTTCGCATGTC
```

Previamente, en el Laboratorio de Biología Periodontal de la Facultad de Odontología UNAM, se ha realizado el trabajo bioinformático para determinar la posibilidad teórica de que CMGN sea una proteína expresada en el humano y que aún no esté caracterizada y/o registrada. Derivado del trabajo bioinformático

previo, se diseñó un vector de expresión con la secuencia codificante de CMGN y se transformaron las células de expresión recombinante. Con la realización del presente proyecto se establecerá el primer paso experimental que ayudará a determinar si la proteína propuesta se expresa en el humano, en particular, si se puede considerar un nuevo marcador del cemento radicular. Si la estrategia experimental que se plantea es positiva, se propondrá el nombre y registro de la proteína CMGN ante el Gene Nomenclature Committee.

PLANTEAMIENTO DEL PROBLEMA

El cemento radicular es un tejido de suma importancia para el mantenimiento y homeostasis del periodonto. En el área clínica, la neoformación del cemento radicular, después de la enfermedad periodontal, constituye la clave para la regeneración periodontal. Pese a la importancia de dicho tejido, el cemento radicular permanece como el tejido mineralizado menos estudiado del humano; debido a la limitación anatómica del tejido, a la dificultad del cultivo *in vitro* y a tener solo dos marcadores moleculares específicos (CEMP1 y HACD1/CAP), lo que ha dificultado la obtención y aplicación de tratamientos predecibles.

JUSTIFICACIÓN

La profundización en el conocimiento de las proteínas cemento específicas ha sido de suma importancia para entender la biología, a nivel molecular, de dicho tejido. La presente investigación plasmará un antecedente para la creación de una línea de producción recombinante que ayude a su futura caracterización bioquímica y la generación de nuevos marcadores moleculares bioactivos que, eventualmente, pueden generar terapias encaminadas a la regeneración del periodonto.

HIPÓTESIS

El sistema recombinante *E. coli Rosetta* permitirá la expresión de la proteína CMGN de forma soluble y en cantidad suficiente para establecer un protocolo de purificación cromatografía líquida de proteínas de alta velocidad.

OBJETIVO GENERAL

Expresar de forma recombinante y purificar a homogeneidad la proteína recombinante humana cementogenina (hrCMGN).

OBJETIVOS ESPECÍFICOS:

1. Realizar el análisis *in silico* de las características fisicoquímicas de la proteína hrCMGN.
2. Inducir la expresión de la proteína hrCMGN en el sistema recombinante *E. coli Rosetta* mediante isopropil- β -D-1-tiogalactopiranosido (IPTG).
3. Diseñar un protocolo de purificación por cromatografía líquida para la proteína hrCMGN.
4. Confirmar mediante inmunodetección (Western blot) la identidad de la proteína hrCMGN.

METODOLOGÍA

Análisis *in silico* de la proteína hrCMGN

Se realizó el análisis bioinformático de la estructura primaria de la proteína hrCMGN con la herramienta informática ProtParam (<https://web.expasy.org/protparam/>). Para realizar el análisis informático se utilizó la secuencia de aminoácidos en código de una letra en formato FASTA de la proteína hrCMGN. Los parámetros para analizar fueron: peso molecular, punto isoeléctrico teórico y composición de aminoácidos.

Predicción de estructura secundaria / terciaria

Se utilizó el servidor informático I-TASSER para realizar la construcción de la estructura tridimensional de la proteína hrCMGN. Se ingresó la secuencia de aminoácidos en formato FASTA y se realizó la visualización y análisis del modelo generado por el servidor a partir de reemplazo molecular.

Predicción de modificaciones post traduccionales

Mediante el servidor informático NetN Glyc 1.0 Server y NetPhos 3. 1, se realizó la predicción *in silico* de los sitios para glicosilación y fosforilación, con el empleo de la secuencia primaria en formato FASTA de hrCMGN.

Cultivo e inducción de la expresión de la proteína hrCMGN

Se crecieron células de *Escherichia coli* de la cepa *Rosetta* (DE3). La clona recombinante fue transformada con el plásmido pET-28a (+) (GenScript, USA), por medio de la técnica de choque térmico. Brevemente, se colocó 100 nanogramos del plásmido en 50 µl de cultivo de células competentes *Rosetta* (DE3) y se incubaron en hielo durante 30 minutos. Posteriormente se realizó el choque térmico al incubar el tubo donde se encontraban las células y el plásmido durante 45 segundos a 42 °C e inmediatamente se enfrió en hielo durante 2 minutos. Después de este proceso, se agregó 200 µl de medio Luria Bertani (LB) y se mantuvo en agitación durante una hora a 37 °C. Las células fueron sembradas en placas de agar LB adicionadas con kanamicina (150 µg/ml) y 24 horas después, se seleccionaron colonias aisladas para su cultivo e inducción

de la expresión de CMGN. Para la expresión inducible, el plásmido pET-28a (+), contiene un promotor inducible con IPTG (promotor lacI), para su selección presenta el gen de resistencia a kanamicina (KanR) y para su purificación e inmunodetección cuenta con una bandera de 6 histidinas (6xHis) en el extremo amino terminal, (figura 1).

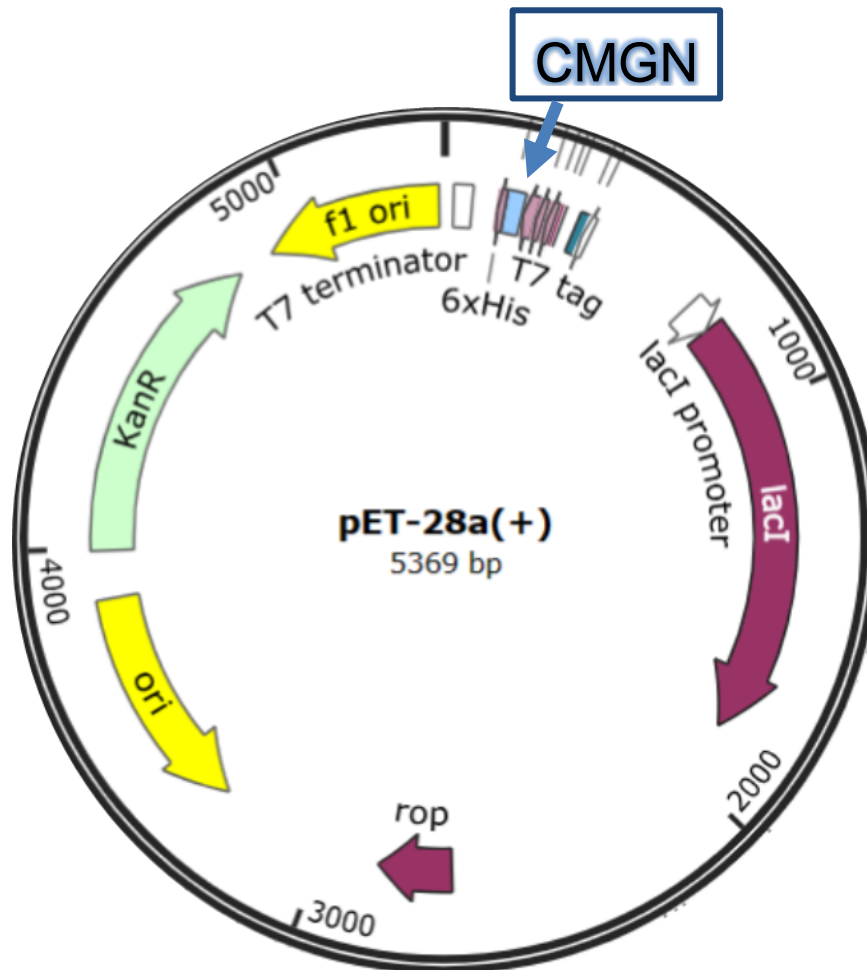


Figura. 1 Plásmido pET- 28a(+)/CMGN

Las células de *Rosetta* con la construcción pET-28a (+) /CMGN, se crecieron a 37°C hasta alcanzar una D.O. de 0.4 a 600 nm, en medio LB adicionado con 100 µg/ml de kanamicina. Para realizar la inducción de la expresión de la proteína hrCMGN, se agregó al medio Isopropil-d-tiogalactopiranosido (IPTG) a una concentración final de 0.5 mM y el cultivo fue incubado a 37°C con agitación a 200 rpm durante una noche. Después del periodo de incubación, las bacterias se recuperaron mediante centrifugación a 7500 rpm a 4°C en ciclos de 10 minutos. El botón celular se resuspendió en amortiguador de ruptura (20mM fosfato de sodio pH 6.8, 100mM cloruro de sodio, 2mM fluoruro de

fenilmetilsulfonilo) 4 ml de amortiguador de ruptura por cada 100 ml de medio LB, y fue sonicado durante cinco ciclos de un minuto con intervalos de un minuto entre cada uno (VirSonic 100, Virtis). Una vez sonicado, se le adicionó inhibidores de proteasas (PMSF, N-etilmaleimida, aprotinina 1 µg/ml, leupeptina y pepstatina 0.5 µg/ml). El sobrenadante se recuperó por medio de dos ciclos de centrifugación a 7,500 rpm durante 15 minutos, a 4°C (Multifuge 3S-R Heraeus). El extracto crudo de proteínas se cuantificó con la técnica colorimétrica de Bradford, utilizando un espectrofotómetro a una longitud de onda UV de 595nm (Thermo scientific GENESYS uv-visible).

Purificación de hrCMGN por FPLC

Para el primer paso de purificación se empleó una cromatografía de afinidad a iones metálicos con el uso de la columna HisPrep FF/16 (GE Healthcare) en el sistema de cromatografía líquida de alta velocidad (FPLC) AKTA star (GE Healthcare). La interacción entre el extracto crudo y la columna se dio en un flujo de 0.5ml/min, posteriormente se realizó un lavado de dos volúmenes y medio de columna (50 ml) para separar la fracción no unida con el amortiguador A (imidazol 10 mM en PBS 1x pH 7.4). Para la separación de la fracción unida de proteínas se utilizó un gradiente lineal con el amortiguador B (imidazol 300 mM en PBS 1x pH 7.4). El volumen para llegar al 100% del amortiguador B fue de 2.5 volúmenes de columna (50 ml), a un flujo de 1ml/min.

