



CENTRO DE INVESTIGACIONES BIOLÓGICAS
DEL NOROESTE, S.C.

Programa de Estudios de Posgrado

ANÁLISIS COMPARATIVO DE LAS PROTEÍNAS INSOLUBLES
INVOLUCRADAS EN EL PROCESO DE FORMACIÓN DE LA
CONCHA DE GASTERÓPODOS.

T E S I S

Que para obtener el grado de

Maestro en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales
(Orientación en Biotecnología)

P r e s e n t a

Daniela Guadalupe Montealegre Mariscal

La Paz, Baja California Sur, agosto de 2021.

ACTA DE LIBERACIÓN DE TESIS

En la Ciudad de La Paz, B. C. S., siendo las 11:00 horas del día 13 del Mes de Agosto del 2021, se procedió por los abajo firmantes, miembros de la Comisión Revisora de Tesis avalada por la Dirección de Estudios de Posgrado y Formación de Recursos Humanos de Centro de Investigaciones Biológicas del Noroeste, S. C., a liberar la Tesis de Grado titulada:

“ANÁLISIS COMPARATIVO DE LAS PROTEÍNAS INSOLUBLES INVOLUCRADAS EN EL PROCESO DE FORMACIÓN DE LA CONCHA DE GASTERÓPODOS”

Presentada por el alumno:

Daniela Guadalupe Montealegre Mariscal

Aspirante al Grado de MAESTRO EN CIENCIAS EN EL USO, MANEJO Y PRESERVACIÓN DE LOS RECURSOS NATURALES CON ORIENTACIÓN EN **BIOTECNOLOGÍA**

Después de intercambiar opiniones los miembros de la Comisión manifestaron su **APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA



Dra. Norma Yolanda Hernández Saavedra
Co-Directora de Tesis



Dra. Crisalejandra Rivera Pérez
Co-Directora de Tesis



Guy Salvesen Ph.D.
Co-Tutor de Tesis



Dra. Gracia Alicia Gómez Anduro,
Directora de Estudios de Posgrado y
Formación de Recursos Humanos.

La Paz, Baja California Sur, a 12 de agosto de 2021.

Los miembros del comité de tesis de la estudiante Daniela Guadalupe Montealegre Mariscal del Programa de Maestría en Ciencias en Uso, Manejo y Preservación de los Recursos Naturales revisamos el contenido de la tesis y otorgamos el VoBo dado que la tesis no representa un plagio de otro documento como lo muestra el reporte de similitud realizado:

- Herramienta antiplagio: iThenticate
 - Filtros utilizados: Citas y Bibliografía
 - Porcentaje de similitud: 5% de similitud
- Se muestra captura de pantalla

Análisis comparativo de las proteínas insolubles involucradas en el proceso de formación de la concha de gasterópodos

ORIGINALITY REPORT

5%

SIMILARITY INDEX

PRIMARY SOURCES

1	hdl.handle.net Internet	59 words — < 1%
2	repositorio.uchile.cl Internet	48 words — < 1%
3	researchonline.jcu.edu.au Internet	39 words — < 1%
4	www.frontiersin.org Internet	29 words — < 1%
5	idus.us.es Internet	26 words — < 1%
6	www.jove.com Internet	26 words — < 1%
7	patents.google.com Internet	25 words — < 1%
8	bdigital.unal.edu.co Internet	20 words — < 1%
9	ocean.kisti.re.kr Internet	20 words — < 1%

Firmas del comité



Dra. Norma Y. Hernández Saavedra
Co-Directora



Dra. Crisalejandra Rivera Pérez
Co-Directora



Guy Salvesen Ph.D.
Co-Tutor



Conformación de Comités

Comité Tutorial

Dra. Norma Yolanda Hernández Saavedra
Centro de Investigaciones Biológicas del Noroeste S.C.
Co-Directora de Tesis

Dra. Crisalejandra Rivera Pérez
CONACyT-Centro de Investigaciones Biológicas del Noroeste S.C.
Co-Directora de Tesis

Ph.D. Guy Salvesen
Sanford Burnham Prebys Medical Discovery Institute
Co-Tutor de Tesis

Comité Revisor de Tesis

Dra. Norma Yolanda Hernández Saavedra
Dra. Crisalejandra Rivera Pérez
Ph.D. Guy Salvesen

Jurado de Examen

Dra. Norma Yolanda Hernández Saavedra
Dra. Crisalejandra Rivera Pérez
Ph.D. Guy Salvesen

Suplente(s)

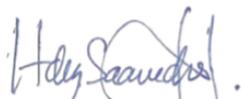
Dra. Elizabeth Monreal Escalante

Resumen

La concha de los moluscos gasterópodos se forma por biomineralización, proceso en el cual moléculas de carbonato de calcio forman cristales polimorfos como calcita y aragonita, los cuales formarán microestructuras como la prismática y la nacarada, respectivamente. Este proceso es controlado por proteínas secretadas por células epiteliales del manto de moluscos denominadas como proteínas de la matriz orgánica (SMP, por sus siglas en inglés). A través de distintas técnicas moleculares, se han aislado y caracterizado proteínas tanto solubles como insolubles, en ácido acético o EDTA, con el objetivo de entender su función en la formación de la concha; sin embargo, estos trabajos han sido enfocados principalmente en bivalvos. Por otra parte, existe un principal interés en la microestructura nacarada, esto debido a sus propiedades mecánicas como la resistencia, ya que su entendimiento podría ser la base para el desarrollo de nuevos biomateriales. En el presente trabajo se determinaron las similitudes y diferencias entre las proteínas insolubles relacionadas a la formación de la concha en gasterópodos, para ello se realizó una base de datos de proteínas insolubles (en su mayoría no caracterizadas) aisladas de las conchas de cuatro organismos gasterópodos: *Haliotis rufescens*, *Haliotis asinina*, *Haliotis laevigata* y *Lottia gigantea*; obtenidas al realizar una revisión bibliográfica y la búsqueda en bases de datos públicas. Así mismo, se realizó un análisis comparativo de las características bioquímicas y funcionales de estas proteínas, utilizando herramientas bioinformáticas. Las proteínas se analizaron y clasificaron de acuerdo a la microestructura, el punto isoeléctrico, los aminoácidos dominantes, modificaciones postraduccionales y dominios funcionales. Se encontró que estas características tienen implicaciones tanto directas como indirectas en el proceso de biomineralización de la concha, tales como la regulación a nivel estructural, funcional y de actividad. Además, se realizó el alineamiento múltiple de las secuencias con el fin de encontrar regiones conservadas que pudieran contribuir con la caracterización de las proteínas. A partir de este análisis, se formaron grupos de proteínas con funciones y dominios homólogos, donde se encontró que las proteínas insolubles aisladas de la concha de gasterópodos no tienen funciones exclusivas de la formación de cristales, sino que también pueden estar implicadas en la protección de la concha, inmunidad, señalización, suministro de precursores, entre otras. Los resultados obtenidos contribuyen al entendimiento del proceso de biomineralización en gasterópodos, e incluso a la identificación de proteínas que no tienen una función asignada.

Palabras clave: biomineralización, gasterópodos, proteínas insolubles, bioinformática.

ORCID: 0000-0002-9582-1594



Vo. Bo.
Dra. Norma Y. Hernández Saavedra
Co-Directora de tesis



Vo. Bo.
Dra. Crisalejandra Rivera Pérez
Co-Directora de tesis

Summary

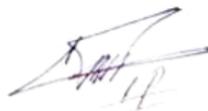
Mollusk shells are formed by biomineralization, a process in which calcium carbonate molecules form polymorphic crystals such as calcite and aragonite, which then form prismatic and nacre microstructures, respectively. This process is controlled by proteins secreted by epithelial cells of the shell of mollusks, called shell matrix proteins (SMP). Through different molecular techniques, both soluble and insoluble proteins, in acetic acid or EDTA, have been isolated and characterized to understand their function in shell formation; however, these works have been mainly focused on bivalves. Moreover, there is a special interest in nacre microstructure, given by its mechanical properties such as resistance since its understanding could be the basis for the development of new biomaterials. In the present work, the similarities and differences between the insoluble proteins related to shell formation in gastropods were determined, for which a database of insoluble proteins (mostly uncharacterized) isolated from the shells of four gastropod organisms was made: *Haliotis rufescens*, *Haliotis asinina*, *Haliotis laevigata*, and *Lottia gigantea*. Insoluble proteins were obtained through bibliographic review and public databases search. Likewise, a comparative analysis of the biochemical and functional characteristics of these proteins was carried out, using bioinformatic tools. Proteins were analyzed and classified according to microstructure, isoelectric point, dominant amino acids, post-translational modifications, and functional domains. These characteristics were found to have both direct and indirect implications in the biomineralization process of the shell, such as regulation at the structural, functional, and activity levels. In addition, multiple sequence alignment was performed to find conserved regions that could contribute to the characterization of the proteins. From this analysis, groups of proteins with homologous functions and domains were formed, where it was found that the insoluble proteins isolated from the gastropod shell do not have exclusive functions of crystal formation but may also be involved in the protection of the shell, immunity, signaling, precursor supply, among others. The results obtained contribute to the understanding of the biomineralization process in gastropods, and even to the identification of proteins that do not have an assigned function.

Keywords: biomineralization, gastropods, insoluble proteins, bioinformatics.

ORCID: 0000-0002-9582-1594



Vo. Bo.
Dra. Norma Y. Hernández Saavedra
Co-Directora de tesis



Vo. Bo.
Dra. Crisalejandra Rivera Pérez
Co-Directora de tesis

Dedicatoria

A mis padres, mi motor, mi mayor ejemplo de amor, esfuerzo y dedicación.

Sin ustedes nada de esto hubiera sido posible.

Los amo con toda mi alma. Este logro es suyo.

Gracias por tanto y por todo.

Agradecimientos

Al Centro de Investigaciones Biológicas del Noroeste S.C. por permitirme realizar mi Tesis de Maestría como parte de su institución.

A CONACyT por otorgar la beca de manutención (No. de registro **743329**).

A mi Comité Tutorial por guiarme e instruirme en el desarrollo y elaboración de la tesis. Gracias por apoyarme y acompañarme en todo momento a pesar de la pandemia y la distancia, gracias por ayudarme y contribuir a mi crecimiento profesional. A la Dra. Crisalejandra por compartir su conocimiento, por la atención y por la paciencia. A la Dra. Norma por sus correcciones enriquecedoras. Al Dr. Salvesen por la valiosa retroalimentación durante todas las evaluaciones de comité.

Al Lic. Horacio Sandoval por todo el apoyo técnico brindado para la realización de las clases, los seminarios, presentaciones y todo lo relacionado con los servicios en línea.

A mi terapeuta, la Psicóloga Deyanira Huerta, por todo su apoyo, su comprensión, por escucharme en esta etapa tan importante en mi vida. Gracias por ayudarme a crear una mejor versión de mí.

A mis compañeros, amigos y futuros maestros: Ayenia, Aldahir, Bárbara, David, Giovanni, Jaqueline, Miranda, Patricia y Kevin. Gracias por su amistad, por sus ánimos, por sus consejos, por hacer mis días de maestría más divertidos. ¡Los quiero, los aprecio mucho! A mis hermanas de maestría Briana e Ivis, con las cuales compartí directoras de tesis. Gracias por estar presente siempre, por los ánimos, por el apoyo, por los consejos y por su amistad. A mis mejores amigas, Eliana y Ximena, que siempre estuvieron conmigo a la distancia. ¡Las quiero mucho y las extraño!

A mi familia por acompañarme a lo largo del camino, por motivarme a seguir adelante y querer llegar cada vez más lejos. A mis padres por darme el amor, la confianza, y la determinación para lograr esta meta. Los amo y les agradezco infinitamente todo lo que hacen por mí. A mis hermanos, por apoyarme, por cuidarme, por mandarme sus buenos deseos. Los amo. A mis abuelos, por mis raíces, por su cariño, sus enseñanzas de vida y por su ejemplo de disciplina, entrega, amor, valentía, resiliencia, fortaleza. Los amo y los extraño.

Contenido

Resumen	i
Summary	ii
Dedicatoria	iii
Agradecimientos	iv
Contenido	v
Lista de figuras	vii
Lista de tablas	viii
Abreviaturas	ix
1. INTRODUCCIÓN	1
2. ANTECEDENTES	5
2.1 Biomineralización	5
2.2 Anatomía de moluscos.....	6
2.3 Composición de las conchas.....	7
2.3.1 Importancia del nácar	8
2.4 Proteínas involucradas en el proceso de Biomineralización.....	8
2.5 Caracterización de proteínas.....	9
2.6 Fraccionamiento de proteínas	10
2.7 Funcionalidad de proteínas.....	13
2.7.1 Dominios	14
2.7.2 Dominios o regiones de baja complejidad.....	15
3. JUSTIFICACIÓN	17
4. HIPÓTESIS	18
5. OBJETIVOS	19
5.1 Objetivo general.....	19
5.2 Objetivos particulares	19
6. MATERIAL Y MÉTODOS	20
6.1 Base de datos de proteínas insolubles a partir de conchas de gasterópodos.....	20
6.1.1 Búsqueda de proteínas	20
6.2 Características de las proteínas insolubles	20
6.2.1 Caracterización de proteínas.....	21
6.3 Comparación de características de las proteínas insolubles de la base de datos generada para gasterópodos.	22
6.4 Identificación de grupos de proteínas usando un árbol guía.....	22
7. RESULTADOS	24
7.1 Base de datos de proteínas insolubles obtenidas de conchas de gasterópodos.....	24
7.1.1 Obtención de proteínas.....	24
7.2 Caracterización de proteínas.....	25
7.2.1 Microestructura	25
7.2.2 Características bioquímicas.....	26
7.2.3 Análisis de contenido de aminoácidos (dominantes)	27
7.2.4 Dominios	29
7.2.5 Motivos y regiones de baja complejidad	31

7.2.6 Modificaciones postraduccionales.....	31
7.2.7 Función.....	33
7.3 Grupos de proteínas.....	35
8. DISCUSIÓN	37
8.1 Proteínas descritas en gasterópodos	37
8.2 Implicaciones de las modificaciones postraduccionales en SMPs	38
8.3 Implicaciones de la carga de las SMPs	41
8.4 Implicaciones de la abundancia de ciertos aminoácidos	42
8.5 Implicaciones de las regiones de baja complejidad	43
8.6 Implicaciones de los dominios en la formación de la concha	44
8.7 Grupos de proteínas.....	46
9. CONCLUSIONES	49
10. LITERATURA CITADA	50

Lista de figuras

Figura 1. Distribución de las proteínas entre las diferentes fracciones de la concha. Diagrama de Venn mostrando las proteínas asociadas a la capa nacarada, prismática, ambas capas o sin microestructura definida.....	26
Figura 2. Distribución de proteínas de acuerdo con límites establecidos para cada característica. Las barras corresponden a las características de las proteínas. El degradado de color en cada una representan los parámetros establecidos.....	27
Figura 3. Aminoácidos dominantes en las secuencias de proteínas. Representación gráfica de la cantidad de proteínas que presentan uno o más de estos siete aminoácidos dominantes.	28
Figura 4. Representación gráfica de la cantidad de proteínas que presentan uno o más aminoácidos dominantes de acuerdo con la microestructura identificada.	29
Figura 5. Representación gráfica de la relación entre los dominios de mayor presencia en las proteínas de la base de datos (de un total de 554) y el número de proteínas en las que fueron identificados.	30
Figura 6. Representación gráfica de la relación entre los dominios de menor presencia en las proteínas de la base de datos (de un total de 554) y el número de proteínas en las que fueron identificados. En el gráfico se muestran únicamente 10 de los 32 dominios de esta clasificación.	31
Figura 7. Identificación de proteínas insolubles con potencial función enzimática. Porcentaje de proteínas que tienen potencial actividad catalítica. Los colores representan la clasificación de enzimas.....	34
Figura 8. Identificación de proteínas con funciones no catalíticas. Clasificación de proteínas de acuerdo con su posible función con su respectiva cantidad de proteínas correspondiente a cada uno.....	35
Figura 9. Grupos de proteínas. Se muestra la cantidad de proteínas pertenecientes a cada uno de los grupos obtenidos.	36

Lista de tablas

Tabla 1. Proteínas de moluscos gasterópodos aisladas de la concha en los últimos 20 años.	11
Tabla 2. Dominios presentes en proteínas aisladas de gasterópodos y sus funciones asociadas.	15
Tabla 3. Herramientas utilizadas para la caracterización de secuencias.....	21
Tabla 4. Parámetros determinados para cada característica.	22
Tabla 5. Información de búsqueda.	24
Tabla 6. Proteínas por especie.....	25
Tabla 7. Potenciales modificaciones postraduccionales de los grupos de proteínas con dominios de mayor presencia.	32
Tabla 8. Potenciales modificaciones postraduccionales de los dominios de menor presencia.	33

Abreviaturas

- SMP:** Shell Matrix Protein
- EDTA:** Ácido etilendiaminotetraacético
- EGF:** Epidermal Growth Factor
- BSMP:** Blue Mussel Shell Protein
- IGF:** Insulin-like Growth Factor
- WAP:** Whey Acidic Protein
- CA:** Carbonic Anhydrase
- VWA:** von Willebrand factor A
- ZP:** Zona Pelucida
- ECM:** Extracellular matrices
- CaCO₃:** Carbonato de calcio
- HCO₃⁻:** Ion Bicarbonato
- LCR:** Low Complexity Regions
- LCD:** Low Complexity Domains
- NCBI:** National Center for Biotechnology Information
- pI:** Punto Isoeléctrico
- MW:** Molecular Weight
- kDa:** kilo Daltons
- A2M:** alpha-2-macroglobulin
- RAB:** Rabhifilin
- SCP:** Secreted Cysteine Rich Protein
- WR1:** Worm-specific Repeat 1
- IF Actin:** Intermediate Filament Actin
- ZnMc:** Zinc Metalloproteinase
- BPTI:** Bovine Pancreatic Trypsin Inhibitor
- GOLD:** Golgi Dynamics
- SUSHI CCP:** Complement Control Protein
- WD40:** Tryptophan (W) Aspartic Acid (D)

SH3/SH2: Src homology

SEA: Sperm protein, Enterokinase and Agrin

eIF5C: Eukaryotic Initiation Factors

PHB: Prohibitin Homologues

RAS: Rat Sarcoma virus

FABP: Fatty Acid Binding Protein

ARF: ADP Ribosylation Factor

ADF-H: Actin Depolymerization Factor Homology

PKS-ER: Polyketide Synthase Enoyl Reductase

KOW: Kyprides, Ouzounis, Woese

SCP: Secreted Cysteine Rich Protein

EC: Enzyme Commission

ADN: Ácido Desoxirribonucleico

ATP: Adenosin trifosfato

GTP: Guanosin trifosfato

PTM: Posttranslational Modifications

SRCR: Scavenger Receptor Cysteine Rich

1. INTRODUCCIÓN

La biomineralización es el proceso que se encarga de la formación de materiales inorgánicos-orgánicos estructurados jerárquicamente generados por organismos vivos (p.e. microorganismos, animales y plantas), como conchas, huesos y dientes, entre otros. Estos materiales están formados por minerales que pueden ser amorfos, policristalinos o cristales individuales (Crichton, 2019). La importancia de este proceso está dada por la capacidad de control que tienen los organismos para formar estructuras funcionales a partir de elementos orgánicos e inorgánicos, tomados selectivamente del medio. Estas estructuras han sido de interés en diferentes áreas interdisciplinarias tales como la biomimética, en donde el entendimiento de procesos biológicos y la aplicación del conocimiento sirven de base para el desarrollo de materiales sintéticos (Lin y Meyers, 2005).

En particular, los moluscos son considerados ideales para estudiar el proceso de biomineralización, dado que presentan una gran diversidad de estructuras y patrones de conchas, donde sus principales funciones son dar soporte, protección contra depredadores, protección contra la desecación o para almacenaje (Simkiss y Wilbur, 1989); además, cuentan con propiedades mecánicas sobresalientes como la rigidez, dureza y alta resistencia (Currey, 1999; Smith *et al.*, 1999). La síntesis a temperatura ambiente de materiales compuestos orgánicos-minerales de altas propiedades mecánicas, el uso de componentes de matriz bioactiva (factores de crecimiento, receptores de factores de crecimiento) para la reparación tisular y el uso de proteínas de la concha como agentes anti-incrustantes biodegradables, son sólo algunas de las aplicaciones que podrían desarrollarse una vez que se comprenda el proceso biológico y los factores involucrados en el proceso de formación de la concha de estos organismos (Marin y Luquet, 2004).

