

**UNIVERSIDAD DE CHILE**  
**FACULTAD DE CIENCIAS - ESCUELA DE POSTGRADO**



**“ESTUDIOS ESTRUCTURALES DE LA HEMOCIANINA DE  
*Concholepas concholepas*: UN ENFOQUE COMBINADO DE  
LAS TÉCNICAS DE CRISTALOGRAFÍA DE RAYOS X Y  
CRYO-EM”**

Tesis entregada a la Universidad de Chile en  
cumplimiento parcial de los requisitos para optar al  
Grado de Magister en Ciencias Biológicas por:

**Sebastián Manuel Muñoz Muñoz**

**Directora de Tesis:**  
**Dra. Victoria Guixé Leguía**

**Co-Director de Tesis:**  
**Dr. Víctor Castro Fernández**

**Agosto, 2022**  
**Santiago, Chile**



## INFORME DE APROBACIÓN TESIS DE MAGÍSTER

Se informa a la Escuela de Postgrado de la Facultad de Ciencias, de la Universidad de Chile que la Tesis de Magíster presentada por el candidato  
**Sebastián Manuel Muñoz Muñoz**

**“ESTUDIOS ESTRUCTURALES DE LA HEMOCIANINA DE *Concholepas concholepas*: UN ENFOQUE COMBINADO DE LAS TÉCNICAS DE CRISTALOGRAFÍA DE RAYOS X Y CRYO-EM”**

Ha sido aprobado por la Comisión de Evaluación, en cumplimiento parcial de los requisitos para optar al grado de Magíster en Ciencias Biológicas.

**Directora de Tesis:**

**Dra. Victoria Guixé Leguía**

---

**Co-Director de Tesis:**

**Dr. Víctor Castro Fernández**

---

**Comisión Revisora y Evaluadora**

**Dr. Ricardo Cabrera**

---

**Dr. Mauricio Báez**

---

Santiago de Chile, Agosto 2022.

*“Donde haya un árbol que plantar, plántalo tú.  
Donde haya un error que enmendar, enmiéndalo tú.  
Donde haya un esfuerzo que todos esquivan, hazlo tú.  
Sé tú el que aparta la piedra del camino.”*

*Gabriela Mistral*

Dedicado a mi tata Manuel y mi mamá Marisol,  
gracias a ustedes soy lo que soy hoy en día.

## BIOGRAFÍA



Mi nombre es Sebastián Manuel Muñoz Muñoz, nací el 19 de enero de 1995 en Punta Arenas, sin embargo, al mes de nacido nos mudamos con mi familia a Puerto Montt, lugar donde pasaría mi niñez y la mayor parte de mi vida. Allí realicé mis estudios escolares en el colegio Salesiano Padre José Fernández Pérez.

Durante mi enseñanza media en el colegio desarrollé interés por la biología, lo cual me llevó a entrar a la carrera de Ingeniería en Biotecnología Molecular en la Universidad de Chile el año 2013. Dentro de la carrera encontré mi vocación por la investigación, además tuve la oportunidad de conocer grandes modelos a seguir como científico, como lo son los profesores Victoria Guixé y Víctor Castro. Por esta razón, entré al laboratorio de Bioquímica y Biología Molecular el año 2017 a realizar una unidad de investigación en el grupo de la profesora Guixé. En este lugar conocí personas increíbles y encontré mi pasión por la biología estructural, defendiendo mi tesis de pregrado titulada “Estructura cristalina del complejo ternario de una fosfofructoquinasa ancestral dependiente de ADP e identificación de residuos claves para la unión de fructosa-6P y para la catálisis” en julio del 2020.

A través de mi trabajo en el laboratorio he podido participar en la publicación de algunos artículos en prestigiosas revistas como *Frontiers in Microbiology* y *JBC*, he podido participar en varias versiones del congreso de la Sociedad Chilena de Bioquímica y Biología Molecular desde el 2018 al 2021, además, he podido realizar algunas pasantías internacionales en el laboratorio del Dr. André Ambrosio en el IFSC de la Universidad de São Paulo, Brasil, siendo la última estadía en septiembre del año 2021, con el objetivo de aprender acerca de la técnica de cryo-EM.

A comienzos del año 2022 ingresé al programa de Doctorado en ciencias mención Biología Molecular, Celular y Neurociencias de la Facultad de Ciencias de la Universidad de Chile para poder seguir perfeccionando mis conocimientos en el área de la biología estructural centrada en proteínas.

## AGRADECIMIENTOS

Partir por agradecer a mi mamá Marisol y a mi tata Manuel, a quienes les dedico y les seguiré dedicando todos mis éxitos tanto personales como académicos. También agradezco al resto de mi numerosa familia, Feñita, Manu, Andrés, Taty, Pelusa, quienes constantemente me brindan su apoyo desde Puerto Montt, pese a que ya no pueda visitarlos como antes, espero que sepan que los llevo en el corazón a cada uno, de igual manera a quienes ya no están, mi nana Raquel y mi tío Javier. También te agradezco a ti Antonieta, me parece increíble ver cómo ha crecido nuestra relación a medida que he ido completando estos procesos, espero que siempre puedas seguir apoyándome, y seguir creciendo juntos en este camino que hemos comenzado.

Agradecer a mis profes Vicky y Víctor, por confiar en mí y embarcarme en esta aventura, permitiendo que siga dedicándome a lo que me apasiona. Agradecer al Nico, mi compañero de batallas, con quien tuvimos la brillante idea de iniciar el programa de doctorado antes de terminar el magíster, siendo un proceso difícil, pero siempre con la frente en alto. Dar gracias a Gabo, Matu y Felipe, los últimos mandriles, quienes han sido y seguirán siendo referentes para mí, no solo como científicos, sino también como personas increíbles. Agradecer al resto de mis compañeros y miembros del laboratorio: Perla y Yerko, por las conversaciones en los “tabaco time”, Profe Ricardo, Exe, Sixto, Consu, Claudia, Isa, Nico P, Dayanna, Sra. Eli, Erwin, Myriam Papers, y especialmente a la gente nueva que ha llegado al laboratorio, Nachito, Nico H, y Erick, son excelentes personas y les deseo éxito en los procesos que quieran emprender.

Me gustaría agradecer también al Dr. André Ambrosio y su grupo, quienes me recibieron en el IFSC en la Universidad de São Paulo, Brasil, lugar donde se llevó a cabo parte importante de esta tesis, agradecer a Matías y especialmente a Flavia, que ayudó a que mi estadía allá fuera de las mejores. También agradecer a la Dra. María Inés Becker y a su grupo en FUCITED, especialmente a Michelle, quien siempre tuvo la disposición para ayudarme con su infinito conocimiento sobre la hemocianina.

Por último, agradecer a mis amigos de Puerto Montt y a mi amigo Javier, a quienes ya no veo mucho, pero me alegro de que estén triunfando en los caminos que emprendieron. A Murillo y Gustavo, con quienes nos seguimos reuniendo para compartir y conversar de la vida. Finalmente, no puedo terminar esta sección sin darle gracias al taca-taca, apodado cariñosamente como “la desgracia de la ciencia en Chile” y que su presencia atrasó en gran medida la entrega de esta tesis, pero que sirvió como punto de encuentro entre miembros antiguos y nuevos del laboratorio, generando momentos muy gratos.

## **FINANCIAMIENTO**

Esta tesis fue desarrollada en el Laboratorio de Bioquímica y Biología Molecular de la Facultad de Ciencias de la Universidad de Chile, con el apoyo de los siguientes proyectos:

Proyecto de colaboración internacional FAPESP 2019/13318-5: Dra. Victoria Guixé (Universidad de Chile), Dr. Víctor Castro (Universidad de Chile), Dra. María Inés Becker (FUCITED) y Dr. André Ambrosio (IFSC-USP).

Proyecto FONDECYT 1191321 (Dra. Victoria Guixé).

Proyecto FONDECYT 1201200 (Dra. María Inés Becker, FUCITED).

## ÍNDICE DE CONTENIDOS

Biografía .....	iii
Agradecimientos .....	iv
Financiamiento .....	v
Índice de Contenidos .....	vi
Índice de Tablas .....	ix
Índice de Figuras .....	x
Lista de Abreviaturas .....	xii
Resumen .....	xiii
Abstract.....	xv
1. Introducción .....	1
1.1 Hemocianinas de molusco: Una proteína gigante .....	1
1.2 Organización estructural de las hemocianinas de molusco .....	2
1.3 Tipos de hemocianinas de molusco .....	7
1.4 Hemocianinas de molusco de tipo 1 .....	10
1.5 Hemocianinas recombinantes .....	13
1.6 Hemocianina de <i>Concholepas concholepas</i> (CCH) .....	14
2. Hipótesis y objetivos .....	18
2.1 Hipótesis .....	18
2.2 Objetivo general .....	18

2.3	Objetivos específicos .....	18
3.	Materiales y Métodos .....	20
3.1	Materiales .....	20
3.1.1	Medios de cultivo .....	20
3.1.2	Reactivos .....	20
3.1.3	Vectores.....	20
3.2	Métodos .....	23
3.2.1	Cuantificación de proteínas .....	23
3.2.2	Expresión y purificación de FUs recombinantes de CCH.....	23
3.2.3	Disociación y purificación de subunidades silvestres de CCH .....	24
3.2.4	Proteólisis y purificación de FU silvestre de CCH.....	24
3.2.5	Electroforesis unidimensional de proteínas en condiciones denaturantes.....	25
3.2.6	Ensayos de cristalización de FU silvestre .....	25
3.2.7	Difracción de rayos X y análisis de datos .....	26
3.2.8	Cryo-EM de CCH.....	28
3.2.9	Procesamiento de imágenes Cryo-EM y reconstrucción 3D de modelo de CCH	28
3.2.10	Generación de modelos de organización de subunidades de CCH .....	29
4.	Resultados .....	31
4.1	OBJETIVO ESPECÍFICO 1.....	31
4.1.1	Identificación de FUs de CCH .....	31



4.1.2	Expresión recombinante y purificación de FU de CCH.....	32
4.1.3	Purificación de FU silvestre de CCH .....	37
4.1.4	Cristalización de FU silvestre de CCH.....	42
4.1.5	Difracción de rayos X y análisis de datos .....	43
4.1.6	Análisis de la estructura de FU silvestre de CCH .....	51
4.2	OBJETIVO ESPECÍFICO 2.....	56
4.2.1	Cryo-EM de CCH y procesamiento de datos .....	56
4.2.2	Análisis del mapa de Cryo-EM de CCH .....	64
5.	Discusión.....	75
5.1	Expresión de FU recombinante .....	75
5.2	Estructura cristalográfica de FU silvestre de CCH.....	78
5.3	Modelo de Cryo-EM de CCH.....	85
6.	Conclusiones .....	92
7.	Bibliografía.....	93
8.	Anexos.....	99
8.1.	Péptidos de FU silvestre purificada obtenidos por espectrometría de masas. ....	99
8.2.	Protocolo de tratamiento de datos de cryo-EM de CCH utilizando el software CryoSPARC .....	101

## ÍNDICE DE TABLAS

Tabla 1. Información cristalográfica y estadística del refinamiento de la estructura de CCHB-g.....	49
Tabla 2. Comparación de distancias de enlaces del centro de unión a cobre de las estructuras CCHB-g y ODH-g.....	55
Tabla 3. Comparación de ángulos de enlaces del centro de unión a cobre de las estructuras CCHB-g y ODH-g .....	55
Tabla 4. Información de colecta y procesamiento de datos de cryo-EM de CCH .....	62
Tabla 5. Valores de coeficiente de correlación para cada uno de los modelos de organización de subunidades de CCH .....	72
Tabla 6. Comparación de parámetros de estructuras similares a CCHB-g.....	81
Tabla 7. Comparación de tecnologías de detector asociadas con la resolución obtenida en mapas de cryo-EM para hemocianinas de molusco de tipo I.....	86

## ÍNDICE DE FIGURAS

Figura 1. Niveles de organización estructural de hemocianinas de molusco.....	3
Figura 2. Estructura de Unidades Funcionales (FU) de hemocianinas de molusco. ....	4
Figura 3. Estructuras de hemocianinas de molusco.....	9
Figura 4. Arquitectura de la pared y collar interno de la hemocianina de molusco tipo 1 .....	12
Figura 5. Secuencias de FU y mapas de vectores.....	22
Figura 6. Identificación de unidades Funcionales (FU) a partir de fragmentos de secuencia de CCH.....	32
Figura 7. Purificación de FUs recombinantes de CCH.....	35
Figura 8. Purificación de las subunidades de CCH .....	38
Figura 9. Purificación de FUs obtenidas desde CCHB.....	40
Figura 10. Espectro de absorción de fracciones de CCH de origen silvestre y recombinante .....	41
Figura 11. Cristalización de FU de CCH. ....	43
Figura 12. Patrón de difracción representativo de datos nativos de FU .....	44
Figura 13. Factores de corrección para dispersión anómala del Cobre .....	44
Figura 14. Modelo inicial de FU obtenido mediante SAD .....	46
Figura 15. Dendrograma de secuencias de FU de RVH y CCH .....	48
Figura 16. Secuencia y estructura cristalográfica de FU-g de CCHB .....	50
Figura 17. N-glicosilaciones presentes en estructuras cristalográficas de FUs de hemocianina de molusco.....	52
Figura 18. Centro de unión a cobre de la estructura CCHB-g .....	53

Figura 19. Resultado de colecta de datos de cryo-EM de CCH y corrección de movimiento. .....	57
Figura 20. Gráfico de ajuste de CTF.....	58
Figura 21. Clases de partículas de CCH seleccionadas. ....	59
Figura 22. Modelo de reconstrucción <i>ab initio</i> de CCH.....	60
Figura 23. Gráfica de estándar dorado de correlación de la capa de Fourier del mapa final de cryo-EM de CCH .....	62
Figura 24. Reconstrucción 3D de CCH obtenida a 3,2 Å.....	63
Figura 25. Comparación del modelo de hemocianina homodidecamérica de <i>Haliotis diversicolor</i> (HDH) con el mapa de cryo-EM de CCH .....	65
Figura 26. Esquema de organización del dímero de subunidades de HDH.....	67
Figura 27. Modelos heterodidecámicos de organización de subunidades de CCH.....	68
Figura 28. Modelos heterodecámicos de organización de FUs y subunidades de CCH .....	71
Figura 29. Comparación de residuos de CCHA-d entre modelos de organización de subunidades de CCH.....	73
Figura 30. Modelo de distribución de subunidades dentro del heterodidecámero de CCH .....	74

## LISTA DE ABREVIATURAS

CCH	Hemocianina de <i>Concholepas concholepas</i>
CCHA	Subunidad A de hemocianina de <i>Concholepas concholepas</i>
CCHB	Subunidad B de hemocianina de <i>Concholepas concholepas</i>
CTF	<i>Contrast transfer function</i>
CC	Coefficiente de correlación
HDH	Hemocianina de <i>Haliotis Diversicolor</i>
FLH	Hemocianina de <i>Fissurella latimarginata</i>
Fuc	Fucosa
GalNAc	N-acetilgalactosamina
GlcNAc	N-acetilglucosamina
GSFSC	<i>Gold-Standard Fourier Shell Correlation</i>
FU	Unidad Funcional
KLH	Hemocianina de “Keyhole Limpet” ( <i>Megathura crenulata</i> )
LacNAc	N-acetillactosamina
LB	Medio Luria Bertani
Man	Manosa
ODH	Hemocianina de <i>Octopus dofleini</i>
PEG	Polietilengicol
PMSF	Fluoruro de fenilmetilsulfonilo
PNGasa F	Péptido:N-glicosidasa F
RMSD	<i>Root Mean Square Deviation</i>
RVH	Hemocianina de <i>Rapana venosa</i>
SAD	<i>Single-wavelength Anomalous Diffraction</i>
SGD	<i>Stochastic Gradient Descent</i>
SUMO	<i>Small Ubiquitin-like Modifier</i>
TPH	Hemocianina de <i>Todarodes pacificus</i>

## RESUMEN

Las hemocianinas son glicoproteínas multiméricas cilíndricas de alto peso molecular que se encuentran disueltas libremente en la hemolinfa de moluscos y artrópodos. Las hemocianinas de molusco se encuentran entre las proteínas más grandes conocidas, con masas moleculares entre 3,3 a 13,5 MDa, y presentan una compleja organización estructural. Su estructura cuaternaria corresponde a un didecámero, formado por una o dos subunidades distintas (homodidecámero y heterodidecámero respectivamente), y, a su vez, cada subunidad está formada por 8 dominios globulares con centros de unión a cobre llamados unidades funcionales (FU).

Actualmente, las hemocianinas son usadas en aplicaciones biomédicas y clínicas como inmunoestimulantes naturales, no tóxicos y no específicos. En este contexto, en Chile se ha determinado que la hemocianina del gastrópodo litoral *Concholepas concholepas* (CCH) posee características que la hacen más efectiva que las de otras especies en inducir respuestas inmunológicas en estudios preclínicos de inmunocastración y vacunas contra cáncer prostático. Sin embargo, a la fecha, no hay información empírica reportada acerca de su estructura. Solo se dispone de caracterizaciones bioquímicas que indican que CCH corresponde a un heterodidecámero, con dos subunidades diferentemente glicosiladas, llamadas CCHA (405 KDa) y CCHB (350 KDa). Más aún, casi la totalidad de la secuencia de CCH permanece desconocida. Por ello se propuso obtener la estructura cristalográfica de al menos una FU de origen recombinante a partir de los 2 fragmentos de secuencias disponibles y determinar la organización tridimensional de las subunidades en un mapa del heterodidecámero de CCH, obtenido por cryo-EM.

Dado que no fue posible obtener una FU expresada de manera recombinante en *E. coli*, se purificó una FU silvestre a partir de la subunidad CCHB de CCH y se determinó su estructura cristalográfica con resolución atómica (1,5 Å). Esto permitió obtener su secuencia, identificándola como CCHB-g, permitió determinar un árbol de N-glicosilaciones identificando la presencia de monómeros de fucosa, siendo la primera estructura tridimensional de FU de hemocianina en presentar este tipo de glicosilación, y, además, permitió analizar en detalle las distancias y ángulos de las interacciones químicas presentes en el centro de unión a cobre. La comparación de este centro con el presente en la estructura cristalográfica de FU de hemocianina de pulpo reportada a 2,3 Å de resolución reveló diferencias de hasta 0,3 Å en distancias y de hasta 11° en ángulos de enlace.

Por otra parte, se obtuvo un mapa de cryo-EM de CCH a resolución cercana a la atómica (3,2 Å) que corresponde al mapa de mayor resolución reportado para una hemocianina de molusco. Además, utilizando la estructura cristalográfica de la FU, generamos el primer modelo de organización de subunidades de una hemocianina heterodidecámerica, situando las subunidades CCHA hacia ambos extremos del cilindro del didecámero y las subunidades CCHB ubicadas hacia el interior del cilindro.

La información estructural presentada en este trabajo no solo será fundamental para las futuras aplicaciones de CCH en medicina y biotecnología, sino que será un aporte para el estudio estructural de las hemocianinas de molusco.

## ABSTRACT

Hemocyanins are giant, cylindrical, multimeric proteins present in the hemolymph of mollusks and arthropods. Mollusk hemocyanins are one of the largest known proteins with a molecular mass of between 3.3 and 13.5 MDa, presenting a complex structural organization. Their quaternary structure corresponds to a didecamer formed by one or two subunits (homodidecamer and heterodidecamer, respectively), each of which is composed of eight globular domains containing copper centers called functional units (FU).

Nowadays, hemocyanins are used as natural, non-toxic, and non-specific immunostimulants in biomedical and clinical applications. In this context, in Chile, it has been determined that the hemocyanin from the gastropod *Concholepas concholepas* (CCH) possesses unique properties that make it more effective than hemocyanins from other species in inducing immunological responses, as observed in preclinical studies of immunocastration and prostatic cancer vaccines.

However, to date, there has been no empirical information reported about its structure. There are only biochemical characterizations indicating that CCH corresponds to a heterodidecamer with two differentially glycosylated subunits called CCHA (405 KDa) and CCHB (350 KDa). Moreover, almost the entire sequence of CCH remains unknown. Therefore, we proposed obtaining the crystallographic structure from at least one FU recombinantly expressed based on the two sequence fragments available and determining the three-dimensional organization of the subunits in a map of the CCH heterodidecamer obtained by cryo-EM.



Because a recombinantly expressed FU could not be obtained from *E. coli*, we purified a wild-type FU from the CCHB subunit of CCH and determined its crystallographic structure at atomic resolution (1.5 Å). The structure allowed us to obtain the sequence, identifying the FU as CCHB-g. It also allowed us to determine an N-glycosylation tree, identifying the presence of fucose monomers, being the first three-dimensional structure of an FU with this type of glycosylation, and finally, allowing us to analyze in detail the distances and angles of the chemical interactions present in the copper-binding center. Its comparison with the copper-binding center of the crystallographic structure of octopus hemocyanin FU reported at 2.3 Å resolution reveals differences of up to 0.3 Å in distance and up to 11° in angles.

On the other hand, a near-atomic resolution cryo-EM map of CCH (3.2 Å) was obtained which corresponds to the highest resolution map reported to date for a mollusk hemocyanin. Moreover, using the FU crystallographic structure, we generate the first subunit organization model of a heterodidecameric hemocyanin, placing the CCHA subunits towards the outer ends of the didecameric cylinder and the CCHB subunits at the inner part of the cylinder.

The structural information presented in this work will not only be essential for future applications of CCH in medicine and biotechnology, but also contribute to the structural study of mollusk hemocyanins.

## 1. INTRODUCCIÓN

### 1.1 Hemocianinas de molusco: Una proteína gigante

La respiración es uno de los sistemas de soporte vital más importantes de los animales, ya que suministra oxígeno a los órganos individuales y tejidos. Para este sistema, los vertebrados poseen eritrocitos, moléculas que contienen hemoglobina, una proteína de 64 kDa de unión a hierro que permite el transporte de moléculas de oxígeno, y que da la coloración roja característica a la sangre. En contraste, moluscos y artrópodos son animales invertebrados que presentan hemolinfa, en la cual circula libremente la hemocianina, una glicoproteína con centros de unión de cobre II, encargada del transporte de oxígeno y que da la coloración azul a la hemolinfa.

Las hemocianinas de moluscos y artrópodos comparten la presencia de un centro activo que contiene dos átomos de cobre coordinado por 6 residuos de histidina, el cual se clasifica como centro de tipo 3 de acuerdo a la clasificación de las proteínas que unen cobre. Cuando este centro de tipo 3 une a una molécula de oxígeno ( $O_2$ ), forma un clúster  $Cu_2O_2$ , correspondiente a una hemocianina oxigenada de color azul, comparada con la hemocianina desoxigenada la cual no presenta coloración. Sin embargo, las estructuras moleculares de las hemocianinas de moluscos y artrópodos son totalmente distintas. Las hemocianinas de artrópodo son multi-hexámeros (aprox. 450 kDa) de subunidades que contienen un solo sitio de unión a oxígeno cada una, mientras que las hemocianinas de molusco son decámeros o multidecámeros formados por subunidades (330-550 kDa) cada una de las cuales contiene al menos 7 unidades funcionales (FU) con un sitio de unión a oxígeno cada una. Las masas moleculares de las hemocianinas de molusco varían desde

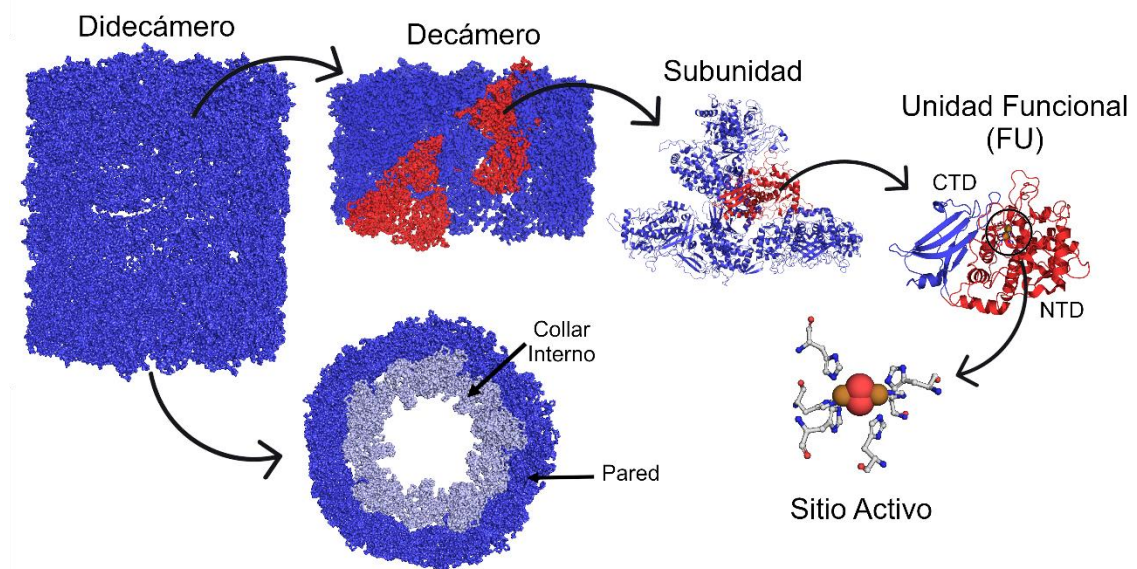
aprox. 3,3 MDa a 13,5 MDa, situándolas como una de las proteínas más grandes dentro del mundo animal (Kato y col., 2020).