Para el segundo paso de purificación se realizó una cromatografía de afinidad a hidroxiapatita, en el sistema de cromatografía líquida de alta velocidad (FPLC) AKTA star (GE Healthcare). Se emplearon, como extracto crudo, las fracciones de la primera cromatografía donde se identificó la presencia de hrCMGN. Para el lavado de la fracción no unida se empleó el amortiguador A (fosfato de sodio 20 mM pH 6.8) y para realizar la elución de la fracción unida se realizó un gradiente lineal con el amortiguador B (fosfato de sodio 400 mM pH 6.8). Para llegar al 100% del amortiguador B se emplearán 10 volúmenes de columna (20 ml), a un flujo de 0.5ml/min.

Para el tercer paso de purificación se emplearon como extracto crudo (dializado contra fosfato de sodio 20 mM y ultrafiltrado por centricon con tamaño de poro

de 10 kDa), las fracciones de la segunda cromatografía donde se identificó la presencia de hrCMGN. Se realizó una nueva cromatografía de afinidad a hiroxiapatita. Se utilizaron amortiguador A (fosfato de sodio 20 mM pH 6.8) y amortiguador B (fosfato de sodio 400 mM pH 6.8), empleamos 10 volúmenes de columna (20 ml) para llegar al 100% de B.

En cada paso de cromatografía, las muestras de interés fueron visualizadas mediante electroforesis SDS-PAGE al 12%. Y al final se realizó inmunodetección por Western Blot, utilizando anticuerpos α -His.

Visualización de cromatografías e inmunodetección de hrCMGN

Para visualizar el resultado de las cromatografías se empleó el cromatograma (absorbancia de aminoácidos aromáticos a 280 nm contra el volumen de elución). Con este parámetro se seleccionaron las fracciones de interés y se realizaron electroforesis en gel de poliacrilamida al 12% en condiciones desnaturalizantes (SDS-PAGE), para cada paso de cromatografía.

Al final se realizó inmunodetección por medio de Western blot. Después de la separación electroforética, las proteínas fueron electro-transferidas a una membrana de PVDF (fluoruro de polivinilideno) durante 1 hora, la membrana se bloqueó para evitar sitios inespecíficos de unión de los anticuerpos (Solución de leche descremada al 1% en PBS 1x pH 7.4). Posteriormente, la membrana fue incubada durante una noche con el anticuerpo primario anti-6xHis (His-probe H-15 Santa Cruz Biotechnology) a una concentración de 1:3000. Transcurrido el periodo de incubación del primer anticuerpo se realizaron 3 lavados de 10 minutos con PBS 1x y se incubó el anticuerpo secundario (IgG conjugado con peroxidasa de rábano) durante 2 horas a una concentración de 1:3000. Pasado el periodo de incubación del anticuerpo secundario se realizó el revelado mediante reacción con DAB (diaminobenzidina), PBS 1x, peróxido de hidrógeno y cloruro de níquel al 8%.

Análisis BLAST

El alineamiento de secuencia local BLAST (Basic Local Alignment Search Tool), permite encontrar homología de secuencia de aminoácidos de nuestra proteína

con una base de datos de proteínas de mamífero (<https://blast.ncbi.nlm.nih.gov>). Se realizó el análisis de la secuencia primaria de hrCMGN en formato FASTA y se seleccionó la base de datos de proteínas no redundantes de UnitProtJK/Swiss-Prot, primero se utilizó el filtro de búsqueda “*Homo sapiens*” y posteriormente un BLAST general en la base de mamíferos.

RESULTADOS

Análisis *in silico* de la proteína hrCMGN

```
ProtParam
User-provided sequence:
      10      20      30      40      50      60
HHHHHHMQTF GEQPGAQCMA ATVAQRTART MACKHLGNNQ GPSAWLPQWL RGRQGPMDGR
      70      80      90     100     110     120
QPREPAEFGD GAVWPSRLPA GLLLTLVTSG ISPAALSSPV RERSRQACGW IQGPLSRVLT
     130     140     150     160     170     180
DEGRRRARLS SLWAREDGQS QGTGRVSSRG AGRGRCMLEG QAPRGPLGAS SGTGPRRTLS
     190     200     210     220     230     240
ALGRPRASDV VTPSGNGAPG PPGPPGFAGP PGAVGTGPPT LGHHPASGL DFLLPQMFAC
QSSRRCSHV
```

Figura. 2 Secuencia de aminoácidos de la proteína hrCMGN ingresada a la herramienta bioinformática ProtParam para su análisis.

La proteína hrCMGN está conformada por 249 aminoácidos (243 aminoácidos propios de la proteína, más 6 histidinas de la bandera en el extremo amino terminal). Se calculó un peso molecular teórico de 26.1 kDa y un punto isoeléctrico teórico de 11.55, este hace referencia al punto donde las cargas positivas y negativas de la proteína se igualan, es decir, representa el valor de pH al cual CMGN tiene carga neta cero. A pH fisiológico la carga de la proteína será positiva.

La composición de aminoácidos de hrCMGN puede observarse en la Tabla 1. donde el aminoácido más abundante es Glicina (Gly) que representa el 14.9% de la composición total (249 residuos), el menos abundante es Lisina (Lys) con 0.4% de la composición total, la proteína no contiene Tirosina (Tyr). El análisis de la secuencia nos permite contabilizar el número de aminoácidos cargados

positiva y negativamente, esto debido a la presencia de aminoácidos como: ácido aspártico y ácido glutámico (carga negativa) o arginina y lisina (carga positiva).

Composición de aminoácidos en hrCMGN		# de aminoácidos
Gly	14.9%	37
Arg	10.8%	27
Pro	10.8%	27
Ala	10%	25
Ser	8.8%	22
Leu	8.4%	21
Gln	6.0%	15
Thr	5.6%	14
Val	4.0%	10
His	4.0%	10
Glu	2.8%	7
Cys	2.4%	6
Asp	2.4%	6
Met	2.4%	6
Phe	2.0%	5
Trp	2.0%	5
Asn	1.2%	3
Ile	0.8%	2
Lys	0.4%	1
Tyr	0.0%	0

Tabla 1. Composición de aminoácidos de CMGN y residuos totales cargados positiva y negativamente. Aminoácidos polares neutros (azul), básicos (verde), ácidos (naranja), apolares (morado), aromáticos (amarillo).

Propiedades físico-químicas de hrCMGN	
Número de aminoácidos	249
Peso molecular	26101.44 Da
pI	11.55

Tabla 2. Propiedades físico-químicas de CMGN.

Predicción de estructura secundaria y terciaria

Mediante el servidor informático I-TASSER se realizó la construcción del modelo tridimensional teórico de hrCMGN. Además, se realizó la cuantificación de la estructura secundaria y las posibles formaciones de puentes disulfuro.

El porcentaje de estructura secundaria (incluyendo la bandera de 6 histidinas) es de: 59% random coil, 41% alfa hélice y 0% lámina beta. (Ver tabla 3).

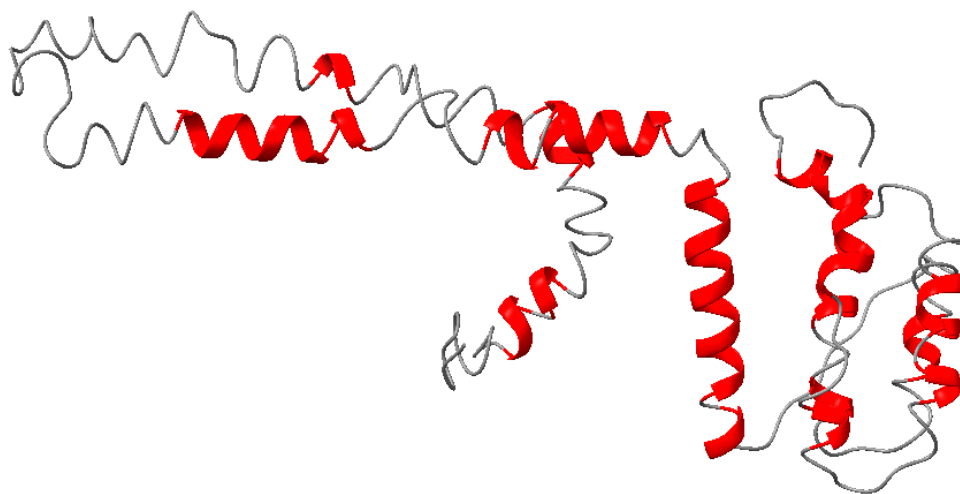


Figura 3. Modelo tridimensional teórico de hrCMGN (gris: random coil, rojo: alfa hélice).