La formación de la concha tiene lugar en el espacio extrapaleal, delimitado por la concha en crecimiento, el periostraco y el manto. Para alcanzar la sobresaturación necesaria para la formación de cristales de carbonato de calcio, los iones de calcio y bicarbonato son transportados a través de la hemolinfa a las células epiteliales, al espacio extrapaleal donde se forma el fluido extrapaleal (fluido sobresaturado). En este espacio, también se encuentra la

matriz orgánica, compuesta de las proteínas que interactúan con los iones minerales dando lugar a la formación de cristales, ya sea calcita, aragonita o ambos, con morfologías definidas. Esta matriz es la encargada de mediar el proceso de biomineralización debido a que regula diferentes aspectos de la deposición de cristales como el inicio de mineralización, ensamblaje en estructuras e inhibición (Marin y Luquet, 2004).

La calcita y la aragonita son los cristales o polimorfos más comunes del carbonato de calcio, y presentan estructuras romboédrica y ortorrómbica, respectivamente. La mayor diferencia entre estos polimorfos está en la organización y la orientación de las moléculas de carbonato. La aragonita puede presentar tres morfologías: tejas, bloques hexagonales o esferas, las cuales se acumulan en forma de pilas o columnas que se empalman con otras (Lin y Meyers, 2005). Por otra parte, la forma espicular de la calcita asemeja prismas triangulares o conos largos y su crecimiento está limitado por la capa externa llamada periostraco (Wilbur y Salenddin, 1983).

Las proteínas que conforman la matriz orgánica de la concha (SMP, por sus siglas en inglés) son las encargadas de dirigir la formación de los cristales de carbonato de calcio y las responsables de las propiedades particulares de las conchas (Falini *et al.*, 1996). Se sabe que estas proteínas presentan regio-especificidad, es decir, las proteínas que son secretadas por el epitelio externo del borde del manto están relacionadas con la formación de la capa prismática de calcita, mientras que las proteínas secretadas por la región dorsal o paleal del manto están relacionadas con la capa nacarada de aragonita (Zhang y Zhang, 2006). Lustrina A (Shen *et al.*, 1997), AP7 y AP24 (Michenfelder *et al.*, 2003), Perlustrina (Weis *et al.*, 2002), Perlucina (Mann *et al.*, 2000), son solo algunas de las proteínas relacionadas a la capa nacarada de aragonita; mientras que Aspeína (Tsukamoto *et al.*, 2004), MSI 31 (Sudo *et al.*, 1997), MSI 7 (Zhang *et al.*, 2003), Primalina-14 (Suzuki *et al.*, 2004) son algunas de las proteínas relacionadas a la capa prismática de calcita.

Estas proteínas pueden clasificarse como solubles o insolubles, de acuerdo con la fracción (altamente variable) en la que se encuentren después de la descalcificación de la concha con un ácido débil o un agente quelante como el EDTA (Marie *et al.*, 2020; Zhang y Zhang, 2006). Esta

clasificación ha sido ampliamente usada para determinar cuál de las fracciones que contienen SMPs controlan el polimorfismo de los cristales de carbonato de calcio (Samata, 1999). Existe la postura que sostiene que la matriz soluble junto con otras moléculas como quitina, pueden seleccionar entre la formación de calcita o aragonita; sin embargo, algunos autores sugieren que la matriz insoluble está involucrada en el control de la selección del polimorfo (Falini *et al.*, 1996; Wilbur y Watabe, 1963). Asimismo, se cree que las proteínas solubles, generalmente ricas en residuos hidrofílicos ácidos como el ácido aspártico, son factores esenciales que determinan las propiedades mineralógicas y cristalográficas; mientras que las proteínas insolubles, las cuales contienen una alta proporción de glicina y alanina, pueden actuar como un marco para la formación y organización de los cristales de la concha y que además, están involucradas en el fortalecimiento de las propiedades mecánicas (Addadi *et al.*, 2006; Lowenstam y Weiner, 1989; Marin *et al.*, 2007; Song *et al.*, 2019).

Particularmente, se ha encontrado que la relación de la fracción insoluble de la matriz orgánica entre las capas o microestructuras presenta diferencias en su composición aminoacídica. Por una parte, la capa prismática es rica en tirosina, prolina y valina; mientras que la capa nacarada contiene alanina y asparagina o ácido aspártico (Song *et al.*, 2019). A pesar de que se pueden encontrar ambas microestructuras, se sabe que la matriz insoluble se caracteriza por estar altamente relacionada a la capa nacarada (Marie *et al.*, 2020). El nácar o madreperla, compuesto por aragonita, es la microestructura más estudiada en los moluscos dadas sus propiedades mecánicas (p.e. dureza, resistencia) y su importancia económica, por lo que se tiene un mayor interés en los genes que codifican para proteínas que están asociadas con éste (Marin y Luquet, 2004).

En general, el estudio de las proteínas de la matriz orgánica puede llevarse a cabo ya sea mediante la extracción de las proteínas de la concha mediante descalcificación (Michenfelder *et al.*, 2003; Sudo *et al.*, 1997; Suzuki *et al.*, 2009; Weiss *et al.*, 2000) o mediante clonación y caracterización de genes relacionados con el proceso de biomineralización (Jackson *et al.*, 2007; Shen *et al.*, 1997; Shi *et al.*, 2019; Zheng *et al.*, 2020).

Aunado a esto, a través de estudios genómicos, transcriptómicos y proteómicos de las conchas de moluscos, se han logrado grandes avances en el conocimiento de las proteínas involucradas en el proceso de biomineralización. La clase Bivalvia ha sido la más estudiada, con 22 géneros y 32 especies, debido al gran interés económico del grupo por la industria perlera. Por otro lado, el estudio de la clase Gastropoda, 13 géneros y 19 especies, se ha concentrado en los géneros *Lottia* y *Haliotis*. Respecto a las demás clases de moluscos, se han reportado pocas publicaciones (Marin *et al.*, 2020).

A diferencia de los moluscos bivalvos, la información disponible sobre las proteínas involucradas en la formación de la concha es limitada en gasterópodos, esto debido a que se le ha dado preferencia a los moluscos formadores de perlas. Además, el conocimiento de proteínas insolubles contra las solubles es todavía menor, debido a su dificultad para ser aisladas, por lo que se requieren estudios enfocados en la búsqueda del entendimiento de este grupo particular de proteínas dada su importancia en el proceso de formación de microestructuras con potenciales aplicaciones. Dado esto, el objetivo del presente trabajo es caracterizar y comprender la función de las proteínas insolubles en el proceso de biomineralización, esto por medio de un estudio bioinformático de las proteínas de gasterópodos reportadas en bases de datos disponibles.

2. ANTECEDENTES

2.1 Biomineralización

La biomineralización es el proceso de formación de una fase mineral biogénica realizada por diferentes seres vivos con fines como protección, soporte, locomoción, entre otros; importante para el estudio tanto de la evolución de muchos organismos como de los compuestos orgánicos-inorgánicos que se forman y que poseen propiedades sobresalientes. Algunos de los organismos, vertebrados e invertebrados, que crean estructuras biomineralizadas son: humanos, puercos, ratas y ratones, los cuales forman huesos y dientes; pollos, los cuales forman huesos y la cáscara del huevo; corales, los cuales forman su esqueleto; fitoplancton y moluscos, los cuales forman conchas. Todas estas estructuras tienen en común el carbonato de calcio, compuesto químico clave para la biomineralización (Evans, 2019).

De manera general, el proceso está compuesto por seis etapas: 1) delimitación y sellado del espacio de nucleación; 2) formación de las matrices compuestas por proteínas, y otras moléculas, que son sintetizadas y exportadas al espacio delimitado para formar el entorno adecuado para comenzar el proceso de nucleación; 3) sobresaturación de iones en el espacio delimitado, creando la insolubilidad del ambiente para la nucleación de un sólido inorgánico; 4) control de la nucleación de biominerales por parte de las proteínas y otras moléculas; 5) control de los biominerales formados para ser depositados; 6) término del proceso de formación de biominerales una vez que se alcanza el requerimiento de estos (aragonita y/o calcita) (Lowenstam y Weiner, 1989).

En particular, la biomineralización ha sido ampliamente estudiada en moluscos, dado que son organismos versátiles respecto a la variedad de tipos y estructuras en sus conchas. Estos organismos pueden formar minerales tanto amorfos como cristalinos. Los minerales de mayor interés son los que se forman a partir del carbonato de calcio, formando cristales ya sea de aragonita o calcita. El filo de los moluscos está compuesto por 7 clases, de las cuales 5 forman conchas. Sin embargo, los estudios se han enfocado mayoritariamente en la clase de los bivalvos y en menor medida en los gasterópodos (Nudelman, 2015).

Con el objetivo de dilucidar el proceso de biomineralización en estos organismos, se han llevado a cabo investigaciones que van desde lo general a lo particular, con enfoques bioquímicos, técnicas de biología molecular y ciencias ómicas (Marin *et al.*, 2016) en moluscos bivalvos marinos de los géneros *Pinctada*, *Pteria* y *Crassostrea*, (por mencionar algunos) y en los gasterópodos marinos *Lottia* y *Haliotis* que son los géneros más estudiados (Marin *et al.*, 2020). A través de ingeniería genética, biología celular y molecular, entre otras áreas de conocimiento, pueden obtenerse proteínas que estén involucradas en la formación de la concha de moluscos, las cuales son de importancia para el entendimiento de este proceso. Aún cuando estos enfoques son altamente eficientes, para poder comprender el proceso de biomineralización de la concha de moluscos primero se requiere tener conocimiento de la biología básica de estos organismos.

2.2 Anatomía de moluscos

Los moluscos están constituidos por: 1) un manto, órgano encargado de depositar los cristales y de la secreción de la matriz orgánica de la concha; 2) el epitelio, que limita los compartimentos del cuerpo del organismo y se compone por dos capas con diferentes funciones: a) la capa interna cubre la superficie del cuerpo (manto, hemolinfa, músculo, tejido conectivo y fibras nerviosas) y acepta iones del medio externo, permitiendo su flujo; b) la capa externa transfiere calcio, bicarbonato, y otros iones desde el epitelio interno hasta el espacio del fluido extrapaleal, además de secretar compuestos orgánicos, lo que provoca la deposición de más de un tipo de estructura de concha; 3) el espacio extrapaleal, que contiene fluido extrapaleal compuesto de sustancias orgánicas e inorgánicas secretadas por el manto, creando el microambiente para la deposición de cristales y la formación de las capas de la concha ya sea nacarada y/o prismática; y 4) la concha, cubierta externamente por una capa fibrosa, flexible y delgada llamada periostracum, la cual protege y sirve como matriz para la deposición inicial de cristales de carbonato de calcio (Saleuddin y Petit, 1983; Wilbur y Saleuddin, 1983).

En particular, las conchas varían en formas y tamaños entre clases. En bivalvos, la concha está formada por dos mitades o valvas, unidas por una bisagra o ligamento flexible, y pueden tener apariencia esférica, aplanada y/o alargada. Por otra parte, los gasterópodos están constituidos

por una sola pieza, una concha dorsal enrollada en espiral, generalmente con apertura en el lado derecho (Brusca y Brusca, 2003; Mollusc of Australia, 2021).

2.3 Composición de las conchas

La concha de los moluscos está compuesta por una matriz mineral y una matriz orgánica. La matriz mineral está compuesta de carbonato de calcio, la cual puede representar más del 95% del peso, depositado ya sea como calcita en la capa prismática o capa externa de la concha; y aragonita, la capa nacarada o interna (Fritz y Morse, 1998). La matriz orgánica de la concha, que representa entre el 0.1 y 5% del peso de la concha, está compuesta por proteínas, glicoproteínas, lípidos, quitina y polisacáridos ácidos. Estas moléculas tienen la capacidad de inducir la nucleación de cristales, seleccionar el polimorfo de carbonato de calcio y controlar el crecimiento y la disposición espacial de los minerales; lo que da lugar a las diferentes formas y estructuras (Falini *et al.*, 1996; Levi-Kalishman *et al.*, 2001). Al ser el principal agente nucleador, la matriz orgánica tiene la capacidad de regular la cristalización, permitiendo o inhibiendo el crecimiento de cristales, según le convenga al organismo (Marin y Luquet, 2004). Los factores que influyen en el proceso son las sustancias orgánicas e inorgánicas presentes en el fluido extrapaleal, el sustrato, la solubilidad, el nivel de supersaturación de carbonato de calcio, y la tasa de crecimiento de los cristales, influenciada por condiciones ambientales (Wilbur y Saleuddin, 1983).

Los cristales, de acuerdo con su crecimiento, pueden formar diferentes estructuras clasificadas como nacarada, foliada, prismática, laminar, esferulítica, cruzada, homogénea y granular (Chen *et al.*, 2012; Currey, 1977; Taylor, 1973; Watabe, 1981). Estas microestructuras pueden combinarse entre sí, formando capas, ya sea únicamente de aragonita, como en las almejas; de calcita, como las vieiras o callos de hacha; o de ambas, como en las ostras perleras y los abulones (Marin *et al.*, 2020). La formación de estas microestructuras es diferente entre moluscos, como es el caso de las dos clases más estudiadas: gasterópodos y bivalvos. Respecto a la capa nacarada o de aragonita, los gasterópodos se caracterizan por formar una estructura de pilas o columnas no contiguas, sin uniformidad, mientras que en los bivalvos los cristales se forman como láminas, asemejando una pared de ladrillos que se ensamblan con orientación uniforme (Wilbur y Saleuddin, 1983).

2.3.1 Importancia del nácar

En particular, el nácar resulta de interés en diferentes áreas de conocimiento debido a su estructura altamente organizada. A través de diferentes estudios tanto *in vitro* como *in vivo* enfocados en ingeniería de tejidos, el nácar ha demostrado diferentes particularidades tales como bioactividad, biocompatibilidad, biodegradabilidad, propiedades osteogénicas y mecánicas, lo que hace de esta microestructura un potencial biomaterial (Flausse *et al.*, 2013; López *et al.*, 2000; Mouriès *et al.*, 2002; Rousseau *et al.*, 2003; Shen *et al.*, 2006; Westbroek y Marin, 1998). Algunas de las aplicaciones más prometedoras son: sustituto óseo, inyección directa en defectos óseos, como recubrimiento en implantes, nanotransportadores enfocados en liberación de drogas o *drug delivery*, y como nanomaterial natural y fluorescente (Gherald *et al.*, 2017).

2.4 Proteínas involucradas en el proceso de Biomineralización

La biomineralización es un proceso complejo que conlleva la regulación por parte de diferentes proteínas a las cuales se les puede clasificar en cuatro grandes grupos: a) proteínas formadoras de la matriz: encargadas de formar el ambiente para la nucleación mediante interacciones entre sí; b) auxiliares de nucleación: estas proteínas están involucradas energéticamente en el ensamblaje de iones, formación de cristales y organización de la fase mineral; c) proteínas comunicadoras: responsables de la comunicación tanto entre células como entre células con la matriz, con el fin de regular el proceso; d) remodeladoras: proteasas o inhibidores de proteasas, encargadas de controlar el entorno de la matriz, ya sea para promover el crecimiento y deposición de minerales, o para poner fin al proceso (Evans, 2019).

En particular, las proteínas encargadas de la formación de la concha se les conoce como proteínas de matriz orgánica o “SMP” (por sus siglas en inglés, Shell Matrix Proteins). Este grupo de proteínas se encarga de la formación de los compuestos minerales de los que se ha hablado anteriormente (p.e. calcita y/o aragonita) los cuales forman las capas o microestructuras, nacarada y prismática, que caracterizan a las conchas de los moluscos. La estructura primaria de estas proteínas, así como su función y su secuencia de aminoácidos ha sido objeto de estudio para entender el proceso de biomineralización (Marin y Luquet, 2004; Zhang y Zhang, 2006).

Con el fin de comprender la función de estas proteínas, se han llevado a cabo diferentes estudios enfocados en transcriptómica y proteómica, en donde se analiza una gran cantidad de información respecto a genes y proteínas. Estos estudios contribuyen significativamente al conocimiento de los potenciales involucrados en el proceso de biomineralización (Shi *et al.*, 2019; Werner *et al.*, 2013). En particular, algunos investigadores se han enfocado en realizar diversos estudios proteómicos, analizando el contenido completo de las proteínas de concha de varios modelos de moluscos (Feng *et al.*, 2017; Marie *et al.*, 2010, 2012; Mann *et al.*, 2012, 2018), por lo que a estos proteomas se les conoce como “shellomas” (Marin *et al.*, 2013).

Si bien la acelerada obtención de información ha beneficiado el campo de investigación, se requiere de un análisis más detallado del conjunto de resultados obtenidos para llegar a conclusiones más precisas sobre el proceso de formación de la concha de estos organismos. Uno de los métodos que han resultado favorables para el estudio de estas proteínas en conjunto ha sido su caracterización bioquímica y clasificación dentro de diferentes grupos con los cuales se puede llegar a determinar cualidades específicas.

2.5 Caracterización de proteínas

Las proteínas ya caracterizadas son una base importante para poder identificar otras proteínas involucradas en la formación de la concha de los moluscos. Estas pueden tener diferentes clasificaciones según el enfoque que se le quiere dar. Una de ellas es la clasificación según la capa de la cual provengan, ya sea nacarada o prismática. Algunas proteínas de la capa nacarada de moluscos son Perlucina y Perlustrina, identificadas en *Haliotis laevis* (Weiss *et al.*, 2000), AP7 y AP24 identificadas en *Haliotis rufescens* (Michenfelder *et al.*, 2003), N66 y N14 identificadas en *Pinctada maxima* (Kono *et al.*, 2000); y de la capa prismática de calcita se ha identificado a la proteína MS31 (Sudo *et al.*, 1997) y la Nacreína (Miyamoto *et al.*, 1996) ambas aisladas de *Pinctada fucata*.

Las proteínas también se pueden clasificar de acuerdo con su punto isoeléctrico (pI) teórico, ya sea en proteínas extremadamente ácidas (pH <4.5), moderadamente ácidas (de pH \geq 4.5 a pH \leq 6) o proteínas básicas (pH >7). Partiendo de esta clasificación, se pueden considerar varios puntos importantes de las proteínas de la concha: a) Respecto a las proteínas extremadamente

ácidas, estas son las más difíciles de purificar y se cree que están mayormente asociadas con la calcita de las conchas de los moluscos (Hare, 1963). Por otra parte, estas proteínas se asocian con un alto contenido de residuos de ácido aspártico y dada su cadena lateral corta, supone la facilidad de unión a iones de calcio, por lo que pueden clasificarse en el grupo de proteínas de baja afinidad y alta capacidad de unión a calcio (Maurer y Hohenester, 1997). b) Respecto a las proteínas moderadamente ácidas, estas están mayormente relacionadas con la formación de aragonita; sin embargo, hay algunas excepciones, ya que se han aislado proteínas básicas relacionadas con esta microestructura (Marin, 2007). En otras palabras, se cree que el punto isoeléctrico está asociado al polimorfismo del cristal de carbonato de calcio, ya sea $pI < 4$ para la capa de calcita, y $pI > 4$ para la capa de aragonita (Zhang y Zhang, 2006).

Las SMPs también se caracterizan por la predominancia de algunos aminoácidos, usualmente tres o cuatro, de los que los más comunes son glicina, ácido aspártico y serina. En menor medida se han encontrado la prolina, cisteína, tirosina, leucina y asparagina. Se ha visto que algunos aminoácidos dominantes, como el ácido aspártico, están asociados con microestructuras calcíticas, mientras que ninguna de las proteínas que se han caracterizado de la capa nacarada es rica en estos aminoácidos (Marin y Luquet, 2004).

Asimismo, las proteínas pueden clasificarse dependiendo de su solubilidad, como fracción soluble e insoluble, después de la descalcificación de la fase mineral, esto al usar soluciones como EDTA, ácido diluido (ácido acético al 10%), agua Milli-Q, urea o detergentes (Marin y Luquet, 2004; Samata, 2004, Zhang y Zhang, 2006). Esta característica en particular se ha utilizado como paso estratégico para la obtención y purificación de las SMPs.

2.6 Fraccionamiento de proteínas

El proceso de descalcificación tiene como resultado dos fracciones orgánicas con una relación altamente variable: la fracción soluble en la solución de descalcificación, la cual representa entre 0.03 y 0.5% del peso de la concha; y la fracción fuertemente insoluble, la cual tiene una mayor variación que va desde 0.01% a 4-5% del peso de la concha (Marin, 2007). Dada esta variabilidad, la obtención de proteínas insolubles es más complicada, por lo que son pocas las proteínas que se han logrado aislar y caracterizar en comparación con las proteínas solubles.