Las hemocianinas de molusco han sido investigadas extensamente en el área de la biomedicina, siendo usadas actualmente como inmunoestimulantes no específicos, naturales y no tóxicos. La hemocianina más estudiada tanto desde el punto de vista bioquímico como en aplicaciones biomédica es la hemocianina del gastrópodo *Megathura crenulata*, comúnmente conocido como “lapa californiana” en inglés “keyhole limpet” (KLH). La KLH ha sido utilizada en diagnóstico para discriminar entre esquistosomiasis humana aguda y crónica, ya que los anticuerpos contra el parásito causante de la enfermedad reaccionan de forma cruzada con esta hemocianina (Beck y col., 2004). Se ha propuesto que las propiedades inmunogénicas de las hemocianinas son derivadas de la presencia de glucanos adheridos en la periferia del complejo proteico y de su gran masa molecular (Harris & Markl, 1999; Geyer y col., 2005 y Siddiqui y col., 2007).

## **1.2 Organización estructural de las hemocianinas de molusco**

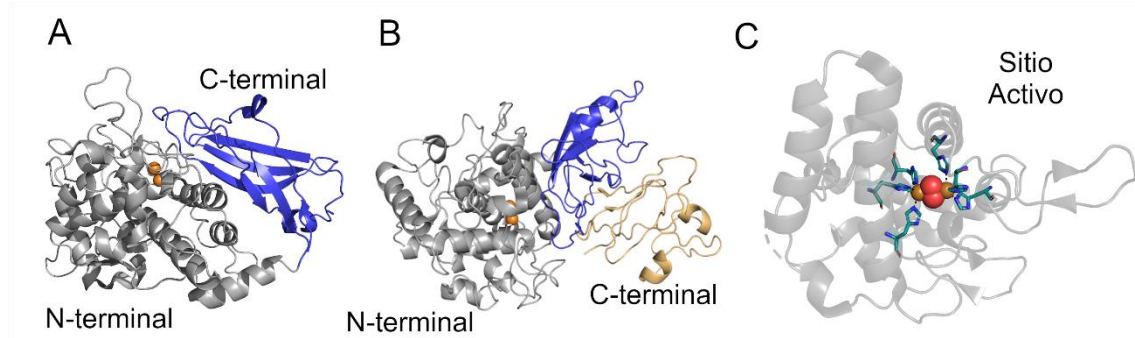
Las características biofísicas, bioquímicas y estructurales de las hemocianinas de molusco solo han sido estudiadas durante las últimas décadas, debido a la gran complejidad de su arquitectura y ensamble biológico. Las hemocianinas de molusco corresponden a cilindros huecos que presentan una pared y una región de collar interno (Kato y col., 2020). Cada cilindro se genera comúnmente por dos decámeros asociados por su eje rotacional formando didecámeros, donde cada decámero está constituido por 10 subunidades (cadena polipeptídica) que pueden representarse como la repetición de 5 dímeros de subunidades. A su vez, cada subunidad contiene 8 dominios globulares de unión a oxígeno llamados

unidades funcionales (FU, nombradas en letra minúscula de la -a hasta la -h) (van Holde & Miller, 1995) (**Figura 1**). Por otra parte, los didecámeros pueden poseer subunidades iguales (homodidecámeros), como el caso de KLH (KLH1 y KLH2) y la hemocianina del molusco *Rapana venosa* (RVH1 y RVH2) que poseen dos isoformas, pero con didecámeros formados por subunidades iguales; o bien, los didecámeros pueden estar formados por subunidades distintas (heterodidecámeros), como el caso de la hemocianina del molusco *Concholepas concholepas* (CCH).



**Figura 1. Niveles de organización estructural de hemocianinas de molusco.** La estructura cuaternaria de decámero, puede asociarse con otro decámero formando didecámeros, los cuales poseen dos regiones, la región de la pared del cilindro coloreada en azul y la región del collar interno mostrada en celeste. El decámero muestra una subunidad coloreada en rojo (dímero). Cada subunidad está compuesta por 8 unidades funcionales (FU) destacada en rojo. Cada unidad funcional se encuentra compuesta por un dominio N-terminal (NTD) destacado en rojo y un dominio C-terminal (CTD) destacado en azul. El centro activo se encuentra en el dominio N-terminal, observándose dos iones de cobre (esferas naranjas) formando un clúster con una molécula O<sub>2</sub> (esferas rojas) coordinado por 6 residuos de histidina. La organización mostrada corresponde en su mayoría a la hemocianina de *Haliotis diversicolor* HDH, exceptuando la imagen de subunidad que corresponde de “Keyhole limpet” KLH (*Megathura crenulata*).

Como se mencionó anteriormente, cada subunidad de la hemocianina de molusco tiene 8 estructuras globulares denominadas FU que se conectan entre ellas como un “collar de perlas” a lo largo de la cadena polipeptídica. Cada FU está separada por regiones denominadas *linker*, de 10-20 residuos de largo. Las FU estándar de molusco (FU-a hasta FU-g) son homólogas entre sí y se pliegan en dos dominios, el dominio N-terminal (NTD) que posee principalmente  $\alpha$ -hélices, presentando el centro activo rodeado por cuatro  $\alpha$ -hélices, y el dominio C-terminal (CTD) que está conformado principalmente por un  $\beta$ -sándwich de 6 hebras, el cual sirve de tapadera del centro activo (**Figura 2A, C**) (Harris & Markl, 1999). Por otra parte, la última FU, denominada tipo h (FU-h), posee una extensión adicional en el extremo C-terminal que presenta un plegamiento *cupredoxin-like*, además de los dominios NTD y CTD (**Figura 2B**) (Jaenicke y col., 2011).



**Figura 2. Estructura de Unidades Funcionales (FU) de hemocianinas de molusco.** **A)** Estructura de unidad funcional tipo -e de *Rapana venosa* (RVH2-e, PDB: 1LNL), destacando los dominios N-terminal en gris y C-terminal en azul. **B)** Estructura de unidad funcional tipo -h de “keyhole limpet” (KLH-h, PDB: 3QJO), la cual posee además un dominio extra *cupredoxin-like* de color crema en el C-terminal. **C)** Centro activo de una unidad funcional tipo -g de *Octopus dofleini* (ODH-g, PDB: 1JS8), destacando los residuos de histidina en cian y los átomos de Cu y O como esferas naranjas y rojas, respectivamente.

Durante mucho tiempo el estudio de la estructura de las hemocianinas de molusco presentó diversas dificultades debido a (i) la inestabilidad de la unión de subunidades, (ii) su gran masa molecular y (iii) su compleja organización estructural (Kato y col., 2020). El primer punto resulta importante ya que puede parecer simple la preparación de una muestra de hemocianina dado que se encuentra dispersa libremente en altas concentraciones en la hemolinfa. Sin embargo, algunas hemocianinas se disocian fácilmente de multidecámeros o decámeros a subunidades, por lo que es fundamental analizar las condiciones en las cuales la hemocianina no se disocie. Respecto al segundo y tercer punto, dado su gran tamaño y la homología presente en las unidades funcionales (FUs), la complejidad estructural de las hemocianinas ha sido investigada principalmente mediante un enfoque que combina dos técnicas diferentes: criomicroscopía electrónica (cryo-EM, según sus siglas en inglés) para el análisis de la molécula completa a baja resolución y cristalografía de rayos X para el análisis de las FU a alta resolución. Existen 6 modelos completos de hemocianina de molusco obtenidos por cryo-EM: “Keyhole limpet” (*Megathura crenulata*; Gatsogiannis & Markl, 2009), *Haliotis diversicolor* (Zhang y col., 2013), *Todarodes pacificus* (Tanaka y col., 2019), *Sepia officinalis* (Boisset & Mouche, 2000), *Nautilus pompilius* (Gatsogiannis y col., 2007) y *Melanoides tuberculata* (Gatsogiannis y col., 2015), mientras que solo hay 3 estructuras cristalográficas de FUs de hemocianinas de molusco, la FU-g de *Octopus dofleini* (Cuff y col., 1998), FU-e de *Rapana venosa* (Perbandt y col., 2003) y la FU-h de *Megathura crenulata* (Jaenicke y col., 2010, 2011). Sin embargo, en general los modelos obtenidos por cryo-EM tienen baja resolución ( $> 5\text{\AA}$ ) por lo que no tienen detalles atómicos, modelando sólo la cadena principal y sin tener en consideración las diferencias en cada

FU, exceptuando el caso de la hemocianina de *Haliotis diversicolor* (HDH) para la cual se reportó una estructura con una resolución de 4,5 Å (Zhang y col., 2013). Además, recientemente se reportó la primera estructura cristalográfica de la hemocianina completa de calamar (*Todarodes pacificus*) a una resolución de 3,0 Å, lo que ha permitido indagar más acerca del detalle atómico de las hemocianinas, la compleja organización estructural y ha permitido tener información más detallada acerca de la disposición de las FU dentro del decámero (Matsuno y col., 2015, Gai y col., 2015, Tanaka y col., 2019). Sin embargo, hasta la fecha, no se cuenta con ninguna estructura de hemocianina obtenida por cryo-EM que presente una resolución cercana a la atómica (3 a 4 Å) que permita analizar en detalle las diferencias entre las FUs y su disposición dentro del didecámero. Otra información relevante, es que todas las estructuras reportadas corresponden a homo-oligómeros y se desconoce la organización de hemocianinas que forman hetero-oligómeros (heterodidecámero) con 2 tipos de subunidades distintas.

### 1.3 Tipos de hemocianinas de molusco

Es importante conocer las bases estructurales de las hemocianinas de molusco que permiten su clasificación, dado que existe una gran cantidad de hemocianinas homólogas de diferentes organismos marinos que tienen un grado importante de diversidad estructural entre ellas, lo que podría esconder las diferencias en el grado de respuesta inmunogénica producida por estas proteínas en terapias inmunológicas. Las hemocianinas de molusco pueden clasificarse en 4 tipos, teniendo en cuenta la disposición de subunidades y su arquitectura general (Kato y col., 2020):

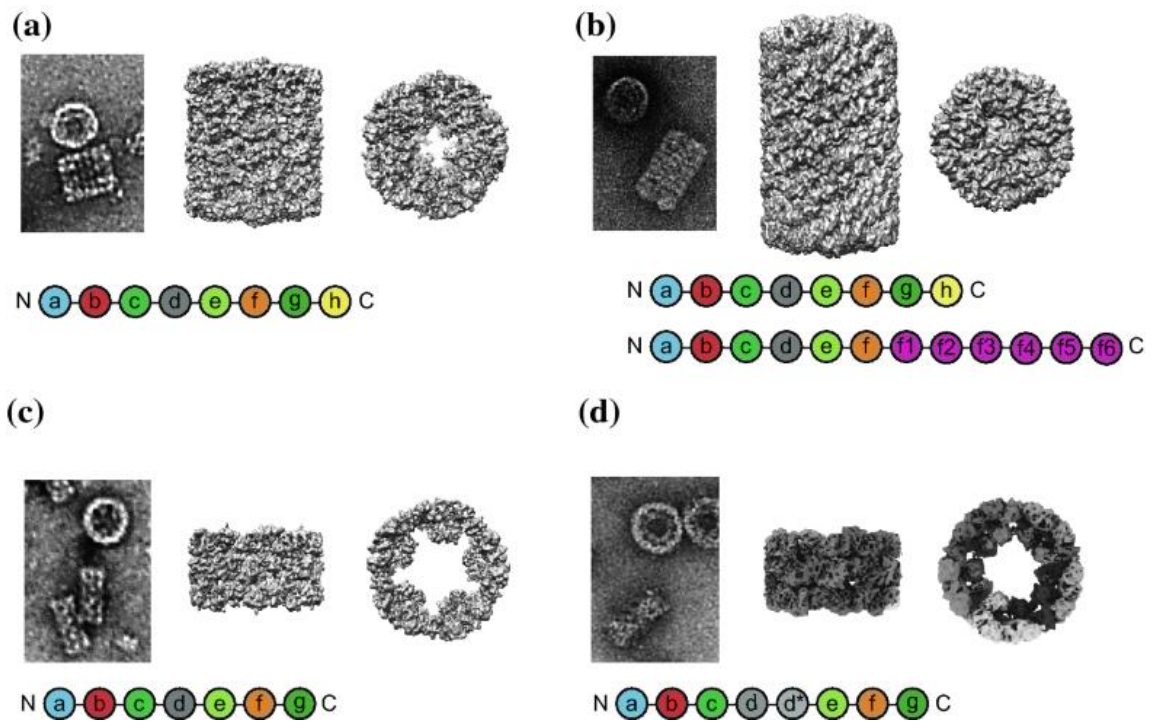
- La hemocianina de molusco tipo 1, es la hemocianina típica de los gastrópodos (lapas o abalones). Este tipo de hemocianina puede ser homo o hetero- didecámeros, con una masa molecular de aprox. 8 MDa. Las subunidades se componen de las FUs basales (FU tipo a-h), la región de la pared del cilindro se compone de las FU-a, -b, -c, -d, -e y -f, mientras que la región del collar interno se compone de las FU-g y -h (Gatsogiannis & Markl, 2009, Zhang y col., 2013) (**Figura 3A**).
- La hemocianina de molusco tipo 2, es la de caracoles ceritioideos del grupo de los gastrópodos. Este tipo de hemocianina es llamada megahemocianina porque su masa molecular es de aprox. 13,5 MDa, presentando un plegamiento tridecámerico. Posee dos subunidades distintas con una masa molecular de 550 KDa (llamada mega subunidad) y 400 KDa (correspondiendo a una subunidad típica). La composición de las FU de la subunidad típica es similar a las de la hemocianina tipo 1. Sin embargo, la composición de la mega subunidad comprende 12 FUs, de las cuales 6 son FU basales desde -a hasta -f y 6 son FU adicionales especiales, denominadas -f1 hasta -f6, localizadas en el C-



terminal de la mega-subunidad. La región de la pared del cilindro se compone de las FU basales de -a hasta -f, mientras que las FU restantes -g, -h y las 6 FU-f especiales componen la región del collar interno, llenando el hueco central del cilindro (Gatsogiannis y col., 2015) (**Figura 3B**).

- La hemocianina de molusco tipo 3, es la de nautilus y pulpos. Este tipo de hemocianina forma un decámero con una masa molecular de aprox, 3,5 MDa. La composición de las FU son las 7 FUs basales de la -a hasta -g, careciendo de la FU-h. Al igual que las hemocianinas tipo 1 y 2, la región de la pared comprende las FU -a hasta -f, mientras que la región del collar interno solo contiene la FU-g, por lo que el hueco central es más grande que las de tipo 1 (Gatsogiannis y col., 2007) (**Figura 3C**).

- La hemocianina de molusco tipo 4, es la de calamares y sepias. Este tipo de hemocianinas forma un decámero con una masa molecular de aprox. 3,8 MDa. La composición de las FU comprende 8 FUs, 7 FUs basales de la -a hasta -g, sin embargo, posee una FU tipo -d\* especial, como quinta FU al lado de la FU tipo -d, la que se piensa que se generó a partir de una duplicación génica de la FU-d y no presenta la FU-h. La región de la pared es igual a la de los otros tipos de hemocianina, mientras que la región del collar interno contiene las FU-d\* y FU-g, por lo que el hueco central se ve más poblado que las de tipo 3, que solo poseen FU-g (Matsuno y col., 2015, Gai y col., 2015, Tanaka y col., 2019) (**Figura 3D**).



**Figura 3. Estructuras de hemocianinas de molusco.** Área representativa de microscopía electrónica de transmisión (TEM) de tinción negativa (izquierda), modelo 3D de los tipos de hemocianina de molusco (derecha) y esquema de tipos de subunidades de hemocianina (abajo). **A)** Hemocianina tipo 1 (Tipo “Keyhole limpet”), imagen TEM de hemocianina de *Nordotis discus hannai*, estructura 3D de cryo-EM de hemocianina de “Keyhole limpet”. **B)** Hemocianina tipo 2 (Mega-hemocianina, *Melanoides tuberculata*). **C)** Hemocianina tipo 3 (Tipo Nautilus), imagen TEM de hemocianina de *Enterooctopus dofleini*, estructura 3D de hemocianina de Nautilus. **D)** Hemocianina tipo 4 (Tipo Calamar), imagen TEM y estructura 3D de cryo-EM de *Todarodes pacificus*. Esquemas de subunidades de hemocianina: *a, b, c, d, e, f, g, h*, corresponden a unidades funcionales (FU) basales, mientras que *d\*, f1, f2, f3, f4, f5, f6* corresponden a FUs especiales. Figura tomada directamente de Kato y col., 2020.

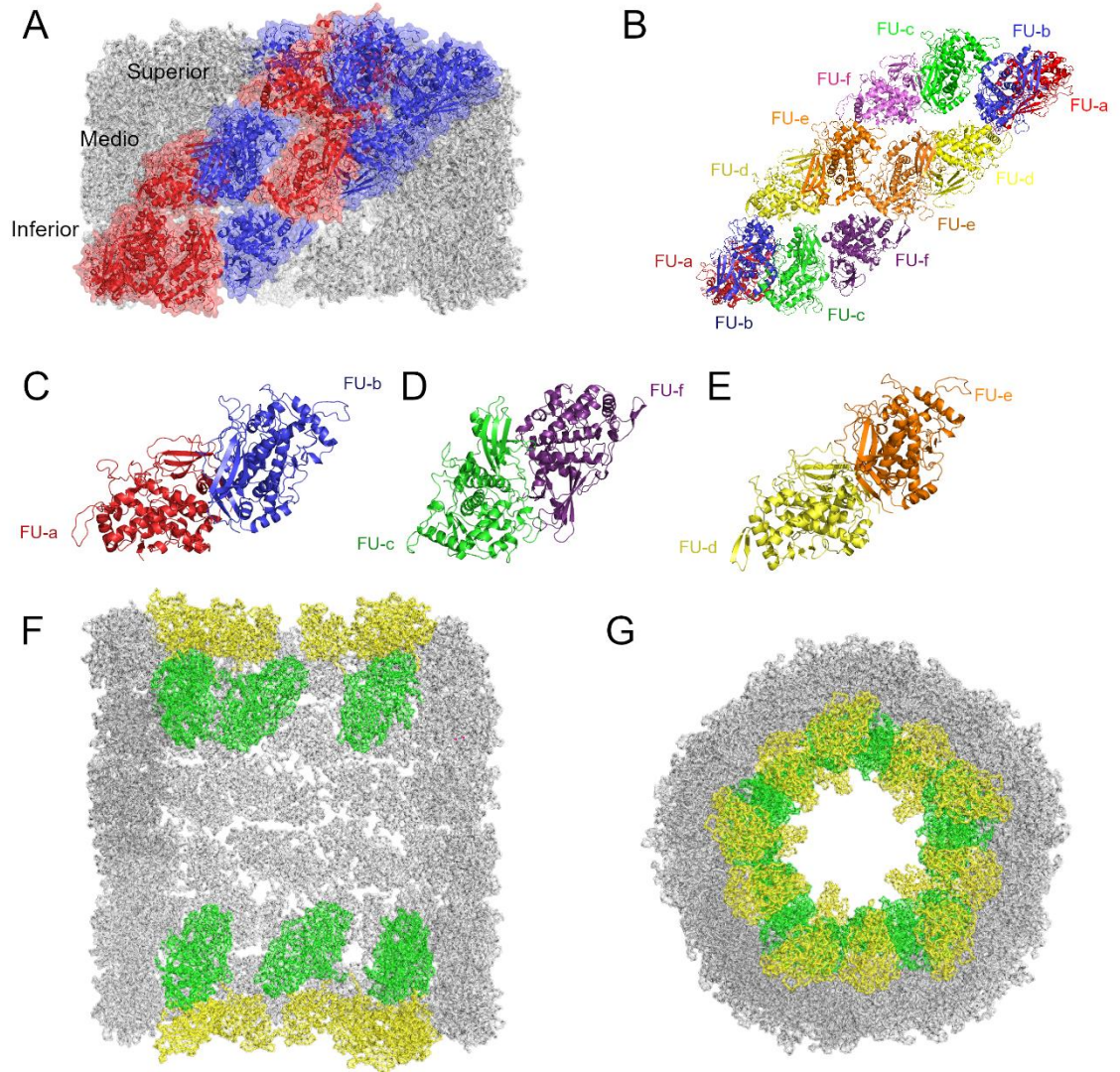
#### 1.4 Hemocianinas de molusco de tipo 1

Históricamente, el modelo de estudio de hemocianina de tipo 1 ha sido la KLH, siendo la más caracterizada hasta la fecha. Sin embargo, aunque se encuentra reportada la estructura de cryo-EM de KLH (PDB: 4BED, Gatsogiannis & Markl, 2009), esta tiene una resolución de 9 Å, la que no permite analizar correctamente la disposición de las FUs dentro del didecámero. La estructura de la hemocianina de *H. diversicolor* (HDH, PDB: 3J32, Zhang y col.,2013) corresponde al modelo de hemocianina obtenido por cryo-EM reportado a mayor resolución hasta la fecha (4,5 Å), lo que ha permitido estudiar de mejor manera la disposición de FUs. Sin embargo, esta estructura solo cuenta con las cadenas laterales de cada subunidad y carece de densidad para ubicar el dominio extra de las FUs tipo -h que forma parte de la región del collar interno de la hemocianina. Además, como se mencionó anteriormente, estos modelos corresponden a homodidecámeros, no existiendo, hasta la fecha, ningún modelo de heterodidecámero.

Tomando la estructura de homodidecámero de HDH, se sabe que las hemocianinas de tipo I forman didecámeros que incluyen todas las FUs basales (-a hasta -h). La región de la pared del cilindro comprende 10 subunidades ensambladas, donde 2 subunidades forman un dímero de subunidades en forma de placa con simetría doble, y 5 de estos dímeros constituyen el cilindro completo con simetría D5 (**Figura 4A**). La pared del cilindro está formada por las FUs -a, -b, -c, -d, -e y -f, y la placa de dímeros de subunidades a su vez está constituida por dímeros de FUs (FU-a-b, FU-c-f y FU-d-e) (**Figura 4B**). En la pared, la placa de dímeros de subunidades puede dividirse en 3 partes: superior, media e inferior (**Figura 4A**), en donde las partes superior e inferior están formadas por los dímeros de

FU-a-b y FU-c-f (**Figura 4C, D**), mientras que la parte media está formada por dos dímeros idénticos (FU-d-e) (**Figura 4E**).

Respecto a la región del collar interno, que no es común para todas las hemocianinas, las FU-g y FU-h forman un dímero FU-g-h entre dos subunidades distintas, cada una situada en una placa de dímeros diferente dentro de la región de la pared (**Figura 4F**). Cinco dímeros de FU-g forman una placa interna dentro del lumen del cilindro, dividida simétricamente en 5 partes dentro de la pared decámera de simetría D5, por lo que el didecámero posee un par de 5 dímeros de FU-g en la parte superior e inferior. Por otra parte, la FU-h cubre la parte superior del dímero FU-g, formando dímeros entre subunidades (**Figura 4G**). De esta manera, el collar interno está formado por los dímeros entre subunidades FU-g y FU-h, los que se encuentran ubicados en ambos extremos exteriores del cilindro (**Figura 4F**).



**Figura 4. Arquitectura de la pared y collar interno de la hemocianina de molusco tipo 1.** **A)** Placa formada por un dímero de subunidades idénticas diferenciadas en colores azul y rojo. Se distinguen las porciones superior, media e inferior de la placa de dímeros. **B)** Arquitectura de la placa de la pared formada por dímeros de FU. **C)-E)** Tres tipos de dímeros de FU que conforman la placa de la pared formada por dímeros de subunidades. **F)** Vista lateral del collar interno de la hemocianina de molusco tipo 1. **G)** Vista frontal del collar interno. La región de la pared se muestra de coloración gris, mientras que las FU-g y FU-h están destacadas de color verde y amarillo respectivamente. Las representaciones mostradas corresponden al homodidecámero de *Haliotis diversicolor* (HDH). Imagen recreada a partir de Kato y col., 2020.

## 1.5 Hemocianinas recombinantes

Resulta importante mencionar que toda la información estructural reportada de hemocianinas de molusco proviene de hemocianinas silvestres, y que prácticamente no existe información estructural reportada proveniente de hemocianinas recombinantes, ya sea de la expresión de hemocianinas completas, de subunidades o de FUs. Al respecto, tras una extensa búsqueda bibliográfica solo se encontró una tesis doctoral (Boisguérin, 2006) en la cual expresaron de forma recombinante en *E. coli* cada una de las FU correspondiente a las dos isoformas de KLH (KLH1 y KLH2). Para esto utilizaron un tag de histidina para purificar las proteínas mediante una columna de afinidad y un tag GST (Glutación-S-transferasa) para mejorar la solubilidad de las proteínas expresadas. Además, se utilizaron diversas metodologías para poder purificar las proteínas desde cuerpos de inclusión. Para evaluar la expresión de las FUs recombinantes solo se realizó un SDS-PAGE, identificando las proteínas mediante inmunoprecipitación con anticuerpos monoclonales anti-His y anticuerpos policlonales anti-KLH1 y anti-KLH2, obteniendo con estos últimos una respuesta débil comparada con la hemocianina silvestre, concluyendo que la imposibilidad de las células bacteriales de agregar glicosilaciones es la razón de la débil respuesta obtenida. Sin embargo, no se utilizó ningún método más específico para comprobar que las proteínas expresadas se encuentren correctamente plegadas y, además, hasta la fecha, esa tesis doctoral no ha sido publicada. Por otra parte, se encontró que la empresa comercial estadounidense CUSABIO® ofrece de forma liofilizada las FU-a y FU-e de la subunidad 2 de la hemocianina de *Rapana venosa* (RVH2), expresada de manera recombinante en levadura, *E. coli*, baculovirus y células de

mamífero, pudiendo comprarse además con un tag de histidina en el N-terminal. Sin embargo, al ser una empresa comercial, no puede accederse a los protocolos utilizados para la expresión y purificación de ambas FUs.