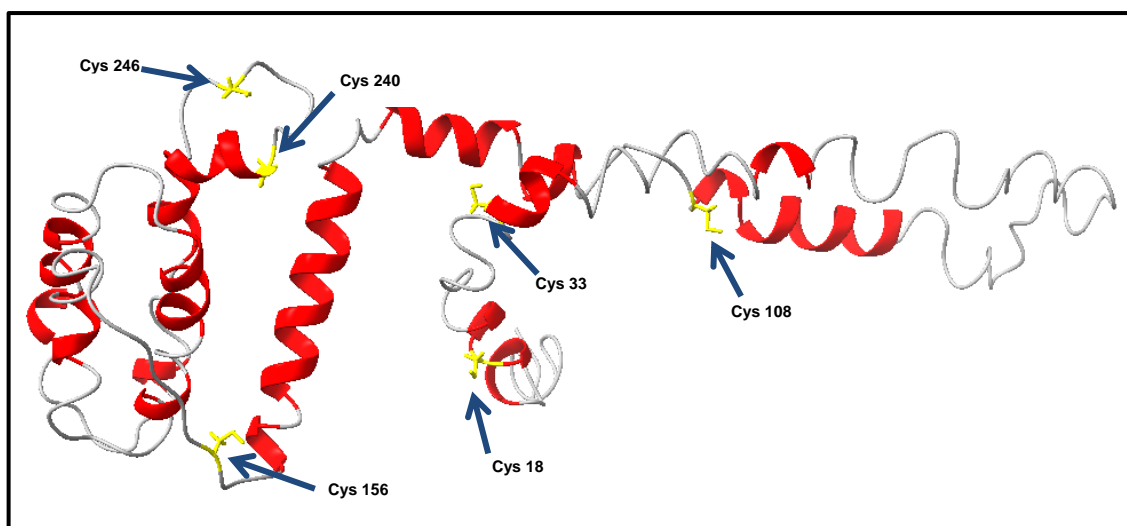
```

1  HHHHHMQTFGEQPGAQCMAATVAQRTARTMACKHLGNNQGPSAWLPQWL
51  RGRQGPMDGRQPREPAEFGDGAVWPSRLPAGLLLTLVTSGISPAALSSPV
101 RERSRQ ACGWIQPLSRVLTDEGRRRARLSSLWAREDGQSQGTGRVSSRG
151 AGRGRCMLEGQAPRGPLGASSGTGPRRTLSALGRPRASDVVTPSGNGAPG
201 PPGPPGFAGPPGAVGTGPPTLGHTHPASGLDFLLPQMFA CQSSRRCSHV

```

Random coil (Gris)	59%
Alfa hélice (Rojo)	41%
Lámina beta (Azul)	0%

Tabla 3. Porcentaje de estructura secundaria de la proteína hrCMGN.



```

1  HHHHHHMQTFGEQPGAQCMAATVAQRTARTMACKHLGNNQGPSAWLPQWL
51  RGRQGPM DGRQPREPAEFGDGAVWPSRLPAGLLLTLVTSGISPAALSSPV
101 RERSRQ ACGWIQGPLSRVLTDEGRRRARLSSLWAREDGQSQGTGRVSSRG
151 AGRGRCMLEGQAPRGPLGASSGTGPRRTLSALGRPRASDVVTPSGNGAPG
201 PPGPPGFAGPPGAVGTGPPTLGHTHPASGLDFLLPQMFA CQSSRRCSHV

```

Figura 4. Modelo tridimensional teórico y estructura primaria de hrCMGN que resaltan las 6 cisteínas (amarillo) presentes en las posiciones 18, 33, 108, 156, 240 y 246. Este modelo permite observar que no hay presencia de puentes disulfuro intracatenarios, pero sí hay cisteínas libres que podrían formar puentes disulfuro. Durante el proceso de purificación se observaron proteínas de mayor

peso molecular que persistían en las electroforesis, se piensa que podrían ser dímeros de hrCMGN.

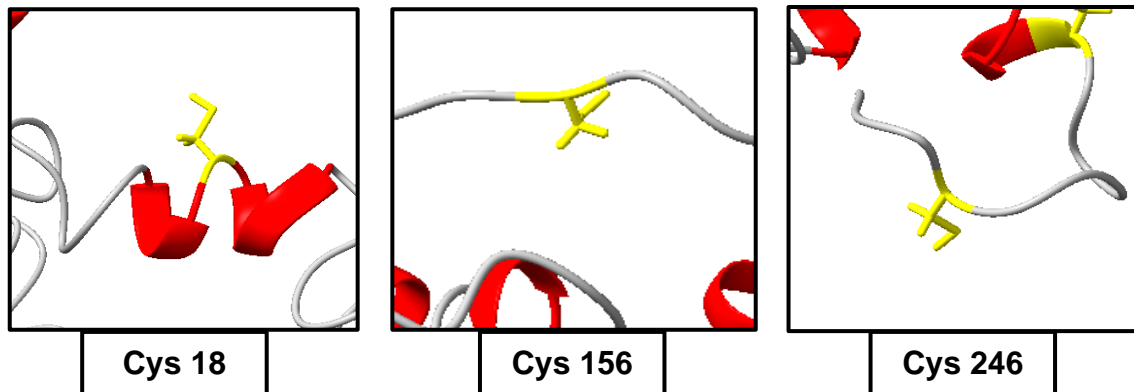
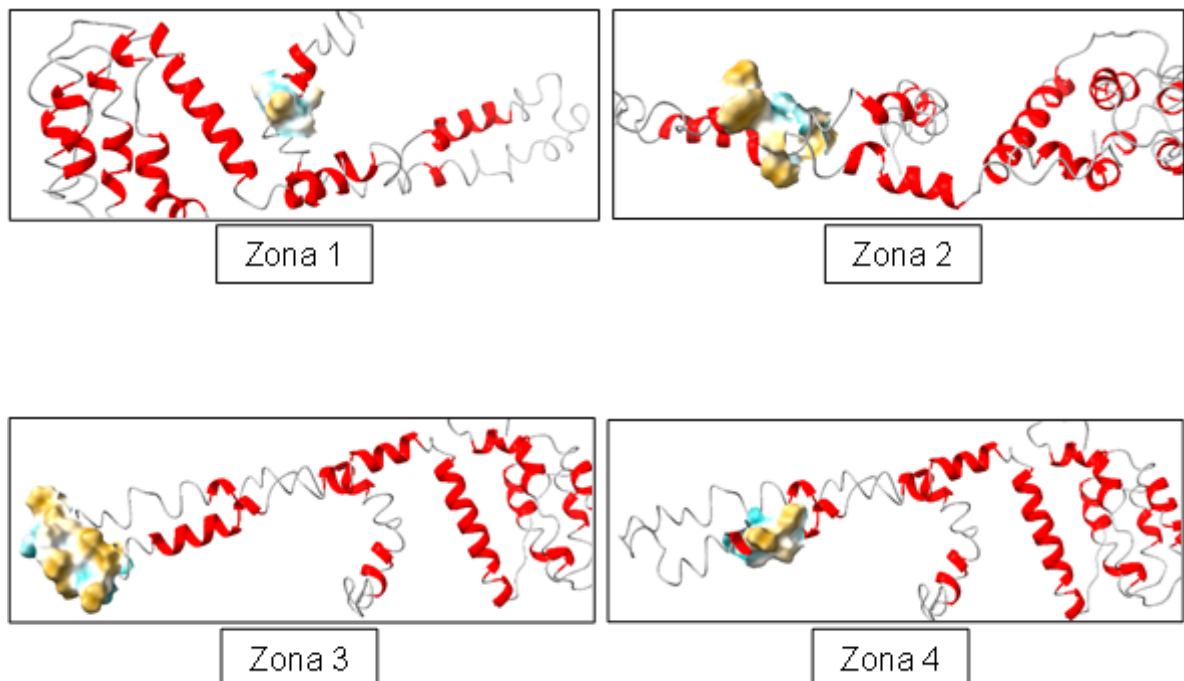
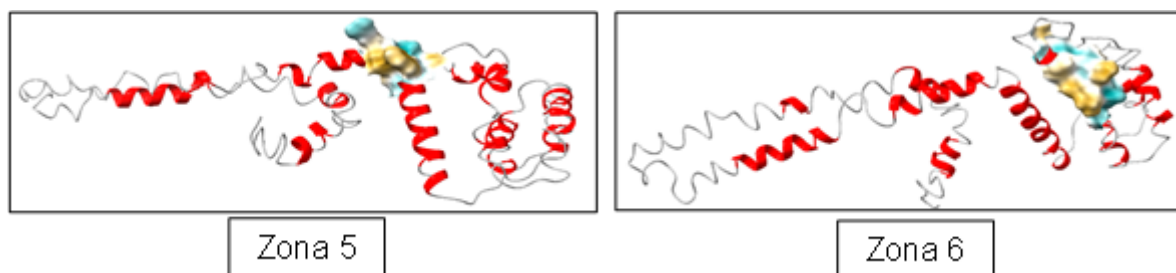


Figura 5. Detalle del modelo para observar las 3 cisteínas libres (18, 156 y 246), que posiblemente pueden formar 3 puentes disulfuro y promover la formación de dímeros durante la purificación de hrCMGN, lo que explicaría las proteínas de mayor peso molecular que nos siguen en las electroforesis durante el proceso de purificación.





```

1  HHHHHHMQTfGEQPGAQCMAATVAQRTARTMACKHLGNNQGPSAWLPQWL
51  RGRQGPMdGRQPREPAEFgDGAVWPSRLPAGLLLTLVTSGISPAALSSPV
101 RERSRQ ACGWIQGPLSRVLTDEgRRRARLSSLWAREDGQSQGTGRVSSRG
151 AGRGRCMLEGGAPRGPLGASSGTGPRRTLSALGRPRASDVVTPSGNGAPG
201 PPGPPGFAGPPGAVGTGPPTLGHThPASGLDFLLPQMFA CQSSRRCSHV

```

Figura 6. Modelo de accesibilidad solvente de hrCMGN y estructura primaria, que muestran las zonas mayormente hidrofóbicas (color arena y subrayadas en amarillo), la zona 1) va de los residuos 20 al 24, zona 2) de los residuos 44 al 50, zona 3) de los residuos 72 al 87, zona 4) de los residuos 94 al 100, zona 5) de los residuos 127 al 134 y la zona 6) de los residuos 227 al 239. De los 249 aminoácidos que forman hrCMGN, hay 71 residuos hidrofóbicos (28.5%) y 178 residuos hidrofílicos (71.5%), por lo que se puede inferir que hrCMGN tiende a ser hidrofílica.

Predicción de modificaciones post traduccionales

Las proteínas cemento específicas (CEMP1 y HACD1/CAP) presentan modificaciones post traduccionales que les permiten la unión de iones calcio a su superficie y regular el crecimiento de la hidroxiapatita. Mientras que las fosforilaciones favorecen la unión de Ca^{2+} a las proteínas y, por lo tanto, pueden desempeñar un papel en las primeras etapas de mineralización durante la formación de fosfato octacálcico. Teniendo esto en cuenta, fue que se investigó la posibilidad teórica de modificación post traduccional en hrCMGN.