Algunas de las proteínas insolubles aisladas y caracterizadas de la capa nacarada, y relacionadas con la formación de aragonita, son: Lustrina A aislada de *Haliotis rufescens* (Shen *et al.*, 1997); MSI60, N16, Nacreína, Pif 97 y Pif 80, aisladas de *Pinctada fucata* (Bahn *et al.*, 2015; Bahn *et al.*, 2017; Miyamoto *et al.*, 1996; Samata *et al.*, 1999; Sudo *et al.*, 1997). Así mismo, la perlina, que es una proteína insoluble aislada de las perlas, que equivale a la capa nacarada dada su composición de cristal de aragonita y matrices orgánicas, es un modelo más simple de estudio (Miyashita *et al.*, 2000).

Algunas de las proteínas insolubles, relacionadas con la microestructura prismática de calcita son MSI31 y MSI7, aisladas de *Pinctada fucata* (Sudo *et al.*, 1997; Zhang *et al.*, 2003). En particular, el conocimiento sobre proteínas insolubles caracterizadas de organismos gasterópodos marinos es menor respecto a las proteínas insolubles de organismos bivalvos. En la Tabla 1 se muestran las proteínas más representativas de organismos gasterópodos en donde se visualiza la falta de información sobre proteínas insolubles a comparación con las solubles.

Tabla 1. Proteínas de moluscos gasterópodos aisladas de la concha en los últimos 20 años.

	Proteína	Función	Referencia
	Perlucin ¹	Acelera la precipitación de CaCO ₃	Weiss <i>et al.</i> , 2000
	Perlustrin ¹	Proteína de unión a IGF	Weiss <i>et al.</i> , 2000
	AP7 ²	Controla nucleación de cristales	Michenfelder <i>et al.</i> , 2003
Solubles	AP24 ²	Posee sitios de unión a Ca ²⁺	Michenfelder <i>et al.</i> , 2003
	AP8 ²	Modifica los cristales de calcita	Fu <i>et al.</i> , 2005
	Perlawapin ¹	Inhibe formación de cristales	Treccani <i>et al.</i> , 2006
	Perlinhibin ¹	Inhibe formación de nácar	Mann <i>et al.</i> , 2007
	ML5A7 ³	Regula el crecimiento de calcita	Marie <i>et al.</i> , 2010
Insolubles	Peroxidasa ⁴	Forma y repara la concha	Marie <i>et al.</i> , 2012
	Perlawapin ¹	Inhibe formación de cristales	Mann <i>et al.</i> , 2018

Notas: Los superíndices indican la especie de la cual fueron identificadas las proteínas. ¹ *Haliotis laevigata*, ² *Haliotis rufescens*, ³ *Haliotis asinina*, ⁴ *Lottia gigantea*. IGF: factor de crecimiento similar a la insulina, por sus siglas en inglés.

La importancia de las proteínas insolubles de la matriz está dada por la función que se les ha atribuido según diferentes estudios. Inicialmente, se creía que las proteínas insolubles de la matriz controlaban los polimorfismos de carbonato de calcio mediante la precipitación de iones (Song 2019; Watabe y Wilbur, 1960; Wilbur y Watabe, 1963); sin embargo, en diversos estudios se ha demostrado que las proteínas solubles de la matriz de la concha son las responsables de esta función (Belcher *et al.*, 1996; Cariolou y Morse, 1988; Falini *et al.*, 1996; Feng *et al.*, 2000; Thompson *et al.*, 2000; Weiner y Hood, 1975; Zaremba *et al.*, 1996).

Posteriormente, se determinó que la fracción insoluble está compuesta por β -quitina y otras proteínas similares a la seda (*silk-like*), relativamente hidrofóbicas y ricas en alanina y glicina, cuya función principal es dar estructura al marco tridimensional en el que se formará la fase mineral (Weiner y Traub, 1980; 1984). Así mismo, se determinó que la fracción soluble está compuesta por proteínas ácidas, ricas en ácido aspártico y/o ácido glutámico cuya función es controlar la nucleación, la morfología y la selección del polimorfo (Young *et al.*, 1977; Weiner y Traub, 1980).

De manera análoga, en estudios más recientes se sostiene que las proteínas de la matriz insoluble son las responsables de crear el entorno necesario para que se lleve a cabo el crecimiento de cristales y proveen, además, la superficie base para la formación de la concha (Marie *et al.*, 2012). Otros autores explican que principalmente las proteínas de la matriz insoluble, así como otras moléculas, proporcionan sitios en los que se lleva a cabo la nucleación; es decir, forman las bases en las que ordenadamente se depositan los iones de calcio y carbonato (Liu *et al.*, 2015). Además de plantear que las proteínas de la fracción insoluble funcionan como marco para la formación de la concha, también se les responsabiliza de acentuar las propiedades mecánicas de las conchas, mientras que a las proteínas de la fracción soluble se les considera esenciales en la determinación de las propiedades de la concha (p.e. selección del polimorfo y la microestructura) (Addadi *et al.*, 2006; Cusack y Freer, 2008; Marie *et al.*, 2012; Marin *et al.*, 2008; Morse *et al.*, 2007). Por tanto, todos estos estudios confirman la importante contribución de las proteínas de la matriz insoluble en el proceso de biomineralización.

2.7 Funcionalidad de proteínas

A través de diferentes estudios, se ha logrado descifrar la funcionalidad de diversas SMPs, tanto solubles como insolubles, tales como AP7, que modula el ensamblaje de la matriz, la nucleación y el crecimiento de cristales (Amos *et al.*, 2009; Wutsman *et al.*, 2004); AP8, que modifica la morfología de los cristales de calcita (Fu *et al.*, 2005); N16, la cual lleva a cabo la estabilización mineral biogénica (Brown *et al.*, 2014); N19, reguladora negativa de la calcificación (Yano *et al.*, 2007); Pif, la cual se une a cristales de aragonita y regula la formación de nácar (Suzuki *et al.*, 2009); Perlwapin, inhibidora del crecimiento de cristales de carbonato de calcio (Treccani *et al.*, 2006); Dermatopontin, la cual participa en la formación de nácar (López *et al.*, 1992); por mencionar algunas.

Dado los hallazgos de una gran diversidad de proteínas, clasificarlas en grupos resulta ser conveniente para su análisis. De acuerdo con su función dentro del proceso de biomineralización, se les puede clasificar como: a) formadoras del marco o revestimiento de la estructura orgánica: proteínas “*silk-like fibroin*” y proteínas formadoras de gel que proveen funciones específicas como ductilidad, elasticidad y resistencia a la fractura; b) proteínas intercrystalinas: proteínas que existen en el interior de la calcita o aragonita, generando porosidad, que aligeran y fortalecen los cristales; c) formadoras de perlas o reparadoras: proteínas que actúan en respuesta a irritantes o virus, recubriendo estas entidades con capas de fase mineral prismática y nácar, creando una perla en el proceso o proteínas que reconstruyen y reparan las capas de la concha en respuesta a daños, respectivamente; y d) proteínas proteolíticas: proteasas específicas de proteínas que inducen su fragmentación para inducir cambios en el proceso de biomineralización, además de inhibidores de proteasas, que detienen el proceso proteolítico en la matriz (Evans, 2019).

Por otra parte, una característica común de muchas de las SMPs de moluscos es su organización modular, donde cada módulo corresponde a un dominio funcional. Por lo tanto, muchas de las proteínas de la matriz extracelular son multifuncionales. Mediante el análisis de las secuencias de cada proteína, se ha logrado deducir e hipotetizar la función de estos dominios. Los dominios más comunes entre estas proteínas son: a) anhidrasas carbónicas, con posible función de

síntesis de bicarbonatos; b) dominios ácidos cortos, con posible función de unión a calcio; c) repeticiones cortas de un aminoácido (Gly, Ala, Asp o Ser), de dos aminoácidos (GlySer O GluAsn), seis o más residuos; d) bloques poli-A o poli-G con posible función estructural; e) dominios ricos en glicina, con posible función estructural de hojas β ; f) dominios de glicina y serina, con posible función de formación de loops; entre otros (Marin, 2004; 2007).

2.7.1 Dominios

Algunas de las proteínas con dominios asociados a una funcionalidad y que están presentes en moluscos gasterópodos se presentan en la Tabla 2. De forma general, se puede observar dominios involucrados directamente en el proceso de biomineralización, ya sea dando estructura, favoreciendo, inhibiendo o modulando la formación de cristales, o sintetizando precursores como lo es el bicarbonato; dominios de unión a moléculas, ya sea carbohidratos como la quitina, a iones como el calcio, a zonas específicas como la matriz extracelular o a factores de crecimiento; y un dominio que presenta actividad enzimática.

Gracias a los diferentes estudios proteómicos realizados por diferentes autores, se han podido determinar y clasificar los dominios funcionales relacionados con la biomineralización. Algunos de ellos son los dominios con función enzimática (p.e., proteasas, peroxidases, tirosinasas), con función de inhibición de proteasas o de inmunidad (p.e., Kunitz-like, IgG-like), con capacidad de aglutinamiento de cationes divalentes (p.e., Fe/Cu, Ca, endonucleasas, EF-hand), dominios con regiones de baja complejidad (p.e., N/Q-rich, A/G-rich, C/M-rich), dominios de matriz extracelular o moléculas de unión a matrices extracelulares (p.e., EGF-like, SUSHI, Fibronectina, Zona Pelúcida), dominios de unión a sacáridos (p.e., quitina, lectina tipo C, quitinasa), por mencionar algunos (Marin *et al.*, 2020).

Tabla 2. Dominios presentes en proteínas aisladas de gasterópodos y sus funciones asociadas.

Proteína	Dominio	Función asociada	Referencia
Lustrin A ¹	C/P, GS	Dar estructura, inhibidor de proteasas	Shen <i>et al.</i> , 1997
Perlucina ²	Lectina tipo C	Unión a carbohidratos	Mann <i>et al.</i> , 2000
Perlustrin ²	Unión a IGF	Interacción de células de unión a IGF	Weiss <i>et al.</i> , 2000
Perlawapin ²	WAP	Inhibe el crecimiento de CaCO ₃	Treccani <i>et al.</i> , 2006
Nacreína ³	α-CA, G-X-N	Síntesis de HCO ₃ ⁻	Miyamoto <i>et al.</i> , 2003
BSMP ⁴	VWA	Modula morfología cristalina	Werner <i>et al.</i> , 2012
BSMP ⁴	Unión a quitina	Unión a quitina	
Ependimina ⁵	Ependimina	Unión a calcio	Marie <i>et al.</i> , 2010
Ferritin-like ⁶	Ferritin	Actividad ferroxidasa	Mann <i>et al.</i> , 2012
LUSP-17 ⁶	EGF, ZP	Unión a ECM, señalización	Marie <i>et al.</i> , 2012
Peroxidasa ⁶	Peroxidasa	Actividad enzimática	Marie <i>et al.</i> , 2012

Notas: Los superíndices indican la especie de la cual fueron identificadas las proteínas. ¹ *Haliotis rufescens*, ² *Haliotis laevigata*, ³ *Turbo marmoratus*, ⁴ *Patella vulgata*, ⁵ *Haliotis asinina*, ⁶ *Lottia gigantea*. C/P: Cisteína/Prolina. GS: Glicina y Serina. IGF: insulin-like Growth Factor. WAP: Whey Acidic Protein. BSMP: blue mussel shell protein. VWA: von Willebrand factor A, por sus siglas en inglés. CA: anhidrasa carbónica. EGF: epidermal growth factor-like. ZP, zona pellucida. ECM: matrices extracelulares.

2.7.2 Dominios o regiones de baja complejidad

Por otra parte, se cree que la funcionalidad de las proteínas puede estar ligada a las regiones o dominios de baja complejidad. Las regiones o dominios de baja complejidad o LCR/LCD (Low Complexity Regions/Low Complexity Domains, por sus siglas en inglés) son secuencias de aminoácidos que se caracterizan por el predominio de uno o pocos residuos; estas regiones pueden ser secuencias cortas (10 aminoácidos aproximadamente) o largas (75 a 200 aminoácidos) y la mayoría son intrínsecamente desordenadas (no exhiben ninguna estructura secundaria o terciaria específica). En modelos eucariontes, se sabe que estas regiones pueden desempeñar funciones como la modulación de las interacciones proteína - proteína, proteína - ácidos nucleicos, localización subcelular de proteína y activación de factores de transcripción, entre otros; sin embargo, estas funciones no se han demostrado en moluscos, por lo que la

mayoría de LCR tienen una función desconocida, respecto a modelos de biomineralización (Marin, 2020; Toll-Riera *et al.*, 2012).

Algunas LCR con funciones conocidas en moluscos son: a) ricas en ácido aspártico, las cuales tienen capacidad de unir iones de calcio, se cree que pueden actuar como nucleadores y que pueden inhibir el crecimiento de cristales cuando están en solución; b) ricos en serina o treonina, los cuales pueden tener funciones similares a la región de ácido aspártico, esto debido a sus potenciales sitios de fosforilación; regiones hidrofóbicas (p.e. glicina, alanina, valina o leucina), las cuales se encuentran en la fracción insoluble de las conchas de moluscos, y desempeñan papeles estructurales en donde crecen los cristales; c) las regiones básicas (p.e. lisina o arginina), se cree que pueden interactuar con las regiones ricas en ácido aspártico, con bicarbonato y con iones de carbonato, dada sus propiedades de carga poli-catiónicas. Por otra parte, la importancia funcional de regiones ricas en metionina, cisteína, asparagina, glutamina y prolina aún no son claras (Marin *et al.*, 2020).

Si bien todos estos estudios, en su conjunto, han contribuido significativamente al entendimiento del proceso de biomineralización, aún falta mucho no solo por descubrir sino por relacionar, para lo cual la información existente y que se genera día a día debe organizarse adecuadamente para contribuir al entendimiento de este proceso tan complejo.

3. JUSTIFICACIÓN

La versatilidad de las estructuras de las conchas, así como sus propiedades (p.e., rigidez, dureza, resistencia), principalmente provistas por la presencia de nácar, hacen de los moluscos gasterópodos un modelo clave para el estudio y entendimiento del proceso biológico de biomineralización. Particularmente, el enfoque de múltiples investigaciones va dirigido a las proteínas involucradas en la formación de esta estructura ya que podrían ser la base para el desarrollo de nuevos biomateriales en diferentes industrias, principalmente en el área de la salud. Sin embargo, los estudios se limitan hacia las proteínas de la fracción soluble de modelos determinados como económicamente importantes, como lo es la ostra perlera y el abulón.

Esta investigación busca contribuir al conocimiento sobre las características que definen a un conjunto de proteínas insolubles de la matriz orgánica de las conchas de moluscos gasterópodos, las cuales han sido poco estudiadas de manera individual. Por medio de herramientas bioinformáticas, se busca analizar y caracterizar a este particular grupo de proteínas, con el objetivo de establecer particularidades que permitan el entendimiento de su papel en el proceso de biomineralización. Este trabajo, facilitará el estudio de futuras proteínas aisladas, permitiendo su comparación para encontrar homologías, o bien, dando paso al descubrimiento de nuevas proteínas.

De manera local, el conocimiento sobre las proteínas involucradas en la formación de microestructuras podría beneficiar a la industria perlera, importante para la Paz, Baja California Sur. Este aporte podría contribuir, en un futuro, al uso y manejo de los organismos utilizados, así como en el mejoramiento de la eficiencia del proceso de formación de perlas.

4. HIPÓTESIS

Si el proceso de biomineralización en moluscos cuenta con un conjunto de proteínas insolubles para la formación de la concha, las proteínas involucradas comparten una relación filogenética, así como características homólogas en su estructura y función, entonces es posible la identificación y caracterización de SMPs insolubles no descritas en gasterópodos en repositorios bioinformáticos.

5. OBJETIVOS

5.1 Objetivo general

Determinar las similitudes y diferencias entre las proteínas insolubles relacionadas a la formación de la concha en gasterópodos.

5.2 Objetivos particulares

- Generar una base de datos de proteínas insolubles relacionadas con la formación de conchas de gasterópodos.
- Analizar las características bioquímicas de las proteínas que integran la base de datos generada.
- Analizar las secuencias reportadas de las proteínas utilizando herramientas bioinformáticas.

6. MATERIAL Y MÉTODOS

6.1 Base de datos de proteínas insolubles a partir de conchas de gasterópodos.

6.1.1 Búsqueda de proteínas

Se utilizaron dos bases de datos (PubMed y Proteínas) del repositorio público Centro Nacional de Información Biotecnológica, NCBI (National Center for Biotechnology Information, por sus siglas en inglés) (Wheeler *et al.*, 2007) para la identificación de proteínas insolubles aisladas de la concha de gasterópodos relacionadas con el proceso de biomineralización. Se utilizaron palabras clave que cumplieran con el objetivo de identificar proteínas insolubles de la matriz de la concha. Los resultados se filtraron por clase (gastropoda) y/o por especie (específicamente organismos marinos). Por lo tanto, se realizó una revisión bibliográfica en la base de datos de literatura PubMed utilizando las siguientes palabras clave: insoluble matrix shell proteins *Lottia gigantea*; shell matrix proteins *Haliotis laevigata*; y shell forming proteins + gastropoda. Por otra parte, se realizó la búsqueda directa en la base de datos de proteínas utilizando las siguientes palabras clave: shell protein *Haliotis rufescens*; insoluble matrix shell protein + gastropoda AND *Haliotis asinina*; y mollusca matrix shell protein AND *Lottia gigantea*. Las palabras clave que no tuvieron resultados satisfactorios o esperados no son mencionadas.

Además de NCBI, se utilizó UniProt (Bateman *et al.*, 2021) y Ensemble Metazoa (Howe *et al.*, 2020) para recabar la información de cada proteína seleccionada: nombre de la proteína, número de acceso y las características bioquímicas reportadas. Así mismo, se obtuvieron las secuencias de aminoácidos, en formato FASTA, de cada proteína.

6.2 Características de las proteínas insolubles

De acuerdo con la literatura reportada, se consideraron como características importantes para estudiar a las proteínas de la matriz orgánica de la concha de moluscos, y que deben formar parte de la base de datos, las siguientes: microestructura; peso molecular (kDa); punto isoeléctrico; número de aminoácidos; presencia de péptido señal; estatus en el registro de la secuencia (completa o incompleta); los tres aminoácidos predominantes en la composición de la proteína; presencia de dominios, motivos y/o regiones de baja complejidad; presencia de

modificaciones postraduccionales (p.e. glicosilaciones, fosforilaciones y/o enlaces disulfuro); y posible función.

6.2.1 Caracterización de proteínas

Se realizó un análisis *in silico* de las secuencias de proteínas obtenidas de las bases de datos utilizando herramientas bioinformáticas para su caracterización. En la Tabla 3 se muestran las características y la herramienta utilizada para su determinación. La única característica que no se determinó *in silico* fue la microestructura ya que esta información se obtuvo directamente de la literatura.

Tabla 3. Herramientas utilizadas para la caracterización de secuencias.

Característica	Herramienta	Referencia
pI, MW, composición aminoacídica	ExPasy ProtParam	Gasteiger <i>et al.</i> , 2005
Longitud de secuencia, pares de bases	NCBI, Ensembl Metazoa	Wheeler <i>et al.</i> , 2007; Howe <i>et al.</i> , 2020
Presencia de péptido señal	SignalP	Almagro <i>et al.</i> , 2019
Estatus de registro de la secuencia	UniProtKB/Swiss-Prot	Bateman <i>et al.</i> , 2021
Glicosilaciones	NetNGlyc, NetOGlyc,	Julenius <i>et al.</i> , 2005
Fosforilaciones	NetPhos	Blom <i>et al.</i> , 1999
Enlaces disulfuro	DISULFIND	Ceroni <i>et al.</i> , 2006
Predicción de motivos	MotifScan	Marchler-Bauer <i>et al.</i> , 2013
Predicción de dominios	PROSITE	Sigrist <i>et al.</i> , 2021
Regiones de baja complejidad	SMART	Letunic <i>et al.</i> , 2021
Predicción de función	InterProScan	Jones <i>et al.</i> , 2014

6.3 Comparación de características de las proteínas insolubles de la base de datos generada para gasterópodos.

Una vez que se completó la base de datos, se realizó un análisis general de la información mediante observación de las tendencias en cada característica, para establecer parámetros y realizar la comparación de características entre las proteínas insolubles seleccionadas, esto por medio de diagramas de Venn, gráficas de barras, gráficos circulares y tablas comparativas (Tabla 4).

Tabla 4. Parámetros determinados para cada característica.