### **1.6 Hemocianina de *Concholepas concholepas* (CCH)**

Como se mencionó anteriormente, la hemocianina de tipo 1 más estudiada, tanto del punto de vista bioquímico como estructural, corresponde a la KLH. Sin embargo, en el último tiempo, han aparecido otras hemocianinas de tipo 1 que se alzan como mejores candidatas para su uso en aplicaciones inmunológicas, como lo son las hemocianinas de los gastrópodos *Fissurella latimarginata* (FLH) (Arancibia y col., 2014) y *Concholepas concholepas* (CCH) (De Ioannes y col., 2004), siendo este último un organismo encontrado exclusivamente en las costas de Perú y Chile, donde se conoce comúnmente con el nombre de “Loco”.

La CCH posee características que la distinguen de otras hemocianinas, como lo son su gran estabilidad y solubilidad. Al contrario de otras hemocianinas, la CCH purificada no requiere de cationes divalentes adicionales como  $\text{Ca}^{2+}$  o  $\text{Mg}^{2+}$  en el medio de almacenamiento para mantener su estructura (De Ioannes y col., 2004). Esta característica facilita las reacciones de acoplamiento químico cuando CCH es utilizada como proteína carrier, por ejemplo, CCH ha sido utilizada exitosamente como carrier de la hormona liberadora de gonadotropina (GnRH) en una vacuna anticonceptiva para controlar la reproducción en ciervos, proporcionando un efecto anticonceptivo más duradero (Miller y col., 2013). Más aún, CCH ha sido evaluada en estudios preclínicos en un modelo experimental murino de cáncer de vejiga superficial y puede considerarse como una

alternativa segura a la terapia con KLH (Molledo y col., 2006; Arancibia y col., 2012). Finalmente, el apoyo más importante para la aplicación clínica de CCH en futuros desarrollos biomédicos se centra en los resultados de un estudio de su uso como adyuvante en una vacuna contra el cáncer basada en células dendríticas (CD) para pacientes con cáncer de próstata resistente a la castración. El estudio demostró que la CCH fue capaz de inducir una respuesta de memoria inmune, medida por la prueba de hipersensibilidad de tipo retardado, y no produjo reacciones adversas tóxicas o alérgicas cuando se administró por vía subcutánea en los pacientes (Reyes y col., 2013). Por lo que, teniendo en cuenta el enorme potencial de CCH en aplicaciones médicas y biotecnológicas, resulta fundamental conocer su estructura molecular y entender en detalle su organización estructural.

Pese al enorme potencial biomédico de la CCH, no hay reportada información empírica acerca de su estructura, habiendo solo caracterizaciones bioquímicas (Oliva y col., 2002; De Ioannes y col., 2004; Becker y col., 2009). Al respecto, la CCH correspondería a una hemocianina de molusco tipo 1, como la mayoría de los gastrópodos. Mediante microscopía electrónica se reportó que la CCH exhibe una estructura predominantemente didecámerica, y utilizando la metodología de filtración en geles se determinó que posee una masa molecular de 8 MDa. Por otra parte, mediante análisis de columna FPLC MonoQ y Western-blots con anticuerpos monoclonales específicos, se demostró que la CCH correspondería a un heterodidecámero, esto quiere decir que posee dos subunidades unidas no covalentemente, llamadas CCHA y CCHB con masas moleculares de 405 y 350 kDa respectivamente (De Ioannes y col., 2004). Acerca del contenido de carbohidratos, se



reportó que cada subunidad de CCH se encuentra diferentemente glicosilada; CCHA (3,6% carbohidratos p/p) posee N-glicosilaciones y O-glicosilaciones, mientras que CCHB (2,5% carbohidratos p/p) posee solo N-glicosilaciones, donde las O-glicosilaciones se encuentran prácticamente ausentes (Becker y col., 2009).

Sin embargo, no existe información de la organización estructural de la CCH y de sus subunidades CCHA y CCHB, debido a que no hay hetero-oligómeros con estructura reportada para hemocianinas de moluscos. Tampoco se dispone de información estructural de las unidades funcionales (FU) de CCH, ni de cómo se encuentran organizadas dentro de las dos subunidades, CCHA y CCHB. Mas aún, su secuencia aminoacídica es desconocida, y solo se cuenta con dos fragmentos parciales de secuencia (llamados CCHX y CCHY), las cuales fueron obtenidos en el laboratorio FUCITED (Dra. María Inés Becker) mediante secuenciación RACE (del inglés Rapid Amplification cDNA Ends), una técnica que permite amplificar el cDNA desde los extremos 5' o 3' de secuencias desconocidas obtenidas a partir de muestras de RNA purificadas desde el hepatopáncreas de *C. concholepas*, donde se ha reportado que se sintetiza la CCH (Manubens y col., 2010).

Como se mencionó anteriormente, la mayoría de la información estructural de hemocianinas de molusco se ha obtenido mediante la combinación de las metodologías de cristalografía de rayos X y Cryo-EM. Desde el año 2013, el progreso en el análisis de partículas individuales mediante Cryo-EM ha sido tan rápido que se ha llamado “la revolución de la resolución” (“resolution revolution”) (Kuhlbrandt, 2014). Esto ha sido posible debido al desarrollo tecnológico en los detectores de microscopios, en el

procesamiento de imágenes y en la preparación de las muestras, entre otros. El uso de detectores directos de electrones, como el que posee el crio-microscopio Titan Krios (LNNano, CNPEM, Brasil) ha permitido la reconstrucción de la estructura de muestras de gran tamaño (~4 MDa) a resoluciones cercanas a la atómica (3 a 4 Å) (Bai y col., 2015).

Dados a los antecedentes mencionados anteriormente, se propone determinar la estructura cristalográfica con resolución a nivel atómico de al menos una FU obtenida de manera recombinante de la hemocianina de interés biomédico CCH, y utilizando dicha estructura, se propone reconstruir un modelo con resolución cercana a la atómica (3 a 4 Å) del heterodidecámero de CCH a partir de un mapa de densidad obtenido mediante cryo-EM y determinar por primera vez la organización estructural de subunidades distintas (CCHA y CCHB) para una hemocianina heterodidecámerica. Finalmente, esta información estructural de CCH no solo será fundamental para las futuras aplicaciones de CCH en medicina y biotecnología, sino que será un aporte para el estudio estructural de las hemocianinas de molusco.

## **2. HIPÓTESIS Y OBJETIVOS**

### **2.1 Hipótesis**

La obtención de la estructura cristalográfica de una FU y el análisis de partículas individuales mediante crio-microscopía electrónica, permitirá conocer la organización tridimensional de las subunidades en heterodidecámero de la hemocianina de *C. concholepas*

### **2.2 Objetivo general**

Obtener la estructura cristalográfica de al menos una FU de origen recombinante y determinar la organización tridimensional de las subunidades por cryo-EM del heterodidecámero de la hemocianina de *C. concholepas*.

### **2.3 Objetivos específicos**

**1. Obtener la estructura cristalográfica de al menos una FU de CCH de origen recombinante.**

- a) Purificar las FU de CCH de secuencia conocida expresadas de manera recombinante en *E. coli*.
- b) Obtener cristales de las FU de CCH purificadas con calidad adecuada para la difracción con rayos X.
- c) Resolver la estructura de la FU de CCH y realizar el refinamiento a partir de datos de difracción de rayos X.

**2. Determinar la organización tridimensional de las subunidades en del heterodidecámero de CCH mediante criomicroscopía electrónica.**

a) Obtener un mapa de densidad electrónica del decámero de CCH a partir del análisis de partículas individuales obtenidas por criomicroscopía electrónica.

b) Obtener un modelo para la organización de las subunidades CCHA y CCHB, acoplando información de la estructura cristalográfica de alguna FU al mapa de densidad obtenido por criomicroscopía electrónica.

### 3. MATERIALES Y MÉTODOS

#### 3.1 Materiales

##### 3.1.1 Medios de cultivo

Para 1 litro de medio de cultivo Luria-Bertani Broth (LB) empleado para la expresión recombinante de proteínas, se mezclaron 10 g de triptona, 10 g de NaCl, 5 g de extracto de levadura, y se aforó a un volumen de 1 litro empleando agua destilada a pH 7,0.

##### 3.1.2 Reactivos

Los reactivos NaCl, Peptona e Imidazol fueron adquiridos en Merck (Kenilworth, NJ, Estados Unidos), mientras que los reactivos extracto de levadura y Tris base fueron adquiridos en Winkler (Valle Grande, Lampa, Santiago de Chile). Por otra parte, los reactivos 2-mercaptoetanol, fluoruro de fenilmetilsulfonilo (PMSF) y elastasa pancreática de cerdo fueron adquiridos en Sigma-Aldrich. Las columnas Histrap HP y Hitrap Q fueron adquiridas en GE Healthcare (Chicago, IL, Estados Unidos).

##### 3.1.3 Vectores

Los genes con las secuencias que codifican para las FUs de CCH (CCHX2 y CCHY1, **Figura 5A**) fueron enviados para su síntesis a la compañía comercial GenScript (USA Inc., NJ, Estados Unidos). Los genes se recibieron subclonados en dos vectores de expresión en *E. coli* con tags de solubilización distintos, haciendo un total de 4 constructos (**Figura 5B**). El vector pET32a (GenScript) presenta un marcador de selección por resistencia a ampicilina y un tag que codifica para tiorredoxina A de *E.coli* (TrxA). Además, se agregó un tag de 6 histidinas y un sitio de reconocimiento para el corte con

proteasa TEV en el N-terminal de las secuencias subclonadas entre los sitios de restricción MscI y XhoI. El vector pHYRSF53 (Addgene) presenta un marcador de selección por resistencia a kanamicina, un tag de 6 histidinas, un tag SUMO (del inglés Small Ubiquitin-like Modifier) que codifica para para la proteína homóloga SMT3 de *S. cerevisiae* en el N-terminal de las secuencias subclonadas entre los sitios de restricción NdeI y XhoI.

**A**

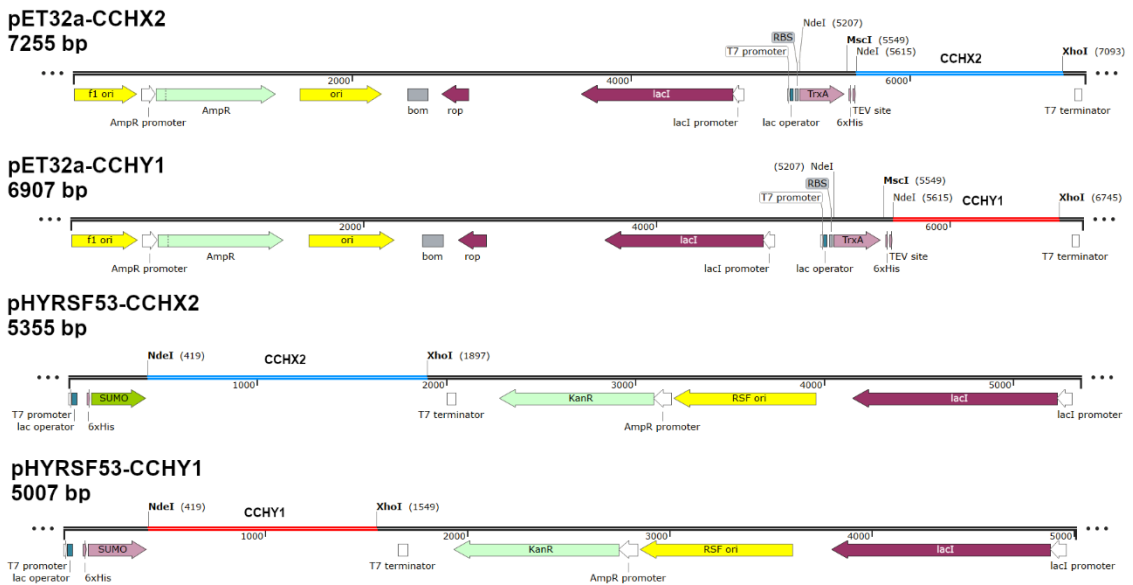
**CCHX2 (CCHB-h)**

DLTRHDLTSLSIPQISNLADALYKLQNDHGPNGF EAIAS YHGAPGLCPENATEHYSCCQHGM P  
 AFPHWHRLTLVQFEHALKEKGA VIGVPYWDWTRPVKGMPTLFTDNSEINPFHTYLMSFNH  
 HIERQVQEEMFTHSAEGDPESLFHQALELLEETNYCDFEVQYEMLHNAVHEL VGGSHYYSMS  
 TLEYSAFDPFFMVHSSIDRIWQIWQTLQKLRHKPFNYAVCAARSMYRPLEPFSYESINPDPL  
 TRENKPAQIFDSHKFHYHYDNLNLNGH SVMELNTMIHAMQAQSRTFAGFVLSGIGTSARVH  
 VDIFKGEETA AVGNFFVLGGSTEMPWAYERIKLDMTEAAKKLGLSGSSHDFDKLTVTKYD  
 GTALDAKFPDPVIVKRRTNAEFDELILPIRKENKLPKIVVRRGTRVLFHPTEEGIAGPIRELGS  
 FTNSEHCAIPPGHAHA YELDVLHVLEPGDY YFVSNLEACKAGSRVQISVDEE

**CCHY1 (CCHA-d)**

SHIRNIEDLTKGERESLRSALHDM EEDGGFEKIAKFHGYPGMCTHDNHPASCCVHGSPTFPH  
 WHRLYVENVENALLNHGSAVSM PYWDWTQPIHHLPLIAEPTY YDSRTHHKEPNAFFRWKV  
 PDTEEYTSRDPRPELYDSHYFLDNVLLALEQTSFCDFEVQFEIVHNGLSFLGGRGQYSMSTL  
 DYSAFDPVFFLHHANVDRIWAIWQALQKFRGLPFDETDCAMNLMNIPLHPFDYDEL DYQYD  
 NLNFHGLTIAQLEHLL EERKSEDR AFLNFMLKGIGTSADVEFELCDKDNHCNFAGTFAVLGG  
 PLEMPWSFDRVFKYDVTKVFKQMLRPDSEYHFEH HIRAVNGTDLNPHLLKPPSVSFPVGRK  
 SY

**B**



**Figura 5. Secuencias de FU y mapas de vectores. A)** Secuencias aminoacídica de las FU identificadas a partir de fragmentos de secuencia de CCH (CCHX2 y CCHY1). **B)** Mapas de los vectores pET32a y pHYRSF53 subclonados con los genes que codifican para CCHX2 (destacado en rojo) y CCHY1 (destacado en azul). Los sitios de restricción utilizados se encuentran destacados en negro.

## **3.2 Métodos**

### **3.2.1 Cuantificación de proteínas**

La determinación de la concentración de proteínas se realizó mediante el método de Bradford (Bradford, 1976) empleando albúmina de suero bovino (BSA) como estándar.

### **3.2.2 Expresión y purificación de FUs recombinantes de CCH**

Células de *E. coli* BL21 fueron transformadas con cada uno de los 4 constructos de FUs de CCH (CCHX2 y CCHY1 en vectores pET32a y pHYRSF53). La estrategia de expresión en donde se obtuvo los mejores rendimientos de purificación para todos los constructos consistió en crecer las células transformadas con 500 mL de medio LB con antibiótico durante toda la noche a 37°C e inducir la expresión de proteínas con 1 mM de IPTG durante 2 horas a 37°C (inducción en estado de crecimiento estacionario de *E. coli*). Para la purificación, las células se centrifugaron a 6.000 rpm por 15 min. y se resuspendieron en amortiguador de purificación (Imidazol 20 mM, NaCl 300 mM, TrisHCl pH 7,8 50 mM, 2-mercaptoetanol 5 mM y PMSF 1 mM). Posteriormente, las células resuspendidas se sonicaron empleando un protocolo de 15 pulsos de 20 segundos de duración a un 40% de amplitud, con descansos de 1 minuto (Digital Sonifier 450, Branson). A continuación, las células lisadas se centrifugaron a 11.000 rpm por 20 minutos para separar la fracción soluble, la cual se cargó en una columna de afinidad por níquel (Histrap HP 5 mL) previamente equilibrada con amortiguador de purificación. Las proteínas fueron eluidas utilizando un gradiente lineal de 100 mL de 20 a 500 mM de imidazol. La pureza de las fracciones colectadas fue evaluada a través de SDS-PAGE.



### **3.2.3 Disociación y purificación de subunidades silvestres de CCH**

La CCH silvestre purificada (Inmunocyanin®) fue facilitada por la Fundación FUCITED, quienes la purifican según lo descrito por Hercowitz y col., con ligeras modificaciones (De Ioannes y col., 2004). La disociación de las subunidades silvestres CCHA y CCHB se llevó a cabo a un pH básico y la purificación de éstas se realizó a través de cromatografía de intercambio aniónico, según lo descrito en Becker y col., 2009. Brevemente, 26 mg de CCH silvestre fueron dializados por 48 horas en 500 mL de amortiguador de disociación (Glicina 130 mM pH 9,6 y EDTA 10 mM), cambiando el amortiguador cada 12 horas aproximadamente. Posteriormente, la muestra dializada se cargó en dos columnas de intercambio aniónico en tándem (Hitrap Q HP 5mL) previamente equilibradas en amortiguador de disociación y las subunidades fueron eluidas incrementando la concentración de NaCl de 400 mM a 600 mM. Los picos correspondientes a las subunidades CCHA y CCHB fueron colectados y concentrados hasta 1 mL utilizando concentradores Amicon 30K MWCO de 50 mL. La pureza de las subunidades colectadas fue evaluada a través de SDS-PAGE.

### **3.2.4 Proteólisis y purificación de FU silvestre de CCH**

Debido al bajo rendimiento de purificación de la subunidad CCHA, se optó por realizar ensayos de proteólisis solo con CCHB, según lo reportado por Becker y col., 2009. Brevemente, al menos 10 mg de CCHB fueron incubados con elastasa pancreática a una concentración de 2% p/p durante 5 horas a 37°C, deteniendo la reacción con la adición de PMSF 1 mM. La muestra de CCHB proteolizada fue cargada en una columna de intercambio aniónico (Hitrap Q HP de 5 mL) previamente equilibrada en amortiguador

que contenía imidazol-HCl 20 mM pH 6,0. Para la purificación de las FUs, se incrementó la concentración de NaCl utilizando un gradiente escalonado de 0 mM hasta 200 mM. La homogeneidad de las fracciones eluidas de la columna fue evaluada a través de SDS-PAGE.

### **3.2.5 Electroforesis unidimensional de proteínas en condiciones denaturantes**

Las muestras se prepararon en solución de carga (SDS 1%, glicerol 10%, TrisHCl pH 6,8 10 mM, beta mercaptoetanol 1 mM y azul de bromofenol) y se calentaron a 95°C por 5 minutos. Las muestras de FU recombinantes y las muestras de FU silvestres obtenidas mediante proteólisis se cargaron en geles de poliacrilamida al 12%, realizando la electroforesis a un voltaje constante de 150 volts por 1 hora. Debido a su gran masa molecular, las subunidades CCHB y CCHA obtenidas mediante disociación de CCH se cargaron en geles de poliacrilamida en gradiente de 5-15% y la electroforesis se realizó a un voltaje constante de 150 volts por 4 horas

Una vez finalizada la electroforesis, los geles se tiñeron toda la noche en solución azul de Coomasie al 0,25 % p/v, para luego ser desteñidos en una solución compuesta por metanol 30%, ácido acético 40% y agua.

### **3.2.6 Ensayos de cristalización de FU silvestre**

Previo a los ensayos de cristalización, la FU de CCH purificada se concentró hasta 8,5 mg/mL en amortiguador de purificación (Imidazol-HCl 20 mM pH 6,0, NaCl 300 mM). El muestreo inicial de cristalización se realizó en placas de 96 pocillos utilizando un equipo automatizado (ARI Crystal Gryphon) mediante el método de difusión de vapor por gota sentada en una mezcla 1:1 de la proteína y las soluciones de 2 *kits* comerciales de

condiciones de cristalización (*kits* Index y SG1, Hampton Research). Se identificó una condición de cristalización favorable en el *kit* Index (PEGMME 2.000 30% p/V y KBr 150 mM). Los cristales crecieron durante 72 horas a 18°C.

Se realizaron ensayos de optimización de cristalización empleando el método de difusión de vapor por gota colgante y un sistema de placa con 24 pocillos. Como estrategia, se mezcló 1 uL de proteína con 1 uL de la solución correspondiente a la condición identificada. En el eje horizontal de la placa se probaron 6 concentraciones del agente precipitante (PEGMME 2000) y en el eje vertical se probaron 4 concentraciones de la sal (KBr). Se utilizaron 2 concentraciones de proteína por condición (5,5 y 3 mg/mL).

### **3.2.7 Difracción de rayos X y análisis de datos**

Los cristales obtenidos de la optimización de CCH-FU fueron colectados, pasados por solución de PEG 400 20% p/V como crioprotector y preservados directamente en nitrógeno líquido. Los cristales fueron difractados de manera remota en la línea de luz MANACA, del sincrotrón de 4ta generación SIRIUS en el Laboratorio Nacional de Luz Sincrotrón (LNLS) en Campinas, Brasil, utilizando un detector PILATUS 2M. Se colectaron conjuntos de datos nativos utilizando una longitud de onda de 0,977 Å y se realizaron experimentos de SAD (del inglés Single-wavelength Anomalous Diffraction) utilizando los mismos cristales y una longitud de onda de 1,37 Å, aprovechando la señal anómala de los dos átomos de cobre presentes en el centro activo de la FU. Para ambos experimentos se utilizó un rango de oscilación de 0,1° colectando un total de 3.600 cuadros.

Los conjuntos de datos anómalos obtenidos mediante experimentos de SAD fueron procesados utilizando la plataforma del programa SHELX (Sheldrick, 2008). Se analizaron las subestructuras de los átomos de cobre en cada *dataset* (SHELXC), se encontraron las posiciones de los dispersores anómalos (SHELXD) y se obtuvieron las fases a partir de la posición de los dispersores anómalos (SHELXE). Como no se contaba con la secuencia de la FU de CCH, se optó por acomodar una cadena principal de polialanina dentro de las densidades electrónicas continuas obtenidas, no obstante, debido a la alta resolución de los conjuntos de datos obtenidos (aprox. 1,7 Å), se pudo asignar cadenas laterales basándose en el análisis visual de las densidades electrónicas, utilizando COOT (Emsley & Cowtan, 2004). Por otra parte, los conjuntos de datos nativos fueron indexados, integrados y escalados utilizando XDS (Kabsch, 2010) y fueron fusionados utilizando Aimless de la plataforma CCP4 (Evans & Murshudov, 2013), obteniéndose un conjunto de datos de alta resolución a 1,5 Å. El posterior refinamiento del modelo obtenido mediante data anómala fue realizado de manera iterativa en COOT y Phenix (Adams y col., 2010) utilizando el conjunto de datos nativos obtenido del mismo cristal a una resolución de 1,5 Å. Las moléculas de agua, bromo y polímeros de PEG fueron colocadas usando COOT, mientras que las N-glicosilaciones fueron colocadas usando la herramienta de adición de N-glicosilaciones de COOT (Emsley & Crispin, 2018). Las moléculas adicionadas fueron refinadas con restricciones generadas en ReadySet/eLBOW en Phenix. Finalmente, la validación de la estructura fue realizada con el servidor web Molprobit (Williams y col., 2018).

### 3.2.8 Cryo-EM de CCH

El experimento de cryo-EM de CCH fue llevado a cabo por el personal del Laboratorio Nacional de Nanotecnología (LNNano) en Campinas, Brasil. La muestra de CCH silvestre purificada facilitada por la Fundación FUCITED fue aplicada a una concentración de 5 mg/mL en una grilla Quantifoil R1,2/1,3 (Quantifoil Micro Tools GmbH, Jena, Alemania), luego fue secada con papel filtro y ultracongelada en etano pre-enfriado usando un Vitrobot Mark IV (ThermoFisher Scientific) con la cámara de muestra mantenida a 4°C y una humedad del 100%. Las grillas posteriormente fueron cargadas en un microscopio electrónico Titan Krios (ThermoFisher Scientific) operado a 300 kV y equipado con una cámara K3 (detector directo de electrones) en modo conteo de electrones a super-resolución. Las micrografías fueron colectadas de manera continua con una dosis total de electrones de 30 e/Å<sup>2</sup> dividida sobre 30 cuadros, con un tamaño de píxel físico de 1,1 Å (0,55 Å en modo super-resolución) y un desenfoque de -0,5, -1,0 y -1,5 μm.