De acuerdo con el servidor informático NetN Glyc 1.0 Server, no se predicen sitios de glicosilación para hrCMGN, mientras que el servidor NetPhos 3. 1ª, predice 36 sitios de fosforilación en las serinas de posición: 43, 76, 89, 92, 97, 98, 104, 116, 130, 131, 140, 147, 148, 170, 171, 180, 188, 194, 228, 242, 243,

247; y en las treoninas de posición: 9, 22, 27, 30, 85, 88, 120, 143, 173, 178, 192, 216, 220, 224.

La N-glicosilación ocurre mediante la formación de un enlace glucosídico entre el grupo amida de una asparagina y un oligosacárido. Para que ocurra la glicosilación, la asparagina debe estar en una secuencia específica de aminoácidos conocida como “motivo de glicosilación”, esto es Asn-X-Ser/Thr, es decir la asparagina (Asn) debe estar seguida de cualquier aminoácido (“X”) excepto prolina (Pro), luego una serina (Ser) o treonina (Thr). Por otro lado, la O-glicosilación se produce en el contexto de la secuencia de aminoácidos Ser/Thr-X-Ser/Thr. Es importante tomar en cuenta que no todas las proteínas que contienen la secuencia serán necesariamente glicosiladas, ya que pueden existir otros factores que influyan, pero es relevante predecir la posibilidad para la investigación de hrCMGN.

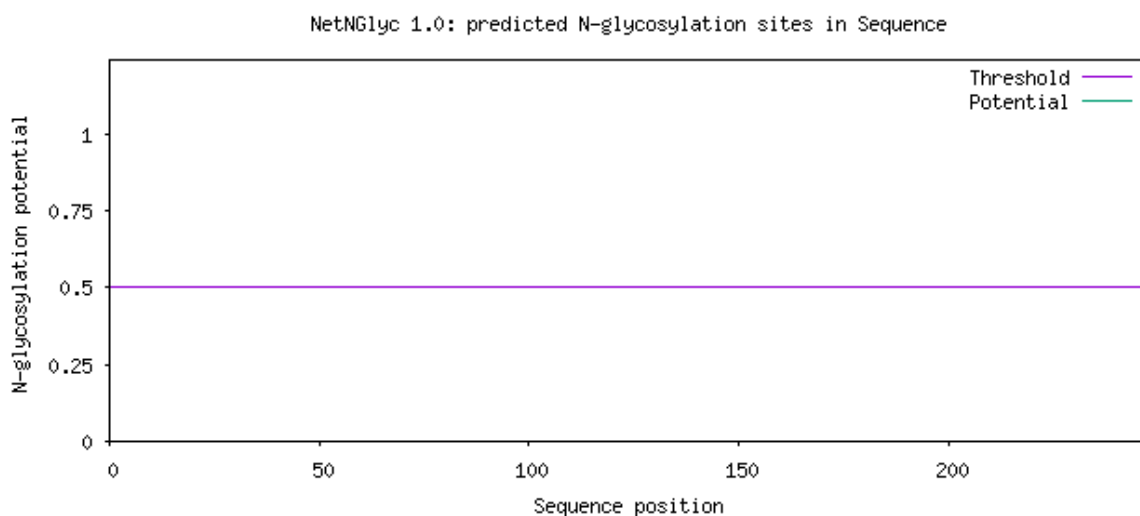


Figura 7. Predicción de posibles sitios de glicosilación en la proteína hrCMGN. El servidor no identificó sitios, un punto a destacar es el bajo porcentaje de asparaginas en hrCMGN (1.2%), lo que reduce las posibilidades de glicosilación.

Las secuencias de aminoácidos que pueden ser fosforiladas, tienen una secuencia específica para que las quinasas (enzimas responsables de la fosforilación) las reconozcan y lleven a cabo la modificación. Existen 3

principales motivos de fosforilación, en los que se añade un grupo fosfato (-PO₄) al grupo hidroxilo (-OH) de los aminoácidos de la proteína.

1. Motivo de fosforilación de serina (Ser): S-X-X-S
2. Motivo de fosforilación de treonina (Thr): T-X-X-T
3. Motivo de fosforilación de tirosina (Tyr): Y-X-X-Y

Donde “S” representa serina, “T” treonina, “Y” tirosina y “X” cualquier aminoácido.

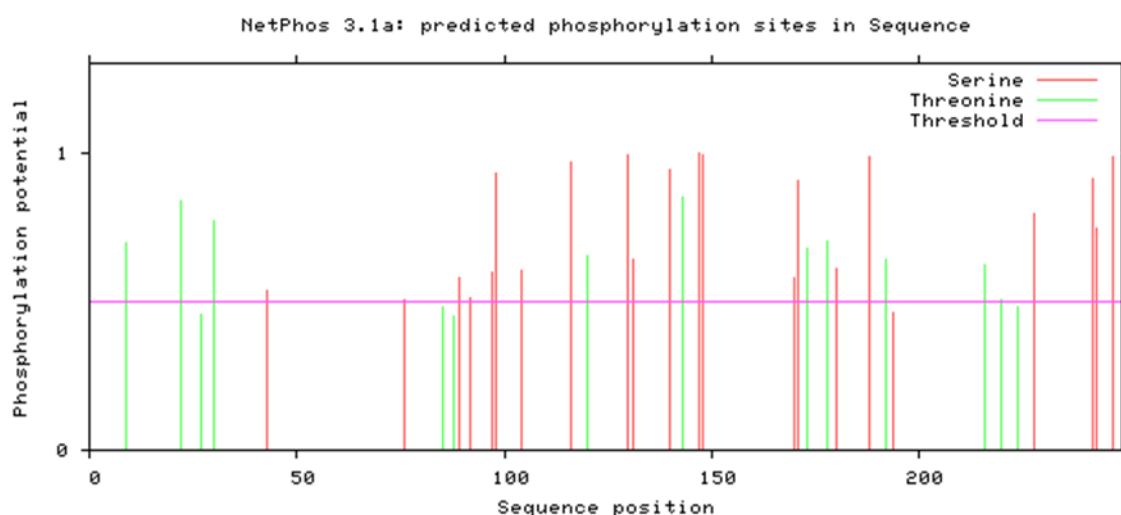


Figura 8. Predicción de posibles sitios de fosforilación en la proteína hrCMGN. El servidor identificó 36 posibles sitios. (<https://services.healthtech.dtu.dk/cgi-bin>). Las serinas forman más fosforilaciones comparado con las treoninas ya que en la composición de aminoácidos hay más serinas (22 residuos) que treoninas (14 residuos). Mientras que no hay fosforilaciones de tirosina porque no hay tirosina en la composición de hrCMGN.

Cultivo y obtención de biomasa

De las células de *E coli Rosetta* con la construcción pET-28a (+) /CMGN en 1L de medio LB adicionado con kanamicina e inducidas con IPTG, se obtuvieron 2.74 g de biomasa, la muestra fue sonicada y el sobrenadante se recuperó y clarifico por medio de centrifugación para comenzar la purificación.

Purificación de hrCMGN por FPLC

Debido a que la hrCMGN cuenta con una bandera de histidinas en el amino terminal, se decidió usar como cromatografía de captura (primer paso) una

cromatografía de afinidad a iones metálicos con el uso de la columna de níquel HisPrep FF/16. En el cromatograma (Fig. 16) se pueden observar: pico 1: la fracción del extracto crudo que no tuvo interacción con la columna, la fracción de proteínas con afinidad al Níquel pico 2 (48% de B) pico 3 (62% de B) y pico 4 (66% de B) con un rango de elución lineal de 10mM (A) - 300 mM (B) de Imidazol en PBS 1x. De acuerdo con los resultados obtenidos en el cromatograma, se realizó electroforesis SDS-PAGE de las fracciones de interés. Ver Figura 9 y 10.

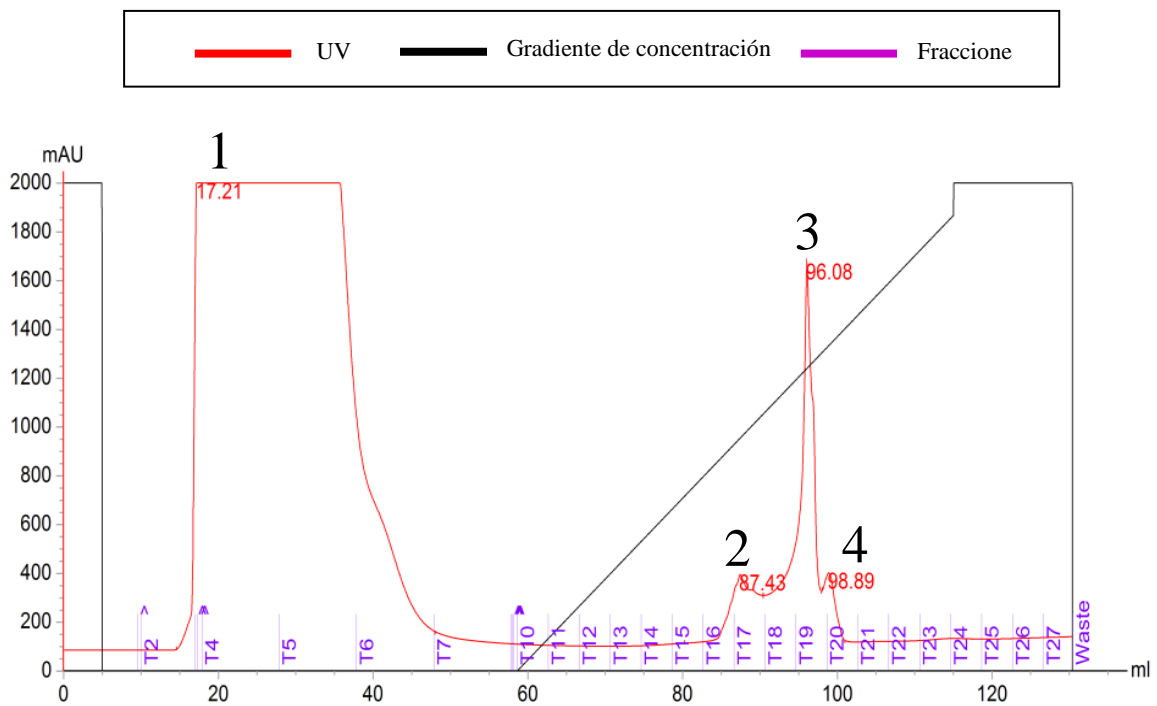


Figura 9. Cromatograma de primer paso de purificación hrCMGN.