Característica	Parámetros
Microestructura	Nacarada, Prismática, Ambas
Peso Molecular (kDa)	<50, 50-99, >100, desconocido
Punto isoelectrico	Ácido (<7), Básico (≥7)
Longitud de secuencia	Pequeña (≤200), Mediana (201-699), Grande (≥700)
Presencia de péptido señal	Si, No
Estatus de registro de secuencia	Completa, Incompleta
Composición aminoacídica	Los 3 aminoácidos con mayor porcentaje
Dominios	Presencia, cantidad y tipo
Motivos	Presencia y cantidad
Regiones de baja complejidad	Presencia y cantidad
Modificaciones postraduccionales	Presencia, cantidad y tipo
Posible función	Enzimática, transporte, regulación, estructural

6.4 Identificación de grupos de proteínas usando un árbol guía

Se realizó el alineamiento múltiple de las secuencias de aminoácidos de las proteínas insolubles completas. Las secuencias se rasuraron eliminando el péptido señal (identificado con la herramienta Signal P 5.0; Almagro *et al.*, 2019) y el alineamiento múltiple se realizó con Clustal Omega (Sievers *et al.*, 2011). Con base en el árbol generado por el programa (guía), se

identificaron grupos preliminares según los clados formados. Una vez identificadas las secuencias de cada clado o grupo, se analizaron por separado con MegaX (Stecher *et al.*, 2020). De los alineamientos obtenidos por cada grupo, se elaboraron árboles filogenéticos utilizando el método de Neighbor Joining usando un Bootstrap de 1000. De cada árbol filogenético, se analizó la relación de los nodos que tuvieran valores de bootstrap > 50%. Así mismo, se analizó la relación que tienen con el proceso de biomineralización en organismos gasterópodos basado en la literatura reportada.

7. RESULTADOS

7.1 Base de datos de proteínas insolubles obtenidas de conchas de gasterópodos.

7.1.1 Obtención de proteínas

Empleando las bases de datos del NCBI (PubMed y Proteínas), se obtuvieron resultados de proteínas insolubles obtenidas de conchas, para cada conjunto de palabras usadas, de los cuales las proteínas se obtuvieron de forma directa. En la Tabla 5 se muestran las palabras clave usadas en cada búsqueda, los resultados totales y los resultados objetivo, así como las proteínas obtenidas en cada uno; en las palabras entre paréntesis se señala la base de datos de la que se obtuvieron los resultados.

Tabla 5. Información de búsqueda.

Palabras clave	Hits totales	Hits blancos	Proteínas
Insoluble matrix shell proteins <i>Lottia gigantea</i>	2 (literatura) ¹	1	427
Shell matrix proteins <i>Haliotis laevigata</i>	5 (literatura) ¹	1	76
Shell forming proteins + gastropoda	12 (literatura) ¹	1	13
Shell protein <i>Haliotis rufescens</i>	4 (proteínas) ²	1	1
Insoluble matrix shell protein + gastropoda AND <i>Haliotis asinina</i>	13 (proteínas) ²	13	13
Mollusca Matrix Shell protein AND <i>Lottia gigantea</i>	34 (proteínas) ²	24	24

Notas: ¹ Resultados obtenidos de la base de datos PubMed de NCBI. ² Resultados obtenidos de la base de datos de proteínas de NCBI.

En total, se encontraron 554 proteínas provenientes de nueve especies de moluscos gasterópodos, una de ellas perteneciente al género *Lottia* y ocho al género *Haliotis*. La distribución de las proteínas según la especie se muestra en la Tabla 6 (en orden descendente). Las secuencias de aminoácidos de *Lottia gigantea* se recuperaron de Ensemble metazoa, mientras que las secuencias de *Haliotis laevigata* se obtuvieron de Uniprot. Todas las demás secuencias se obtuvieron de la base de datos de proteínas de NCBI.

Tabla 6. Proteínas por especie.

No.	Especie	Proteínas
1	<i>Lottia gigantea</i>	451
2	<i>Haliotis laevigata</i>	79
3	<i>Haliotis asinina</i>	15
4	<i>Haliotis rubra</i>	2
5	<i>Haliotis varia</i>	2
6	<i>Haliotis ovina</i>	2
7	<i>Haliotis rufescens</i>	1
8	<i>Haliotis discus hannai</i>	1
9	<i>Haliotis tuberculata</i>	1
Total		554

7.2 Caracterización de proteínas

La caracterización de proteínas se llevó a cabo de acuerdo con los parámetros previamente establecidos. A continuación, se describen los resultados obtenidos para cada característica.

7.2.1 Microestructura

De acuerdo con la información reportada en literatura, se encontró que, de las 554 proteínas, 60 están relacionadas con la capa nacarada, 24 con la capa prismática y 34 están relacionadas con ambas capas. Por otra parte, el resto de las proteínas encontradas (436) no cuentan con una asociación definida a alguna microestructura (Fig. 1).



Figura 1. Distribución de las proteínas entre las diferentes fracciones de la concha. Diagrama de Venn mostrando las proteínas asociadas a la capa nacarada, prismática, ambas capas o sin microestructura definida.

7.2.2 Características bioquímicas

Se determinaron las características bioquímicas de las proteínas de la base de datos creada, p.e. peso molecular, el punto isoeléctrico, la longitud aminoacídica, la presencia de péptido señal y el estatus de secuencia (Fig. 2).

Peso molecular. Se determinaron tres rangos para su análisis y se encontró que, de las 554 proteínas, 413 pesan menos de 50 kDa, 105 pesan entre 50 y 99 kDa, y 36 pesan más de 100 kDa.

Punto isoeléctrico. Para su clasificación, se establecieron dos intervalos, de acuerdo con los valores observados previamente. Se encontró que, de las 554 proteínas, 335 proteínas pueden considerarse ácidas, ya que tienen un valor de pI menor a 7, y 219 básicas, ya que tienen un valor de pI mayor o igual a 7.

Longitud de la secuencia de aminoácidos. Se establecieron tres parámetros para su análisis, encontrándose que, de las 554 proteínas, 148 son consideradas pequeñas (secuencia constituida por 200 aminoácidos o menos), 343 medianas (secuencia constituida entre 201 a 699 aminoácidos) y 63 secuencias grandes (proteínas compuestas por 700 aminoácidos o más).

Presencia de péptido señal. De las 554 proteínas, 198 presentan péptido señal, mientras que 356 no.

Estatus de registro de la secuencia. Para esta característica se consideró si el registro de la proteína consistió en la secuencia de aminoácidos completa o incompleta (fragmento). De las 554 proteínas, se encontró que 501 registros de proteínas cuentan con las secuencias completas, mientras que 53 registros de proteínas consisten únicamente de un fragmento de la proteína; en la Figura 2 se muestra el concentrado de los resultados antes mencionados.

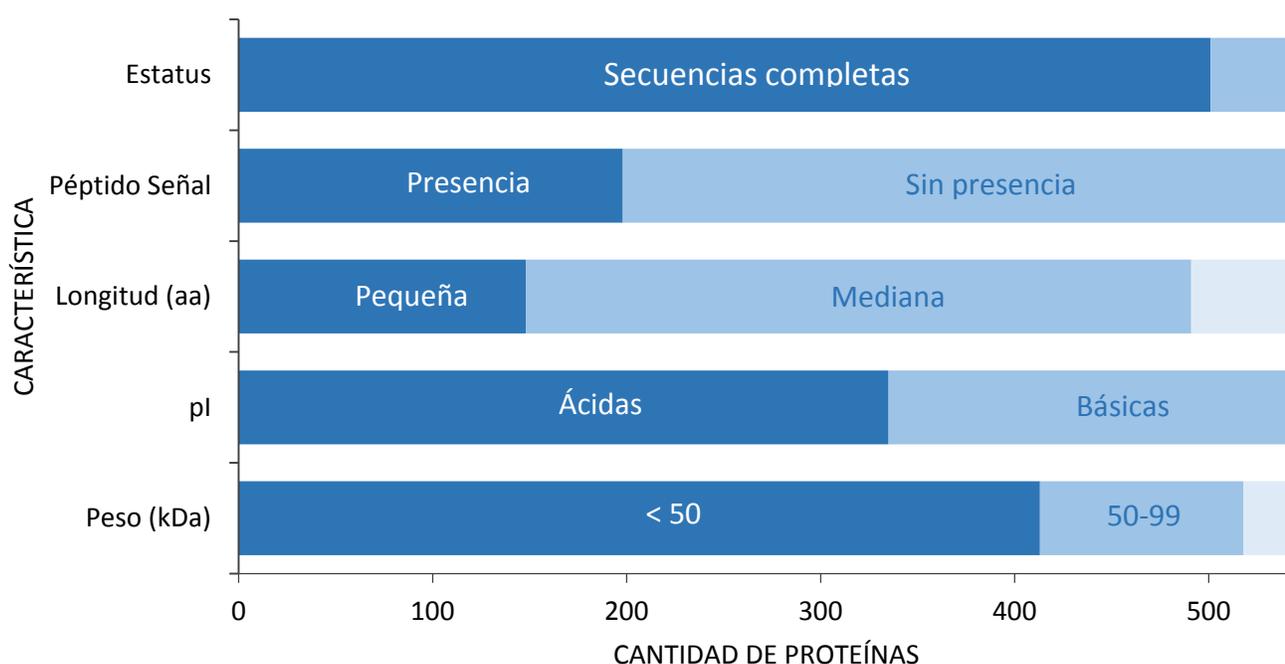


Figura 2. Distribución de proteínas de acuerdo con límites establecidos para cada característica. Las barras corresponden a las características de las proteínas. El degradado de color en cada una representa los parámetros establecidos.

7.2.3 Análisis de contenido de aminoácidos (dominantes)

De acuerdo con los tres aminoácidos con mayor frecuencia (en porcentaje de composición) en las secuencias de aminoácidos de las proteínas analizadas, se identificaron siete aminoácidos dominantes. Estos aminoácidos considerados “**dominantes**” representan entre el 7 y 35% de la composición total de las secuencias de proteínas que contienen uno o más de estos. Los aminoácidos dominantes son: treonina (Thr), presente en 106 proteínas; valina (Val), presente

en 158 proteínas; alanina (Ala), presente en 161 proteínas; lisina (Lys), presente en 170 proteínas; serina (Ser), presente en 184 proteínas; glicina (Gly), presente en 218 proteínas; y leucina (Leu), presente en 262 proteínas (Fig. 3).

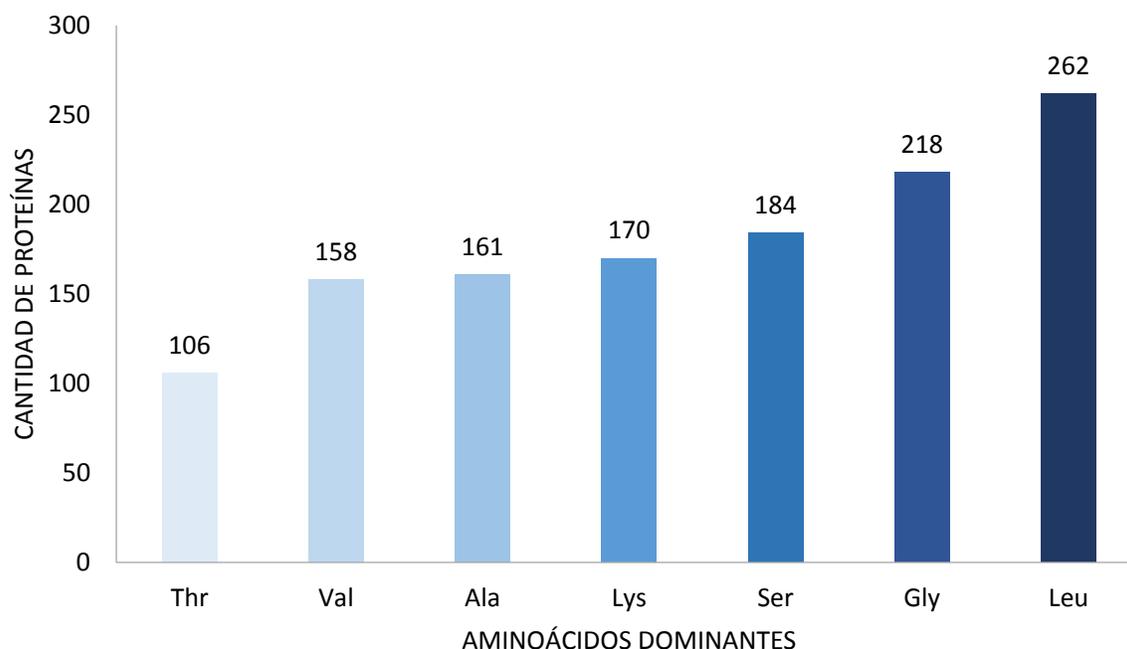


Figura 3. Aminoácidos dominantes en las secuencias de proteínas. Representación gráfica de la cantidad de proteínas que presentan uno o más de estos siete aminoácidos dominantes.

De igual manera, se determinaron los aminoácidos dominantes de las proteínas en relación con una microestructura definida (Fig. 4). De las 60 proteínas relacionadas a la **capa nacarada**, los aminoácidos dominantes resultaron ser: serina (Ser), presente en 22 proteínas; prolina (Pro), presente en 26 proteínas; y glicina (Gly), presente en 33 proteínas. De las 24 proteínas relacionadas a la **capa prismática**, los aminoácidos dominantes son: glicina (Gly), presente en 10 proteínas; serina (Ser), presente en 12 proteínas; y treonina (Thr), presente en 14 proteínas. Respecto a las 34 proteínas relacionadas con **ambas capas**, los aminoácidos dominantes son: leucina (Leu), presente en 12 proteínas; alanina (Ala), presente en 13 proteínas; y glicina (Gly), presente en 21 proteínas.

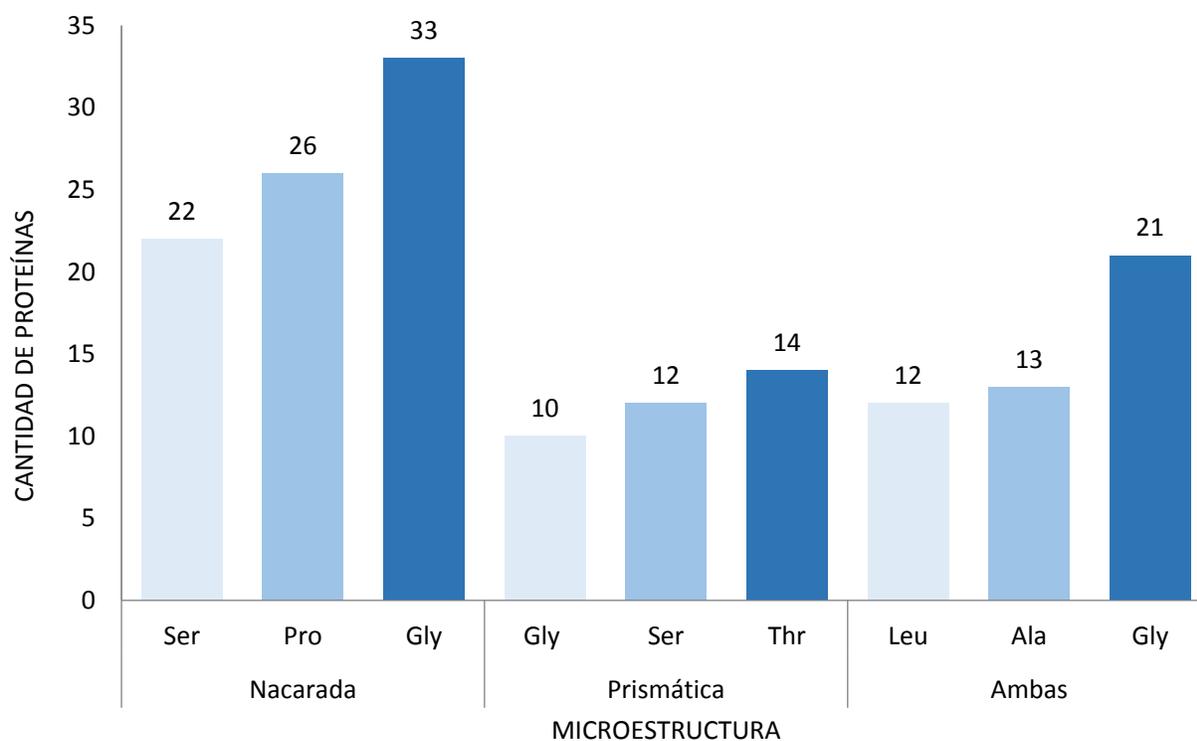


Figura 4. Representación gráfica de la cantidad de proteínas que presentan uno o más aminoácidos dominantes de acuerdo con la microestructura identificada.

7.2.4 Dominios

Mediante el análisis de predicción de dominios, se encontró que de las 554 proteínas 340 presentan al menos un dominio conocido, mientras que en 214 proteínas no se encontró coincidencia con algún dominio reportado. Entre las proteínas de la base de datos, la clasificación de dominios, de acuerdo con su presencia y número, se estableció como dominios más representados y dominios menos representados.

Los **dominios más representados** para la base de datos creada son: A2M (α -2-macroglobulin), Annexin, Gelsolin-like, dominios ricos en glicina (Gly), Ciclofilina PPlasa, RAB, dominios ricos en alanina (Ala), SCP-like (Secreted Cysteine Rich Protein, por sus siglas en inglés), ZP (Zona Pelúcida), EF hand, EGF-like (Epidermal Growth Factor, por sus siglas en inglés), Anhidrasa carbónica, Proteosoma alfa, Small GTP binding, Tioredoxin, WR1 (Worm-specific repeat type 1, por sus siglas en inglés), WAP (Whey Acidic Proteins, por sus siglas en inglés), Ependymin-like y dominio de unión a quitina (Fig. 5).

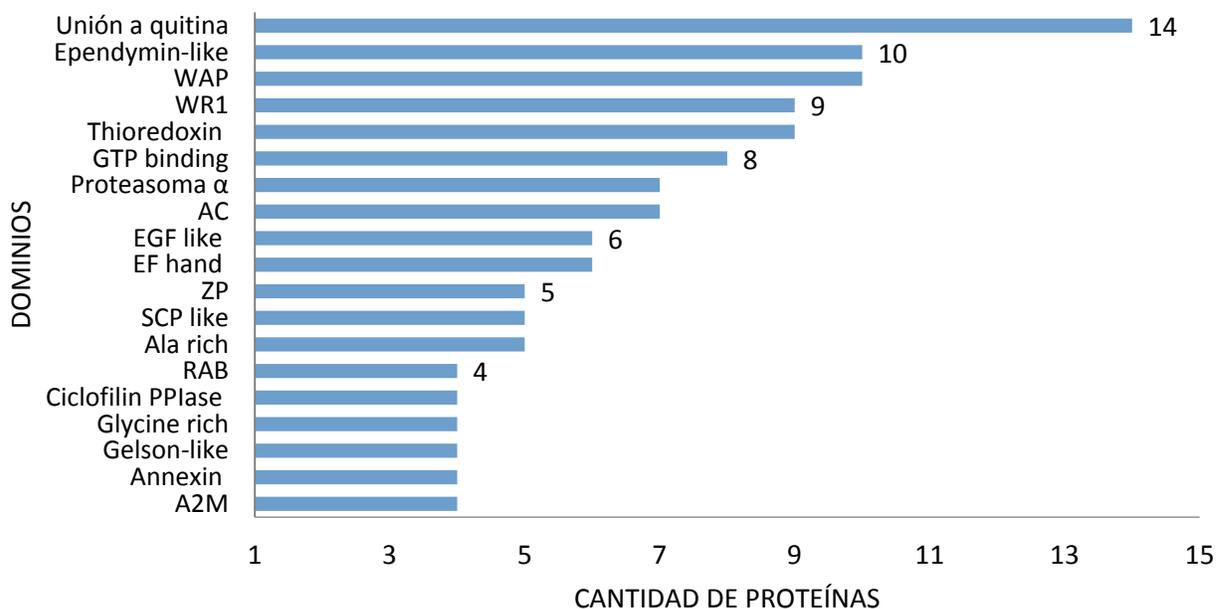


Figura 5. Representación gráfica de la relación entre los dominios de mayor presencia en las proteínas de la base de datos (de un total de 554) y el número de proteínas en las que fueron identificados.