### 3.2.9 Procesamiento de imágenes Cryo-EM y reconstrucción 3D de modelo de CCH

Las películas colectadas fueron importadas al software cryoSPARC v3.2.0 (Punjani y col., 2017), en donde se realizó la corrección de movimiento y estimación de CTF (del inglés *Contrast Transfer Function*). Las películas fueron filtradas dependiendo de la resolución de ajuste de CTF, usando un valor de corte de 8 Å. De las películas restantes, a partir de 100 micrografías iniciales se seleccionaron automáticamente partículas de CCH para ser usadas en una primera ronda de alineación y clasificación 2D supervisada. Los mejores promedios de las clases 2D resultantes fueron seleccionados por inspección visual y luego fueron usados como referencia para una nueva ronda de selección de partículas basadas

en la correlación. Las nuevas partículas seleccionadas fueron usadas para una nueva ronda de alineación y clasificación 2D para encontrar los mejores grupos de partículas para el procesamiento. Las partículas seleccionadas se extrajeron para una nueva ronda de alineación y clasificación 2D, quedando con un total de 199.405 imágenes de partículas de 600 x 600 píxeles para el procesamiento 3D.

Para el procesamiento 3D, se utilizó una reconstrucción 3D *ab initio* de simetría D5 con el objetivo de generar 3 reconstrucciones 3D, seguidas del refinamiento de cada una. Las imágenes de partículas de la mejor reconstrucción 3D generada (130.952 imágenes) fueron seleccionadas para efectuar rondas de refinamiento 2D y 3D de tipo homogéneo y heterogéneo. La calidad del mapa final de cryo-EM fue evaluada estimando la correlación de la capa de Fourier entre dos mitades del mapa. La resolución global fue estimada a 3,2 Å basándose en el criterio de corte 0,143 (Rosenthal & Henderson, 2003).

### **3.2.10 Generación de modelos de organización de subunidades de CCH**

Utilizando el software MODELLER (Webb & Sali, 2016) se generaron modelos por homología de las FU-a, -b, -c, -e y -f utilizando como templatado las secuencias de FU de ambas subunidades de RVH (RVH1 y RVH2) y utilizando la estructura de CCHB-g. Para el caso del modelo por homología de la FU-d se usó la secuencia de CCHA-d (**Figura 5A, Pág. 22**), para la FU-g, se utilizó la estructura cristalográfica de CCHB-g, mientras que para el modelo por homología de FU-h se usó la secuencia disponible de CCHX2 (**Figura 5A, Pág. 22**), y como templatado se utilizó la FU de KLV-h (PDB: 3QJO). Como templatado de la organización de FUs dentro de las subunidades, se utilizó el modelo de la

hemocianina de *Haliotis diversicolor* (HDH). Se generaron dos modelos I y II, de organización de las subunidades CCHA y CCHB.

El mapa de cryo-EM de CCH fue importado al software ChimeraX 1.3 (Pettersen y col., 2021) y los modelos I y II de organización de subunidades de CCH fueron alineados y ubicados dentro del mapa de CCH mediante *docking* rígido utilizando el mismo software. Ambos modelos fueron refinados utilizando la herramienta refinamiento de espacio real para cryo-EM del software Phenix y COOT, acomodando regiones *loops* dentro de la densidad del mapa de cryo-EM y completando regiones *linker* entre FU mediante cadenas de polialanina. Ambos modelos fueron evaluados utilizando coeficientes de correlación para determinar el grado de ajuste los modelos a la densidad del mapa de cryo-EM de CCH.

## 4. RESULTADOS

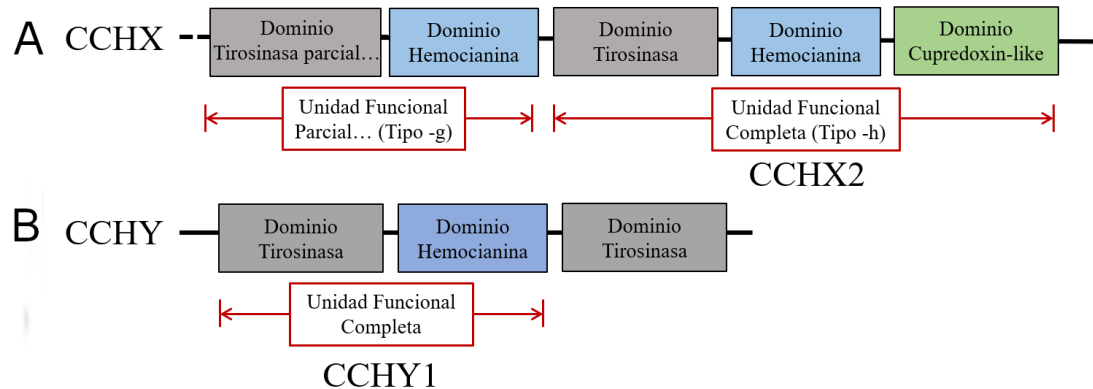
### 4.1 OBJETIVO ESPECÍFICO 1

#### 4.1.1 Identificación de FUs de CCH

Actualmente no se cuenta con la secuencia aminoacídica de CCH, sin embargo, el grupo de investigación de la Dra. María Inés Becker (FUCITED) fue capaz de identificar dos fragmentos parciales de secuencia de CCH (aquí llamados CCHX y CCHY). Estos fragmentos de secuencia fueron analizados utilizando el servidor web CDD (Conserved Domain Database; Lu y col., 2020), mostrando que contenían los dominios N-terminal (Dominio tirosinasa) y C-terminal (Dominio hemocianina) esperados para una FU (**Figura 6**). Mas aún, las secuencias CCHX y CCHY fueron alineadas con todas las FU de las hemocianinas de la isoforma 1 de *Megathura crenulata* (KLH1) y las dos subunidades de *Rapana venosa* (RVH1 y RVH2). Se encontró que el fragmento CCHX posee una FU completa hacia el C-terminal (ahora llamada CCHX2), en la cual, además de reconocer los dominios esperados anteriormente, se reconoce un tercer dominio *cupredoxin-like*, el cual está presente solamente en las FU de tipo h. En el mismo fragmento CCHX, hacia el N-terminal se encuentra una FU parcial, ya que faltan aminoácidos hacia el N-terminal comúnmente encontrados en el dominio tirosinasa. Esta FU sería de tipo g, ya que se encuentra contigua a la FU-h, separada por 19 residuos que corresponderían al linker -gh (**Figura 6A**). Por otra parte, el fragmento CCHY posee una FU completa en el N-terminal (aquí llamada CCHY1) y una FU parcial hacia el C-terminal, puesto que solo presenta un dominio tirosinasa, faltando el dominio hemocianina (**Figura 6B**). Esta FU podría tratarse de una FU-d, ya que posee una identidad de



secuencia de 51% con KLH1-d, y un 71% con FU-d de la subunidad 2 de la hemocianina de *Rapana venosa* (RVH2-d), comparado con el promedio de 40% de identidad de secuencia obtenido con el resto de las FUs de KLH y RVH2.



**Figura 6. Identificación de unidades Funcionales (FU) a partir de fragmentos de secuencia de CCH. A)** Análisis de fragmento CCHX identificando una FU parcial en el N-terminal y una FU completa en el C-terminal denominada CCHX2 (FU-h debido a la presencia de un dominio extra *cupredoxin-like*). **B)** Análisis de fragmento CCHY identificando una FU completa en el N-terminal. Los recuadros rojos delimitan FUs y los rectángulos coloreados representan dominios tirosinasa (color gris), hemocianina (color azul) y *cupredoxin-like* (color verde).

#### 4.1.2 Expresión recombinante y purificación de FU de CCH

Después de identificar dos secuencias de FU completas (CCHX2 y CCHY1, **Figura 5A, Pág. 22**) estas se enviaron para su síntesis a la compañía comercial GenScript USA Inc. Debido a la presencia de glicosilaciones en estas FU (Becker y col., 2009), se contempló la utilización de *tags* de solubilización que ayudaran en el correcto plegamiento de las FUs expresadas de manera recombinante. Para esto, se optó por la utilización de dos sistemas *tag* para los que se reportaba un gran incremento en la solubilidad de las proteínas fusionadas: el *tag* tiorredoxina (TrxA) y el *tag* SUMO (Rosano & Ceccarelli, 2014). El *tag* tiorredoxina (gen TrxA de *E. coli*) actúa como antioxidante, facilitando la reducción

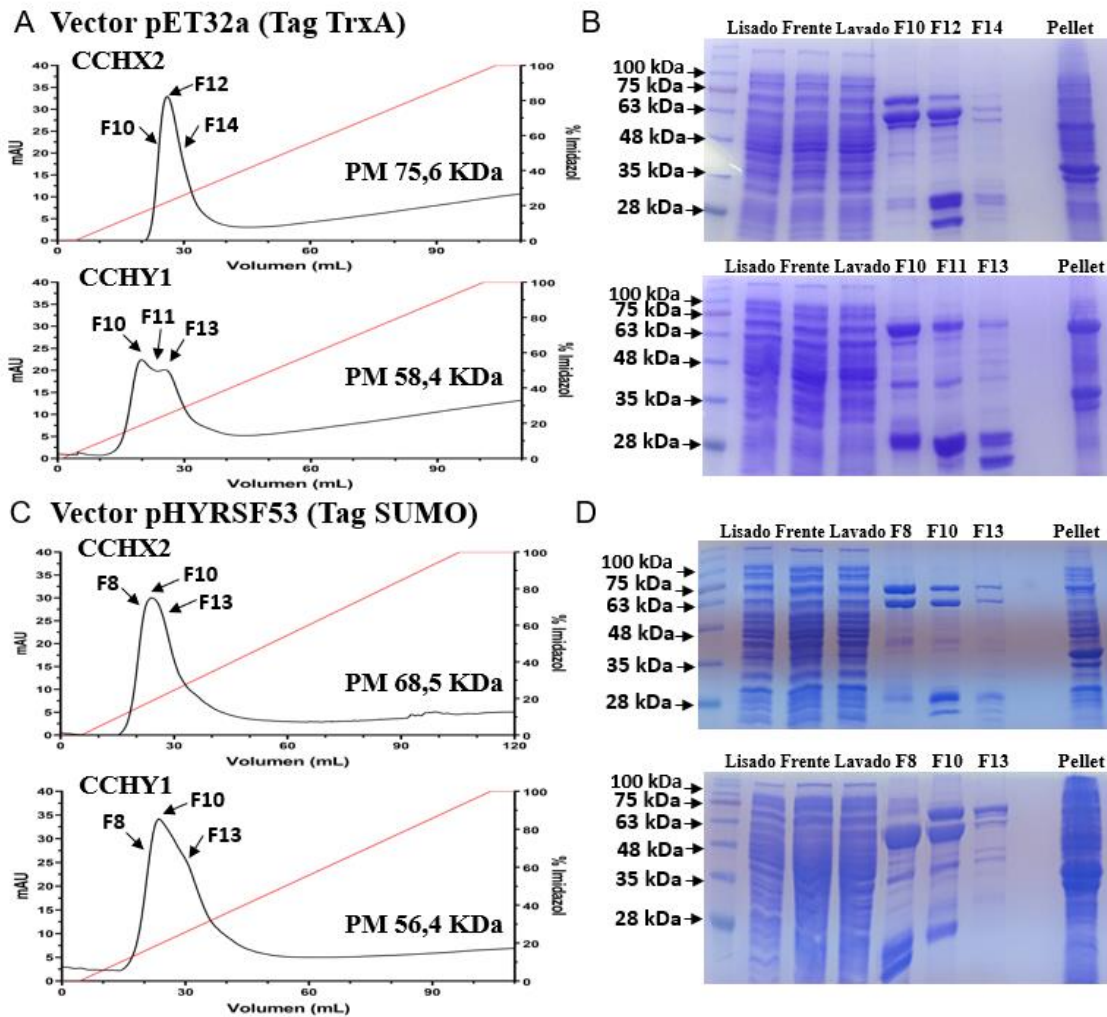
de otras proteínas a través de un intercambio tiol-disulfuro en las cisteínas, facilitando la solubilización (LaVallie y col., 1993 Nordberg & Arnér, 2001). Por otra parte, el *tag* SUMO (gen SMT3 de *S. cerevisiae*) ayuda en la expresión y solubilización de la proteína deseada, dado que se ha sugerido que actúa como un sitio de nucleación para el plegamiento de la proteína fusionada (Guerrero y col., 2015, Butt y col., 2005).

Con el propósito de hacer compatibles ambos sistemas de expresión con las secuencias de FUs, para el caso del vector pET-32a (con el *tag* de tiorredoxina) se solicitó el subclonamiento de los constructos con un 6-*tag* de histidina, un sitio de corte para la proteasa TEV y la secuencia que codifica para las FU entre los sitios MscI y XhoI (**Figura 5B, Pág. 22**). Para el caso del vector pHYRSF53 (con el *tag* SUMO), se solicitó el subclonamiento del constructo en el sitio de corte NdeI, dado que el vector ya incluye un *tag* de histidina en el N-terminal y el *tag* SUMO puede escindirarse a través de la utilización de la proteasa Ulp1, que reconoce específicamente la estructura terciaria del *tag* SUMO (Guerrero y col., 2015) (**Figura 5B, Pág. 22**). Finalmente, se cuenta con 4 proteínas, las dos FUs CCHX2 y CCHY1 con sus respectivos *tags* Tiorredoxina y SUMO.

Se probaron distintas estrategias para la expresión de las 4 proteínas en *E. coli*, variando principalmente el tiempo de expresión. Como consenso para las 4 proteínas, al aumentar el tiempo de expresión, se obtenía menor cantidad de proteína soluble, apreciándose un mayor enriquecimiento de bandas en el tamaño esperado al realizar un SDS-PAGE con muestra de *pellet* de bacterias sonicadas después de la expresión de proteínas, sugiriendo que las proteínas se irían a cuerpos de inclusión. La estrategia con la que se obtuvo mejor rendimiento de purificación para todas las proteínas fue inducir la expresión después de

12 horas de crecimiento (inducción en estado estacionario de *E. coli*) durante 2 horas a 37°C.

El detalle del procedimiento de purificación para las FUs recombinantes se describe en la sección de Materiales y Métodos. En la **Figura 7**, se muestran los cromatogramas de elución de las FU CCHX2 y CCHY1 fusionadas con el *tag* TrxA **Figura 7A, B**) y SUMO (**Figura 7C, D**), con sus respectivas masas moleculares esperadas considerando los *tags* fusionados, además se muestran los respectivos SDS-PAGE con cada uno de los pasos de purificación de las proteínas, en conjunto con las fracciones colectadas de los picos de elución. En el caso de CCHX2 se observa que, independiente del *tag* utilizado, se observa un solo pico en el cromatograma, sin embargo, en las fracciones colectadas se observan bandas enriquecidas en 28 kDa y dos bandas cercanas a los 63 kDa, lo que se encuentra cercano a la masa molecular estimada de CCHX2 considerando los *tags*. Para cada uno de los dos *tags* se obtuvieron alrededor de 3,5 mg totales de CCHX2 a partir de 500 mL de medio LB. Por otra parte, en el caso de CCHY1, independiente del *tag* utilizado, en los cromatogramas se observan dos picos de elución superpuestos, mientras que en el caso de CCHY1 con *tag* SUMO se observan las mismas bandas enriquecidas a 28 y 63 kDa, aunque CCHY1 con el *tag* TrxA muestra una sola banda enriquecida a los 63kDa, cercano a la masa molecular esperada (58,4 kDa). Más aún, en el carril *pellet* con muestra insoluble de células lisadas, se observa la banda de 63 kDa enriquecida. Para cada uno de los *tags* fusionados con CCHY1 se obtuvieron alrededor de 2,8 mg totales a partir de 500 mL de medio LB.

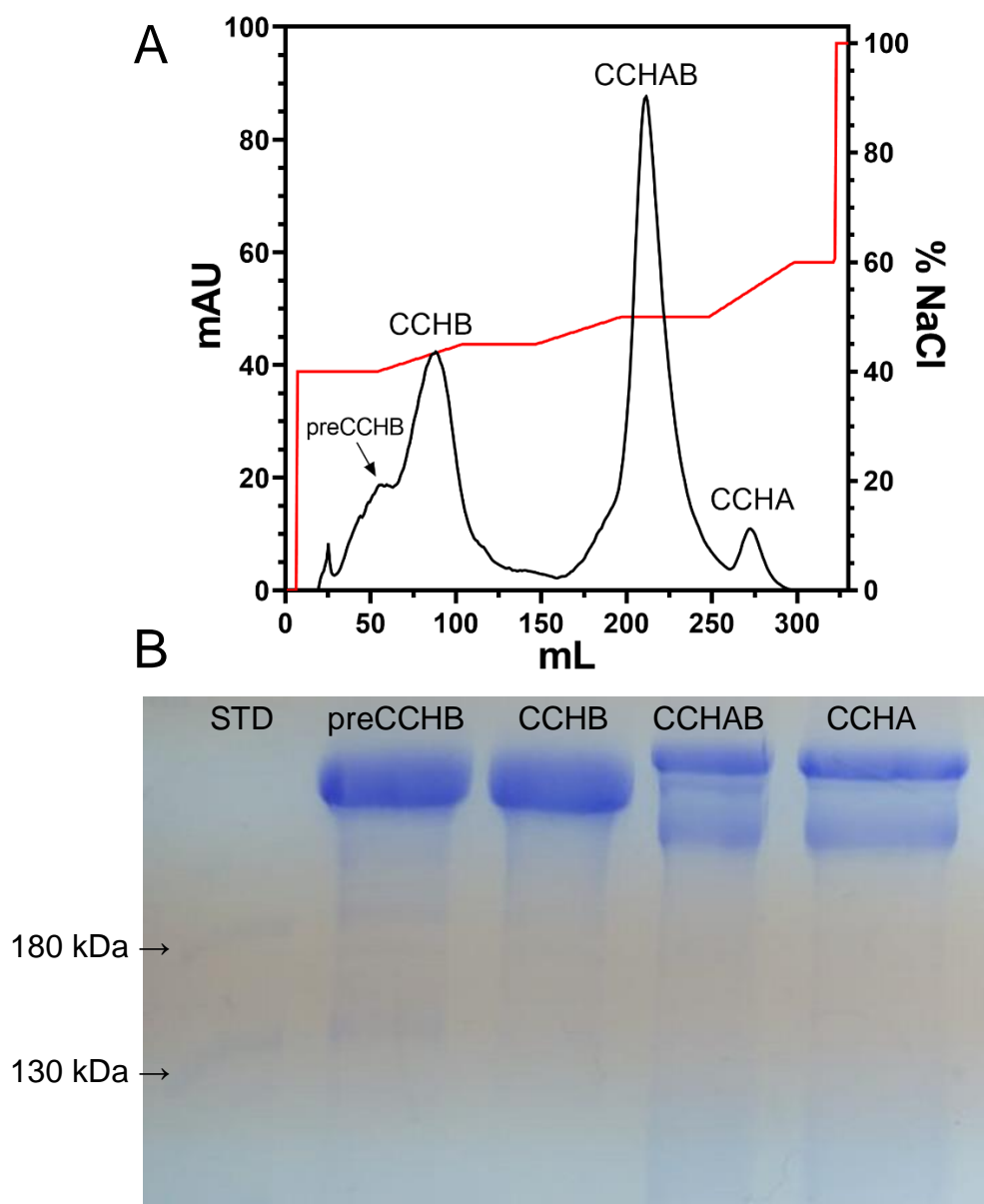


**Figura 7. Purificación de FUs recombinantes de CCH.** A) y C) Cromatogramas de elución de las FUs CCHX2 y CCHY1 subclonadas en los respectivos vectores pET32a (*Tag TrxA*) y pHYRSF53 (*Tag SUMO*), se indican las masas moleculares esperadas de cada constructo y las fracciones correspondientes al pico de elución de cada uno. Las líneas rojas indican el porcentaje de concentración de imidazol en donde 100% corresponde a 500 mM de Imidazol. B) y D) SDS-PAGE con el esquema de purificación de las FUs recombinantes subclonadas en los vectores pET32a y pHYRSF53 respectivamente. Los carriles indican estándar de peso molecular, lisado celular soluble, frente y lavado de columna HisTrap HP, fracciones con proteína eluída de la columna y *pellet* de bacterias lisadas.

Los rendimientos de purificación obtenidos para ambas FU recombinantes fueron bastante bajos teniendo en cuenta que presentan un bajo grado de pureza según lo observado en los SDS-PAGE. Más aún, se realizaron mediciones de absorbancia a las fracciones obtenidas, dado que está reportado que las subunidades de CCH muestran 3 picos característicos (De Ioannes y col., 2004); un pico a 280 nm correspondiente a residuos aromáticos, uno a 345 nm correspondiente al complejo Cu-O<sub>2</sub>-Cu y uno a los 550 nm correspondiente al complejo de coordinación Cu-histidina (**Figura 10B, Pág. 41**). Sin embargo, las muestras purificadas de CCHX2 y CCHY1 solo presentan el pico de absorbancia a 280 nm, lo que sugiere que las muestras purificadas no son capaces de unir cobre.

### 4.1.3 Purificación de FU silvestre de CCH

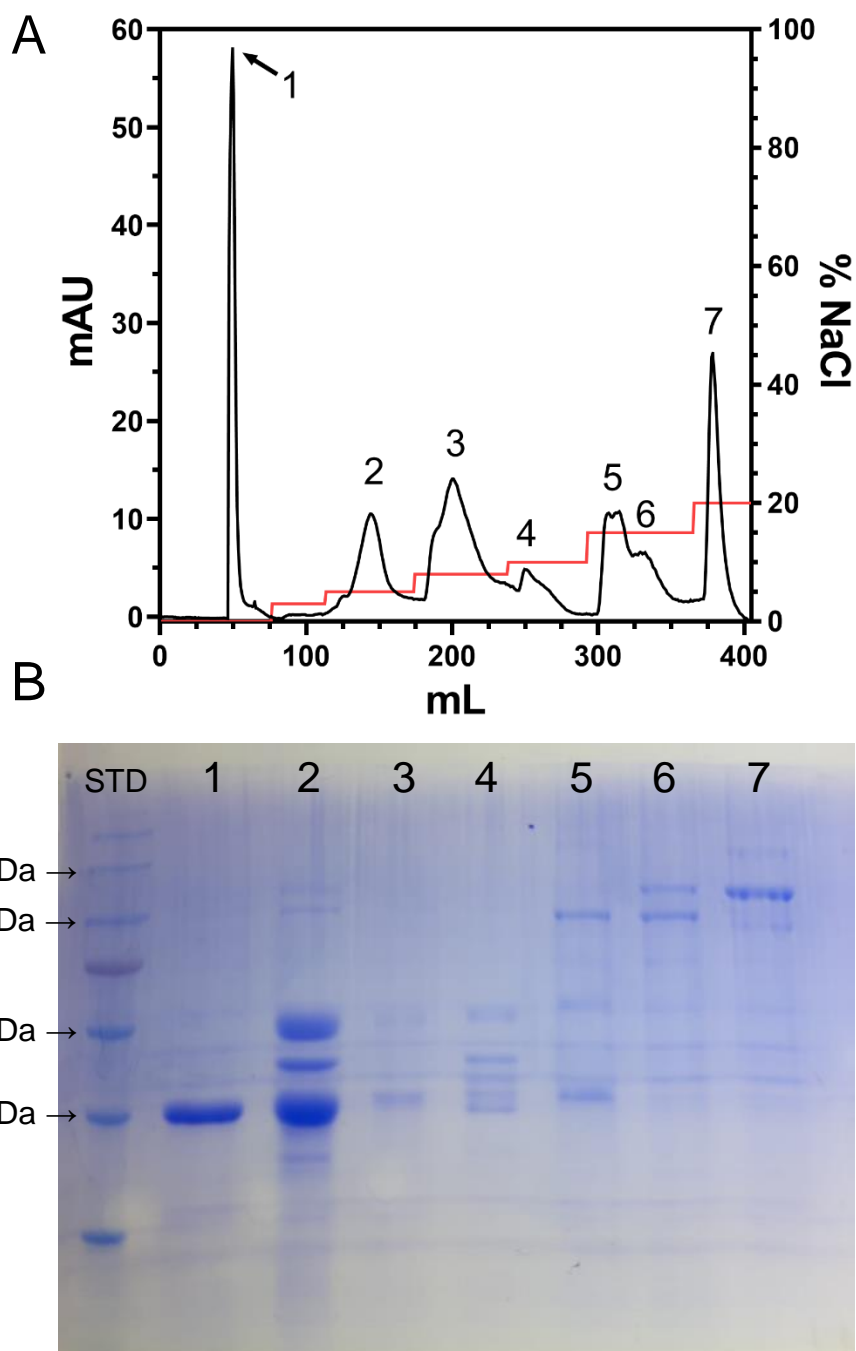
Considerando los resultados obtenidos con la expresión recombinante de FU de CCH, se optó por intentar purificar una FU silvestre de CCH mediante proteólisis con elastasa. Para esto, la fundación FUCITED proporcionó viales de 26 mg de CCH silvestre. El primer paso corresponde a la disociación de las subunidades de CCH (CCHA y CCHB) mediante diálisis en solución alcalina para luego separar las subunidades en una columna de intercambio aniónico Hitrap Q HP. En la **Figura 8A**, se observa el cromatograma de elución de la muestra de CCH disociada. En el primer gradiente lineal desde 40% NaCl (400 mM) hasta 45% NaCl (450 mM) se observan dos picos, ambos correspondientes a la subunidad CCHB, en 50% NaCl (500 mM) se observa un gran pico correspondiente a la co-elución de CCHB y CCHA, mientras que en el último gradiente lineal desde 50% NaCl (500 mM) hasta 60% NaCl (600 mM) se observa un pico correspondiente a la subunidad CCHA, según lo reportado en literatura (De Ioannes y col., 2004). En la **Figura 8B**, se observa el SDS-PAGE de los picos colectados; los dos primeros picos, correspondientes a la subunidad CCHB, se observan como una sola banda de alto peso molecular, mientras que, en el caso del último pico correspondiente a CCHA, se observan dos bandas de alto peso molecular, lo que corresponde a lo reportado, ya que CCHA posee un sitio auto-proteolítico que se activa en las condiciones denaturantes del SDS-PAGE (De Ioannes y col., 2004). Finalmente, en el tercer pico correspondiente a la co-elución de CCHB y CCHA, se observan las 3 bandas combinadas correspondientes a la mezcla de las dos subunidades.