Cromatografía de afinidad a iones metálicos.

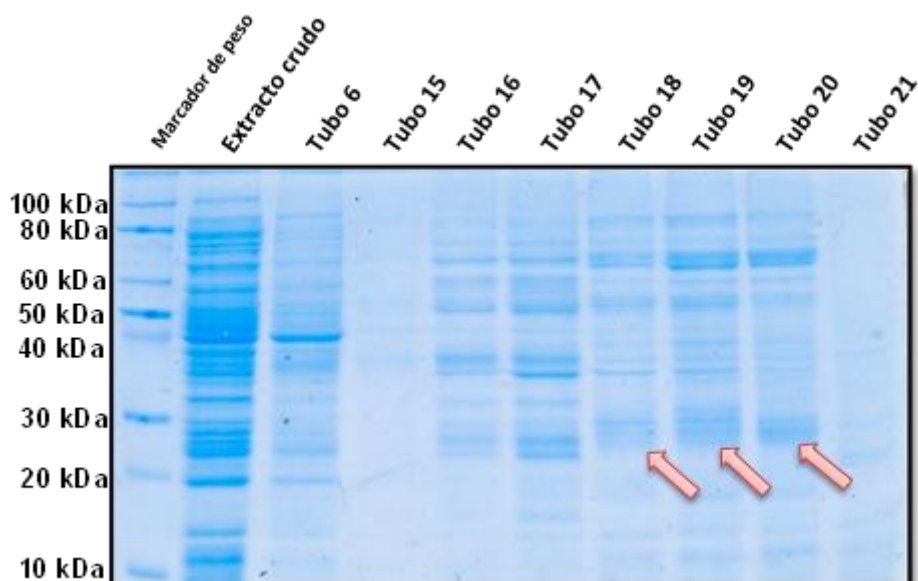


Figura 10. Electroforesis SDS-PAGE del primer paso de purificación de hrCMGN. Marcador de peso (carril 1), extracto crudo (carril 2), pico 2 del cromatograma (carril 3 a 6), pico 3 del cromatograma (carril 7 y 8), pico 4 del cromatograma (carril 9), tubo 21 (carril 10). Bandas correspondientes a hrCMGN señaladas con flechas rojas.

Segundo paso de purificación. Cromatografía de afinidad a hidroxiapatita, con un gradiente lineal de (A) 20mM – (B) 400mM de fosfato de sodio pH 6.8. Se utilizaron las fracciones correspondientes al pico 4 de la cromatografía previa, la muestra se dializó utilizando el amortiguador A y se concentró mediante filtros para centrífuga de 10kDa de tamaño de poro (Amicon Ultra 0.5 mL). En el cromatograma del segundo paso de purificación se puede observar la fracción de proteínas no unidas a la columna (picos 1 y 2), la fracción de proteínas con afinidad a hidroxiapatita (pico 3 con 52% de B). Se realizó electroforesis SDS-PAGE del extracto final obtenido en la primera cromatografía (carril 1), pico 3 del cromatograma (carril 2 a 6). Ver figuras 11 y 12.

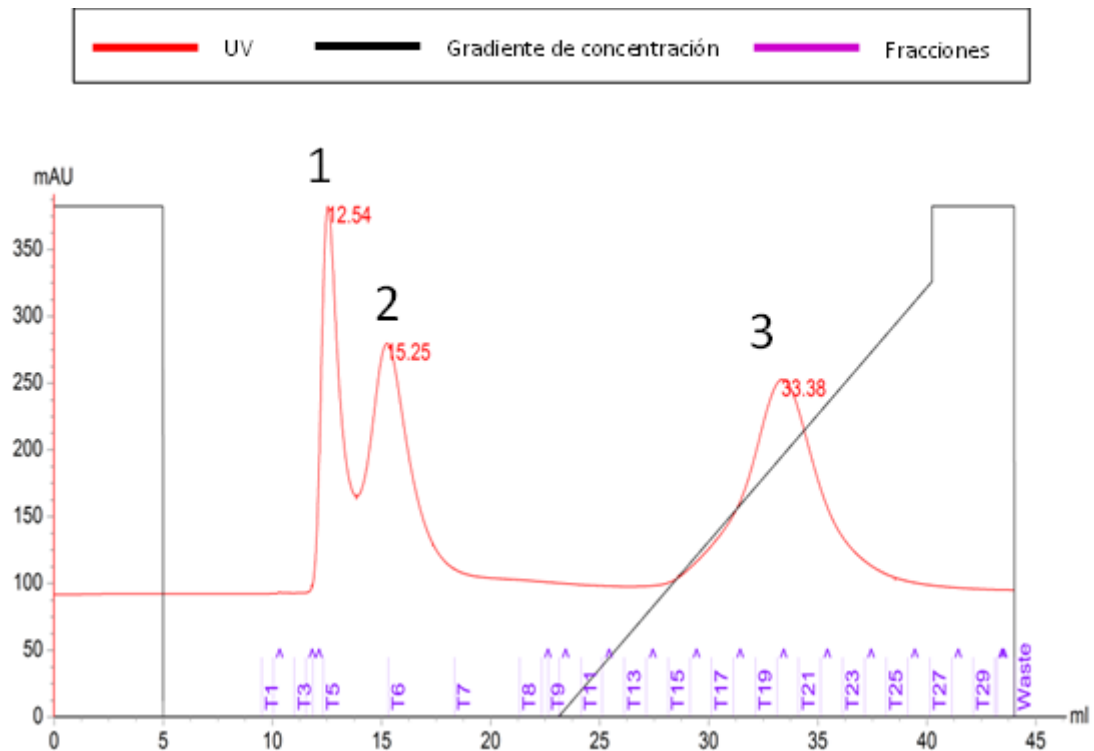


Figura 11. Cromatograma del segundo paso de purificación. Columna de hidroxiapatita. Fracción de proteínas no unidas a la columna (picos 1 y 2), fracción de proteínas con afinidad a hidroxiapatita (pico 3).

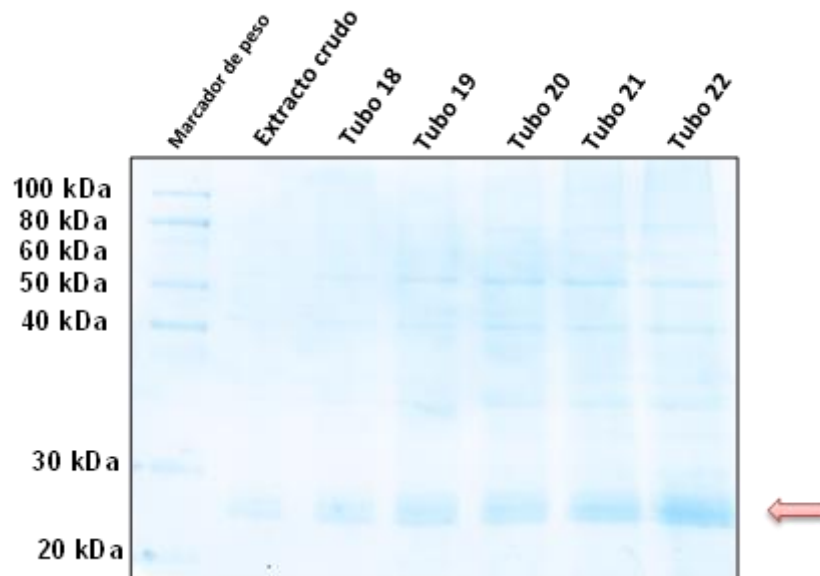


Figura 12. Electroforesis SDS-PAGE del segundo paso de purificación de hrCMGN por cromatografía de afinidad a hidroxiapatita. Marcador de peso (carril

1), extracto crudo (carril 2), pico 3 del cromatograma (carriles 3 al 7). Flecha roja indica la posición de la proteína en todos los carriles.

Para el tercer paso de purificación se realizó una cromatografía de afinidad a hidroxiapatita, con un gradiente lineal de 20mM (A) - 400mM (B) de fosfato de sodio pH 6.8. Se concentraron y dializaron las fracciones de la segunda cromatografía mediante filtros para centrífuga de 10kDa de tamaño de poro (Amicon Ultra 0.5 mL). En el cromatograma del tercer paso se puede observar la fracción de proteínas no unidas a la columna (1), el intervalo de lavado de las proteínas no unidas a la columna (2), la fracción de proteínas con afinidad a hidroxiapatita con un rango de elución de 20mM a 400mM de fosfato de sodio (3). Se realizó electroforesis SDS-PAGE del extracto crudo (carril 2), pico 2 del cromatograma (carril 3) y del pico 3 del cromatograma al 42% de B (carril 4 a 10). Ver figuras 13 y 14.