Por otra parte, los **dominios menos representados** fueron: IF Actine, Ferritin-like, G alpha (Guanine nucleotide binding proteins), dominios ricos en ciertos aminoácidos (Pro, Asp, Met, Gln), Zinc metaloproteinasas (ZnMc, por sus siglas en inglés), Inhibidor BPTI Kunitz 1 y 2, GOLD (Golgi dynamics), SUSHI CCP (Complement control protein), WD40 (Trp-Asp ending), ROD (Intermediate filament), SH3/SH2 (Src homology), SEA, eIF5C (Eukaryotic initiation factors), Ubiquitin-like, PHB (Prohibitin homologues), 14_3_3, Peroxidasa, IGF-BP (Insulin-like Growth Factor Binding Proteins), RAS, FABP (Fatty Acid Binding Protein), Cadherina, Vitogelina, ARF (ADP ribosylation factor), 2 ADF-H (Actin depolymerisation factor homology), PKS ER (Polyketide synthase, enoyl reductase), Reelin, Lectina tipo C, Calponin-like, RAS y Plastocyanin-like (Fig. 6).

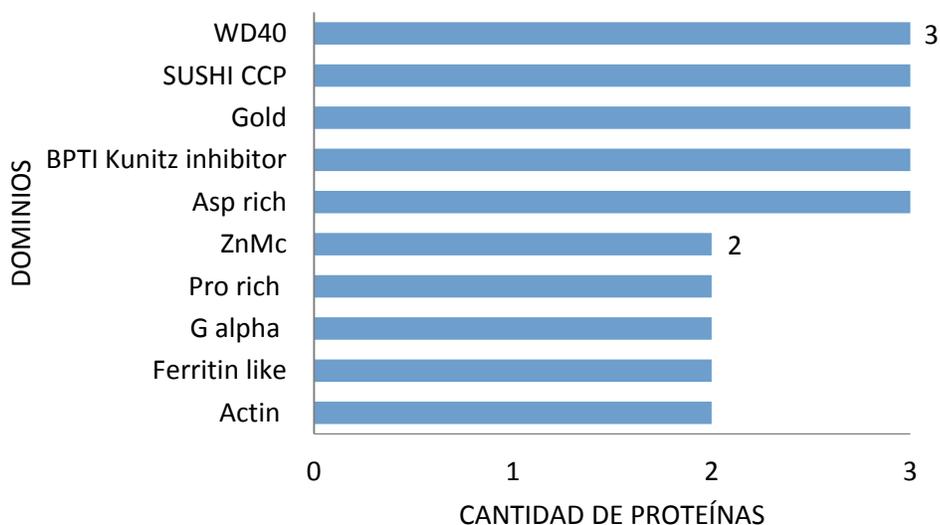


Figura 6. Representación gráfica de la relación entre los dominios de menor presencia en las proteínas de la base de datos (de un total de 554) y el número de proteínas en las que fueron identificados. En el gráfico se muestran únicamente 10 de los 32 dominios de esta clasificación.

7.2.5 Motivos y regiones de baja complejidad

En el análisis de predicción de motivos, se encontró que sólo tres proteínas cuentan con un motivo conocido. Los motivos encontrados fueron: **KOW** (Kyprides, Ouzounis, Woese), motivo de unión a RNA; **SAP**, motivo de unión putativa a ADN; y **EFh**, motivo de unión a calcio. Por otra parte, se encontró que 302 proteínas cuentan con al menos una región de baja complejidad, mientras que 252 proteínas no presentan ninguna región de este tipo.

7.2.6 Modificaciones postraduccionales

El análisis de predicción de modificaciones postraduccionales arrojó que de las 554 proteínas que conforman la base de datos, el 56% presentan posibles sitios de glicosilación, el 100% de las proteínas presentan al menos un posible sitio de fosforilación, y el 8% de las proteínas presentan un puente disulfuro. Por otra parte, se determinó la presencia de modificaciones postraduccionales en los grupos de proteínas con los dominios clasificados como de mayor y menor presencia, según la base de datos generada. En el grupo de proteínas con mayor presencia de dominios, se encontró que todas presentan potenciales sitios de fosforilación; 11 grupos de proteínas cuentan con potenciales N-glicosilaciones (p.e., A2M, SCP like Y ZP); 8

grupos de proteínas cuentan con O-glicosilaciones (p.e., ZP, WR1, WAP), y 3 grupos de proteínas cuentan con potenciales enlaces disulfuro (ZP, EGF-like, Ependymin-like) (Tabla 7).

Tabla 7. Potenciales modificaciones postraduccionales de los grupos de proteínas con dominios de mayor presencia.

Dominio	Fosforilación	Glicosilación		Enlace disulfuro
		N-Glic	O-Glic	
A2M	●	●	○	○
Annexin	●	○	○	○
Gelson-like	●	○	○	○
Glycine rich	●	○	●	○
Ciclofilin PPlase	●	●	○	○
RAB	●	○	○	○
Alanine-rich	●	○	●	○
SCP like	●	●	●	○
ZP	●	●	●	●
EF hand	●	○	○	○
EGF like	●	●	●	●
Anhidrasa carbónica	●	●	○	○
Proteasoma α	●	●	○	○
Small GTP binding	●	○	○	○
Thioredoxin	●	○	○	○
WR1	●	●	●	○
WAP	●	●	●	○
Ependymin-like	●	●	○	●
Unión a quitina	●	●	●	○

Por otro lado, respecto a los grupos de proteínas con dominios de menor presencia, se encontró que todos los grupos cuentan con potenciales sitios de fosforilación; 4 grupos (Actina, ZnMc, Gold y SUSHI CCP) cuentan con posibles N-glicosilaciones, mientras que 4 grupos de proteínas (Proline-rich, inhibidor BPTI Kunitz, Gold y SUSHI CCP) cuentan con posibles sitios de O-glicosilaciones; y solo un grupo (SUSHI CCP) tiene posibles enlaces disulfuro (Tabla 8).

Tabla 8. Potenciales modificaciones postraduccionales de los grupos de proteínas con dominios de menor presencia.

Dominio	Fosforilación	Glicosilación		Enlace disulfuro
		N-Glic	O-Glic	
Actina	●	●	○	○
Ferritin-like	●	○	○	○
G alpha	●	○	○	○
Proline-rich	●	○	●	○
ZnMc	●	●	○	○
Asp rich	●	○	○	○
Inhibidor BPTI Kunitz	●	○	●	○
Gold	●	●	●	○
SUSHI CCP	●	●	●	●
WD40	●	○	○	○

Nota: Únicamente se presentan 10 de los 32 grupos con menor presencia.

7.2.7 Función

Se encontró que de las 544 proteínas que conforman la base de datos, 319 proteínas tienen una función conocida o potencial, mientras que se desconoce la posible función de 235 proteínas.

En un análisis más detallado, se encontró que 179 proteínas cuentan con una posible función enzimática por lo que en la Figura 7 se muestra su clasificación de acuerdo con la comisión de enzimas (E.C., por sus siglas en inglés). Por lo tanto, las proteínas se clasificaron de la siguiente manera: 92 hidrolasas (52.3%), 34 oxidorreductasas (19%), 27 transferasas (14.4%), 11 liasas (5.7%), 11 isomerasas (5.7%) y 5 translocasas (2.9%).

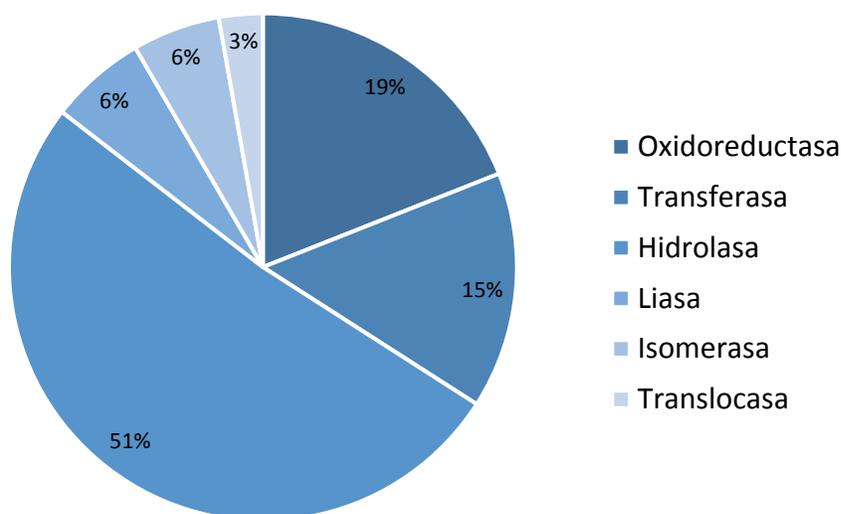


Figura 7. Identificación de proteínas insolubles con potencial función enzimática. Porcentaje de proteínas que tienen potencial actividad catalítica. Los colores representan la clasificación de enzimas.

Además de la potencial función enzimática, se encontró que 132 proteínas cuentan con otras funciones potenciales tales como de unión, de inhibición, y estructurales (Fig. 8). Respecto a las proteínas con posible función de unión, se encontró que 24 de ellas se unen a iones de calcio; 19 a actina; 11 de ellas se unen a otras proteínas; 14 a moléculas de quitina; 21 se unen a otros iones metálicos tales como zinc, cobre, magnesio, hierro o ion férrico, y potasio; 3 proteínas tienen capacidad de unión a ADN y 4 a ATP. Respecto a las proteínas con capacidad de inhibición, 18 de ellas inhiben peptidasas y 3 inhiben endopeptidasas. Por último, respecto a las proteínas estructurales, se encontró que 15 proteínas son constituyentes estructurales de los ribosomas.

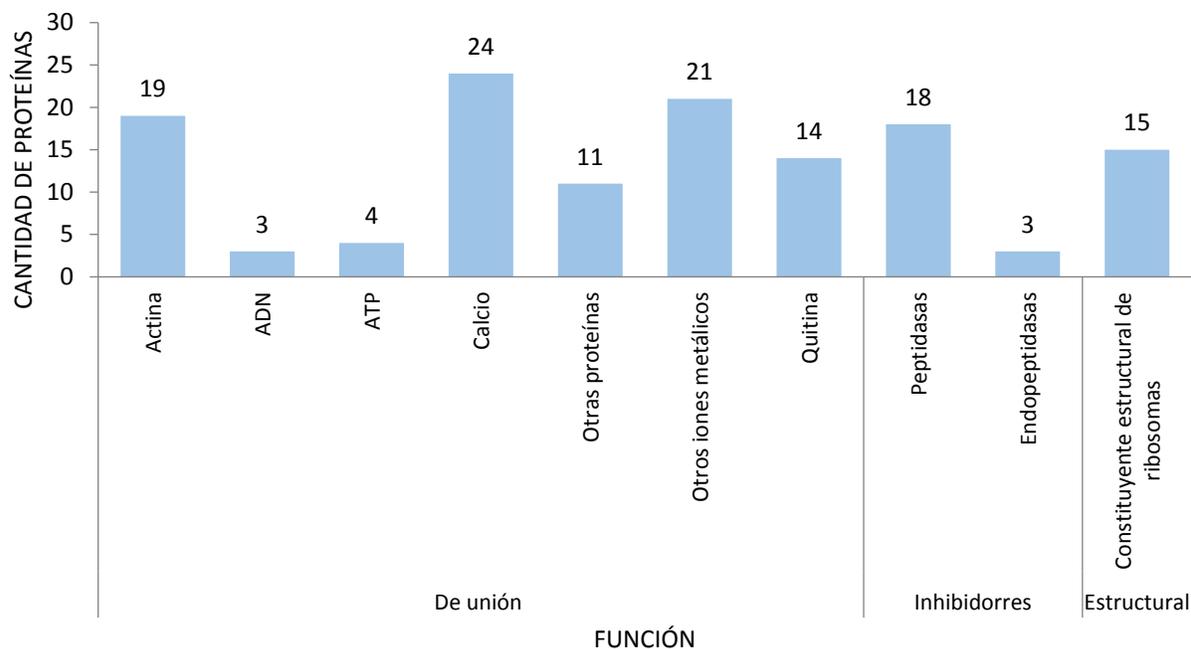


Figura 8. Identificación de proteínas con funciones no catalíticas. Clasificación de proteínas de acuerdo con su posible función con su respectiva cantidad de proteínas correspondiente a cada uno.

7.3 Grupos de proteínas

De acuerdo con el árbol guía, obtenido mediante el uso del programa Clustal Omega, se obtuvieron 35 grupos preliminares. A partir de los árboles filogenéticos que se realizaron por cada grupo preliminar, se obtuvieron 22 grupos de proteínas que presentaron valores de bootstrap mayores al 50% (Fig. 9). Posteriormente, se obtuvieron 10 grupos que están relacionados con la formación de la concha en moluscos: zinc metaloproteasas, EGF-like y zona pelúcida, las anhidrasas carbónicas, ependymin-like, los inhibidores de proteasas, serina/treonina fosfatasas, proteínas con dominio SCP y las proteínas ricas en residuos de prolina, glicina y tirosina.

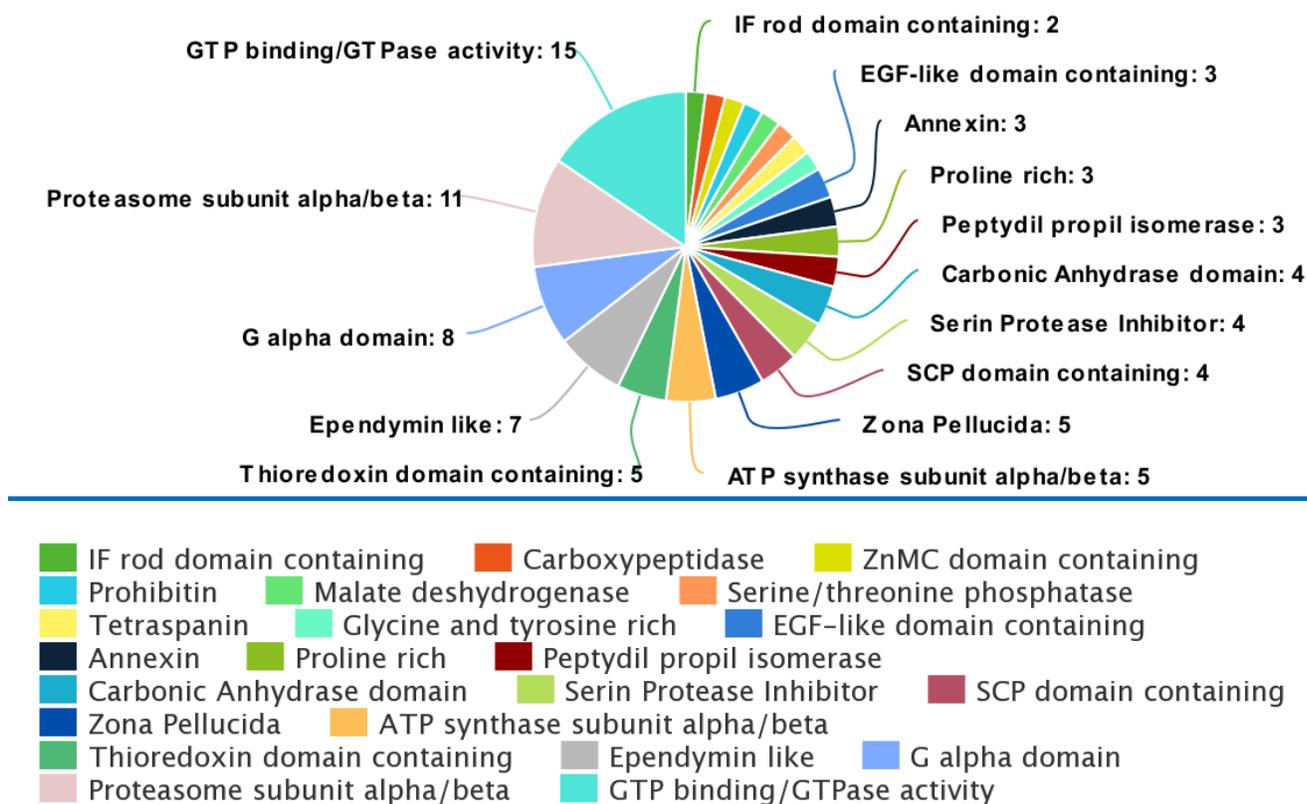


Figura 9. Grupos de proteínas. Se muestra la cantidad de proteínas pertenecientes a cada uno de los grupos obtenidos.

8. DISCUSIÓN

El proceso de biomineralización de moluscos se ha estudiado mediante la caracterización de las proteínas que conforman la matriz orgánica de la concha y que participan en dicho proceso biológico; se han caracterizado proteínas y genes de manera individual, pero también se han realizado estudios de transcriptomas y proteomas de diferentes especies de moluscos (Feng *et al.*, 2017; Mann *et al.*, 2012; Marie *et al.*, 2010; Shi *et al.*, 2019; Werner *et al.*, 2013), siendo los moluscos bivalvos los más estudiados dada su importancia comercial. Debido a que las investigaciones enfocadas en moluscos gasterópodos han sido limitadas, el estudio de las proteínas de la matriz de la concha de este grupo en particular, representa una contribución al estado del arte del proceso de biomineralización. Por lo anterior, se elaboró una base de datos con las secuencias de las proteínas de la matriz de gasterópodos disponibles a la fecha y se realizó un análisis bioinformático para caracterizarlas y determinar su relación con la formación de la concha.

8.1 Proteínas descritas en gasterópodos

Las proteínas formadoras de las microestructuras de las conchas de moluscos (capa nacarada: aragonita, capa prismática: calcita) han sido estudiadas por diversos autores (Michenfelder *et al.*, 2003; Miyamoto *et al.*, 1996; Shen *et al.*, 1997; Samata *et al.*, 1999; Tsukamoto *et al.*, 2004; Weiss *et al.*, 2000; Zhang *et al.*, 2003). Debido a su importancia económica en la industria perlera y a las propiedades mecánicas (resistencia, dureza, rigidez) que les confieren a las conchas de moluscos, la mayoría de los estudios se han dirigido a la obtención de las proteínas formadoras de nácar, tanto en bivalvos como en gasterópodos (Currey, 1999; Smith *et al.*, 1999). Recientemente, Liu y Zhang (2021) reportaron 1024 proteínas de bivalvos y 700 de gasterópodos, derivadas de estudios proteómicos realizados (durante las últimas dos décadas) a partir de las conchas. En este estudio, se analizaron 554 de las proteínas de gasterópodos reportadas en bases de datos (secuencias de proteínas o ácidos nucleicos) que están asociadas a la fracción insoluble resultante de la descalcificación de las conchas. De la base de datos creada a partir de estas proteínas, 118 están relacionadas con la formación de una microestructura definida. La mayoría de las proteínas están relacionadas con la formación de la capa nacarada

(51%), seguido de las proteínas que están presentes en ambas capas (p.e. nacarada y prismática; 29%), y proteínas relacionadas con la formación de la capa prismática (20%).

El resultado de esta distribución puede deberse a diferentes factores, ya sean biológicos o técnicos. Respecto a los factores biológicos, puede deberse a los diferentes arreglos estructurales que se forman (Currey, 1997). Se ha visto que la mayoría de estas estructuras están compuestas por cristales de aragonita; por tanto, hay una mayor cantidad de proteínas que forman y regulan el crecimiento de este tipo de cristales (Chen *et al.*, 2012). También pueden influir factores como el tamaño de la concha, la preservación y el grosor de las capas, los cuales pueden impactar en la obtención del tipo y cantidad de proteínas (Zhang, 2006). Respecto a las causas técnicas, puede deberse a diferentes razones, tales como: a) la búsqueda específica de las proteínas formadoras de la capa nacarada, por lo que la capa prismática es removida (Mann *et al.*, 2012); b) el proceso de extracción de las proteínas, ya que al tratarlas para remover contaminantes de las conchas, generalmente con hipoclorito de sodio, la capa prismática se ve afectada (Mann *et al.*, 2018). Por otra parte, el hecho de que para las 436 proteínas restantes no esté definido un tipo estructura particular, se debe a que estas no forman estructuras nacaradas como lo es el caso de *Lottia gigantea* (Marie *et al.*, 2012); o debido a que los estudios realizados para su obtención tienen como base el enfoque proteómico con el que se busca encontrar, de forma general, todas las proteínas presentes en la concha (Marie *et al.*, 2012).

8.2 Implicaciones de las modificaciones postraduccionales en SMPs

Las modificaciones postraduccionales (PTM, por sus siglas en inglés) son cambios químicos que inducen una diversidad de alteraciones tanto estructurales como funcionales, y son un mecanismo de regulación de la actividad biológica de las proteínas (Pérez-Sala, 2012). En las proteínas de la matriz orgánica de las conchas, las PTMs más comunes son glicosilación, fosforilación y formación de enlaces disulfuro, las cuales podrían jugar un papel relevante en el proceso de biomineralización de las conchas (Nudelman *et al.*, 2006).