**Figura 8. Purificación de las subunidades de CCH.** **A)** Cromatograma representativo de la separación de CCHA y CCHB mediante intercambio aniónico (Hitrap Q HP). La línea roja representa el gradiente de NaCl en donde 100% corresponde a 1M NaCl. La línea negra representa la absorbancia a 280 nm, en donde los dos primeros picos corresponden a subunidad CCHB, el tercer pico corresponde a la coelución de CCHA y CCHB y el último pico corresponde a la subunidad CCHA. **B)** SDS-PAGE representativo de los picos obtenidos de la columna de intercambio aniónico.

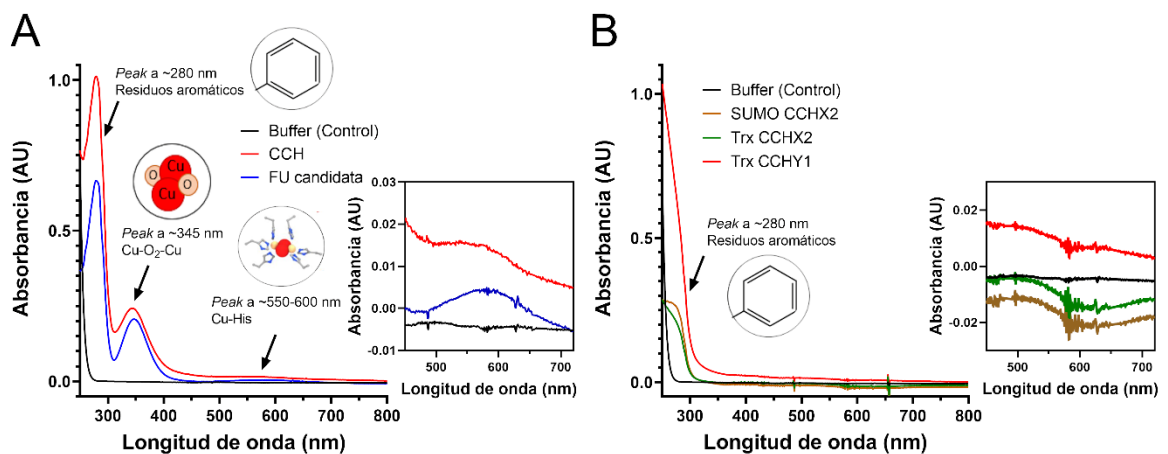
Como resultado de cada proceso de purificación de subunidades de CCH, se obtuvieron alrededor de 6,5 mg totales de CCHB, 0,5 mg totales de CCHA y 9 mg totales de coelución de CCHA y CCHB. Debido al bajo rendimiento de purificación de CCHA, se optó por realizar ensayos de proteólisis solo con CCHB, de la cual se incubaron alrededor de 10 mg junto con elastasa pancreática. Posteriormente, con el fin de lograr aislar una sola FU, se hicieron ensayos con la subunidad digerida dializada en diferentes rangos de pH antes de cargar las muestras en la columna de intercambio aniónico. La condición que generó mejores resultados fue dializar la muestra digerida contra amortiguador imidazol-HCl 20 mM pH 6,0, dado que en la elución de la columna de intercambio aniónico se observa un pico de gran tamaño en el primer lavado sin NaCl (**Figura 9A**). Al evaluar la fracción colectada en un SDS-PAGE, se observa una sola banda homogénea situada a los 48 kDa (**Figura 9B**), masa molecular que corresponde a lo reportado para otras FUs en literatura (de 45 a 55 kDa) (De Ioannes y col., 2004), convirtiéndose en una clara candidata a FU aislada. Por otra parte, se observa que hasta el pico 4 a 10% NaCl (100 mM) eluyen fracciones con masa molecular principalmente entre 48 y 63 kDa, mientras que los picos 5 y 6 a 15% NaCl (150 mM) y el pico 7 a 20% NaCl (200 mM) presentan fracciones de mayor masa molecular entre 100 y 130 kDa (**Figura 9**). Es decir, al aumentar la concentración de NaCl, eluyen fracciones con mayor masa molecular que pueden corresponder a varias FU unidas entre sí por la región *linker*. Solo eluyeron fracciones de la columna hasta una concentración de 200 mM, aun cuando se llegó hasta una concentración de 1 M NaCl.





**Figura 9. Purificación de FUs obtenidas desde CCHB.** A) Cromatograma representativo de columna de intercambio aniónico (Hitrap Q HP) para separación de FUs a partir de una muestra de CCHB digerida con elastasa. La línea roja representa el gradiente de NaCl donde 20% corresponde a 200 mM de NaCl. La línea negra representa la absorbancia a 280 nm, en donde los picos de elución de enumeran del 1-7. B) SDS-PAGE representativo de los picos obtenidos de la columna de intercambio aniónico. Pico 1 corresponde a “FU candidata”.

A partir de la digestión de 10 mg de CCHB, se obtuvo alrededor de 1 mg de la fracción candidata a FU aislada. Para corroborar que esta muestra corresponde efectivamente a una FU, se realizó un espectro de absorción para determinar si presenta los 3 picos de absorción reportados para hemocianina (De Ioannes y col., 2004). En la **Figura 10A** se observa que la muestra “FU candidata” presenta los 3 picos. Se muestra, además, un acercamiento al pico cercano a los ~550 nm, correspondiente a la coordinación de Cu con histidinas.



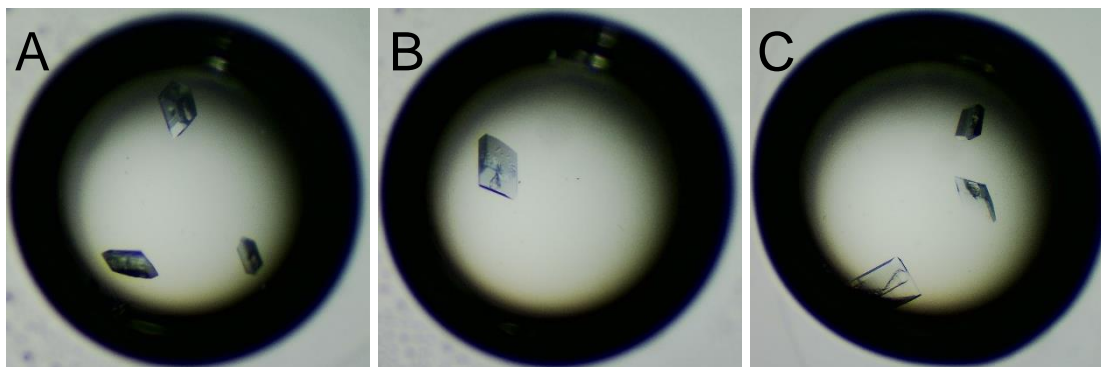
**Figura 10. Espectro de absorción de fracciones de CCH de origen silvestre y recombinante. A)** Espectro de absorción de fracciones silvestres de CCH. Se observan 3 picos característicos, a ~280 nm correspondiente a residuos aromáticos, a ~345 nm correspondiente a Cu-O<sub>2</sub>-Cu y a ~550-600 nm correspondiente al complejo de coordinación Cu-histidina. El recuadro indica un acercamiento a ambos espectros (CCH y FU candidata) con un máximo de absorbancia a 580 nm correspondiente al complejo de coordinación Cu-histidina. **B)** Espectro de absorción de FU expresadas de manera recombinante en *E. coli*, solo se observa el pico de absorción a 280 nm correspondiente a residuos aromáticos. El recuadro indica un acercamiento a los espectros mostrando la ausencia del pico a ~550-600 nm. El amortiguador (control) en cada uno de los gráficos corresponde a una solución con 20 mM Imidazol y 300 mM NaCl.

Más aún, la muestra presente en el gel de la “FU candidata” y la muestra de CCHB en solución fueron enviadas a un servicio de espectrometría de masas (*Protein Identification*, BGI Americas, MA, Estados Unidos) junto con la información de secuencia de los fragmentos parciales de CCH (CCHX y CCHY) y las secuencias de subunidades de la hemocianina de *Rapana venosa* (RVH1 y RVH2), dado que RVH presenta altos niveles de identidad de secuencia con los fragmentos de secuencia obtenidos para CCH (**Anexo 8.1**). Los resultados indicaron que la “FU candidata” posee un 27% de cobertura con la secuencia CCHX, identificándose 19 péptidos (**Anexo 8.1A**), además, mostró un 11% de cobertura con la secuencia de RVH1 identificándose 25 péptidos (**Anexo 8.1B**). Los resultados obtenidos para CCHB en solución mostraron una cobertura de 22% con la secuencia de RVH1 identificándose 69 péptidos comparado con un porcentaje de cobertura de 13% con la secuencia de RVH2. A partir de estos resultados, se tiene evidencia que la muestra purificada corresponde efectivamente a una FU de CCHB, y además se pudo comprobar que la subunidad CCHB de *C. concholepas* corresponde a la subunidad RVH1 de *Rapana venosa*.

#### **4.1.4 Cristalización de FU silvestre de CCH**

Con la muestra de FU candidata purificada a homogeneidad se realizaron ensayos de cristalización mediante difusión de vapor por gota sentada utilizando 2 *kits* comerciales (*kits* Index y SG1, Hampton Research), identificándose una condición de cristalización favorable en el *kit* Index (PEGMME 2.000 30% p/V y KBr 150 mM). Para optimizar la condición inicial se realizaron ensayos de cristalización por gota colgante, empleando como estrategia el uso de una matriz de optimización variando el porcentaje de agente

precipitante (PEGMME 2000) y el porcentaje de agente co-precipitante (KBr). Los mejores cristales se obtuvieron a una menor concentración de precipitante (entre 20 a 25% p/V) y a una mayor concentración de co-precipitante (entre 100 a 200 mM) (**Figura 11**).

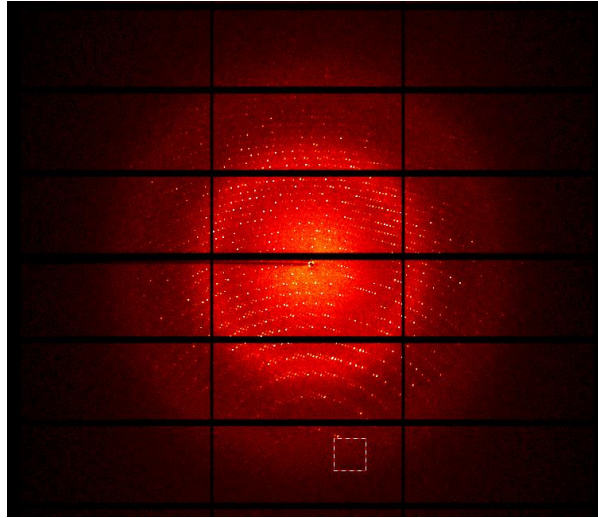


**Figura 11. Cristalización de FU de CCH.** (A-C) Ejemplos de cristales obtenidos en la optimización efectuada mediante el método de difusión de vapor por gota colgante a partir de la condición original: PEGMME 2000 30% p/V y KBr 150 mM.

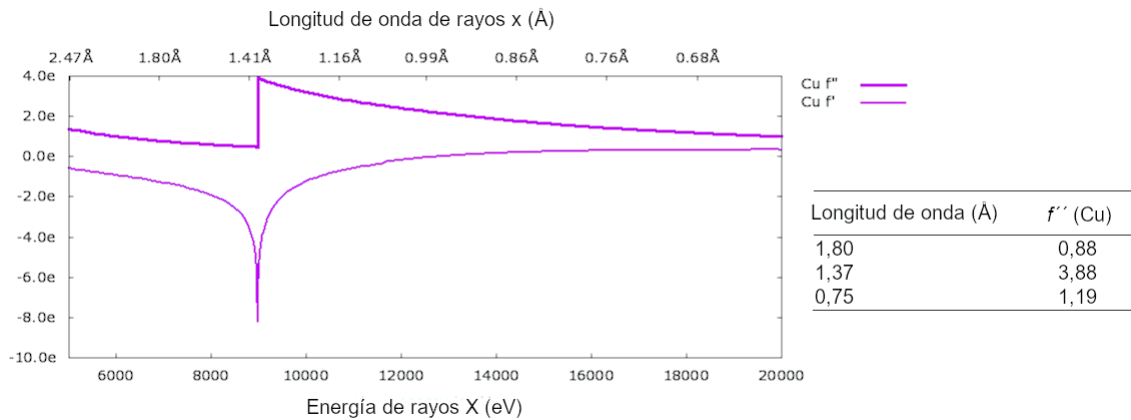
#### 4.1.5 Difracción de rayos X y análisis de datos

Los cristales obtenidos del proceso de optimización fueron difractados de manera remota (Santiago-Campinas) en la línea de luz MANACA del sincrotrón de 4<sup>ta</sup> generación SIRIUS (LNLS, Campinas, Brasil). Se colectaron 8 conjunto de datos nativos utilizando una longitud de onda de 0,977 Å, entre estos, se colectó un conjunto de datos con una resolución cercana a los 1,5 Å (**Figura 12**). Además, como estrategia para resolver el problema de las fases, se realizaron experimentos de SAD (*Single-wavelength Anomalous Diffraction*) utilizando una longitud de onda de 1,37 Å aprovechando la presencia de los átomos de cobre presentes en el centro activo de la FU cristalizada, en la cual el cobre presenta un factor de corrección imaginario de dispersión anómala ( $f''$ ) de 3,88 (**Figura 13**), lo cual permite estimar experimentalmente las fases de las reflexiones colectadas. Se

colectaron 4 conjuntos de datos anómalos utilizando los mismos cristales previamente difractados; entre estos, el cristal que había difractado a  $1,5 \text{ \AA}$  produjo un conjunto de datos anómalos a una resolución de  $1,7 \text{ \AA}$ .

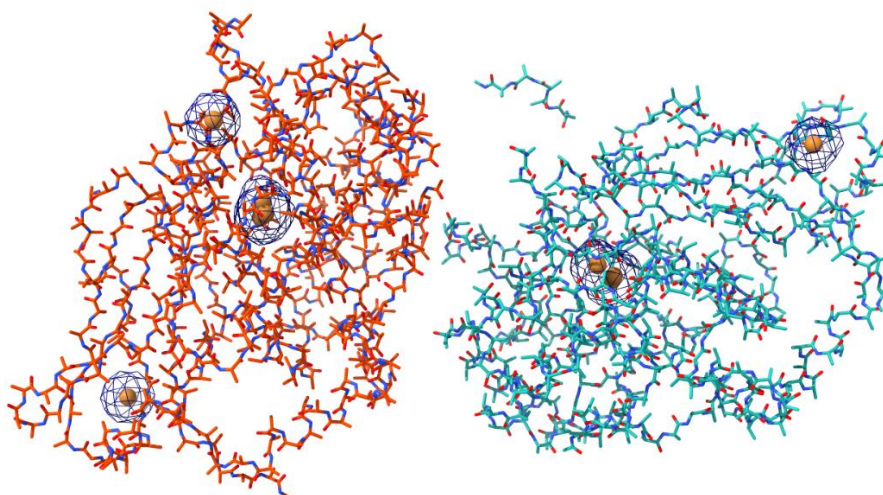


**Figura 12. Patrón de difracción representativo de datos nativos de FU.** Patrón de difracción obtenido de manera remota en la línea MANACA del sincrotrón de 4ta generación SIRIUS (LNLS, Campinas, Brasil) con una resolución cercana a los  $1,5 \text{ \AA}$ .



**Figura 13. Factores de corrección para dispersión anómala del Cobre.** Gráfico de valores teóricos de factores de dispersión anómalos real ( $f'$ ) e imaginario ( $f''$ ) para cobre respecto a la longitud de onda de rayos X. La diferencia entre estos factores se hace máxima a una longitud de onda de  $1,37 \text{ \AA}$ , la cual es adecuada para experimentos de SAD con átomos de cobre.

El conjunto de datos anómalos ( $\sim 1,7 \text{ \AA}$ ) para el mejor cristal fue integrado en el programa XDS, determinando que el grupo espacial era P 21 21 21. Utilizando la plataforma del programa SHELX (Sheldrick, 2008), se ubicaron las posiciones de los dispersores anómalos y se obtuvieron las fases a partir de la posición de dichos dispersores anómalos. Como no se conocía la secuencia de la FU candidata, se trazó un polímero de polialanina en la densidad electrónica obtenida. En la **Figura 14** se muestra la cadena principal de polialanina generada a partir de la densidad electrónica obtenida, donde se observa la presencia de dos moléculas de FU en la unidad asimétrica y el mapa anómalo de diferencia, mostrando la ubicación de átomos pesados que dieron origen a la dispersión anómala. Se observa que, además de las señales anómalas correspondientes a los pares de átomos de cobre ubicados en cada centro activo, se aprecian 3 señales extra, 2 ubicadas en la cadena A (una cercana al centro activo y otra ubicada en el dominio N-terminal) y la otra señal ubicada en el dominio N-terminal de la cadena B, las que fueron asignadas inicialmente a átomos de cobre (**Figura 14**). Sin embargo, estos átomos no presentaban una coordinación relacionada a un metal de transición como el cobre. A partir de esto, se propuso que estos átomos pudieran corresponder a iones bromuro ( $\text{Br}^-$ ), dado que la condición de cristalización contenía KBr 150 mM. Además, el bromo presenta una dispersión anómala significativa ( $f'' = 1,08$ ) a la longitud de onda en la que se realizó el experimento de SAD ( $1,37 \text{ \AA}$ ). Finalmente, se refinaron estos iones bromuro, encontrando que explicaban de mejor manera la densidad electrónica en estas posiciones (**Figura 16B**).

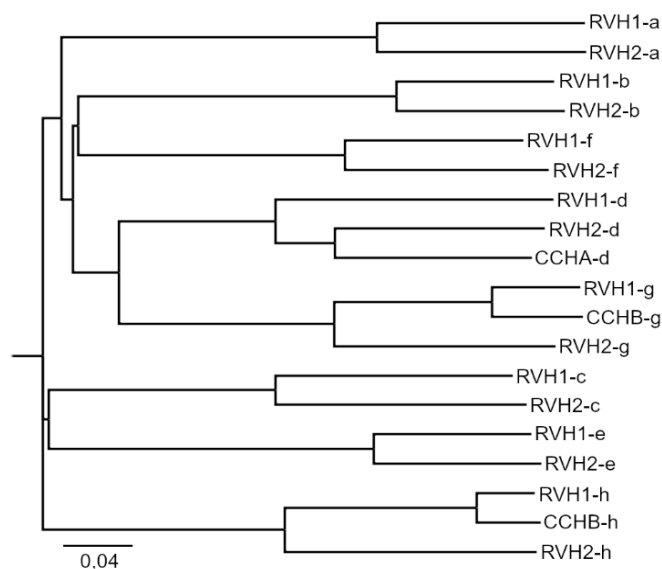


**Figura 14. Modelo inicial de FU obtenido mediante SAD.** Cadenas de polialanina obtenidas desde la densidad electrónica resultante de la estimación de fases con dispersores anómalos. La cadena A de polialanina se representa en color rojo, mientras que la cadena B se representa en color azul. El mapa de dispersión anómalo se muestra como densidad electrónica de color azul rodeando a los átomos dispersores anómalos, que inicialmente se asignaron como cobre.

Dado que no se contaba con la secuencia de la FU cristalizada y gracias a la alta resolución del conjunto de datos anómalos ( $\sim 1,7 \text{ \AA}$ ), se identificaron los residuos de la proteína asignando cadenas laterales a través del análisis visual de las densidades electrónicas utilizando COOT (Emsley & Cowtan, 2004). La secuencia inicial de la FU candidata de CCHB obtenida mediante el análisis de la densidad de las cadenas laterales, fue alineada en primer lugar con los péptidos obtenidos mediante espectrometría de masas (**Anexo 8.1**) para corregir errores en la colocación de residuos que presentaban ambigüedad en la asignación por densidad electrónica (Aspártico/Asparagina, Glutámico/Glutamina, Histidina/Fenilalanina, entre otros). Posteriormente, se alineó la secuencia corregida con las secuencias disponibles de CCH (CCHX y CCHY, **Figura 5, Pág. 22**), observándose

que la secuencia de la FU candidata cristalizada poseía una identidad de secuencia de 98% con el fragmento de FU parcial tipo -g del fragmento CCHX, diferenciándose solo en 5 residuos en los que existía ambigüedad en la asignación (como en los casos mencionados anteriormente), por lo cual la FU cristalizada se identificó como CCHB-g (**Figura 16A**). Después de corregir los 5 residuos, se logró completar la secuencia parcial de la FU-g del fragmento CCHX al agregar los 110 residuos faltantes del dominio N-terminal tirosinasa (**Figura 6, Pág. 32**). Por otra parte, por espectrometría de masas se determinó que la subunidad CCHB es equivalente a la subunidad 1 de la hemocianina de *Rapana venosa* (RVH1), por tanto, utilizando las secuencias de FU disponibles de las 2 subunidades de RVH (RVH1 y RVH2) se generó un dendrograma de distancia basado en el porcentaje de identidad de secuencia de las FUs (**Figura 15**), el cual se utilizó para comparar la secuencias de FUs de CCH disponibles. Se observa que cada tipo de FU de las subunidades de RVH forman su respectivo grupo, las FUs CCHB-g y CCHB-h forman grupos con las FUs RVH1-g y RVH1-h respectivamente (**Figura 15**), dada la alta identidad de secuencia entre las subunidades CCHB y RVH2. Por otra parte, se observa que la FU del fragmento CCHY correspondería a una FU-d (**Figura 6, Pág. 32**) y considerando que forma un grupo con la RVH2-d (**Figura 15**), se infiere que correspondería a la CCHA-d, considerando nuevamente la alta identidad de secuencia entre las subunidades CCHA y RVH1. Finalmente, podemos decir que contamos con 3 secuencias identificadas de FU de CCH: CCHA-d, CCHB-g y CCHB-h.





**Figura 15. Dendrograma de secuencias de FU de RVH y CCH.** Árbol de distancia basado en identidad de secuencias de FUs de las subunidades 1 y 2 de RVH (RVH1 y RVH2). Se ubican las secuencias de FU de subunidades CCHA (CCHA-d) y CCHB (CCHB-g y CCHB-h). La barra indica 0,04 cambios por sitio.

Después de identificada la secuencia del modelo obtenido por datos anómalos, el modelo fue utilizado para refinar una nueva estructura del conjunto de datos nativos para el mismo cristal que presentaba una mayor resolución (1,5 Å) (**Tabla 1**). El proceso posterior de refinamiento del nuevo modelo también se realizó en COOT obteniéndose mayores detalles, como el agregar las moléculas de oxígeno (O<sub>2</sub>) correspondientes a los *clusters* de Cu<sub>2</sub>O<sub>2</sub> del centro activo de las cadenas de FU, los que se refinaron adecuadamente a partir de la densidad electrónica obtenida. También se identificaron N-glicosilaciones, y se agregaron aguas cristalográficas y otras moléculas como polietilenglicol (PEG) (**Figura 16B**). El modelo final presentó un valor de *R-work/R-free* de 0,16/0,18, observándose 393 residuos en cada una de las cadenas de FU identificadas en la unidad asimétrica, y un

98,47% de los residuos dentro de regiones favorables en el gráfico de Ramachandran y ninguno en regiones no permitidas (**Tabla 1**).