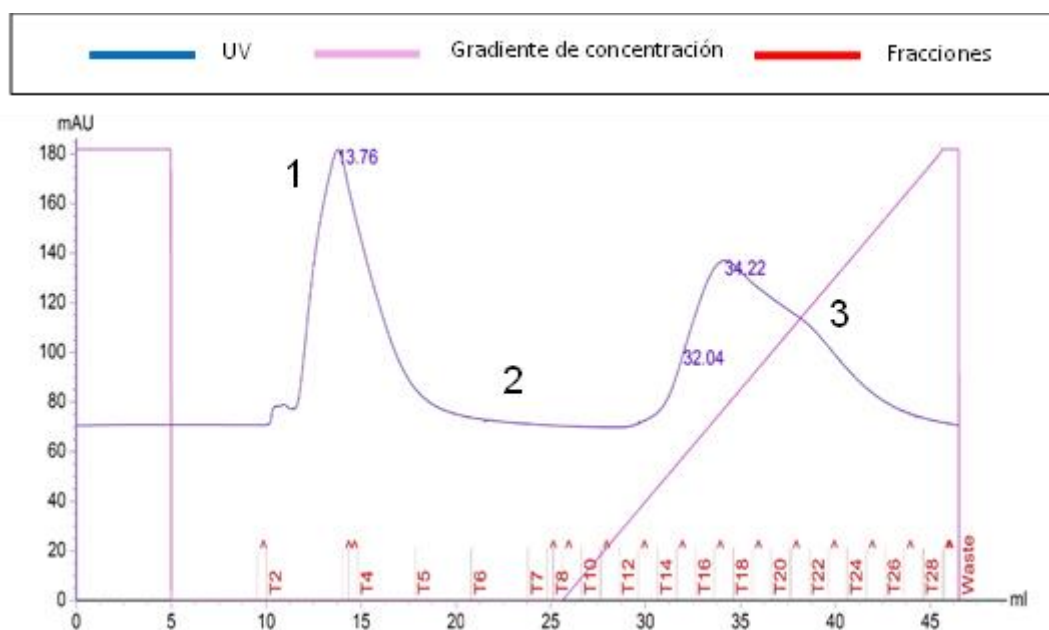


Figura 13. Cromatograma del tercer paso de purificación de hrCMGN. Fracción de proteínas no unidas a la columna (1), intervalo de lavado de las proteínas no unidas a la columna (2), fracción de proteínas con afinidad a hidroxiapatita con un rango de (pico 3). Cromatografía de afinidad a hidroxiapatita.

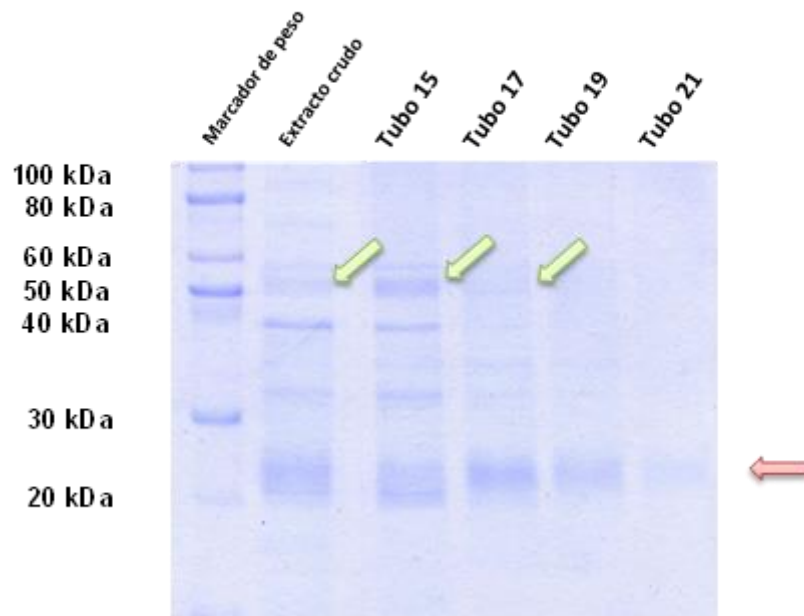


Figura 14. Electroforesis SDS-PAGE del cromatograma del tercer paso de purificación de hrCMGN por cromatografía de afinidad a hidroxiapatita. Marcador de peso (carril 1), extracto obtenido del segundo paso de cromatografía (carril 2), pico 3 del cromatograma (carril 3 al 6). Se observan posibles dímeros de hrCMGN formados por puentes disulfuro (flechas de verde).

El extracto crudo de proteínas se cuantificó con la técnica colorimétrica de Bradford, que consiste en la cuantificación de la unión de un colorante (Azul de Coomassie G-250) a la proteína, comparando esta unión con la de diferentes cantidades de una proteína estándar (Albúmina de Suero Bovino BSA). La cuantificación se hace midiendo la absorbancia en un espectrofotómetro, a 595 nm, y se grafica la absorbancia vs la concentración de proteínas, se obtiene la curva de calibración de la proteína estándar. Con esta curva de calibración, se puede interpolar la concentración de proteínas en una muestra al medir su absorbancia a 595 nm. El rendimiento final de la purificación de hrCMGN es equivalente a 1.7 mg de proteína por cada litro de cultivo.

Inmunodetección de hrCMGN

La identificación de la proteína hrCMGN se realizó mediante inmunodetección por Western Blot, utilizando los anticuerpo primario policlonal anti-6xHis (His-probe H-15 Santa Cruz Biotechnology) a una concentración de 1:3000. Para el anticuerpo secundario se utilizó un IgG conjugado con peroxidasa de rábano. Debido a que en los geles SDS- PAGE mostraron posibles dímeros de la proteína, la muestra se trató con reductores ditioneitol (DTT) y urea 8M.

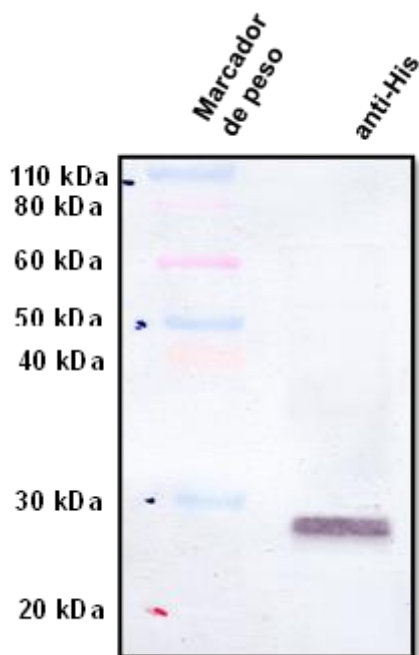


Figura 15. Inmunodetección de hrCMGN por Western Blot, utilizando anti-6xHis

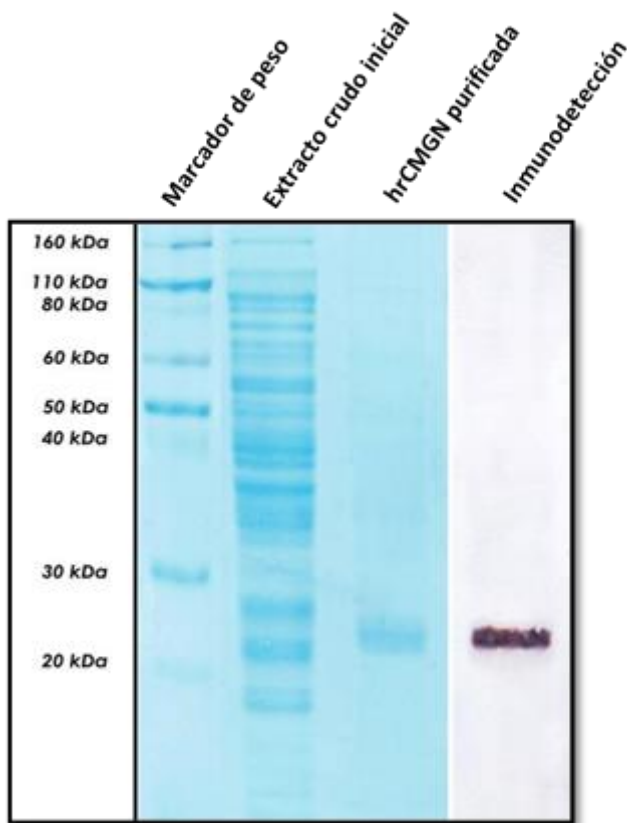


Figura 16. Comparativa de la muestra a través del proceso de purificación. Marcador de peso (carril 1), extracto crudo inicial (carril 2), hrCMGN después del tercer paso de purificación (carril 3), inmunodetección de hrCMGN con anti-6xHis (carril 4).

BLAST

De acuerdo con el análisis obtenido en la plataforma, la secuencia de hrCMGN comparte un 50% de homología con una región de 121 aminoácidos de la proteína EMI del dominio 1. Tiene 37.6% de homología con una región de 115 aminoácidos de la proteína colágena II en la isoforma 1 de la cadena alfa 1. Con la cadena alfa 1 de la colágena VIII resultó un 32.6% en una región de 109 aminoácidos. Y con la una isoforma de la ectodisplasina A hubo homología del 29.5% con una región de 105 aminoácidos. Por otra parte, los resultados obtenidos por taxonomía mostraron que CMGN tiene homología en un rango de 69–82% en proteínas no caracterizadas de algunos miembros de la familia de los homínidos (*Gorilla gorilla*, *Pan paniscus*, *Pan troglodytes*, *Nomascus leucogenys*, *Symphalangus sindactylus*). Este resultado es interesante, ya que

se demuestra que CMGN y/o variantes están presentes en diferentes homínidos y que aún no ha sido caracterizada y registrada.

Sequences producing significant alignments		Download	Select columns	Show	100			
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> LOW QUALITY PROTEIN: uncharacterized protein LOC100982282 [Pan paniscus]	Pan paniscus	296	296	83%	3e-97	82.35%	301	XP_024782396.1
<input checked="" type="checkbox"/> LOW QUALITY PROTEIN: uncharacterized protein LOC100609279 [Pan troglodytes]	Pan troglodytes	296	296	83%	3e-95	82.35%	416	XP_054529925.1
<input checked="" type="checkbox"/> uncharacterized protein LOC101127504 [Gorilla gorilla gorilla]	Gorilla gorilla gorilla	270	270	75%	3e-88	82.61%	195	XP_018872123.2
<input checked="" type="checkbox"/> uncharacterized protein LOC101176683 [Nomascus leucogenys]	Nomascus leucogenys	251	251	88%	3e-78	72.09%	376	XP_012364727.2
<input checked="" type="checkbox"/> uncharacterized protein LOC129460117 [Symphalangus syndactylus]	Symphalangus syndactylus	236	236	88%	5e-74	69.77%	278	XP_055093711.1

Entry	Entry Name	Protein Names	Organism	Length	Alignment
<input type="checkbox"/> B0QYK4	B0QYK4_HUMAN	EMI domain containing 1	Homo sapiens (Human)	422 AA	50% 121 [6e-7]
<input type="checkbox"/> P02458-1	CO2A1_HUMAN	Isoform 1 of Collagen alpha-1(II) chain[...]	Homo sapiens (Human)	1,418 AA	37.6% 115 [0.0000047]
<input type="checkbox"/> P27658	CO8A1_HUMAN	Collagen alpha-1(VIII) chain[...]	Homo sapiens (Human)	744 AA	32.6% 109 [0.000024]
<input type="checkbox"/> Q92838-9	EDA_HUMAN	Isoform 8 of Ectodysplasin-A[...]	Homo sapiens (Human)	386 AA	29.5% 105 [0.000062]

Figura 17. Resultados del BLAST obtenidos de las plataformas NCBI Blast:Protein Sequence (nih.gov) y UnitProt Blast (www.uniprot.org).