Mediante diversos estudios tanto *in silico* como *in vitro*, se ha visto que las glicosilaciones son las modificaciones más abundantes en las SMPs (Mann *et al.*, 2000; Marie *et al.*, 2013;

Michenfelder *et al.*, 2003; Treccani *et al.*, 2006; Weiss *et al.*, 2000). Algunas proteínas aisladas de la fracción insoluble que presentan este tipo de modificación son *Perlina* (Montagnani *et al.*, 2011), *Asp-rich* (Marie *et al.*, 2013) y *Ps19* (Arroyo-Loranca *et al.*, 2020). Dentro del proceso de biomineralización de la concha, algunas funciones que se han atribuido a las glicoproteínas son: la orientación de los cristales que se forman (Addadi *et al.*, 1987; Marie *et al.*, 2007), ser un mediador del soporte orgánico insoluble para que se lleve a cabo la nucleación (Albeck *et al.*, 1996), y la selección del polimorfismo, ya sea para formar calcita o aragonita (Falini *et al.*, 1996). Dadas estas funciones, se puede decir que esta modificación se asocia con la regulación a nivel estructural por lo cual resulta ser de gran importancia para la formación de la concha. En la predicción de glicosilaciones de las 554 proteínas, se encontró que el 56% tienen al menos un sitio potencial. Se puede interpretar que poco más de la mitad de las proteínas de esta base de datos podrían estar relacionadas con la regulación estructural de la formación de la concha. Por otra parte, si bien es posible la predicción de posibles glicosilaciones, los análisis *in silico* no determinan con exactitud la cantidad de sitios respecto a las proteínas nativas, ya que debido al plegamiento de sus estructuras no es posible la detección precisa, al igual que la determinación de la complejidad de los glucanos. Sin embargo, este tipo de estudios son una aproximación muy útil para estudiar y caracterizar proteínas (Nakae *et al.*, 2018).

Al igual que las glicosilaciones, las fosforilaciones son modificaciones comunes en las SMPs, sin embargo, las proteínas con esta característica que se han identificado en estudios *in vitro* e *in vivo* han sido pocas (Álvarez *et al.*, 2009; Du *et al.*, 2018; Mann *et al.*, 2010; Kong *et al.*, 2019). Algunas de las funciones atribuidas a las fosfoproteínas, según el nivel de fosforilación, son: la regulación del crecimiento de los cristales de carbonato de calcio (Borbas *et al.*, 1991), alteraciones en la nucleación, tamaño y morfología de los cristales (Du *et al.*, 2018), y la atracción de calcio por parte de los grupos fosforilados, lo que facilita la interacción entre proteínas y cristales, lo que ocasiona la inhibición del crecimiento de los cristales (Du *et al.*, 2018). En este estudio, la predicción de sitios de fosforilación arrojó que las 554 proteínas cuentan con al menos uno de ellos. Con ello, se podría decir que es una modificación importante en las SMPs, y que su presencia podría influenciar directamente la regulación del proceso de biomineralización. Esta modificación podría estar regulando en sí la actividad de las

proteínas, esto por medio de la activación o inactivación, lo que contribuiría evitando que se sintetizen más proteínas. Al estar presentes en todas las proteínas, nos habla de un nivel de regulación muy alto debido al complicado proceso de formación de las conchas.

La formación de enlaces disulfuro es una modificación postraducciona, muy importante, involucrada en el plegamiento, la estabilización e integridad de la estructura terciaria de las proteínas (Servier y Kaiser, 2002). Esta modificación se ha encontrado en algunas SMPs relacionadas a dominios que contienen residuos de cisteínas. Algunos de estos dominios son: Lectina tipo C (Mann *et al.*, 2000), dominio de unión a quitina (Miyamoto *et al.*, 2013), dominio WAP o dominio de proteína ácida de suero (Marie *et al.*, 2010, Treccani *et al.*, 2006), Dominio EGF o factor de crecimiento epidérmico (Marie *et al.*, 2013), por mencionar algunos. Se ha reportado que la formación de enlaces disulfuro, así como su reversibilidad, podrían regular la estructura de las proteínas para lograr una función en particular asociada a la formación de la concha, como lo es controlar la nucleación mineral (Rivera-Pérez, 2020). En este trabajo, al realizar el análisis de predicción de enlaces disulfuro, se encontró que 44 proteínas presentan al menos una de estas modificaciones. Al analizar las secuencias de estas proteínas, se encontraron los siguientes dominios con residuos de cisteína: WAP y Kunitz BPTI, dominios inhibidores de proteasas; C-lectina, EGF-like y ependymin-like, dominios de unión a calcio; thioredoxin, dominio oxidoreductasa; SRCR (Scavenger Receptor Cysteine Rich), dominio de unión a ligandos específicos; dominios de unión a quitina, así como regiones repetidas de cisteínas (Regiones de Baja Complejidad). Podría decirse que esta modificación confiere estabilidad a proteínas que cumplen funciones importantes asociadas a la formación de las conchas, evitando su degradación.

El péptido señal es una secuencia corta de aminoácidos ubicada en el extremo N-terminal de las proteínas. Su función es marcar la vía de secreción y dar una dirección destino específica, para después ser eliminada una vez que se cumpla su función de direccionamiento (Heijne, 1990; Owji *et al.*, 2018). Esta señalización dirige a las proteínas que tienen funciones específicas en sitios diferentes al citosol, por lo que no todas las proteínas cuentan con uno (Alberts *et al.*, 2002). En particular, las proteínas extracelulares tienen péptidos que les permiten salir de la

célula, como es el caso de las SMPs. Las células del epitelio externo del manto secretan estas proteínas hacia el fluido extrapaleal, en donde se encuentra la matriz orgánica responsable de la formación de las capas de la concha (Marin y Luquet, 2007). De las secuencias analizadas en este estudio, se determinó el péptido señal del 64% de las proteínas, mientras que el 36% no presentaron esta característica. La obtención de esta información está sujeta a alguno de los métodos de obtención de las proteínas (utilizados por los autores que reportaron las secuencias). Por ejemplo, en estudios proteómicos, en donde las muestras (p.e. conchas) son pulverizadas y desmineralizadas, se obtiene un gran número de secuencias de proteínas expresadas y funcionales, por lo que estas proteínas ya no presentan péptido señal (Mann *et al.*, 2012). Por otra parte, las secuencias reportadas pueden presentar péptido señal cuando se obtienen péptidos a partir de las conchas, que luego serán comparados con una base de datos de secuencias de nucleótidos o transcritos que se traducen previamente a proteínas; o bien, comparados contra bases de datos de proteínas o librerías genómicas (Marie *et al.*, 2010; 2012; Mann *et al.*, 2018). Otro método común es la deducción de la secuencia aminoacídica de las proteínas mediante clonación del ADN complementario, en donde el péptido señal puede observarse (Jackson *et al.*, 2017; Shen *et al.*, 1997). Conocer la presencia o ausencia de esta región es importante al momento de realizar la alineación múltiple, permitiendo un alineamiento más acertado, el cual ayuda a determinar el grado de homología entre secuencias, ya sea dominios o regiones conservadas, que contribuyan a la caracterización de las proteínas.

8.3 Implicaciones de la carga de las SMPs

El punto isoeléctrico de una proteína es el pH al que la proteína no tiene carga neta (Xia, 2007). Esta característica ha sido objeto de estudio en las SMPs ya que se ha relacionado con la formación de los polimorfos de carbonato de calcio y, por tanto, con una microestructura específica (Zhang y Zhang, 2006). Debido a esto, se han determinado parámetros para caracterizar a las proteínas según la microestructura en la que se ubican, siendo la siguiente la clasificación más común: proteínas altamente ácidas ($pI < 4.5$), proteínas moderadamente ácidas ($4.5 \leq pI \leq 7$), y proteínas básicas ($pI > 7$) (Marin *et al.*, 2007). Las proteínas extremadamente ácidas se han reportado, en su mayoría, en la capa prismática de calcita, mientras que las proteínas moderadamente ácidas y las básicas se han encontrado generalmente en la capa nacarada de aragonita (Marin *et al.*, 2007; Furashi *et al.*, 2010). En este estudio, se utilizaron dos

parámetros (proteínas ácidas $pI < 7$ y proteínas básicas $pI \geq 7$) para clasificar a las 554 proteínas. Se encontró que el 60% son ácidas ($pI < 7$), mientras que el 40% son básicas ($pI \geq 7$). Sin embargo, al analizar 118 proteínas de acuerdo con su microestructura (aquellas proteínas en las que se reportó una microestructura particular, ver Figs. 2 y 4), se encontró que los rangos establecidos por Marin *et al.* (2007) no se cumplen en su mayoría. De 60 proteínas relacionadas a la formación de nácar, 54 tienen valores mayores a 4.5; mientras que, de 35 proteínas relacionadas a la formación de la capa prismática, solo dos se clasifican como extremadamente ácidas. Por otra parte, de las 34 proteínas encontradas en ambas capas, 2 son extremadamente ácidas, 13 moderadamente ácidas, y 19 son básicas. Por tanto, sólo se puede decir que la relación entre un punto isoeléctrico moderadamente ácido y básico y la capa nacarada se cumple. Sin embargo, el parámetro de las proteínas de la capa de calcita no se obedece. Se puede argumentar entonces que los parámetros establecidos ya no son tan acertados y que no aplican para una cantidad grande de proteínas. Los rangos establecidos son ambiguos, ya que las proteínas no muestran un patrón claro y definido que pueda determinarse como universal para todo el conjunto de proteínas de la matriz orgánica de las conchas.

8.4 Implicaciones de la abundancia de ciertos aminoácidos

Las SMPs se caracterizan por tener predominancia de entre dos a cuatro aminoácidos. Los aminoácidos más comunes son glicina, ácido aspártico, serina y, en menor medida, prolina, cisteína, tirosina, leucina y asparagina (Marin y Luquet, 2004).

Por una parte, ciertos aminoácidos se relacionan con las proteínas de las fracciones que se obtienen después de la desmineralización. Respecto a las proteínas de la fracción soluble, estas se caracterizan por ser ácidas, ricas en ácido aspártico; mientras que las proteínas de la matriz insoluble generalmente son hidrófobas, ricas en alanina y glicina (Marin y Luquet, 2004). En particular, las proteínas de la fracción insoluble son las que resultan de interés, dada la falta de información que existe y su potencial función en la formación de nácar. En el análisis de diversas proteínas de esta fracción se han obtenido resultados variados. En general, se ha visto que la glicina es el aminoácido dominante en las proteínas de la fracción insoluble, con una presencia del 30% (Marie *et al.*, 2012).

Por otra parte, se ha encontrado relación de ciertos aminoácidos con las microestructuras de la concha, algunos autores sostienen que los aminoácidos dominantes de la capa nacarada son alanina y glicina (Akiyama, 1966; Grégoire, 1972; Hare, 1963); otros, que es rica en alanina, asparagina o ácido aspártico (Marie *et al.*, 2012). Por otra parte, los aminoácidos dominantes que se han observado en las proteínas insolubles de la capa prismática son tirosina, prolina y valina (Marie *et al.*, 2012); mientras que otros autores han observado que la glicina es el aminoácido dominante (Marin *et al.*, 2007; Sudo *et al.*, 1997; Suzuki *et al.*, 2004).

En el análisis de la composición de aminoácidos de las 554 proteínas pertenecientes a la fracción insoluble, se encontraron como aminoácidos dominantes a: la treonina, valina, alanina, lisina, serina, glicina y leucina. Respecto a las proteínas relacionadas a una microestructura, se encontró que los aminoácidos dominantes de las proteínas de la capa nacarada son serina, glicina y prolina; las proteínas de la capa prismática son ricas en glicina, serina y treonina; mientras que las proteínas relacionadas a ambas capas son ricas en leucina, alanina y glicina (Fig. 4). Por tanto, se observaron 3 aminoácidos hidrofóbicos, 3 hidrófilos y 1 básico. Los resultados concuerdan con la información reportada por diferentes autores, en donde la glicina es el aminoácido más abundante en las proteínas de la fracción insoluble. Así mismo, respecto a las microestructuras, solo la dominancia de la glicina en ambas capas coincide con las observaciones reportadas por diferentes autores (Grégoire, 1972; Marin *et al.*, 2007; 2012; Suzuki *et al.*, 2004). La glicina es un aminoácido pequeño, hidrofóbico, y se ha observado como una ventaja en la estructura de las proteínas ya que confiere flexibilidad entre estructuras secundarias (Hall, 1998). En general, la presencia de los aminoácidos dominantes se ve relacionada con las regiones específicas denominadas regiones o dominios de baja complejidad las cuales se han asociado con el proceso de biomineralización.

8.5 Implicaciones de las regiones de baja complejidad

Diferentes funciones relacionadas a la formación de la concha, tales como promover o inhibir la formación de cristales, se han asociado a las regiones de baja complejidad de las proteínas de los moluscos (Marin *et al.*, 2020). Aproximadamente el 30% de las proteínas de la matriz orgánica identificadas en *Lottia*, *Haliotis* y *Pinctada* contienen tales repeticiones (Kocot *et al.*, 2012). Se ha visto que las proteínas de la matriz orgánica que contienen LCR son especie-

específicas, por tanto, se puede pensar que han evolucionado de forma independiente en los diferentes modelos de moluscos (Marie *et al.*, 2012). El origen y la evolución de este tipo de proteínas son poco claros, sin embargo, su presencia en una gran cantidad de proteínas, así como su gran diversidad, sugiere que son componentes importantes del secretoma del manto de los moluscos (Kocot *et al.*, 2012). Existe la probabilidad de que los mecanismos de evolución de las conchas sean altamente dinámicos, caracterizándose por expansiones independientes de familias de genes, cambios de dominios y deslizamiento de la replicación; lo cual explicaría la gran diversidad de proteínas, estructuras y formas de las conchas entre los diferentes organismos que comparten un tejido homólogo evolutivo: el manto (Jackson *et al.*, 2009; Kocot *et al.*, 2012).

Diferentes autores sostienen la teoría de que las proteínas que contienen regiones intrínsecamente desordenadas, como son las proteínas que contienen LCR, no pueden estudiarse individualmente, ya que las diferentes funciones son llevadas a cabo en conjunto. Por lo tanto, sugieren estudiar el proceso de biomineralización a nivel supramolecular y no con el enfoque molecular “uno por uno” (Marin *et al.*, 2007; Pancsa *et al.*, 2019). En este estudio, se observó la presencia de regiones de baja complejidad en 302 proteínas de gasterópodos, de las cuales 82 están asociadas a una o ambas capas de la concha (nácar 48, calcita 13 y ambas 25). En el análisis de estas regiones, al ser tan variadas, no se encontró ninguna asociación hacia algún aminoácido en específico. Para el estudio particular de estas regiones, se requiere un análisis más detallado usando una herramienta que arroje resultados de un grupo grande de proteínas en un tiempo corto.

8.6 Implicaciones de los dominios en la formación de la concha

Además de la formación de la concha, las proteínas pueden tener diversas funciones involucradas directa o indirectamente en el proceso de biomineralización. Para estudiar la función de las proteínas, es más factible hacerlo mediante la asociación a dominios funcionales de acuerdo con lo observado en diferentes grupos de SMPs ya caracterizadas. Las clasificaciones más comunes son: dominios con función catalítica, de protección o inmunidad, de unión a sacáridos o cationes divalentes, dominios de matriz extracelular o moléculas de unión a matrices extracelulares, y dominios con regiones de baja complejidad (Marin, 2020).

En un estudio reciente se encontró que los principales dominios reportados en SMPs son: dominios con LCR (glicina, lisina, metionina, asparagina, tirosina, treonina, ácido aspártico, glutamina, cisteína, valina, prolina, serina, alanina, y leucina); dominios extracelulares como von Willebrand factor tipo A (vWA, por sus siglas en inglés), vWD, fibronectin tipo III (FN3, por sus siglas en inglés), SUSHI CCP, Laminin_G, factor de crecimiento epidérmico (EGF, por sus siglas en inglés), thrombospondin, dominios transmembranales, SRCR (Scavenger Receptor CYsteine Rich), Receptor LDL, Zona Pellucida y Protocadherin; dominios de carácter inmunológico como alpha-2-macroglobulin (A2M), EF-hand; dominios de interacción con polisacáridos como unión a quitina, lectina de tipo C y unión a galactosa; dominios con función enzimática: anhidrasa de carbónica (CA), peroxidasa, quitinasa, tirosinasa, oxidorreductasa, glucocinasa, lacasa; inhibidores de proteasas como inhibidor de serina proteasa, Kunitz-like, Kazal-like, proteína ácida de suero (WAP), metaloproteasa e inhibidor de tripsina (Liu y Zhang, 2021).

En este trabajo, mediante el análisis de predicción de dominios, se encontró que 340 proteínas presentan al menos un dominio conocido, de las cuales, 251 tienen una función asociada. Los dominios más destacados en la clasificación enzimática son los de tipo: a) proteasa (e.g. SCP-like) los cuales podrían estar involucrados en la degradación, remodelación y maduración de la matriz; b) peroxidasas (p.e. Thioredoxin), las cuales podrían estar implicadas en la formación del periostraco con el objetivo de formar polímeros insolubles resistentes a las proteasas (Herlitze *et al.*, 2018); y c) las anhidrasas carbónicas, las cuales están involucradas en la conversión de dióxido de carbono en bicarbonato (Miyamoto *et al.*, 1996; Le Roy *et al.*, 2014). Respecto a los dominios de inmunidad, se encontró el Alpha-2-macroglobulin (A2M), el cual podría estar relacionado con el mecanismo de protección de las conchas (Arivalagan *et al.*, 2016). Por otra parte, los dominios destacados en la clasificación de unión a cationes divalentes son EF-hand y de tipo ependimina, los cuales se unen a calcio; si bien no se conoce su función, podrían estar involucrados en el suministro de precursores catiónicos del carbonato de calcio o bien, en la señalización o interacción con otros miembros de la matriz (Huang *et al.*, 2007; Marin *et al.*, 2020). Respecto a los dominios de unión a sacáridos, el más destacado es el de unión a quitina, ya que juega un papel esencial en la estructura de la concha (Weiss *et al.*, 2019). Por último, los

dominios más destacados en la clasificación de matriz extracelular son Zona Pelúcida y EGF-like (Epidermal Growth Factor, por sus siglas en inglés), los cuales juegan un papel en la estructuración del marco tridimensional, ensamblándose en láminas, fibras o geles (Jovin *et al.*, 2005; Marin, 2020). Estos resultados concuerdan con los principales dominios encontrados en las proteínas de la matriz orgánica, y tienen una fuerte relación con el proceso de biomineralización dada la funcionalidad que se les atribuye. En general, podemos decir que las proteínas que se encuentran en las conchas no sólo son responsables de formar las microestructuras, sino que también pueden tener diferentes funciones que están relacionadas indirectamente con el proceso de biomineralización, tales como regulación de la matriz orgánica, protección, señalización, suministro de precursores, entre otras.

8.7 Grupos de proteínas

Se han identificado y descrito una cantidad considerable de proteínas de la matriz orgánica de las conchas de moluscos (Clarck *et al.*, 2020; Evans, 2019; Kocot *et al.*, 2016; Liu *et al.*, 2021; Marin *et al.*, 2007; Murdock *et al.*, 2020; Song *et al.*, 2019). Con el fin de asociar grandes repertorios de secuencias de estas proteínas a funciones relacionadas con el proceso de biomineralización, se han utilizado diversas herramientas de alineamiento, sin embargo, la poca homología entre secuencias ha dificultado el trabajo (Marin *et al.*, 2020).

A pesar de no poder caracterizar grandes grupos de proteínas, es posible caracterizar grupos reducidos. En este estudio, al tener un conjunto grande de proteínas, se realizó la división por grupos de acuerdo con los clados obtenidos en un árbol guía (no se muestra). De acuerdo con los parámetros aplicados para el análisis, se obtuvieron 10 grupos de proteínas que tienen cierta homología y que están relacionados con el proceso de biomineralización. Estos grupos se caracterizan por tener dominios o funciones que se encuentran dentro de las más destacadas dentro de las proteínas de la matriz orgánica de las conchas (Liu *et al.*, 2021; Marin *et al.*, 2020). Los grupos de proteínas y las funciones o características asociadas respecto al proceso de biomineralización son: zinc metaloproteasas, los cuales regulan la composición de la matriz orgánica (Rawlings y Barrett, 1995); proteínas EGF-like y ependymin-like, las cuales tienen dominios de unión a calcio (Appella *et al.*, 1988); zona pelúcida, el cual es un dominio de unión a la matriz extracelular (Jovine *et al.*, 2002; 2005); las anhidrasas carbónicas, las cuales catalizan

la formación de bicarbonato (Deutsch, 1987); los inhibidores de proteasas, inhiben la degradación de enzimas como método de regulación y protección de las proteínas encargadas de la formación de la concha (Bedouet 2007); serino/treonina fosfatasas, que son reguladores de la actividad de las proteínas formadoras de cristales (Cohen *et al.*, 1990); proteínas con dominio SCP, las cuales son inhibidores de proteasas y regulan la matriz extracelular (Gibbs *et al.*, 2008); y las proteínas con regiones ricas en prolina, glicina y tirosina, las cuales pueden ser dominios de unión para el comienzo de la nucleación, como marcos estructurales (Shen *et al.*, 1997; Zhang y Zhang, 2006; Zheng *et al.*, 2020).