**Tabla 1. Información cristalográfica y estadística del refinamiento de la estructura de CCHB-g.**

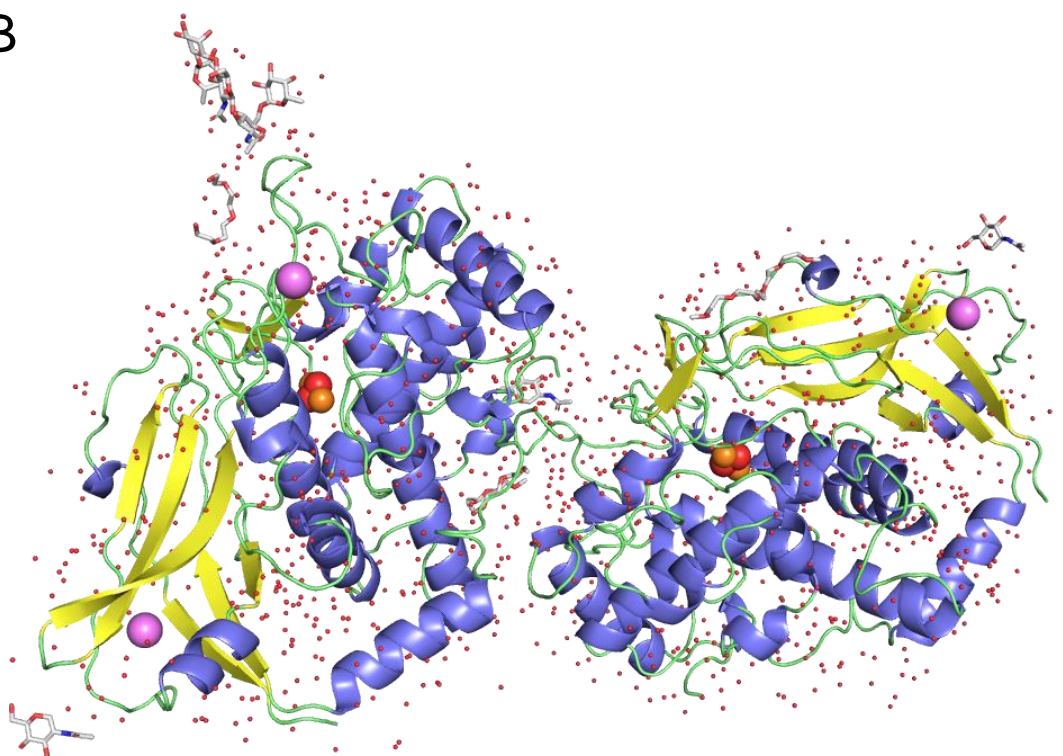
<b>Colecta de datos y procesamiento</b>	
Longitud de onda (Å)	0,977
Rango de resolución (Å)	47,41 – 1,50
Grupo espacial	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Celda unitaria (Å)	a = 75,68 b = 103,37 c =
(°)	119,05
	$\alpha = \beta = \gamma = 90$
Reflexiones totales/ reflexiones únicas	1.972.506 / 149.421
Multiplicidad	13,2 (13,3)
Compleitud	99,92 (99,99)
Promedio I/sigma	15,01 (2,43)
R-merge	0,1211 (1,39)
R-meas	0,12621 (1,443)
R-pim	0,03473 (0,3867)
CC1/2	0,999 (0,768)
CC*	1 (0,932)
<b>Refinamiento</b>	
R-work/R-free	0,1659 / 0,1870
RMS (enlaces) (Å)	0,010
RMS (ángulos) (°)	0,99
Residuos de proteína	786
Promedio de factor B (Å <sup>2</sup> )	21,73
Macromoléculas	19,30
Ligandos	38,47
Solvente	33,69
Número de átomos no hidrógeno	7.551
Macromoléculas	6.333
Ligandos	161
Agua	1.057
Gráfico de Ramachandran (%)	
Favorecidos	98,47
Permitidos	1,53
No permitidos	0,00
Rotámeros fuera (%)	1,20
Clash score	5,52

A

**CCHB-g**

```
ASAVRKDINTLTAAEVTDLRDALRKVQAGIGKETYNFIAGAHGYAGCSKGDMLVACCQHG  
MASFPGWHRVFTRQMEIALSWEGASVGLPYWDWTEAFTMLPTLVSEESLNPFFHHGHIPGKA  
ENITTTTRAPRQLFKDPEHGEESFFFRQALLAFEQRDFCDFEVQFEVLHNALHSWGIGTSPYG  
MSTLEYAA YDPTFFIHHSNVDRQFAIWQELQKHRGLDYNTANCHIQDLRKPPEFNRANNPV  
LVTRVHSRAIDAFNYDQYGYQYDHLHFHGLTVDKLDEKLEKRKEQDRVFLNFMLRGIKMSA  
DVVFDLCNAQGT CNFAGTFAILGGPLEMPWNFDRVFKYDVTKIFQQMRLRPDSNYTIPIRIRA  
VNGMQLDPNLLEPPSVTFVPGK
```

B

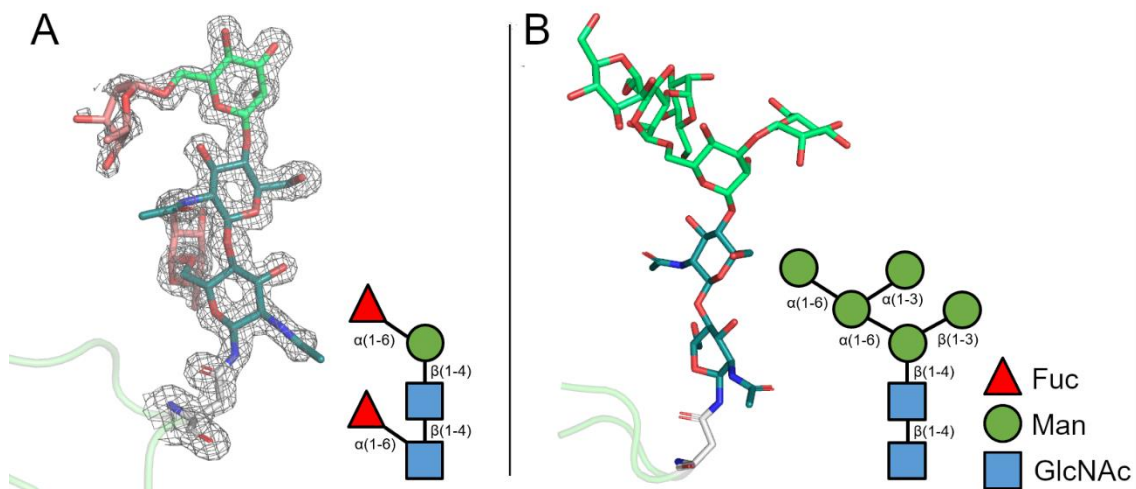


**Figura 16. Secuencia y estructura cristalográfica de FU-g de CCHB.** A) Secuencia de FU-g de CCHB obtenida a través de identificación de la densidad electrónica de cadenas laterales y alineamiento de secuencia con el fragmento CCHX. B) Modelo cristalográfico obtenido después del refinamiento y coloreado de acuerdo con la estructura secundaria:  $\alpha$ -hélices en azul, hebras  $\beta$  en amarillo y *loops* en verde. Los átomos de Cu, O y Br, se representan como esferas de colores naranja, rojo y rosado respectivamente. Las estructuras de monómeros correspondientes a N-glicosilaciones en conjunto con polímeros de PEG se representan en color blanco. Las aguas cristalográficas se representan como esferas de color rojo.

#### 4.1.6 Análisis de la estructura de FU silvestre de CCH

La FU-g de CCHB presenta el clásico plegamiento descrito para una FU de molusco (FU-a hasta FU-g), observándose 2 dominios, el N-terminal que presenta solo  $\alpha$ -hélices y el dominio C-terminal que presenta un  $\beta$ -sándwich de 6 hebras (**Figura 16B**).

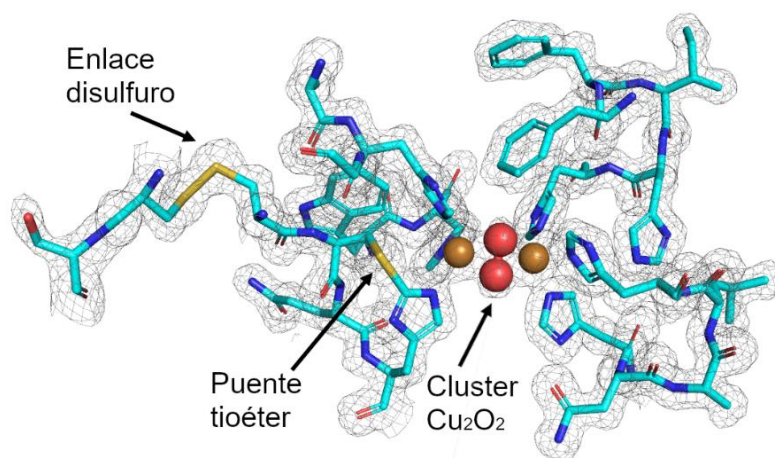
Al analizar la estructura, se observa la presencia de N-glicosilaciones en 4 puntos, dos en cada cadena, correspondiente a los residuos Asn123 y Asn361. En el caso de la cadena A, en el residuo Asn123 se observa un árbol de glicosilación correspondiente a dos monómeros de N-acetilglucosamina (GlcNAc) y un monómero de manosa (Man) unidos mediante enlace ( $\beta$ 1-4), y dos monómeros de fucosa (Fuc) unidos mediante enlace ( $\alpha$ 1-6), es decir Fuc( $\alpha$ 1-6)Man( $\beta$ 1-4)GlcNAc( $\beta$ 1-4)[Fuc( $\alpha$ 1-6)]GlcNAc (**Figura 17A**). En cuanto a los otros residuos, solo hay densidad electrónica disponible para asignar monómeros de GlcNAc. Al comparar el árbol de N-glicosilaciones presente en el residuo Asn123 (ubicado en una vuelta  $\beta$ ) de CCHB-g (**Figura 17A**) con un árbol de glicosilaciones presente en el residuo Asn53 (también ubicado en una vuelta  $\beta$ ) de ODH-g de pulpo (2,3 Å de resolución) (**Figura 17B**) se observa en esta última que, junto al residuo de asparagina, se encuentran dos monómeros de GlcNAc unidos a un monómero de manosa en enlace  $\beta$ (1-4), como se espera en una secuencia consenso de N-glicosilación (Emsley & Crispin, 2018). Sin embargo, en el árbol de N-glicosilación de CCHB-g se observan dos monómeros de fucosa enlazados  $\alpha$ (1-6), mientras que en el árbol de N-glicosilación de ODH-g se observan solamente monómeros de manosa enlazados de forma  $\alpha$ (1-3) y  $\alpha$ (1-6) (**Figura 17**).



**Figura 17. N-glicosilaciones presentes en estructuras cristalográficas de FUs de hemocianina de molusco. A)** Árbol de N-glicosilaciones correspondiente a la estructura CCHB-g. Se muestra un mapa de densidad electrónica de doble diferencia ( $2mF_0-DF_c$ ) contorneado a  $1 \sigma$ . **B)** Árbol de N-glicosilaciones correspondiente a la estructura ODH-g de pulpo (PDB 1JS8). Los residuos de Asparagina se muestran en color blanco, mientras que los monómeros de GlcNAc, Manosa y Fucosa se muestran en colores azul, verde y rojo respectivamente. Además, se muestra un esquema de figuras con los monómeros de azúcares, en el cual se indican los tipos de enlaces.

Respecto al centro de unión a cobre, la estructura de CCHB-g presenta densidad electrónica suficiente para asignar el clúster  $Cu_2O_2$  en conjunto con los residuos cercanos a  $5 \text{ \AA}$ , en los cuales se destacan los 6 residuos de histidina que coordinan los átomos de cobre. Se observa un puente tioéter entre los residuos Cys57 e His60, el cual se encuentra conservado en las estructuras de FU disponibles: se observa en la estructura de ODH-g de pulpo (Cuff y col., 1998) dada su resolución ( $2,3 \text{ \AA}$ ), mientras que en las estructuras de FU de RVH2-e (Perbandt y col., 2003) y KLH-h (Jaenicke y col., 2011) ( $3,3$  y  $4 \text{ \AA}$ , respectivamente) solo se aprecia el grupo tiol de la cadena lateral de cisteína orientado hacia la histidina, dado que las estructuras no poseen la resolución necesaria para sustentar el enlace, aunque se puede inferir que está presente. El puente tioéter también se observa

en todas las FU de la estructura cristalográfica de la hemocianina de calamar (TPH) determinada a una resolución de 3,0 Å (Gai y col., 2015). Por otra parte, en la estructura de CCHB-g, se presenta un enlace disulfuro entre los residuos Cys47 y Cys56, que también se encuentra altamente conservado y que se puede observar en todas las estructuras mencionadas anteriormente (**Figura 18**).



**Figura 18. Centro de unión a cobre de la estructura CCHB-g.** Los átomos correspondientes al clúster  $\text{Cu}_2\text{O}_2$  se muestran como esferas de color naranja para el cobre y rojo para el oxígeno. Los residuos dentro de un radio de 5 Å se muestran de color cian, destacando un enlace disulfuro y un puente tioéter. Se muestra un mapa de densidad electrónica de doble diferencia ( $2mF_0-DF_c$ ) contorneado a  $1 \sigma$ .

Respecto al centro de unión a cobre, se compararon las distancias entre las histidinas y los átomos del clúster  $\text{Cu}_2\text{O}_2$ , (**Tabla 2**) y los ángulos de enlace entre átomos del clúster (**Tabla 3**), entre las estructuras CCHB-g y ODH-g. Esta última es la estructura de mayor resolución de una FU reportada hasta la fecha (2,3 Å) y la única que posee densidad electrónica para asignar los átomos de oxígeno dentro del clúster (Cuff y col.,1998). La medición de distancias y ángulos de enlaces observadas en CCHB-g muestran diferencias importantes con las mediciones del centro de unión a cobre previamente determinadas con

la estructura de ODH-g. En primer lugar, hay una diferencia importante entre las distancias de un par de histidinas con átomos de Cu (H41 y H170, en numeración de CCHB-g); en CCHB-g presentan la misma distancia en ambas cadenas (2,07 Å), mientras que en las cadenas de ODH-g, H41 presenta un importante aumento de 0,2-0,3Å y en H169 un aumento de 0,16-0,23Å. La distancia entre los átomos de Cu es prácticamente la misma en las dos cadenas de ambas estructuras (3,5 Å) y no se registran diferencias importantes en las distancias entre átomos de Cu y oxígeno. No obstante, en la distancia entre los átomos de oxígeno hay una importante diferencia dado que en las cadenas de CCHB-g esta distancia presenta valores de 1,59 y 1,68 Å, mientras que en las cadenas de ODH-g presenta valores de 1,36 y 1,38 Å, es decir 0,2-0,3 Å menos (**Tabla 2**).

Por otra parte, en el caso de los ángulos también hay importantes diferencias. En los ángulos con el vértice centrado en los átomos de oxígeno, se observa que en CCHB-g los ángulos centrados en O1 (130,63° y 127,21°) son de 1° y hasta 9° menores que los ángulos centrados en O1 de ODH-g (136,47° y 131,19°), mientras que los ángulos centrados en O2 de CCHB-g miden prácticamente lo mismo (126°) y son de 8° a 11° menores que los ángulos de ODH-g (137,25° y 134,77). Respecto, de los ángulos con vértice centrado en los átomos de cobre, en el caso de CCHB-g los ángulos de cada cadena son similares para cada átomo de cobre (entre 48° y 51°) siendo de 5° hasta 10° menores que los ángulos de ODH-g (entre 41° y 43°). Finalmente, el ángulo diedro Cu1-O1-O2-Cu2 que da cuenta del grado de torsión entre cada enlace Cu-O, presentó un valor similar entre las cadenas de CCHB (161° y 163,57°) mientras que en el caso de ODH-g las dos cadenas muestran valores bastante distintos (173,06° y 158,14°) (**Tabla 3**).

**Tabla 2. Comparación de distancias de enlaces del centro de unión a cobre de las estructuras CCHB-g y ODH-g.**

CCHB-g	CCHB-g (Cadena A) Distancia (Å)	CCHB-g (Cadena B) Distancia (Å)	ODH-g	ODH-g (Cadena A) Distancia (Å)	ODH-g (Cadena B) Distancia (Å)
d(H41(N)-Cu)	2,07	2,07	d(H41(N)-Cu)	2,27	2,37
d(H59(N)-Cu)	2,14	2,15	d(H60(N)-Cu)	2,21	2,26
d(H68(N)-Cu)	2,24	2,31	d(H69(N)-Cu)	2,34	2,33
d(H170(N)-Cu)	2,07	2,07	d(H169(N)-Cu)	2,23	2,30
d(H174(N)-Cu)	2,18	2,21	d(H173(N)-Cu)	2,15	2,17
d(H201(N)-Cu)	2,13	2,14	d(H200(N)-Cu)	2,12	2,23
d(Cu1-Cu2)	3,50	3,50	d(Cu1-Cu2)	3,54	3,50
d(Cu1-O1)	1,93	1,95	d(Cu1-O1)	1,93	1,92
d(Cu1-O2)	1,97	1,97	d(Cu1-O2)	1,98	1,91
d(Cu2-O1)	1,93	1,95	d(Cu2-O1)	1,88	1,92
d(Cu2-O2)	1,97	1,96	d(Cu2-O2)	1,82	1,88
d(O1-O2)	1,59	1,68	d(O1-O2)	1,38	1,36

**Tabla 3. Comparación de ángulos de enlaces del centro de unión a cobre de las estructuras CCHB-g y ODH-g.**

CCHB-g	CCHB-g (Cadena A) Ángulo (°)	CCHB-g (Cadena B) Ángulo (°)	ODH-g	ODH-g (Cadena A) Ángulo (°)	ODH-g (Cadena B) Ángulo (°)
$\alpha(\text{Cu1-O1-Cu2})$	130,63	127,21	$\alpha(\text{Cu1-O1-Cu2})$	136,47	131,19
$\alpha(\text{Cu1-O2-Cu2})$	126,12	126,06	$\alpha(\text{Cu1-O2-Cu2})$	137,53	134,77
$\alpha(\text{O1-Cu1-O2})$	48,22	50,87	$\alpha(\text{O1-Cu1-O2})$	41,21	41,71
$\alpha(\text{O1-Cu2-O2})$	48,22	51,03	$\alpha(\text{O1-Cu2-O2})$	43,71	42,04
$\alpha(\text{Cu1-O1-O2-Cu2})$	161,00	163,57	$\alpha(\text{Cu1-O1-O2-Cu2})$	173,06	158,14

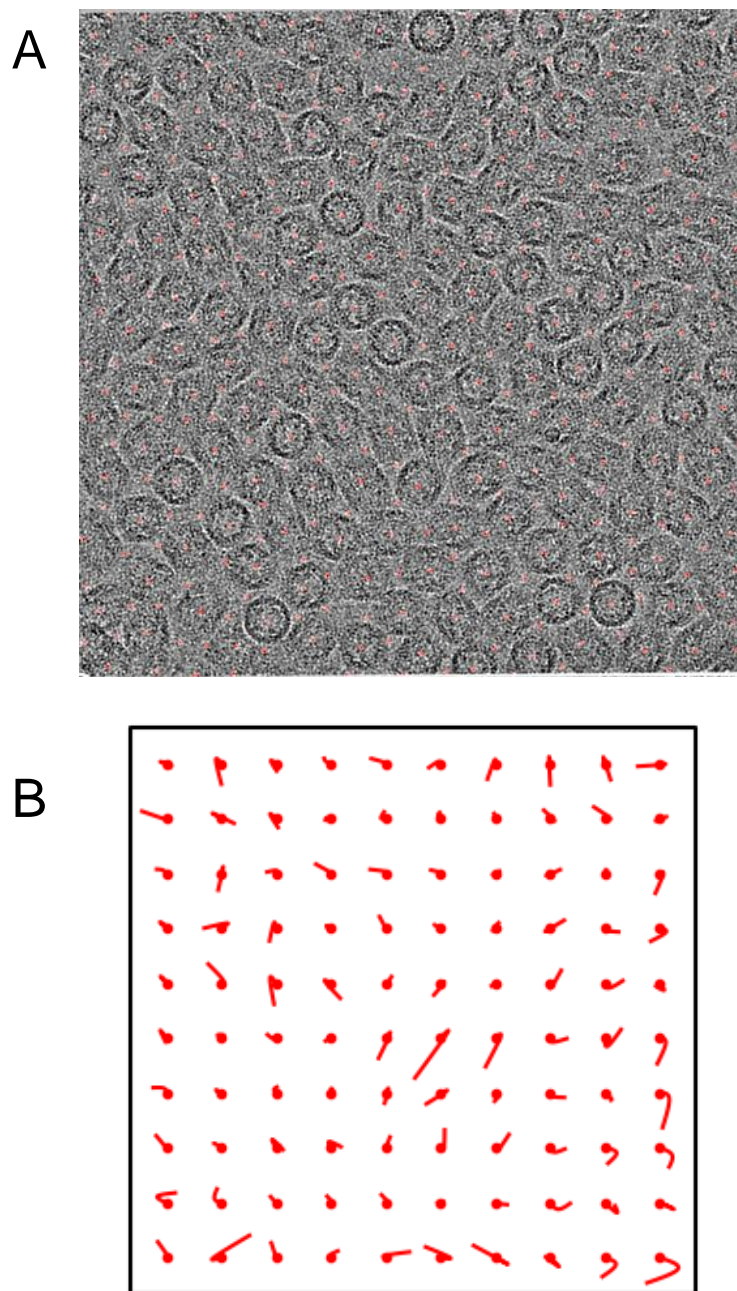


## 4.2 OBJETIVO ESPECÍFICO 2

### 4.2.1 Cryo-EM de CCH y procesamiento de datos

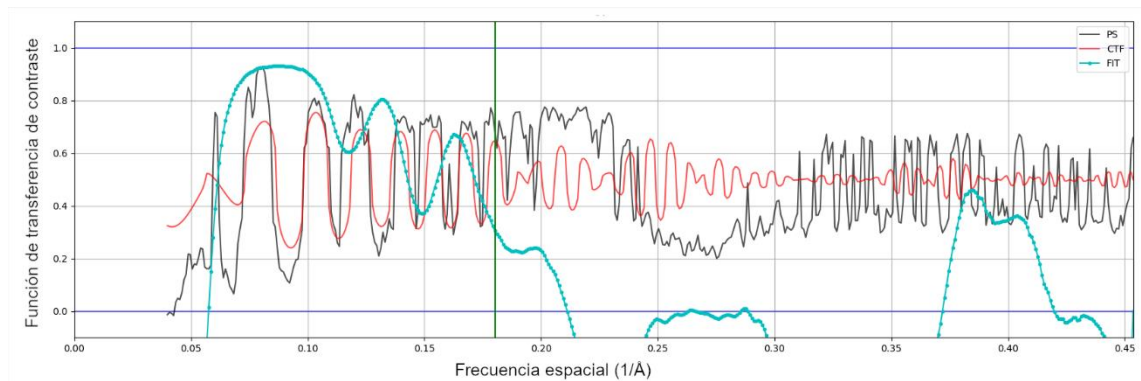
El experimento de Cryo-EM fue llevado a cabo por el personal de LNNano, en Campinas, Brasil, utilizando un microscopio electrónico Titan Krios equipado con una cámara K3 en modo conteo de electrones a super-resolución. La muestra de CCH fue preparada a una concentración de 5 mg/mL y como estrategia se colectaron micrografías de manera continua con una dosis total de electrones de  $30 \text{ e}/\text{Å}^2$  dividida sobre 30 cuadros, con un tamaño de píxel físico de  $1,1 \text{ Å}$  ( $0,55 \text{ Å}$  en modo super-resolución) y un desenfoque de  $-0,5$ ,  $-1,0$  y  $-1,5 \text{ μm}$ .

Durante el experimento se colectaron 2.060 películas (3,6 TB de información), las cuales fueron cargadas en el software cryoSPARC v3.2.0 para realizar el primer paso que corresponde a la corrección de movimiento (**Figura 19A**). En un gráfico de movimiento de parche representativo de una de las micrografías se definen sectores (parches) de la micrografía y se representa el movimiento de las partículas de ese sector producto de la energía depositada por el haz de electrones (deformación anisotrópica) (**Figura 19B**).



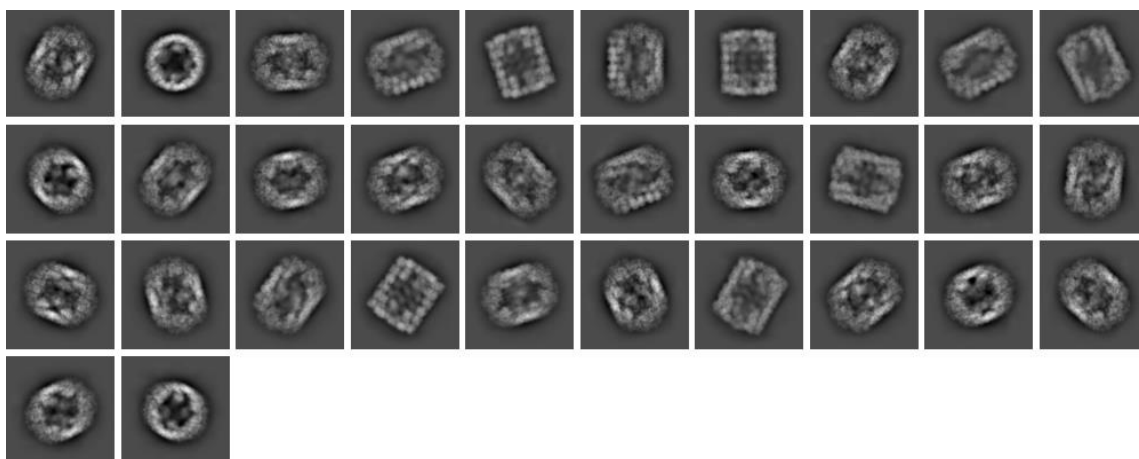
**Figura 19. Resultado de colecta de datos de cryo-EM de CCH y corrección de movimiento.** **A)** Micrografía representativa de cryo-EM de CCH colectada en un microscopio electrónico Titan Krios, después del proceso de corrección de movimiento. Los puntos rojos indican los sitios de corrección de movimiento. **B)** Gráfico de movimiento de parche representativo de una de las películas colectadas de CCH. El rectángulo representa el área de las micrografías, mientras que los puntos rojos representan los parches en los que se divide la película, mientras que las líneas rojas representan el movimiento anisotrópico promedio de las partículas ubicadas dentro de cada parche de la película.