DISCUSIÓN

La regeneración de hueso alveolar, cemento y ligamento periodontal ha sido el principal objetivo de la terapia periodontal. Sin embargo, existen pocos estudios sobre los mecanismos celulares y moleculares que pueden desencadenar su regeneración, debido a diversas limitaciones como son la poca disponibilidad de tejido y el laborioso procedimiento para su obtención y para su procesamiento y cultivo *in vitro*.⁴⁴ Actualmente, la expresión recombinante de proteínas en células *E. coli* ha demostrado ser una opción rápida y efectiva de obtener mayor cantidad de proteína de interés. Debido a su larga historia de uso en biotecnología, *Escherichia coli* se ha convertido en el organismo de primera elección para la expresión de proteínas recombinantes. Sin embargo, la producción de proteínas sigue siendo una tarea desafiante y las condiciones óptimas de cultivo son difíciles de establecer para grandes conjuntos de muestras. Las condiciones que funcionan para un producto genético podrían no necesariamente trabajar para otro.⁵¹

En el presente estudio comprobamos la efectividad de esta herramienta biotecnológica, ya que se obtuvo un rendimiento de 1.7mg de hrCMGN purificada por cada litro de cultivo. Este rendimiento es mayor al rendimiento obtenido por Romo 2016 en la purificación de hrCEMP-1 expresada en la levadura *Pichia pastoris* (0.940mg/litro)²⁸ y menor al obtenido por Nidome 2018 de hrHACD/CAP expresada en *E. coli* BL21 (2.89mg/litro del monómero y 3.14mg/litro del dímero).⁴⁵ Es relevante señalar que se ha logrado generar un sistema de expresión recombinante y un protocolo de purificación para hrCMGN, que es similar a lo obtenido en la purificación de las otras proteínas cemento específicas.

Para establecer el protocolo de purificación de hrCMGN, se utilizó la afinidad de la bandera de histidinas al níquel y la propiedad que comparten las proteínas cemento específicas que es su alta afinidad a la hidroxiapatita.⁵² La purificación de hrCMGN se realizó mediante tres pasos de cromatografía y estos demostraron ser efectivos. En el primer paso de cromatografía (captura) se empleó la afinidad al níquel, se demostró que la bandera de histidinas fue útil para la captura de hrCMGN y para su correcta purificación. Para la etapa intermedia de cromatografía se pudo confirmar la afinidad de hrCMGN a la

hidroxiapatita, como se ha comprobado en otros estudios de proteínas HACD1/CAP y CEMP-1 (Montoya et al 2015; Romo et al, 2016; Nidome 2018), donde se inyectaron 50 µg de CEMP-1 a una columna de hidroxiapatita y se eluyó a una concentración de 300mM de fosfato de sodio. Proteínas no relacionadas al proceso de biomineralización eluyen a bajas concentraciones de fosfato de sodio (entre 10mM y 120mM), y las proteínas relacionadas a procesos de biomineralización eluyen a concentraciones de hasta 200mM de fosfato de sodio, lo que demuestra que CEMP1 posee alta afinidad a la hidroxiapatita debido a la interacción de los dominios cargados positiva y negativamente.⁵³ En este estudio se comprobó que para realizar el pulido de la purificación (tercer paso de cromatografía) el uso de la columna de hidroxiapatita, permitió separar fracciones débilmente unidas a la columna.

Una herramienta importante y de mucha utilidad en el proceso de caracterización de proteínas y que establece, generalmente, la primera herramienta de análisis es el estudio *in silico*. Esta herramienta nos permite predecir y analizar propiedades de las proteínas utilizando únicamente la secuencia de aminoácidos y nos puede ayudar a predecir el comportamiento químico de la proteína de trabajo, así como el posible protocolo de inicio del proceso de purificación.^{47, 48} De acuerdo con los resultados obtenidos en el análisis *in silico*, CMGN tiene un punto isoeléctrico de 11.55, lo que la hace una proteína con carga positiva a pH fisiológico. Comparando hrCMGN con las otras proteínas cemento específicas (CEMP-1 y HACD1/CAP), podemos observar que CEMP-1 tiene un pI de 9.73 (proteína con carga positiva a pH fisiológico) y HACD1/CAP tiene un pI de 7.82 (muy cercano al pH fisiológico), se ha descrito que las proteínas cemento específicas a pesar de no ser ácidas, interactúan fuertemente con la hidroxiapatita, se sugiere que esto es debido a sus implicaciones biológicas durante el proceso de biomineralización.

Para realizar la identificación de hrCMGN se empleó la inmunodetección por Western Blot mediante el anticuerpo monoclonal anti 6xHis. El resultado del western blot es contundente y se reconoce con alta especificidad a la banda purificada y esta tiene una movilidad relativa con el peso molecular calculado para hrCMGN.

Como resultado del BLAST se obtuvo la homología con la cadena alfa 1 de la colágena II, que se expresa en la mayoría de los tejidos conectivos y juega un rol importante en el desarrollo embrional del esqueleto y para el correcto desarrollo del cartílago para soportar fuerzas compresivas.⁴⁹ También mostró homología con la cadena alfa 1 de la colágena VIII, que es un componente macromolecular del endotelio, necesaria para la migración y proliferación de las células del músculo liso vascular, además de mantener la integridad estructural de vasos sanguíneos. Por último, la ectodisplastina A mostró homología con nuestra secuencia, y esta es una citocina involucrada en la señalización durante la morfogénesis de los órganos derivados del ectodermo.⁵⁰

Por otra parte, la homología de hrCMGN con proteínas no caracterizadas presentes en algunos miembros de la familia de los homínidos, puede relacionarse con la presencia de hrCMGN como una proteína que puede estar en el cemento radicular pero que no ha sido descrita o estudiada, por lo que nuestro trabajo toma relevancia como el primer estudio enfocado en determinar si se trata de una proteína cemento específica.

El análisis de estructura secundaria mostró que hrCMGN es una proteína altamente desorganizada (59% random coil), seguido de estructura alfa hélice (40%) mostrando gran similitud a la estructura secundaria de HACD1/CAP (45.9% random coil), mientras que CEMP1 está conformada principalmente de random coil (62.5%). Las proteínas intrínsecamente desordenadas están compuestas principalmente de regiones aleatorias, con frecuencia son hidrofílicas y están fosforiladas o glicosiladas, han demostrado ser multifuncionales y poseen propiedades de unión fuerte a la hidroxiapatita (tal es el caso de las proteínas SIBILING (small integrin-binding ligand, N-linked glycoprotein).^{28, 29, 30}

En este estudio se reveló que hrCMGN comparte algunas características antes mencionadas: es una proteína altamente fosforilada, presenta una alta afinidad a hidroxiapatita y su estructura es principalmente aleatoria, por lo que podemos asociarla funcionalmente a las proteínas sibilin.

CONCLUSIÓN

En el presente estudio se propone a CMGN como una posible proteína cemento específica que no ha sido descrita y que de acuerdo con los resultados obtenidos en relación a su similitud con HACD1/CAP y CEMP, proteínas cemento específicas ya descritas, sería de gran importancia continuar con su estudio para profundizar y mejorar el entendimiento de la biología del cemento radicular y en un futuro, aplicar dichos conocimientos en el desarrollo de nuevas terapéuticas enfocadas en la regeneración periodontal.

Logramos establecer un protocolo para su producción recombinante y purificación mediante un sistema eficaz y confiable (sin depender de su fuente nativa), se asentó un precedente para futuras investigaciones, que además ampliará el conocimiento molecular del cemento radicular.

REFERENCIAS

1. Rosano GL, Morales ES, Ceccarelli EA. New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Sci.* 2019 Aug;28(8):1412-1422.
2. Hayat SMG, Farahani N, Golichenari B, Sahebkar A. Recombinant Protein Expression in *Escherichia coli* (*E. coli*): What We Need to Know. *Curr Pharm Des.* 2018;24(6):718-725.
3. Gardner DG, Baxter JD. Recombinant DNA. *Am J Med.* 1982 Apr;72(4):551-3.
4. Ferrer-Miralles, N., Saccardo, P., Corchero, J. L., Xu, Z., & García-Fruitós, E. (2014). General Introduction: Recombinant Protein Production and Purification of Insoluble Proteins. *Insoluble Proteins*, 1–24.
5. Correa, R., Arenas, J., Montoya, G., Hoz, L., López, S., Salgado, F., Arroyo, R., Salmeron, N., Romo, E., Zeichner-David, M. and Arzate, H. (2019), Synthetic cementum protein 1–derived peptide regulates mineralization *in vitro* and promotes bone regeneration *in vivo*. *The FASEB Journal*, 33: 1167-1178.
6. Montoya, G., Arenas, J., Romo, E., Zeichner-David, M., Alvarez, M., Narayanan, A. S., Arzate, H. (2014). Human recombinant cementum attachment protein (hrPTPLa/CAP) promotes hydroxyapatite crystal formation *in vitro* and bone healing *in vivo*. *Bone*, 69, 154–164.
7. Serrano J, Romo E, Bermúdez M, Narayanan AS, Zeichner-David M, Santos L, Arzate H. Bone regeneration in rat cranium critical-size defects induced by Cementum Protein 1 (CEMP1).