En estos 10 grupos de proteínas, 36 proteínas en total, se observa la diversidad de funciones que estas presentan y como se relacionan para poder llevar a cabo el proceso de formación de la concha de manera regulada. Estas proteínas presentan funciones tanto directas como indirectas (es decir, que no forman cristales de aragonita o calcita, pero están involucradas en el proceso previo a su formación) que pueden considerarse como representativas y fundamentales. Por una parte, están las proteasas y los inhibidores de proteasas, importantes para regular los componentes de la matriz orgánica de la concha de los moluscos, ya sea permitiendo o inhibiendo que las proteínas actúen según las necesidades del organismo. Si bien no se conoce la proporción, debe haber un equilibrio entre estos dos tipos de proteínas para asegurar un correcto funcionamiento (Liu *et al.*, 2015). Otro regulador importante son las fosfatasas, las cuales están involucradas directamente en la actividad de las proteínas, activando funciones importantes tales como el control de la nucleación, el polimorfismo, la morfología y la organización de los cristales de carbonato de calcio durante la formación de la concha (Du *et al.*, 2018).

Las proteínas con dominios de anhidrasa carbónica, por su parte, son intermediarias importantes para la formación de carbonato de calcio. Además, se cree que la funcionalidad de las proteínas podría variar entre especies, ya que en diferentes estudios se ha visto que tienen tanto la capacidad de inhibir la precipitación del carbonato de calcio como la propiedad de calcificación sobresaliente de la concha (Kono *et al.*, 2000; Miyamoto *et al.*, 1996; 2005). Si bien

la función de este dominio no está completamente definida, es claro que su interacción con el calcio (Ca^{2+}) resulta de importancia para el proceso de biomineralización (Zhang et al., 2006).

Por otra parte, están las proteínas involucradas en diferentes procesos biológicos que implican la señalización y el transporte a nivel célula/matriz o bien, en la estructuración del marco de la concha (Marin 2020; Liu *et al.*, 2021).

9. CONCLUSIONES

El análisis de las proteínas de la matriz mediante análisis bioinformático de este estudio permitió dar una visión general de las proteínas insolubles de los moluscos gasterópodos, ya que no existen muchos estudios específicos, en cuanto a proteínas insolubles, que contribuyan a determinar el conjunto de proteínas que intervienen en la formación de la concha.

De forma particular, dada la diversidad de microestructuras, existen una gran diversidad de proteínas que las forman, en particular las relacionadas con la formación del nácar. Por tanto, la formación de la concha de los moluscos es un proceso muy complejo que necesita altos niveles de regulación ya sea por medio de modificaciones postraduccionales o por proteínas con funciones diferentes. Con los nuevos descubrimientos de proteínas tan diversas, es necesario una reestructuración de las clasificaciones. En particular, la de punto isoeléctrico, ya que los parámetros establecidos ya no se adaptan a un conjunto grande de proteínas y debe delimitarse a grupos pequeños, ya sea entre clases o especies. Por otra parte, a pesar de que las regiones de baja complejidad han sido poco estudiadas, se ha visto que su función resulta ser muy importante para la nucleación y desarrollo de los cristales, proceso inicial de la formación de la concha, por lo que son un área de oportunidad para poder entender el proceso. Respecto a los dominios sin una relación conocida o evidente con el proceso de biomineralización, resulta importante analizarlos para poder hipotetizar cuál sería su función en este proceso.

Si bien todas las proteínas de la base de datos fueron aisladas de la concha (por diferentes autores), estas no tienen funciones exclusivas de formación de cristales. Hay muchas otras funciones que están implicadas en el proceso de biomineralización como protección, inmunidad, señalización, regulación de la matriz orgánica, suministro de precursores, entre otras. Finalmente, una base de datos de proteínas, en este caso de la fracción insoluble, favorece la identificación de nuevas proteínas que no tienen una función asignada.

10. LITERATURA CITADA

Addadi, L., Joester, D., Nudelman, F. y Weiner, S. (2006). Mollusk shell formation: A source of new concepts for understanding biomineralization processes. *Chemistry - A European Journal*, 12(4), pp. 980–987. doi: 10.1002/chem.200500980.

Akiyama, M. (1966). Conchiolin-Constituent Amino Acids and Shell Structures of Bivalve Shells. *Proceedings of the Japan Academy*, 42(7), pp. 800–805. doi: 10.2183/pjab1945.42.800.

Albeck, S., Addadi, L. y Weiner, S. (1996). Regulation of Calcite Crystal Morphology by Intracrystalline Acidic Proteins and Glycoproteins. *Connective Tissue Research*, 35(1–4), pp. 365–370. doi: 10.3109/03008209609029213.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. y Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37(4), pp. 420–423. doi: 10.1038/s41587-019-0036-z.

Alvares, K., Dixit, S. N., Lux, E. y Veis, A. (2009). Echinoderm Phosphorylated Matrix Proteins UTMP16 and UTMP19 Have Different Functions in Sea Urchin Tooth Mineralization. *Journal of Biological Chemistry*, 284(38), pp. 26149–26160. doi: 10.1074/jbc.M109.024018.

Amos, F. F. y Evans, J. S. (2009). AP7, a Partially Disordered Pseudo C-RING Protein, Is Capable of Forming Stabilized Aragonite in Vitro. *Biochemistry*, 48(6), pp. 1332–1339. doi: 10.1021/bi802148r.

Appella, E., Weber, I. T. y Blasi, F. (1988). Structure and function of epidermal growth factor-like regions in proteins. *FEBS Letters*, 231(1), pp. 1–4. doi: 10.1016/0014-5793(88)80690-2.

Arivalagan, J., Marie, B., Sleight, V. A., Clark, M. S., Berland, S. y Marie, A. (2016). Shell matrix proteins of the clam, *Mya truncata*: Roles beyond shell formation through proteomic study. *Marine Genomics*, 27, pp. 69–74. doi: 10.1016/j.margen.2016.03.005.

Arroyo-Loranca, R. G., Hernandez-Saavedra, N. Y., Hernandez-Adame, L. y Rivera-Perez, C. (2020). Ps19, a novel chitin binding protein from *Pteria sterna* capable to mineralize aragonite plates in vitro. *PLOS ONE*. Editado por G. J. Vermeij, 15(3), p. e0230431. doi: 10.1371/journal.pone.0230431.

Bahn, S. Y., Jo, B. H., Choi, Y. S. y Cha, H. J. (2017). Control of nacre biomineralization by Pif80 in pearl oyster. *Science Advances*, 3(8), p. e1700765. doi: 10.1126/sciadv.1700765.

Bahn, S. Y., Jo, B. H., Hwang, B. H., Choi, Y. S. y Cha, H. J. (2015). Role of Pif97 in Nacre Biomineralization: In Vitro Characterization of Recombinant Pif97 as a Framework Protein for the Association of Organic-Inorganic Layers in Nacre. *Crystal Growth and Design*, 15(8), pp. 3666–3673. doi: 10.1021/acs.cgd.5b00275.

Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., *et al.* (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), pp. D480–D489. doi: 10.1093/nar/gkaa1100.

Bédouet, L., Duplat, D., Marie, A., Dubost, L., Berland, S., Rousseau, M., Milet, C. y Lopez, E. (2007). Heterogeneity of proteinase inhibitors in the water-soluble organic matrix from the oyster nacre. *Marine Biotechnology*, 9(4), pp. 437–449. doi: 10.1007/s10126-007-7120-y.

Blom, N., Gammeltoft, S. y Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of molecular biology*, 294(5), pp. 1351–62. doi: 10.1006/jmbi.1999.3310.

Brown, A. H., Rodger, P. M., Evans, J. S. y Walsh, T. R. (2014). Equilibrium Conformational Ensemble of the Intrinsically Disordered Peptide n16N: Linking Subdomain Structures and Function in Nacre. *Biomacromolecules*, 15(12), pp. 4467–4479. doi: 10.1021/bm501263s.

Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, y P. W. (2002). *Molecular Biology of the Cell*. 4th ed. New York: Garland Science.

Brusca, G. J. y Brusca, R. C. (2003). *Invertebrates*. 2nd ed. Cariolou, M. A. y Morse, D. E. (1988) “Purification and characterization of calcium-binding conchiolin shell peptides from the mollusc, *Haliotis rufescens*, as a function of development”, *Journal of Comparative Physiology B: Biochemical, Systemic and Environmental Physiology*, 157(6), pp. 717–729. doi: 10.1007/BF00691002.

Ceroni, A., Passerini, A., Vullo, A. y Frascioni, P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Research*, 34(Web Server), pp. W177–W181. doi: 10.1093/nar/gkl266.

Chen, P. Y., McKittrick, J. y Meyers, M. A. (2012). Biological materials: Functional adaptations and bioinspired designs. *Progress in Materials Science*, 57(8), pp. 1492–1704. doi: 10.1016/j.pmatsci.2012.03.001.

Clark, M. S., Peck, L. S., Arivalagan, J., Backeljau, T., Berland, S., Cardoso, J. C. R., Caurcel, C., Chapelle, G., De Noia, M., Dupont, S., Gharbi, K., Hoffman, J. I., Last, K. S., Marie, A., Melzner, F., Michalek, K., Morris, J. y Power, D. M. (2020). Deciphering mollusc shell production: the roles of genetic mechanisms through to ecology, aquaculture and biomimetics”, *Biological Reviews*, 95(6), pp. 1812–1837. doi: 10.1111/brv.12640.

Cohen, P., Brewis, N. D., Hughes, V. y Mann, D. J. (1990). Protein serine / threonine an expanding family. *Febs*, 268(2), pp. 355–359. Crichton, R. (2019) “Biom mineralization”, *Biological Inorganic Chemistry*, pp. 517–544. doi: 10.1016/b978-0-12-811741-5.00019-9.

Currey, J. D. (1999). The design of mineralised hard tissues for their mechanical functions. *The Journal of experimental biology*, 202(Pt 23), pp. 3285–94. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/10562511>. [Consulta: 15 de mayo 2021]

Currey, J. D. (1977). Mechanical properties of mother of pearl in tension. *Proceedings of the Royal Society of London - Biological Sciences*, 196(1125), pp. 443–463. doi: 10.1098/rspb.1977.0050.

Cusack, M. y Freer, A. (2008). Biomineralization: Elemental and organic influence in carbonate systems. *Chemical Reviews*, 108(11), pp. 4433–4454. doi: 10.1021/cr078270o.

Deutsch, H. F. (1987). Carbonic anhydrases. *International Journal of Biochemistry*, 19(2), pp. 101–113. doi: 10.1016/0020-711X(87)90320-X.

Du, J., Xu, G., Liu, C. y Zhang, R. (2018). The role of phosphorylation and dephosphorylation of shell matrix proteins in shell formation: an in vivo and in vitro study. *CrystEngComm*, 20(27), pp. 3905–3916. doi: 10.1039/c8ce00755a.

Evans, J. S. (2019). The Biomineralization Proteome: Protein Complexity for a Complex Bioceramic Assembly Process. *Proteomics*, 19(16), pp. 1–12. doi: 10.1002/pmic.201900036.

Falini, G., Albeck, S., Weiner, S. y Addadi, L. (1996). Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science*, 271(5245), pp. 67–69. doi: 10.1126/science.271.5245.67.

Feng, D., Li, Q., Yu, H., Kong, L. y Du, S. (2017). Identification of conserved proteins from diverse shell matrix proteome in *Crassostrea gigas*: Characterization of genetic bases regulating shell formation. *Scientific Reports*, 7(March), pp. 1–12. doi: 10.1038/srep45754.

Feng, Q. L., Cui, F. Z., Pu, G., Wang, R. Z. y Li, H. D. (2000). Crystal orientation, toughening mechanisms and a mimic of nacre. *Materials Science and Engineering: C*, 11(1), pp. 19–25. doi: 10.1016/S0928-4931(00)00138-7.

Flausse, A., Henrionnet, C., Dossot, M., Dumas, D., Hupont, S., Pinzano, A., Mainard, D., Galois, L., Magdalou, J., Lopez, E., Gillet, P. y Rousseau, M. (2013). Osteogenic differentiation of human bone marrow mesenchymal stem cells in hydrogel containing nacre powder. *Journal of Biomedical Materials Research Part A*, p. n/a-n/a. doi: 10.1002/jbm.a.34629.

Fritz, M. y Morse, D. E. (1998). The formation of highly organized biogenic polymer/ceramic composite materials: The high-performance microaluminate of molluscan nacre. *Current Opinion in Colloid & Interface Science*, 3(1), pp. 55–62. doi: 10.1016/S1359-0294(98)80042-3.

Fu, G., Valiyaveettil, S., Wopenka, B. y Morse, D. E. (2005). CaCO₃ biomineralization: Acidic 8-kDa proteins isolated from aragonitic abalone shell nacre can specifically modify calcite crystal morphology. *Biomacromolecules*, 6(3), pp. 1289–1298. doi: 10.1021/bm049314v.

Furuhashi, T., Miksik, I., Smrz, M., Germann, B., Nebija, D., Lachmann, B. y Noe, C. (2010). Comparison of aragonitic molluscan shell proteins. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, 155(2), pp. 195–200. doi: 10.1016/j.cbpb.2009.11.007.

Gerhard, E. M., Wang, W., Li, C., Guo, J., Ozbolat, I. T., Rahn, K. M., Armstrong, A. D., Xia, J., Qian, G. y Yang, J. (2017). Design strategies and applications of nacre-based biomaterials. *Acta Biomaterialia*, 54, pp. 21–34. doi: 10.1016/j.actbio.2017.03.003.

Gibbs, G. M., Roelants, K. y O'Bryan, M. K. (2008). The CAP superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins - Roles in reproduction, cancer, and immune defense. *Endocrine Reviews*, 29(7), pp. 865–897. doi: 10.1210/er.2008-0032.

Grégoire, C. (1972). Structure of the molluscan shell. *Chemical Zoology*, VII, pp. 45–102. Hare, E. P. (1963) "Amino Acids in the Proteins from Aragonite and Calcite in the shells of *Mytilus californianus*", 139(4), pp. 5–6.

Hall, J. C. (1998). Review: Glycine. *Journal of Parenteral and Enteral Nutrition*, 22(6), pp. 393–398. doi: 10.1177/0148607198022006393.

Herlitze, I., Marie, B., Marin, F. y Jackson, D. J. (2018). Molecular modularity and asymmetry of the molluscan mantle revealed by a gene expression atlas. *GigaScience*. doi: 10.1093/gigascience/giy056.

Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., *et al.* (2020). Ensembl Genomes 2020—enabling non-vertebrate genomic research. *Nucleic Acids Research*, 48(D1), pp. D689–D695. doi: 10.1093/nar/gkz890.

Huang, J., Zhang, C., Ma, Z., Xie, L. y Zhang, R. (2007). A novel extracellular EF-hand protein involved in the shell formation of pearl oyster. *Biochimica et Biophysica Acta - General Subjects*, 1770(7), pp. 1037–1044. doi: 10.1016/j.bbagen.2007.03.006.

Jackson, D. J., Reim, L., Randow, C., Cerveau, N., Degnan, B. M. y Fleck, C. (2017). Variation in orthologous shell-forming proteins contribute to molluscan shell diversity. *Molecular Biology and Evolution*, 34(11), pp. 2959–2969. doi: 10.1093/molbev/msx232.

Jackson, D. J., Wörheide, G. y Degnan, B. M. (2007). Dynamic expression of ancient and novel molluscan shell genes during ecological transitions. *BMC Evolutionary Biology*, 7, pp. 1–17. doi: 10.1186/1471-2148-7-160.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R. y Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), pp. 1236–1240. doi: 10.1093/bioinformatics/btu031.

Jovine, L., Darie, C. C., Litscher, E. S. y Wassarman, P. M. (2005). Zona pellucida domain proteins. *Annual Review of Biochemistry*, 74, pp. 83–114. doi: 10.1146/annurev.biochem.74.082803.133039.

Jovine, L., Qi, H., Williams, Z., Litscher, E. y Wassarman, P. M. (2002). The ZP domain is a conserved module for polymerization of extracellular proteins. *Nature Cell Biology*, 4(6), pp. 457–461. doi: 10.1038/ncb802.

Julenius, K., Mølgaard, A., Gupta, R. y Brunak, S. (2005). Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology*, 15(2), pp. 153–64. doi: 10.1093/glycob/cwh151.

Kocot, K. M., Aguilera, F., McDougall, C., Jackson, D. J. y Degnan, B. M. (2016). Sea shell diversity and rapidly evolving secretomes: Insights into the evolution of biomineralization. *Frontiers in Zoology*, 13(1), pp. 1–10. doi: 10.1186/s12983-016-0155-z.

Kong, J., Liu, C., Yang, D., Yan, Y., Chen, Y., Liu, Y., Zheng, G., Xie, L. y Zhang, R. (2019). A novel basic matrix protein of *Pinctada fucata*, PNU9, functions as inhibitor during crystallization of aragonite. *CrystEngComm*, 21(8), pp. 1250–1261. doi: 10.1039/C8CE02194E.

Kono, M., Hayashi, N. y Samata, T. (2000). Molecular mechanism of the nacreous layer formation in *Pinctada maxima*. *Biochemical and Biophysical Research Communications*, 269(1), pp. 213–218. doi: 10.1006/bbrc.2000.2274.

Le Roy, N., Jackson, D. J., Marie, B., Ramos-Silva, P. y Marin, F. (2014). The evolution of metazoan α -carbonic anhydrases and their roles in calcium carbonate biomineralization. *Frontiers in Zoology*, 11(1), pp. 1–16. doi: 10.1186/s12983-014-0075-8.

Letunic, I., Khedkar, S. y Bork, P. (2021). SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research*, 49(D1), pp. D458–D460. doi: 10.1093/nar/gkaa937.

Levi-Kalishman, Y., Falini, G., Addadi, L. y Weiner, S. (2001). Structure of the Nacreous Organic Matrix of a Bivalve Mollusk Shell Examined in the Hydrated State Using Cryo-TEM. *Journal of Structural Biology*, 135(1), pp. 8–17. doi: 10.1006/jsbi.2001.4372.

Lin, A. y Meyers, M. A. (2005). Growth and structure in abalone shell. *Materials Science and Engineering A*, 390(1–2), pp. 27–41. doi: 10.1016/j.msea.2004.06.072.

Liu, C., Li, S., Kong, J., Liu, Y., Wang, T., Xie, L. y Zhang, R. (2015). In-depth proteomic analysis of shell matrix proteins of *Pinctada fucata*. *Scientific Reports*, 5, pp. 1–14. doi: 10.1038/srep17269.

Liu, C. y Zhang, R. (2021). Biomineral proteomics: A tool for multiple disciplinary studies. *Journal of Proteomics*, 238(December 2020), p. 104171. doi: 10.1016/j.jprot.2021.104171.

Liu, X., Dong, S., Jin, C., Bai, Z., Wang, G. y Li, J. (2015). Silkmapin of *Hyriopsis cumingii*, a novel silk-like shell matrix protein involved in nacre formation. *Gene*, 555(2), pp. 217–222. doi: 10.1016/j.gene.2014.11.006.

Lopez, E., Le Faou, A., Borzeix, S. y Berland, S. (2000). Stimulation of rat cutaneous fibroblasts and their synthetic activity by implants of powdered nacre (mother of pearl). *Tissue and Cell*, 32(1), pp. 95–101. doi: 10.1054/tice.1999.0091.

Lopez, E., Vidal, B., Berland, S., Camprasse, S., Camprasse, G. y Silve, C. (1992). Demonstration of the capacity of nacre to induce bone formation by human osteoblasts maintained in vitro. *Tissue and Cell*, 24(5), pp. 667–679. doi: 10.1016/0040-8166(92)90037-8.