El segundo paso en el procesamiento de datos corresponde a la estimación de la función de transferencia de contraste (CTF, del inglés *Contrast Transfer Function*), que corresponde a una función que modela el grado de desenfoco en el microscopio electrónico para producir un nivel de contraste de fase adecuado para poder diferenciar las partículas del fondo. Es fundamental estimar esta función para cada una de las películas y así poder filtrar películas que presenten un CTF ajustado a resoluciones muy bajas que puedan afectar la resolución de la reconstrucción 3D. En un gráfico representativo de estimación de CTF en una película, se observa el espectro de potencia producido por las fases de los electrones al momento de la colecta de micrografías (color negro), el CTF calculado para esas fases (color rojo) y una función de correlación (color cyan) la cual utiliza un valor umbral de 0,3 como parámetro de ajuste entre ambos espectros, que en este caso corresponde a una resolución de 5,549 Å ( $1/\text{Å} = 0,18$ ) (**Figura 20**). Con los datos trabajados, se opta por filtrar las películas que presenten una resolución de ajuste de CTF mayores a 8 Å, quedando con un total de 1.793 películas.



**Figura 20. Gráfico de ajuste de CTF.** Gráfico de frecuencia espacial en donde la línea negra corresponde al espectro de poder observado en las micrografías, la línea roja corresponde al CTF calculado respecto al espectro de poder, la línea cyan corresponde al ajuste de correlación entre ambas funciones y la línea verde vertical indica la frecuencia en la que el ajuste de correlación cae por debajo de un valor umbral de 0,3.

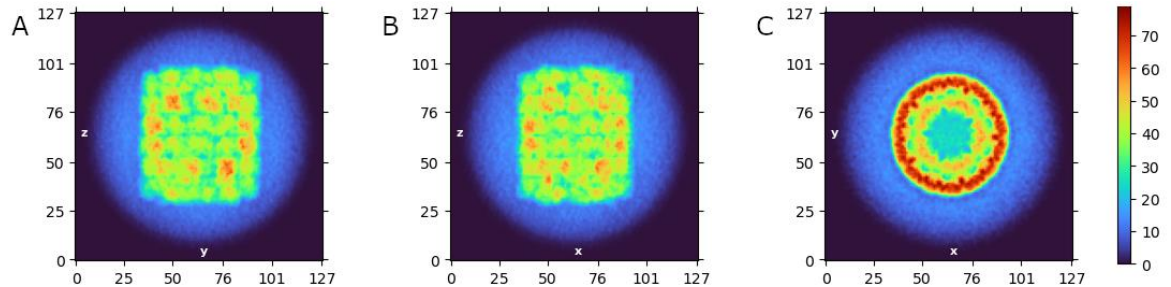
De estas 1.793 películas, se seleccionaron automáticamente partículas de CCH a partir de 100 micrografías iniciales para ser usadas como moldes en tres rondas de alineación y clasificación 2D supervisada, para seleccionar los mejores grupos de partículas, extraerlas y generar templados 2D. Inicialmente, se recolectaron 558.149 imágenes de partículas seleccionadas automáticamente, pero posteriormente luego de las 3 rondas de análisis, se seleccionaron 199.045 imágenes de partículas agrupadas en 32 clases distintas, las cuales fueron seleccionadas a través de inspección visual considerando resolución límite de 8 Å (Figura 21).



**Figura 21. Clases de partículas de CCH seleccionadas.** 32 grupos de imágenes 2D de CCH seleccionadas luego de 3 rondas de alineación y clasificación 2D supervisada.

Con los grupos de imágenes de partículas seleccionadas, se utilizó una reconstrucción 3D *ab initio* de simetría D5 con el objetivo de generar 3 reconstrucciones 3D, de las cuales se seleccionó la mejor reconstrucción 3D mediante inspección visual (Figura 22). Para esto, el software cryoSPARC utiliza dos algoritmos, el descenso de gradiente estocástico (SGD, del inglés *Stochastic Gradient Descent*) y el de optimización de máxima probabilidad de ramificación y límite. El primer algoritmo, SGD permite realizar el proceso de

reconstrucción *ab initio* 3D sin necesidad de poseer una estructura inicial homóloga a la que se quiere reconstruir, dado que utiliza diferentes subconjuntos de imágenes cryo-EM elegidos al azar. Esto produce iteraciones en cada paso, explorando un mayor espacio de soluciones posibles de mapas 3D antes de llegar a la solución más probable, teniendo en cuenta todas las imágenes extraídas de las fotografías. El algoritmo de optimización de máxima probabilidad de ramificación y límite permite acelerar la búsqueda de alineación óptima del mapa 3D, dado que calcula un límite inferior de soluciones posibles de alineaciones, descartando de plano las soluciones inferiores a ese límite, optimizando el gasto computacional del proceso. Ambos algoritmos permiten reducir en gran medida (a horas o minutos) los pasos principales en la determinación de la densidad cryo-EM (Punjani y col., 2017).

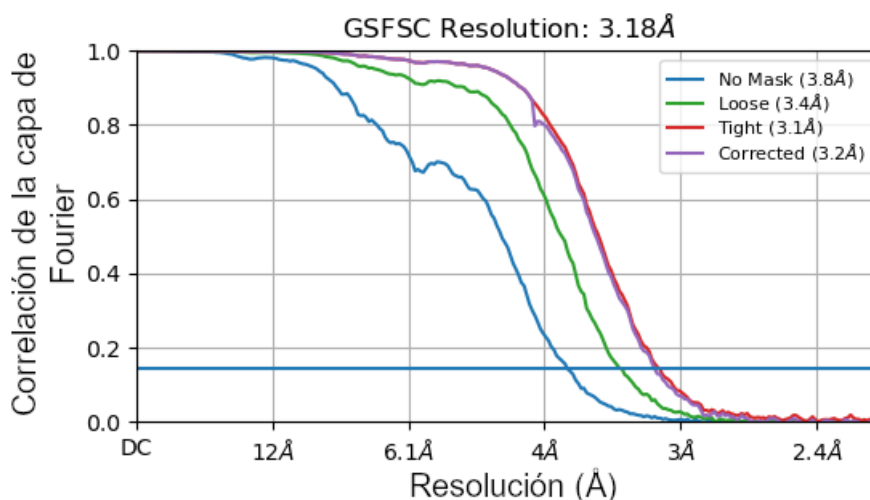


**Figura 22. Modelo de reconstrucción *ab initio* de CCH.** Vistas ortogonales (A) Frontal, (B) Lateral y (C) Superior del mejor modelo de densidad de CCH reconstruida *ab initio*. Se muestra un mapa de calor referente a zonas de densidad del modelo reconstruido. El mapa de color de la derecha muestra unidades arbitrarias de densidad estimadas a partir de parámetros estructurales de obtención de las imágenes, en donde rojo representa zonas de mayor densidad y azul zonas de menor densidad.

Las imágenes de partículas de la mejor reconstrucción 3D generada (130.952 imágenes) fueron seleccionadas para efectuar rondas de refinamiento 2D en donde se re-seleccionan partículas con rondas de alineación y clasificación 2D supervisada utilizando las imágenes

de partículas seleccionadas del mejor modelo 3D. Posteriormente, se realizan rondas de refinamiento 3D de tipo heterogéneo y homogéneo. En el refinamiento heterogéneo se clasifican las partículas de acuerdo con la estructura inicial obtenida en la reconstrucción *ab initio*, facilitando la búsqueda de pequeñas diferencias entre estructuras que no son obvias, sobre todo a baja resolución. Mientras que en el refinamiento homogéneo se usan las partículas seleccionadas del mejor modelo 3D, corrigiendo aberraciones ópticas y refinando el desenfoque de las partículas mientras se genera el modelo (Punjani y col., 2017). La calidad del mapa de densidad final de cryo-EM fue evaluada utilizando el estándar dorado de correlación de la capa de Fourier (GSFSC, del inglés *Gold-Standard Fourier Shell Correlation*) entre dos mitades del mapa (**Figura 23**). La resolución global fue estimada a 3,18 Å utilizando la curva corregida (*corrected*) (**Figura 23**) basándose en el criterio de corte 0,143, el cual indica que la frecuencia en la cual la GSFSC cae por debajo de un valor de ajuste de 0,143 es un buen indicador de la verdadera resolución del mapa. Se ha propuesto este criterio para hacer la estimación de medición de la resolución de datos de cryo-EM comparable a las mediciones usadas en cristalografía de rayos X (Rosenthal & Henderson, 2003).

Los valores de parámetros respecto al experimento de colecta y el posterior procesamiento de datos de Cryo-EM de CCH se presentan en la **Tabla 4**. De la misma forma, en el **Anexo 8.2** se presenta el protocolo detallado con los procesos y criterios utilizados en el tratamiento de datos colectados de cryo-EM de CCH utilizando el software cryoSPARC.



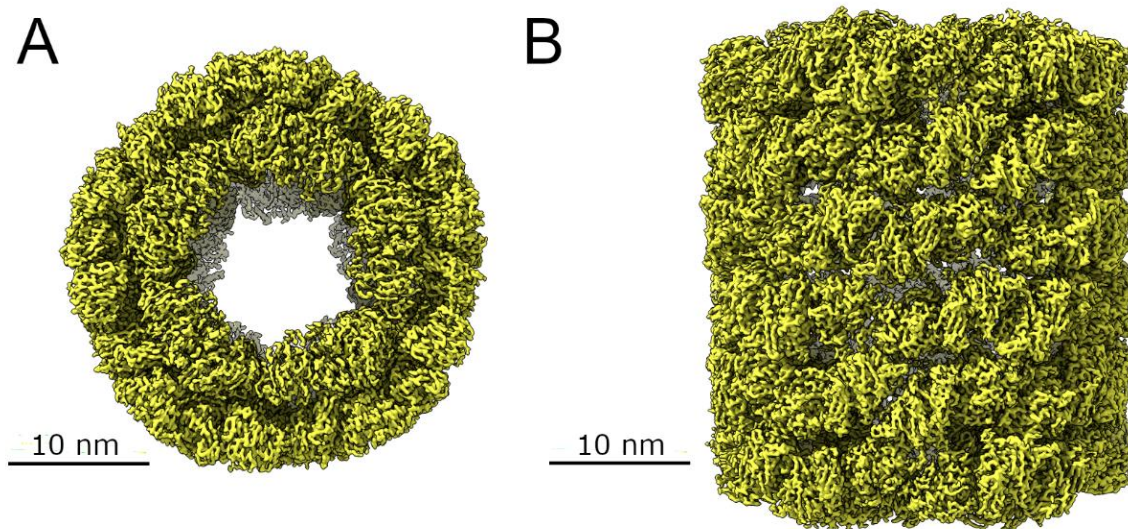
**Figura 23. Gráfica de estándar dorado de correlación de la capa de Fourier del mapa final de cryo-EM de CCH.** Se muestra el ajuste de las correlaciones en función de la resolución. La curva azul (*No mask*) indica la correlación “cruda” entre dos mitades independientes del mapa, sin aplicar máscara, incluyendo la estructura y el solvente. La curva verde (*Loose*) indica la correlación después de aplicar una máscara tenue sobre el solvente de ambas mitades de mapa. La curva roja (*Tight*) utiliza la misma máscara aplicada en la curva verde (*Loose*) excepto que reduce las distancias de dilatación en angstrom (Å) para calcular la máscara de solvente diferenciando la estructura de solvente (de 25 a 6 Å). La curva morada (*Corrected*) indica la misma máscara aplicada en la curva roja (*Tight*) pero ajustada por corrección por sustitución de ruido. Esta última curva se utilizó para calcular la resolución final del mapa, utilizando el criterio de corte de 0,145.

**Tabla 4. Información de colecta y procesamiento de datos de cryo-EM de CCH.**

Microscopio	Titan Krios
Cámara	K3 (Detección directa)
Voltaje (kV)	300
Exposición de electrones ( $e^-/\text{Å}^2$ )	30
Rango de desenfoque ( $\mu\text{m}$ )	-0,5 a -1,5
Aberración esférica (mm)	0,01
Tamaño de píxel (Å)	1,1 (0,55)
Simetría impuesta	D5
Imágenes de partículas iniciales (N°)	558.149
Imágenes de partículas finales (N°)	130.952
Resolución del mapa (Å)	3,18
Limite FSC	0,143



Finalmente, el modelo de mapa de cryo-EM de CCH fue extraído del software cryoSPARC y cargado en el software ChimeraX en donde se analizó y donde se pueden apreciar dos diferentes vistas, superior y frontal (**Figura 24**).



**Figura 24. Reconstrucción 3D de CCH obtenida a 3,2 Å.** Vistas A) Superior, del collar interno y B) Frontal, de la pared del mapa 3D de densidad del heterodidecámero de CCH obtenido mediante cryo-EM, coloreado en amarillo.

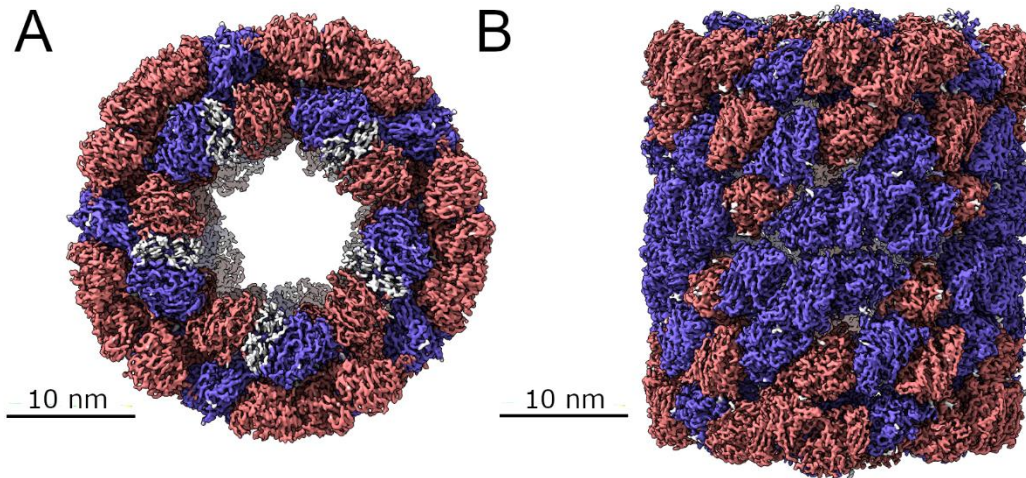


#### 4.2.2 Análisis del mapa de Cryo-EM de CCH

Respecto al mapa de densidad de cryo-EM obtenido para CCH, se observa que este presenta la estructura didecámerica esperada para una hemocianina de molusco de tipo I. El mapa presenta un diámetro externo de aprox. 313,9 Å, mientras que el diámetro del collar interno es de aprox. 220,6 Å. Por otra parte, cada decámero por separado mide alrededor de 158,5 Å de alto, indicando que el didecámero completo mide alrededor de 317 Å. Las medidas obtenidas son comparables con las medidas observadas en el modelo de cryo-EM obtenido para la hemocianina de *Haliotis diversicolor* (HDH), las cuales son 315,1 Å aprox. de diámetro externo, 223,6 Å aprox. de diámetro del collar interno, mientras que el decámero por separado mide alrededor de 158,0 Å aprox. de alto, lo que indica una altura total del didecámero de 316 Å. Esta similitud en las medidas del mapa de cryo-EM de CCH y el modelo de cryo-EM de HDH se hace más evidente al realizar un *docking* rígido del modelo de HDH dentro del mapa de densidad de cryo-EM de CCH, coloreando este último según la proximidad de los átomos del modelo de las subunidades asimétricas diméricas (subunidades de extremo e internas) descrito en el homodidecámero de HDH (**Figura 25A y B**).

Interesantemente, la región de la pared del didecámero encaja bastante bien entre el modelo del homodidecámero de HDH y el mapa de densidad de CCH, evidenciado por el coloreado continuo en el mapa de densidad de CCH (ambas subunidades asimétricas del dímero) (**Figura 25B**). Sin embargo, al observar la región del collar interno, se observan regiones no coloreadas en la ubicación de las FU-h. Esto se debe a que el modelo de HDH no presenta el característico dominio extra *cupredoxin-like* ubicado en el extremo C-

terminal de las FU-h de molusco, lo que indicaría que el mapa de CCH si presenta densidad asociada a la presencia de este dominio (**Figura 25A**).

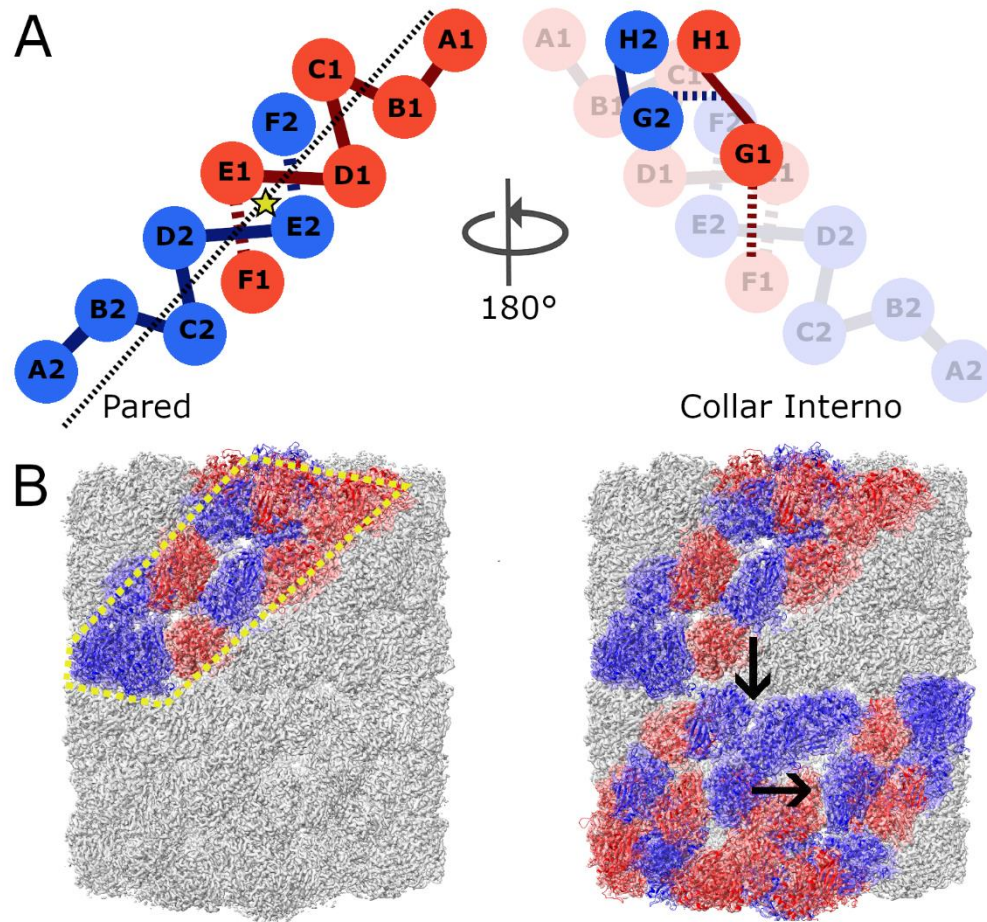


**Figura 25. Comparación del modelo de hemocianina homodidecamérica de *Haliotis diversicolor* (HDH) con el mapa de cryo-EM de CCH. Vistas A) Superior, del collar interno y B) Frontal, de la pared del modelo de *docking* rígido de HDH dentro del mapa de densidad de cryo-EM de CCH. El volumen del mapa de CCH se coloreó mediante proximidad de los átomos del modelo de HDH. Se coloreó de rosa las subunidades asimétricas de los extremos y de morado las subunidades asimétricas del interior de HDH, mientras que las zonas grises indican que no hay proximidad de átomos del modelo de HDH con el mapa de cryo-EM de CCH. Se utilizó el criterio de proximidad calculado por ChimeraX para colorear el mapa.**

Uno de los principales problemas para la generación de modelos de organización de subunidades en el heterodidecámero es la falta de información de hemocianinas homólogas, debido a que todas las estructuras reportadas corresponden a homooligómeros, por lo que la ubicación espacial de 2 subunidades diferentes en un mismo oligómero es totalmente desconocida a la fecha. Sumado a esto, para CCH se desconoce la secuencia de las subunidades CCHA y CCHB, por lo que se utilizaron las secuencias de FU disponibles de las subunidades RVH1 (correspondiente a CCHB) y RVH2 (correspondiente a CCHA) de la hemocianina de *Rapana venosa* (RVH) la cual es bastante

cercana filogenéticamente. Se generaron modelos por homología de las FU-a, -b, -c, -e y -f utilizando las secuencias de FU de ambas subunidades de RVH (RVH1 y RVH2) y utilizando la estructura de CCHB-g como templado. En el caso del modelo por homología de la FU-d se utilizó la secuencia de CCHA-d (CCHY1, **Figura 5A, Pág. 22**), para la FU-g, se utilizó la estructura cristalográfica de CCHB-g y para la FU-h se usó la secuencia disponible de CCHB-h (**Figura 5A, Pág. 22**) utilizando como templado la estructura de KLH-h (PDB: 3QJO). Para estas 3 últimas FUs se utilizó el mismo modelo para cada una de las dos subunidades.

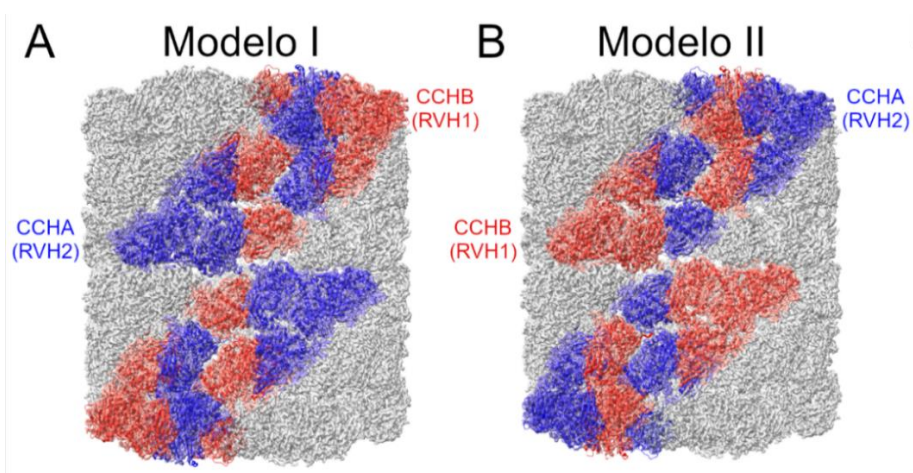
Para la organización de las FUs dentro de las subunidades, se utilizó el modelo de la hemocianina de *Haliothis diversicolor* (HDH), el cual consiste en un homodidecámero determinado a una resolución de 4,5 Å, donde el dímero de subunidades se ha descrito como una unidad asimétrica mínima dentro del cilindro (Zhang y col., 2013). Al analizar la distribución de FUs dentro del dímero de HDH, se puede observar que existe un eje de simetría en la región de la pared (**Figura 26A**). Esta simetría se pierde en la región del collar interno, debido a la ubicación asimétrica de las FUs de tipo -g respecto a las de tipo -h de cada dímero (**Figura 26A**). Por consiguiente, el dímero de subunidades se considera como una unidad asimétrica mínima, ya que puede utilizarse para recrear el cilindro completo mediante operaciones de traslación y rotación (**Figura 26B**).



**Figura 26. Esquema de organización del dímero de subunidades de HDH. A)** En la región de la pared se observa un eje de simetría formado por las dos subunidades, coloreadas de rojo y azul con las unidades funcionales denotadas de A1-G1 y A2-G2 respectivamente. La línea negra punteada denota el eje de simetría a nivel de la pared en el dímero y la estrella denota el punto de rotación y traslación con el cual es posible reconstruir el didecámero. En la parte superior derecha se muestra que la simetría se pierde a nivel del collar interno. **B)** Considerando la asimetría del modelo de HDH se construyó el modelo de organización de FUs dentro de mapa de cryo-EM de CCH. La unidad asimétrica destacada en línea amarilla punteada permite reconstruir todo el cilindro que conforma el didecámero mediante operaciones de rotación y traslación representadas por las flechas negras.

Considerando que el mapa de CCH tiene la misma simetría que el modelo de HDH, se generaron los dos únicos modelos posibles de organización para el dímero de subunidades CCHA y CCHB (llamados modelo I y II). Ambos modelos fueron ubicados mediante

*docking* rígido dentro del mapa heterodidecámero de CCH. En la **Figura 27** se ve la diferencia en la ubicación de las subunidades a nivel didecámero para cada modelo: el modelo I ubica las subunidades CCHB hacía ambos extremos del didecámero, cercano al poro de la estructura, mientras que las subunidades CCHA se ubican hacia la parte interior del didecámero (**Figura 27A**). En el modelo II, las subunidades CCHB se ubican hacia el interior del didecámero, mientras que las subunidades CCHA se ubican hacia la parte exterior del didecámero (**Figura 27B**).



**Figura 27. Modelos heterodidecámericos de organización de subunidades de CCH.** *Docking* rígido de los modelos de organización dentro del mapa del didecámero de CCH. **A)** Modelo I. **B)** Modelo II. Los modelos de subunidades se colorearon en rojo para CCHB y azul para CCHA y el mapa de CCH se coloreó según la proximidad de las subunidades.