8. Bosshardt DD, Selvig KA. Dental cementum: the dynamic tissue covering of the root. *Periodontol* 2000. 1997 Feb;13:41-75.
9. SAYGIN, N.E., GIANNOBILE, W.V. and SOMERMAN, M.J. (2000), Molecular and cell biology of cementum. *Periodontology* 2000, 24: 73-98.
10. Ivanovski S, Komaki M, Bartold PM, Narayanan AS. Periodontal-derived cells attach to cementum attachment protein via $\alpha 5 \beta 1$ integrin. *J Period Res* 1999; 34: 154-159
11. Arzate H, Portilla-Robertson J, Aguilar-Mendoza ME. Recombination of epithelial root sheath and dental papilla cells *in vitro*. *Arch Med Res* 1996; 27: 573-577

12. Valdés De Hoyos A, Hoz-Rodríguez L, Arzate H, Narayanan AS. Isolation of Protein-Tyrosine Phosphatase-like Member-a Variant from Cementum. *Journal of Dental Research*. 2012;91(2):203-209.
13. Ivarez-Pérez MA, Pitaru S, Álvarez-Fregoso O, Reyes Gasga J, Arzate H. Anti-cementoblastoma-derived protein antibody partially inhibits mineralization on a cementoblastic cell line. *J Struct Biol* 2003; 143: 1-13.
14. Hoz L, Romo E, Zeichner-David M, Sanz M, Nuñez J, Gaitán L, Mercado G, Arzate H. Cementum protein 1 (CEMP1) induces differentiation by human periodontal ligament cells under three-dimensional culture conditions. *Cell Biol Int*. 2012 Feb;36(2):129-36.
15. Arzate, H., Zeichner-David, M., & Mercado-Celis, G. (2014). Cementum proteins: role in cementogenesis, biomineralization, periodontium formation and regeneration. *Periodontology* 2000, 67(1), 211–233.
16. HASSELL, T. M. (1993). Tissues and cells of the periodontium. *Periodontology* 2000, 3(1), 9–38. doi:10.1111/j.1600-0757.1993.
17. CHO, M.-I., & GARANT, P. R. (2000). Development and general structure of the periodontium. *Periodontology* 2000, 24(1), 9–27.
18. Narayanan, A. S., & Bartold, P. M. (1996). Biochemistry of Periodontal Connective Tissues and their Regeneration: A Current Perspective. *Connective Tissue Research*, 34(3), 191–201.
19. Iwata, T., Yamato, M., Ishikawa, I., Ando, T. and Okano, T. (2014), Tissue Engineering in Periodontal Tissue. *Anat. Rec.*, 297: 16-25
20. Arzate, H., Olson, S. W., Page, R. C., Gown, A. M., & Narayanan, A. S. (1992). Production of a monoclonal antibody to an attachment protein derived from human cementum. *The FASEB Journal*, 6(11), 2990–2995.
21. Diekwisch TG. The developmental biology of cementum. *Int J Dev Biol*. 2001 Sep;45(5-6):695-706.
22. Gottlieb B. Biology of the cementum. *J Periodontol* 1942: 13: 13–19.
23. Gottlieb B. The new concept of periodontoclasia. *J Periodontol* 1946: 17: 7.
24. Villareal-Ramírez E, Moreno A, Mas-Oliva J, Chávez-Pacheco J, Narayanan AS, Gil-Chavarría I, Zeichner-David M, Arzate H. Characterization of recombinant human cementum protein 1 (hrCEMP1): Primary role in biomineralization. *BBRC* 2009; 384: 49-54

25. Carmona B, Alvarez M, Narayanan AS, Zeichner-David M, Reyes J, Molina J, Garcia AL, Suarez JL, Chavarria IG, Villarreal E, Arzate H. Human Cementum Protein 1 induces expression of bone and cementum proteins by human gingival fibroblasts. *Biochem Biophys Res Commun* 2007; 358: 763–769.
26. Park JB, Matsuura M, Han KY, Norderyd O, Lin WL, Genco RJ, Cho MI. Periodontal regeneration in class III furcation defects of beagle dogs using guided tissue regenerative therapy with platelet-derived growth factor. *J Periodontol* 1995; 66: 462–477
27. Paula-Silva FW, Ghosh A, Arzate H, Kapila S, da Silva LA, Kapila YL. Calcium hydroxide promotes cementogenesis and induces cementoblastic differentiation of mesenchymal periodontal ligament cells in a CEMP1- and ERK-dependent manner. *Calcif Tissue Int* 2010; 87: 144– 157.
28. Romo Arévalo, Enrique sustentante Papel inductor de la proteína del cemento 1 recombinante humana (hrCEMP1) en el proceso de mineralización *in vitro* / 2010. TESIUNAM
29. Forman-Kay JD, Mittag T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* 2013;21:1492–9
30. He G, Dahl T, Veis A, George A. Nucleation of apatite crystals *in vitro* by self-assembled dentin matrix protein 1. *Nat Mater* 2003;2:552–8
31. Weiner S, Addadi L. Acidic macromolecules of mineralized tissues: the controllers of crystal formation. *Trends Biochem Sci* 1991;16:252–6.
32. Delak K, Harcup C, Lakshminarayanan R, Sun Z, Fan Y, Moradian-Oldak J, et al. The tooth enamel protein, porcine amelogenin, is an intrinsically disordered protein with an extended molecular configuration in the monomeric form. *Biochemistry* 2009;48:2272–81.
33. Peysselon F, Xue B, Uversky VN, Ricard-Blum S. Intrinsic disorder of the extracellular matrix. *Mol Biosyst* 2011;7:3353–65.
34. Olson S, Arzate H, Narayanan AS, Page RC. Cell attachment activity of cementum proteins and mechanism of endotoxin inhibition. *J Dent Res* 1991; 70: 1272–1277.

35. Komaki M, Iwasaki K, Arzate H, Narayanan AS, Izumi Y, Morita I. Cementum protein 1 (CEMP1) induces a cementoblastic phenotype and reduces osteoblastic differentiation in periodontal ligament cells. *J Cell Physiol* 2012;227: 649–57
36. Liu HW, Yacobi R, Savion N, Narayanan AS, Pitaru S. A collagenous cementum-derived attachment protein is a marker for progenitors of the mineralized tissue-forming cell lineage of the periodontal ligament. *J Bone Miner Res* 1997; 12: 1691–1699
37. Alvarez-Pérez, M. A., Narayanan, S., Zeichner-David, M., Rodríguez Carmona, B., & Arzate, H. (2006). Molecular cloning, expression and immunolocalization of a novel human cementum-derived protein (CP-23). *Bone*, 38(3), 409–419.
38. Arzate H, Jiménez-García LF, Álvarez-Pérez MA, Landa A, Bar-Kana I, Pitaru S. Immunolocalization of a Human Cementoblastoma-conditioned Medium-derived Protein. *Journal of Dental Research*. 2002;81(8):541-546.
39. Maldonado, S., Romo, E., Serrano, J., Pérez, A., Guerra, C., Zeichner-David, M., Mercado, G. and Arzate, H. (2014), Cementum protein 1 (CEMP1) activates p38 and JNK during the mineralisation process by cementoblast-like cells *in vitro*. *Cell Biol Int Rep*, 21: 8-16.
40. Romo, E., Arzate, H., Montoya, G. Rodríguez, Adela. (2016). “High-level expression and characterization of a glycosylated human cementum Protein 1 with lectin activity” *FEBS Lett.*, 590 (1), pp. 129- 138.
41. Pitaru, S., Narayanan, S.A., Olson, S., Savion, N., Hekmati, H., Alt, I. and Metzger, Z. (1995), Specific cementum attachment protein enhances selectively the attachment and migration of periodontal cells to root surfaces. *Journal of Periodontal Research*, 30: 360-368.
42. Wain HM, Bruford EA, Lovering RC, Lush MJ, Wright MW, Povey S. Guidelines for human gene nomenclature. *Genomics*. 2002 Apr;79(4):464-70.
43. Arzate H, Chimal-Monroy J, Hernández-Lagunas L, Díaz de León L. Human cementum protein extract promotes chondrogenesis and mineralization in mesenchymal cells. *J Periodontal Res*. 1996 Feb;31(2):144-8.

44. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol.* 2014 Apr 17;5:172
45. Nidome Campos, Mikado Alejandro, sustentante. Estudio de plegamiento y estabilidad estructural de la proteína de adhesión del cemento radicular (HACD1/CAP) / 2018. TesiUNAM
46. Savojardo, C., Babbi, G., & Martelli, P. L. (2020). Computational methods for protein function prediction: a practical overview. *Current Opinion in Structural Biology*, 62, 161-168.
47. Pitaru S, Savion N, Hekmati H, Olson S, Narayanan AS. Binding of a cementum attachment protein to extracellular matrix components and to dental surfaces. *J Periodontal Res* 1992; 27: 640–6
48. Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.
49. Manon-Jensen T, Kjeld NG, Karsdal MA. Collagen-mediated hemostasis. *J Thromb Haemost.* 2016 Mar;14(3):438-48.
50. Kere J, Srivastava AK, Montonen O, Zonana J, Thomas N, Ferguson B, Munoz F, Morgan D, Clarke A, Baybayan P, Chen EY, Ezer S, Saarialho-Kere U, de la Chapelle A, Schlessinger D. X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nat Genet.* 1996 Aug;13(4):409-16. doi: 10.1038/ng0895-409.
51. Stephen F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schäffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", *FEBS J.* 272:5101-5109.
52. Carmignotto, G. P., & Azzoni, A. R. (2019). On the expression of recombinant Cas9 protein in *E. coli* BL21 (DE3) and BL21 (DE3) Rosetta strains. *Journal of biotechnology*, 306, 62-70.
53. Tiselius, A. (1956). "Protein chromatography on calcium phosphate columns." *Arch Biochem Biophys*, 65, pp. 132-155.

54. Wutteke M, Müller S, Nitsche DP, Paulsson M, Hanish FG, Maurer P. Structural characterization of human recombinant and bone derived sialoprotein. Functional implications for cell attachment and hydroxyapatite binding. *J Biol Chem* 276, 36839-48 (2001).