Lowenstam, H. A. y Weiner, S. (1989). *On Biomineralization*. Oxford University Press, New York. Mann, K., Cerveau, N., Gummich, M., Fritz, M., Mann, M. y Jackson, D. J. (2018) “In-depth proteomic analyses of *Haliotis laevis* (greenlip abalone) nacre and prismatic organic shell matrix”, *Proteome Science*, 16(1). doi: 10.1186/s12953-018-0139-3.

Mann, K., Edsinger-Gonzales, E. y Mann, M. (2012). In-depth proteomic analysis of a mollusc shell: Acid-soluble and acid-insoluble matrix of the limpet *Lottia gigantea*. *Proteome Science*, 10(1), pp. 1–18. doi: 10.1186/1477-5956-10-28.

Mann, K., Poustka, A. J. y Mann, M. (2010). Phosphoproteomes of *Strongylocentrotus purpuratus* shell and tooth matrix: identification of a major acidic sea urchin tooth phosphoprotein, phosphodontin. *Proteome Science*, 8(1), p. 6. doi: 10.1186/1477-5956-8-6.

Mann, K., Siedler, F., Treccani, L., Heinemann, F. y Fritz, M. (2007). Perlinhibin, a cysteine-, histidine-, and arginine-rich miniprotein from abalone (*Haliotis laevis*) nacre, inhibits in vitro calcium carbonate crystallization. *Biophysical Journal*, 93(4), pp. 1246–1254. doi: 10.1529/biophysj.106.100636.

Mann, K., Weiss, I. M., André, S., Gabius, H. J. y Fritz, M. (2000). The amino-acid sequence of the abalone (*Haliotis laevis*) nacre protein perlucin: Detection of a functional C-type lectin domain with galactose/mannose specificity. *European Journal of Biochemistry*, 267(16), pp. 5257–5264. doi: 10.1046/j.1432-1327.2000.01602.x.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D. y Bryant, S. H. (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic acids research*, 41(Database issue), pp. D348-52. doi: 10.1093/nar/gks1243.

Marie, B., Jackson, D. J., Ramos-Silva, P., Zanella-Cléon, I., Guichard, N. y Marin, F. (2013). The shell-forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-specific novelties. *FEBS Journal*, 280(1), pp. 214–232. doi: 10.1111/febs.12062.

Marie, B., Joubert, C., Tayalé, A., Zanella-Cleón, I., Belliard, C., Piquemal, D., Cochennec-Laureau, N., Marin, F., Gueguen, Y. y Montagnani, C. (2012). Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. *Proceedings of the National Academy of Sciences of the United States of America*, 109(51), pp. 20986–20991. doi: 10.1073/pnas.1210552109.

Marie, B., Marie, A., Jackson, D. J., Dubost, L., Degnan, B. M., Milet, C. y Marin, F. (2010). Proteomic analysis of the organic matrix of the abalone *Haliotis asinina* calcified shell. *Proteome Science*, 8, pp. 1–11. doi: 10.1186/1477-5956-8-54.

Marin, F. (2020). Mollusc shellomes: Past, present and future. *Journal of Structural Biology*, 212(1). doi: 10.1016/j.jsb.2020.107583.

Marin, F., Bundeleva, I., Takeuchi, T., Immel, F. y Medakovic, D. (2016). Organic matrices in metazoan calcium carbonate skeletons: Composition, functions, evolution. *Journal of Structural Biology*, 196(2), pp. 98–106. doi: 10.1016/j.jsb.2016.04.006.

Marin, F. y Luquet, G. (2004). Molluscan shell proteins. *Comptes Rendus - Palevol*, 3(6-7 SPEC.ISS.), pp. 469–492. doi: 10.1016/j.crpv.2004.07.009.

Marin, F., Luquet, G., Marie, B. y Medakovic, D. (2007). Molluscan Shell Proteins: Primary Structure, Origin, and Evolution. *Current Topics in Developmental Biology*, 80(07), pp. 209–276. doi: 10.1016/S0070-2153(07)80006-8.

Marin, F., Marie, B., Hamada, S. Ben, Silva, P., Roy, N. Le, Wolf, S. E., Montagnani, C., Joubert, C., Piquemal, D. y Marie, B. (2013). Shellome: proteins involved in mollusc shell biomineralization – diversity, functions. *Recent Advances in Pearl Research*, (33), pp. 149–166. Disponible en: <http://hal.archives-ouvertes.fr/hal-00793668/>. [Consulta: 30 de junio 2021]

Maurer, P. y Hohenester, E. (1997). Structural and functional aspects of calcium binding in extracellular matrix proteins. *Matrix Biology*, 15(8–9), pp. 569–580. doi: 10.1016/S0945-053X(97)90033-0.

Michenfelder, M., Fu, G., Lawrence, C., Weaver, J. C., Wustman, B. A., Taranto, L., Evans, J. S. y Morse, D. E. (2003). Erratum: Characterization of Two Molluscan Crystal-Modulating Biomineralization Proteins and Identification of Putative Mineral Binding Domains (Biopolymers (2003) 70:4 (522-533)). *Biopolymers*, 73(2), p. 291. doi: 10.1002/bip.20020.

Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T. y Matsushiro, A. (1996). A carbonic anhydrase from the nacreous layer in oyster pearls. *Proceedings of the National*

Academy of Sciences of the United States of America, 93(18), pp. 9657–9660. doi: 10.1073/pnas.93.18.9657.

Miyamoto, H., Miyoshi, F. y Kohno, J. (2005). The carbonic anhydrase domain protein nacrein is expressed in the epithelial cells of the mantle and acts as a negative regulator in calcification in the mollusc *Pinctada fucata*. *Zoological Science*, 22(3), pp. 311–315. doi: 10.2108/zsj.22.311.

Miyashita, T., Takagi, R., Nakano, S., Miyamoto, H., Nishikawa, E. y Matsushiro, A. (2000). Complementary DNA Cloning and Characterization of Pearlin, a New Class of Matrix Protein in the Nacreous Layer of Oyster Pearls. *Marine Biotechnology*, 2(4), pp. 348–351. doi:10.1007/PL00021687.

Molluscs of Australia (2021) *Gastropods*. Disponible en: http://www.mesa.edu.au/molluscs/molluscs_01.asp [Consulta: 12 de Mayo 2021]

Montagnani, C., Marie, B., Marin, F., Belliard, C., Riquet, F., Tayalé, A., Zanella-Cléon, I., Fleury, E., Gueguen, Y., Piquemal, D. y Cochenec-Laureau, N. (2011). Pmarg-pearlin is a matrix protein involved in nacre framework formation in the pearl oyster *Pinctada margaritifera*. *ChemBioChem*, 12(13), pp. 2033–2043. doi: 10.1002/cbic.201100216.

Morse, J. W., Arvidson, R. S. y Lüttge, A. (2007). Calcium carbonate formation and dissolution. *Chemical Reviews*, 107(2), pp. 342–381. doi: 10.1021/cr050358j.

Murdock, D. J. E. (2020). The ‘biomineralization toolkit’ and the origin of animal skeletons. *Biological Reviews*, 95(5), pp. 1372–1392. doi: 10.1111/brv.12614.

Nakae, S., Shionyu, M., Ogawa, T. y Shirai, T. (2018). Structures of jacalin-related lectin PPL3 regulating pearl shell biomineralization. *Proteins: Structure, Function, and Bioinformatics*, 86(6), pp. 644–653. doi: 10.1002/prot.25491.

Nudelman, F. (2015). Nacre biomineralisation: A review on the mechanisms of crystal nucleation. *Seminars in Cell and Developmental Biology*, 46, pp. 2–10. doi: 10.1016/j.semcd.2015.07.004.

Nudelman, F., Gotliv, B. A., Addadi, L. y Weiner, S. (2006). Mollusk shell formation: Mapping the distribution of organic matrix components underlying a single aragonitic tablet in nacre. *Journal of Structural Biology*, 153(2), pp. 176–187. doi: 10.1016/j.jsb.2005.09.009.

Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A. y Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. *European Journal of Cell Biology*, 97(6), pp. 422–441. doi: 10.1016/j.ejcb.2018.06.003.

Panca, R., Schad, E., Tantos, A. y Tompa, P. (2019). Emergent functions of proteins in non-stoichiometric supramolecular assemblies. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1867(10), pp. 970–979. doi: 10.1016/j.bbapap.2019.02.007.

Pereira Mouriès, L., Almeida, M.-J., Milet, C., Berland, S. y Lopez, E. (2002). Bioactivity of nacre water-soluble organic matrix from the bivalve mollusk *Pinctada maxima* in three mammalian cell types: fibroblasts, bone marrow stromal cells and osteoblasts. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 132(1), pp. 217–229. doi: 10.1016/S1096-4959(01)00524-3.

Pérez Sala, M. D. (2012). Modificaciones postraduccionales de proteínas: mecanismos clave en el control de su actividad. *Centro de Investigaciones Biológicas (CIB), CSIC.*, 2(4), pp. 1–2. Disponible en: https://www.sebbm.es/web/images/archivos/archivos_tinymce/perezsaladolores-pdf-enero_2012.pdf. [Consulta: 20 de Junio 2021]

Rawlings, B. N. D. y Barrett, A. J. (1995). Evolutionary families of metallopeptidases In *Proteolytic Enzymes: Aspartic and Metallo Peptidases* en *Methods in enzymology*, pp. 183–228.

Rivera-Perez, C., Flores-Sánchez, I. A., Ojeda Ramírez de Areyano, J. J., Rojas Posadas, D. I. y Hernández-Saavedra, N. Y. (2020). A shell matrix protein of *Pinctada mazatlanica* produces nacre platelets in vitro. *Scientific Reports*, 10(1), pp. 1–11. doi: 10.1038/s41598-020-77320-7.

Rousseau, M., Pereira-Mouriès, L., Almeida, M. J., Milet, C. y Lopez, E. (2003). The water-soluble matrix fraction from the nacre of *Pinctada maxima* produces earlier mineralization of MC3T3-E1 mouse pre-osteoblasts. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology*, 135(1), pp. 1–7. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/12781967>.

Saleuddin, A. S. M. y Petit, H. P. (1983). *The Mode of Formation and the Structure of the Periostracum. The Mollusca*. ACADEMIC PRESS, INC. doi: 10.1016/b978-0-12-751404-8.50013-x.

Samata, T. (2004). Recent Advances in Studies on Nacreous Layer Biomineralization Molecular and Cellular Aspects. *Thalassas: An international journal of marine sciences*, 20(1), pp. 25–44.

Samata, T., Hayashi, N., Kono, M., Hasegawa, K., Horita, C. y Akera, S. (1999). A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*. *FEBS Letters*, 462(1–2), pp. 225–229. doi: 10.1016/S0014-5793(99)01387-3.

Sevier, C. S. y Kaiser, C. A. (2002). Formation and transfer of disulphide bonds in living cells. *Nature Reviews Molecular Cell Biology*, 3(11), pp. 836–847. doi: 10.1038/nrm954.

Shen, X., Belcher, A. M., Hansma, P. K., Stucky, G. D. y Morse, D. E. (1997). Molecular cloning and characterization of Lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *Journal of Biological Chemistry*, 272(51), pp. 32472–32481. doi: 10.1074/jbc.272.51.32472.

Shen, Y., Zhu, J., Zhang, H. Y Zhao, F. (2006). *In vitro* osteogenetic activity of pearl. *Biomaterials*, 27(2), pp. 281–287. doi: 10.1016/j.biomaterials.2005.05.088.

Shi, Y., Xu, M., Huang, J., Zhang, H., Liu, W., Ou, Z. y He, M. (2019). Transcriptome analysis of mantle tissues reveals potential biomineralization-related genes in *Tectus pyramis* Born. *Comparative Biochemistry and Physiology - Part D: Genomics and Proteomics*, 29(November 2018), pp. 131–144. doi: 10.1016/j.cbd.2018.11.010.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. y Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), p. 539. doi: 10.1038/msb.2011.75.

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., Bougueleret, L. y Xenarios, I. (2012). New and continuing developments at PROSITE. *Nucleic Acids Research*, 41(D1), pp. D344–D347. doi: 10.1093/nar/gks1067.

Simkiss, K. y Wilbur, K. (1989). *Biomineralization*. New York: ACADEMIC PRESS, INC.S

mith, B. L., Schäffer, T. E., Vlani, M., Thompson, J. B., Frederick, N. A., Klndt, J., Belcher, A., Stucky, G. D., Morse, D. E. y Hansma, P. K. (1999). Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. *Nature*, 399(6738), pp. 761–763. doi: 10.1038/21607.

Song, X., Liu, Z., Wang, L. y Song, L. (2019). Recent advances of shell matrix proteins and cellular orchestration in marine molluscan shell biomineralization. *Frontiers in Marine Science*, 6(FEB). doi: 10.3389/fmars.2019.00041.

Stecher, G., Tamura, K. y Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution*. Editado por C. Russo, 37(4), pp. 1237–1239. doi: 10.1093/molbev/msz312.

Sudo Shu, Fujikawa Takahiko, Nagakura Takeshi, Sakaguchi Kesami, Ohkubo Takeshi, Tanaka Minoru, K. N. (1997). Structures of mollusc shell framework proteins. *Nature*, 387, pp. 563–564.

Suzuki, M., Murayama, E., Inoue, H., Ozaki, N., Tohse, H., Kogure, T. y Nagasawa, H. (2004). Characterization of Prismalin-14, a novel matrix protein from the prismatic layer of the Japanese pearl oyster (*Pinctada fucata*). *Biochemical Journal*, 382(1), pp. 205–213. doi: 10.1042/BJ20040319.

Suzuki, M., Saruwatari, K., Kogure, T., Yamamoto, Y., Nishimura, T., Kato, T. y Nagasawa, H. (2009). An acidic matrix protein, Pif, is a key macromolecule for nacre formation. *Science*, 325(5946), pp. 1388–1390. doi: 10.1126/science.1173793.

Taylor, J. (1973). The structural evolution of the bivalve shell. *Palaeontology*, pp. 519–534.

Thompson, J. B., Paloczi, G. T., Kindt, J. H., Michenfelder, M., Smith, B. L., Stucky, G., Morse, D. E. y Hansma, P. K. (2000). Direct observation of the transition from calcite to aragonite growth as

induced by abalone shell proteins. *Biophysical Journal*, 79(6), pp. 3307–3312. doi: 10.1016/S0006-3495(00)76562-3.

Toll-Riera, M., Radó-Trilla, N., Martys, F. y Albá, M. M. (2012). Role of low-complexity sequences in the formation of novel protein coding sequences. *Molecular Biology and Evolution*, 29(3), pp. 883–886. doi: 10.1093/molbev/msr263.

Treccani, L., Mann, K., Heinemann, F. y Fritz, M. (2006). Perlwapin, an abalone nacre protein with three four-disulfide core (whey acidic protein) domains, inhibits the growth of calcium carbonate crystals. *Biophysical Journal*, 91(7), pp. 2601–2608. doi: 10.1529/biophysj.106.086108.

Tsukamoto, D., Sarashina, I. y Endo, K. (2004). Structure and expression of an unusually acidic matrix protein of pearl oyster shells. *Biochemical and Biophysical Research Communications*, 320(4), pp. 1175–1180. doi: 10.1016/j.bbrc.2004.06.072.

Von Heijne, G. (1990). Protein targeting signals. *Current Opinion in Cell Biology*, 2(4), pp. 604–608. doi: 10.1016/0955-0674(90)90100-5.

Watabe, N. (1981). Crystal growth of calcium carbonate in the invertebrates. *Progress in Crystal Growth and Characterization*, 4(1–2), pp. 99–147. doi: 10.1016/0146-3535(81)90050-2.

Watabe, N. y Wilbur, K. M. (1960). Influence of the Organic Matrix on Crystal Type in Molluscs. *Nature*, 188(4747), pp. 334–334. doi: 10.1038/188334a0.

Weiner, S. y Traub, W. (1984). Macromolecules in mollusc shells and their functions in biomineralization. *Philosophical Transactions of the Royal Society of London, Series B*, 304, pp. 425–434. doi: 10.1098/rstb.1984.0036.

Weiner, S. y Hood, L. (1975). Soluble protein of the organic matrix of mollusk shells: A potential template for shell formation. *Science*, 190(4218), pp. 987–989. doi: 10.1126/science.1188379.

Weiner, S. y Traub, W. (1980). X-Ray Diffraction Study of the Insoluble Organic Matrix of Mollusk Shells. 111(2), pp. 311–316.

Weiss, I. M. (2019). *Mineral-chitin composites in molluscs*. Editado por H. Cohen, E., Merzendorfer. Springer US.

Weiss, I. M., Kaufmann, S., Mann, K. y Fritz, M. (2000). Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc *Haliotis laevigata*. *Biochemical and Biophysical Research Communications*, 267(1), pp. 17–21. doi: 10.1006/bbrc.1999.1907.

Weiss, I. M., Tuross, N., Addadi, L. y Weiner, S. (2002). Mollusc larval shell formation: Amorphous calcium carbonate is a precursor phase for aragonite. *Journal of Experimental Zoology*, 293(5), pp. 478–491. doi: 10.1002/jez.90004.

Werner, G. D. A., Gemmell, P., Grosser, S., Hamer, R. y Shimeld, S. M. (2013). Analysis of a deep transcriptome from the mantle tissue of *Patella vulgata* Linnaeus (Mollusca: Gastropoda: Patellidae) reveals candidate biomineralising genes. *Marine Biotechnology*, 15(2), pp. 230–243. doi: 10.1007/s10126-012-9481-0.

Westbroek, P. y Marin, F. (1998). A marriage of bone and nacre. *Nature*, 392(6679), pp. 861–862. doi: 10.1038/31798.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D. J., Madden, T. L., et al. (2007). Database resources of the National Center for Biotechnology Information”, *Nucleic Acids Research*, 36(Database), pp. D13–D21. doi: 10.1093/nar/gkm1000.

Wilbur, K. M. Y Saleuddin, A. S. M. (1983). *Shell Formation, The Mollusca*. ACADEMIC PRESS, INC. doi: 10.1016/b978-0-12-751404-8.50014-1.

Wilbur, K. M. y Watabe, N. (1963). EXPERIMENTAL STUDIES ON CALCIFICATION IN MOLLUSCS AND THE ALGA *Coccolithus huxleyi*. *Annals of the New York Academy of Sciences*, 109(1), pp. 82–112. doi: 10.1111/j.1749-6632.1963.tb13463.x.

Wustman, B. A., Morse, D. E. y Evans, J. S. (2004). Structural characterization of the N-terminal mineral modification domains from the molluscan crystal-modulating biomineralization proteins, AP7 and AP24. *Biopolymers*, 74(5), pp. 363–376. doi: 10.1002/bip.20086.

Xia, X. (2007). Protein isoelectric point. *Bioinformatics and the Cell*, pp. 207–219. doi: 10.1007/978-0-387-71337-3_10.

Yano, M., Nagai, K., Morimoto, K. y Miyamoto, H. (2007). A novel nacre protein N19 in the pearl oyster *Pinctada fucata*. *Biochemical and Biophysical Research Communications*, 362(1), pp. 158–163. doi: 10.1016/j.bbrc.2007.07.172.

Young, S. D., Crenshaw, M. A. y King, D. B. (1977). Mantle protein excretion and calcification in the hardshell clam *Mercenaria mercenaria*. II. Protein synthesis and excretion by the isolated mantle. *Marine Biology*, 41(3), pp. 259–262. doi: 10.1007/BF00394912.

Zaremba, C. M., Belcher, A. M., Fritz, M., Li, Y., Mann, S., Hansma, P. K., Morse, D. E., Speck, J. S. y Stucky, G. D. (1996) “Critical Transitions in the Biofabrication of Abalone Shells and Flat Pearls”, *Chemistry of Materials*, 8(3), pp. 679–690. doi: 10.1021/cm9503285.

Zhang, C. y Zhang, R. (2006). Matrix proteins in the outer shells of molluscs. *Marine Biotechnology*, 8(6), pp. 572–586. doi: 10.1007/s10126-005-6029-6.

Zhang, Y., Xie, L., Meng, Q., Jiang, T., Pu, R., Chen, L. y Zhang, R. (2003). A novel matrix protein participating in the nacre framework formation of pearl oyster, *Pinctada fucata*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 135(3), pp. 565–573. doi: 10.1016/S1096-4959(03)00138-6.

Zheng, X., Zhao, S., Lei, S., Ma, R., Liu, L., Xie, Y., Shi, X. y Chen, J. (2020). Cloning and characterization of a novel Lustrin A gene from *Haliotis discus hannai*. *Comparative Biochemistry and Physiology Part - B: Biochemistry and Molecular Biology*, 240(October 2019), p. 110385. doi: 10.1016/j.cbpb.2019.110385.