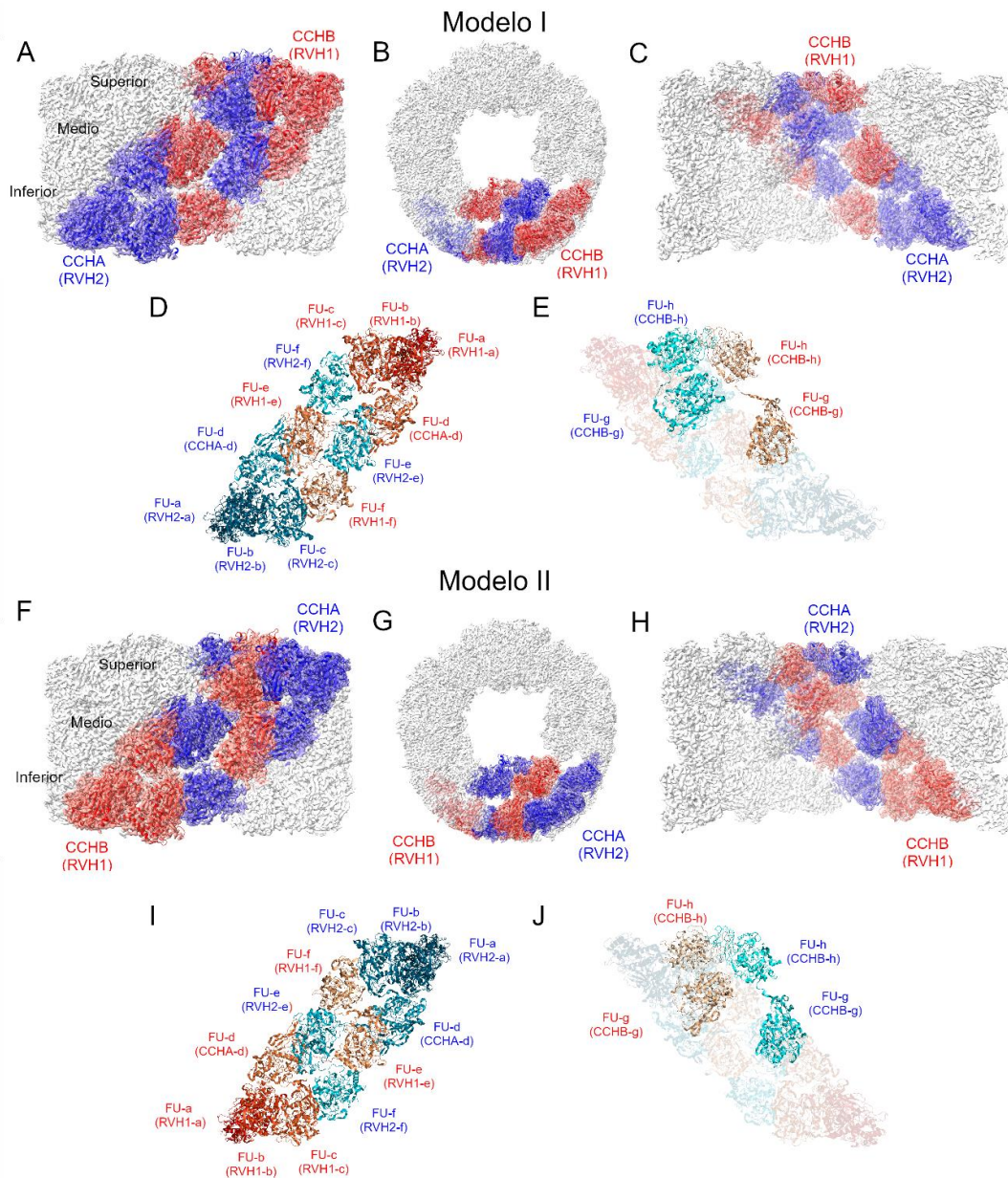
Posteriormente, los modelos fueron procesados a nivel de dímero, ya que el modelo didecámero puede obtenerse mediante operaciones simétricas a partir del dímero de subunidades. Ambos modelos fueron refinados en el espacio real utilizando los softwares Phenix y COOT, acomodando regiones *loops* dentro de la densidad del mapa de cryo-EM y completando las regiones *linker* mediante cadenas de polialanina. La **Figura 28** muestra de forma detallada la disposición de las subunidades y de las FU dentro del decámero para

cada modelo después de un ciclo de refinamiento: Modelo I (**Figura 28A-E**) y Modelo II (**Figura 28F-J**).

Para evaluar cuál de los dos modelos ajusta mejor con la densidad de cryo-EM de CCH, se compararon los valores de coeficiente de correlación (CC) entre cada una de las subunidades de ambos modelos y la densidad del mapa de cryo-EM obtenidos después de cada ciclo de refinamiento. Inicialmente, se refinaron los modelos correspondientes a la distribución de subunidades dentro del didecámero (**Figura 27**) obteniendo valores de CC bastante similares para los dos modelos, sin embargo, con una ligera preferencia hacia la opción del modelo II (**Tabla 5**). Al momento de acotar los modelos a solo las subunidades correspondientes al dímero (**Figura 28A y F**), se observa prácticamente la misma tendencia que en el caso anterior (**Tabla 5**). Considerando que las secuencias de los modelos del dímero no corresponden exactamente a CCH, sino a RVH, lo que pudiese interferir significativamente en el análisis de correlación, se realizó el análisis solo con FUs de los modelos para los cuales se tenía certeza de la secuencia de CCH. Solo se refinaron las FUs CCHB-g y CCHB-h, ubicándolas en la cadena CCHB de cada modelo (**Figura 28E y J**). En este caso se observa que la diferencia entre los CC se hace más dramática, apuntando a que el modelo II es donde se acomodan de mejor manera estas FUs (**Tabla 5**). Sin embargo, este resultado coincide con el hecho que en el modelo I, las FUs CCHB-g y CCHB-h se ubican en la zona del collar interno, en donde ambas FU están separadas por una región *linker* más extensa que en el caso del modelo II (**Figura 28E y J**), por lo que probablemente las FUs en esa zona sean más móviles y por consiguiente esa zona presente una densidad de mapa de menor calidad, lo que también pudiese sesgar

el valor del coeficiente de correlación. Finalmente, para eliminar el sesgo mencionado anteriormente, se acotaron los modelos solo a la CCHA-d, de la cual también se conoce la secuencia y para la cual mediante análisis de identidad de secuencia hay evidencia que sería una FU-d que es equivalente con RVH2-d (**Figura 15, Pág. 48**). Esta FU-d se ubica en el área de la pared del didecámero, siendo esta zona la que presenta la mejor calidad del mapa de densidad de CCH y lo que elimina el sesgo de la calidad del mapa (**Figura 28D e I**). Los resultados del refinamiento de este modelo indican nuevamente que el modelo que se ajusta mejor es el modelo II (**Tabla 5**).





**Figura 28. Modelos heterodecámicos de organización de FUs y subunidades de CCH.** **A) y F)** Vista frontal de la pared del decámero, se muestran las regiones superior, media e inferior de la placa de dímeros de subunidades. **B) y G)** Vista superior del collar interno del decámero. **C) y H)** Vista frontal del collar interno, hacia el interior del decámero. Los modelos de subunidad CCHA y CCHB se colorearon de color azul y rojo respectivamente, el mapa de densidad se presenta de color gris, coloreado por proximidad con las subunidades. **D) e I)** Detalle de la posición de FUs de la pared del decámero. **E) y J)** Detalle de la posición de FUs del collar interno. Las FUs correspondientes al modelo de subunidad CCHA y CCHB se colorearon con degradados del color azul y rojo respectivamente. En paréntesis se indica la secuencia con la que se realizó el modelo de cada FU.

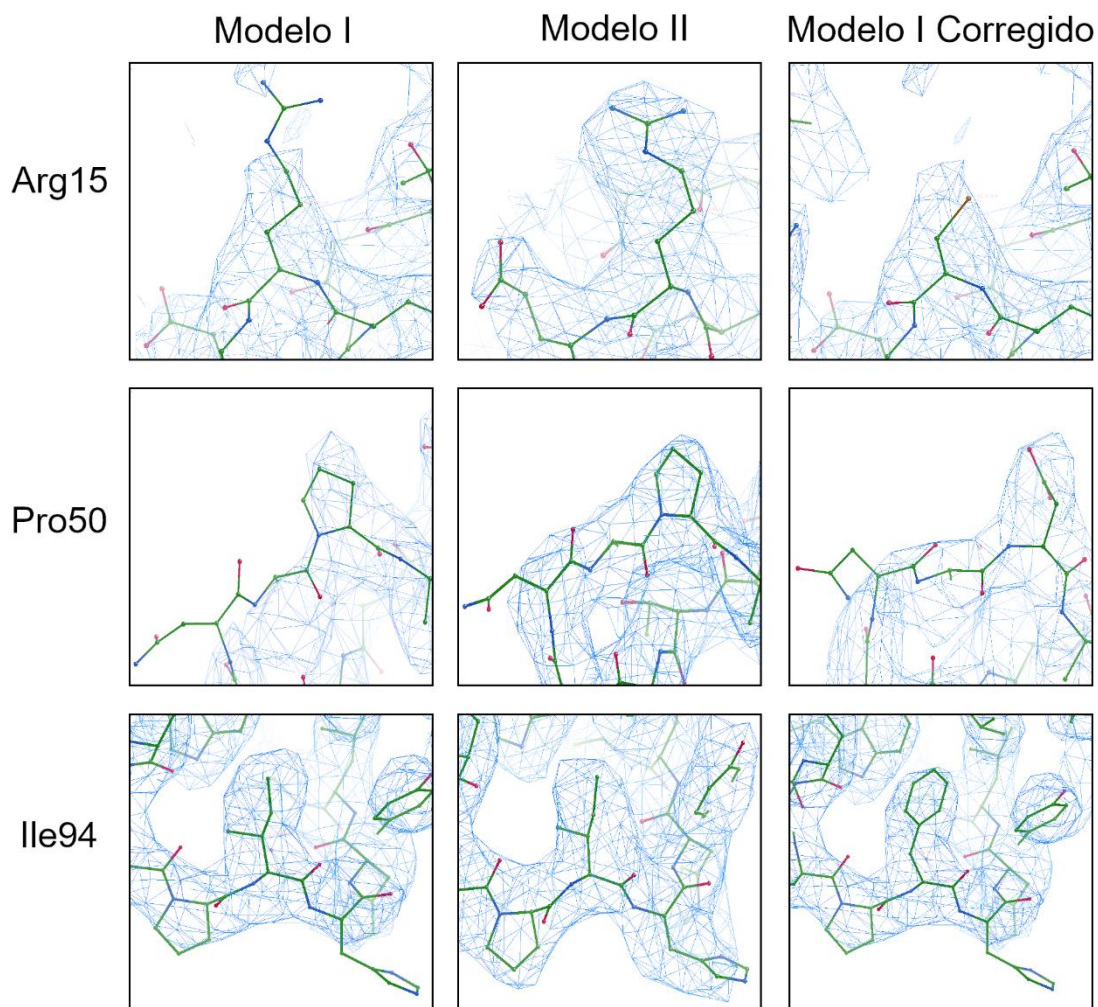


**Tabla 5. Valores de coeficiente de correlación para cada uno de los modelos de organización de subunidades de CCH.**

	Modelo I (CC)	Modelo II (CC)
Modelo didecámero	0,80	0,81
Modelo dímero refinado	0,80	0,82
Modelo solo CCHB-g y CCHB-h	0,68	0,81
Modelo solo CCHA-d	0,76	0,79

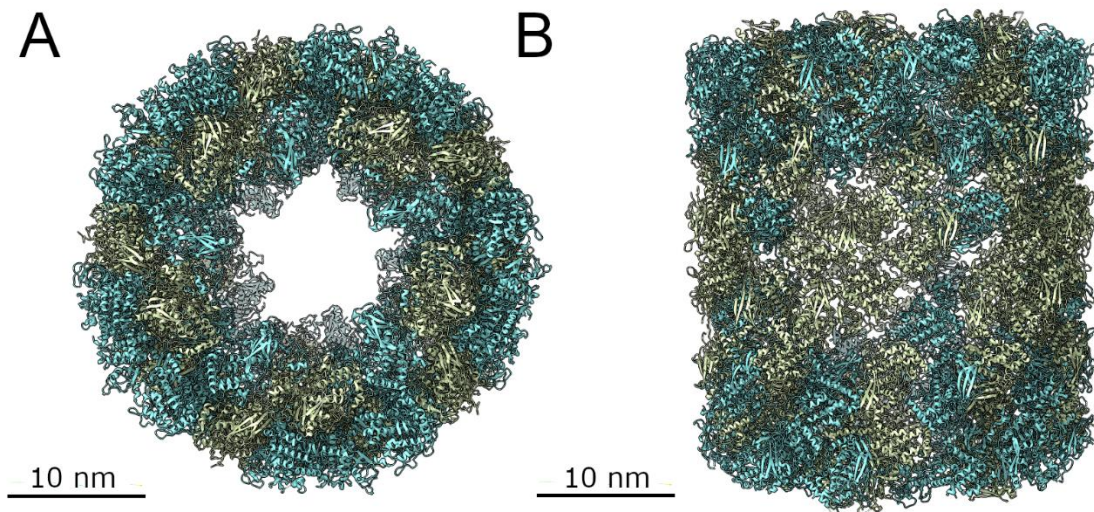
Además de comparar los coeficientes de correlación de cada modelo, se buscaron residuos no conservados entre hemocianinas correspondientes al dominio N-terminal de CCHA-d, dado que esta zona posee densidad de buena calidad en el mapa. Además, dentro del alineamiento de secuencias de las FU disponibles de CCH y las FUs de ambas subunidades de RVH se encuentran posiciones no conservadas, lo que permitirá apreciar diferencias notables en la densidad de cadenas laterales de residuos entre los modelos I y II. Se presentan ejemplos de 3 residuos identificados: en el caso de Arg15 (numeración de CCHA-d) el modelo I no tiene densidad cercana al grupo guanidino, al contrario del modelo II. Más aun, en la FUs RVH1-d y RVH2-d el residuo Arg15 se encuentra reemplazado por una cisteína, y al corregir este residuo por cisteína en el modelo I se observa que ajusta adecuadamente a la densidad circundante (**Figura 29**). En los casos de los residuos Pro50 e Ile94 sucede algo similar, aunque estos residuos se encuentran conservados en las FUs de RVH1-d y RVH2-d, por lo que al analizar la posición de estos residuos en RVH1-g y RVH2-g (que presentan alta identidad de secuencia con las FU-d, **Figura 15, Pág. 48**) se observa que se encuentran reemplazados por residuos de aspártico

y fenilalanina respectivamente, los cuales, al ser corregidos, calzan adecuadamente con la densidad (**Figura 29**). Estos resultados sugieren que en realidad la disposición de la subunidad CCHA no corresponde a la postulada en el modelo I.



**Figura 29. Comparación de residuos de CCHA-d entre modelos de organización de subunidades de CCH.** En los dos primeros recuadros se muestran la posición de residuos Arg15, Pro50 e Ile94 de CCHA-d dentro del mapa de densidad correspondiente a la posición postulada en el modelo I y II. En el tercer recuadro se muestra la corrección del respectivo residuo dentro del modelo I, por residuos de cisteína, aspártico y fenilalanina respectivamente, los cuales fueron observados en FUs de tipo -d y-g de subunidades de RVH.

Finalmente, todos los análisis realizados sugieren que el modelo II describe adecuadamente la organización de las subunidades dentro del didecámero de CCH, siendo esta con la subunidad CCHA hacia ambos extremos del cilindro del didecámero y las subunidades CCHB ubicadas hacia el interior del cilindro del didecámero. Utilizando esta información, se propone el primer modelo a resolución cercana a la atómica (3,2 Å) de una hemocianina heterodidecámerica que plantea la distribución de las distintas subunidades dentro de la estructura del didecámero (**Figura 30**). Este modelo de la hemocianina heterodidecámerica CCH, junto con la estructura 3D de alta resolución (1,5 Å) de la FU CCHB-g, representan un gran avance en el conocimiento de la hemocianina CCH de alto interés biomédico y para el estudio de hemocianinas de molusco en general.



**Figura 30. Modelo de distribución de subunidades dentro del heterodidecámero de CCH.** Se presentan las vistas **A**) Superior del collar interno y **B**) Frontal de la pared del modelo propuesto de heterodidecámero de CCH. La estructura se representa como *cartoon* en donde las subunidades equivalentes a CCHA se colorean de celeste y las subunidades equivalentes a CCHB se colorean de verde pálido.

## 5. DISCUSIÓN

### 5.1 Expresión de FU recombinante

En literatura prácticamente no existe información respecto de hemocianinas de molusco expresadas de manera recombinante. Esto principalmente debido a la complejidad de su estructura, su gran tamaño y la presencia de glicosilaciones en su superficie. Al respecto, existen varios trabajos en donde se expone la importancia estructural que tendrían estas glicosilaciones para la mantención de la estructura decámerica: En la hemocianina de calamar *Todarodes pacificus*, la cual corresponde a la única estructura de hemocianina de molusco resuelta en su totalidad mediante cristalografía de rayos X a una resolución de 3 Å, se observaron la formación de *clusters* de glicosilaciones y 5 potenciales motivos de N-glicosilación conservados entre sitios de unión de unidades funcionales (FUs). Además, estudios con hemocianina de calamar desglucosilada enzimáticamente mostraron, a través de microscopia electrónica de transmisión (MET), que la formación de decámeros disminuyó significativamente (Gai y col., 2015). Sin embargo, los 5 sitios potenciales de N-glicosilación descritos para la hemocianina de calamar se encuentran conservados en hemocianinas de cefalópodos (Hemocianinas tipo 4), 4 de estos sitios se encuentran conservados en hemocianinas de nautilus (Hemocianinas tipo 3) y solo 2 de estos sitios se encuentran conservados en hemocianinas de gastrópodos (Hemocianinas tipo 1), lo cual podría sugerir que la característica de utilizar las glicosilaciones para la estabilización de estructura cuaternaria podría ser exclusiva de las hemocianinas de cefalópodos (Gai y col., 2015).

Posteriormente, se reportó que las N-glicosilaciones de hemocianinas de tipo I de los gastrópodos *Megathura crenulata* (KLH), *Fissurella latimarginata* (FLH) y *Concholepas concholepas* (CCH) también contribuyen a la estabilidad de la estructura cuaternaria didecámerica. Las 3 hemocianinas al ser desglicosiladas enzimáticamente y cargadas en una cromatografía de exclusión molecular mostraron la aparición de especies mixtas parcialmente disociadas con estructuras bajo 1 MDa comparada con los 8 MDa de las especies nativas. Además, estudios de microscopía electrónica de transmisión (MET) revelaron que ninguna de las hemocianinas conservó la forma de cilindro hueco. Más aun, estudios de dicroísmo circular en el UV lejano mostraron que la desglicosilación de estas hemocianinas no afecta la proporción de estructuras secundarias, aunque las curvas de desplegamiento utilizando cloruro de guanidinio mostraron que el plegamiento de las hemocianinas KLH y FLH si se ve afectado en las especies desglicosiladas. Este último análisis no pudo realizarse en CCH dado que se formaron agregados proteicos durante el proceso de replegamiento (Salazar y col., 2019).

Considerando los análisis presentados anteriormente, podría inferirse que la presencia de glicosilaciones resulta fundamental en el plegamiento de las unidades funcionales (FUs) de CCH en el proceso de expresión recombinante de estas. Teniendo en consideración que en el sistema de expresión utilizado (*E. coli*) no se producen glicosilaciones, se incluyó la utilización de los *tags* de solubilización TrxA y SUMO en el N-terminal de las secuencias de FU de CCH expresadas (CCHX y CCHY, **Figura 5, Pág. 22**), dado que, por ejemplo, se ha reportado que el *tag* SUMO ha sido utilizado para la expresión recombinante en *E. coli* de proteínas pertenecientes al virus SARS-CoV, entre ellas la glicoproteína Spike,

cuyo fragmento C-terminal (Spk C) fue fusionado con un tag SUMO en el N-terminal y expresado en una cepa de *E. coli* BL21 (Peroutka y col., 2011).

En este trabajo, a pesar de la utilización de *tags* de solubilización, no se logró la expresión recombinante y purificación de FUs de CCH en *E. coli*, evidenciado en los bajos rendimientos de purificación obtenidos (~ 3 mg totales), el bajo nivel de pureza observado en las fracciones de proteína cargadas en SDS-PAGE (**Figura 7, Pág. 35**) y principalmente, en la ausencia de picos característicos de absorbancia relacionados a los centros de unión a cobre en hemocianinas de molusco (**Figura 10, Pág. 41**) indicando que las muestras purificadas no serían capaces de unir cobre. En vista de que se cuenta con las secuencias codificantes para 2 FUs de CCH de secuencia conocida (**Figura 5A, Pág. 22**) y aun cuando se encuentran sintetizadas con optimización de codones para su expresión en *E. coli*, se podría plantear a futuro el utilizar un sistema de expresión en levadura como *Pichia pastoris* el cual ha sido ampliamente utilizado para la expresión de glicoproteínas recombinantes como fitasas (Guo y col., 2008), endopoligalacturonasas (Ogawa y col., 2009) y elastasas (Han y col., 2014), solo por mencionar algunos ejemplos. Sin embargo, se tiene la desventaja de que las células de levadura solo utilizan manosas como la única unidad de azúcar en el proceso de agregar cadenas externas al árbol de N-glicosilación, por lo que no se podrían adicionar monómeros de galactosa y fucosa a las FU recombinantes en este tipo de sistemas (De Pourcq y col., 2010).

## 5.2 Estructura cristalográfica de FU silvestre de CCH

Una de las principales razones por las que se optó por intentar expresar una FU de origen recombinante a partir de las secuencias disponibles (**Figura 5A, Pág. 22**), fue por el bajo rendimiento de purificación de una FU silvestre. Para la obtención de 1 mg total de FU purificada se debían digerir al menos 10 mg totales de subunidad CCHB, de la cual se obtenían en promedio 6,5 mg totales purificados a partir de un stock de 26 mg totales de CCH purificada correspondientes a un vial proporcionado por la fundación FUCITED. Sin embargo, después de realizar alrededor de 10 purificaciones fue posible obtener proteína suficiente para realizar un ensayo de cristalización inicial en el cual se obtuvieron cristales de tamaño mayor a 0,2 mm.

Posteriormente, se realizó la difracción remota de los cristales obtenidos, colectando tanto conjuntos de datos nativos como conjunto de datos anómalos mediante experimentos de SAD, para resolver el problema de las fases. La integración de los datos cristalográficos mostró dos moléculas de FU en la unidad asimétrica, y 2 señales de dispersión anómala en cada una de las moléculas de la unidad asimétrica, las cuales corresponden a los átomos de cobre del centro activo. Sin embargo, se encontraron otras 3 señales adicionales, las cuales inicialmente se atribuyeron también a átomos de cobre (**Figura 14, Pág. 46**). Los átomos de cobre que corresponden a un metal de transición no permiten explicar la coordinación que presentaban los átomos correspondientes a estas señales, pero los iones de bromuro ( $\text{Br}^-$ ) si explican de mejor manera tanto la coordinación del entorno químico como la densidad electrónica presente. En este punto cabe destacar la importancia del análisis de las condiciones de cristalización, ya que la solución de cristalización del cristal

cuyos datos fueron analizados contenía 150 mM de KBr. Además, el bromo presenta un grado importante de dispersión anómala a la longitud de onda que se realizó el experimento de SAD (1,373 Å) y finalmente, en literatura se ha descrito que históricamente los haluros (bromuro y yoduro) corresponden a una excelente opción para hacer experimentos de SAD (Dauter & Dauter, 2007), incluso llegando a proponerse su utilización para la determinación de estructuras en cristales de proteínas de membrana (Melnikov y col., 2017).

Una vez resuelto el problema de las fases y dado que no poseíamos la secuencia de la FU cristalizada, se trazó un polímero de polialanina en la densidad electrónica obtenida y gracias a la alta resolución del conjunto de datos anómalos ( $\sim 1,7$  Å), se pudieron identificar los residuos de la proteína completando las cadenas laterales mediante análisis visual de las densidades electrónicas. En este punto resulta importante recordar que la cristalografía de rayos X a alta resolución ( $< 2,0$  Å) permite secuenciar proteínas las cuales cuentan con una condición de cristalización. Por ejemplo, en literatura se ha descrito la herramienta findMySequence, diseñada para la identificación de secuencias de proteínas desde datos cristalográficos y reconstrucciones de Cryo-EM, siendo utilizada para la identificación de una proteína purificada y cristalizada proveniente del veneno de la serpiente venenosa *Bothrops atrox* (Chojnowski y col., 2021). En esta tesis se logró identificar la secuencia de la FU cristalizada solo con el apoyo de secuencias de péptidos obtenidos a través de espectrometría de masas. Más aun, se logró confirmar que la secuencia obtenida a través de cristalografía coincidía con la parte N-terminal de la secuencia disponible de CCH correspondiente a un fragmento parcial de FU tipo -g



(CCHX, **Figura 5, Pág. 22**), diferenciándose solo en 5 residuos ambiguos a nivel de densidad electrónica de cadena lateral, lo que habla de una alta confianza en la estrategia de secuenciación a través de la cristalografía. Finalmente, de esta forma se logró completar los 110 residuos faltantes de la secuencia parcial de FU tipo -g, posteriormente identificada como CCHB-g.

El modelo de CCHB-g obtenido a partir del conjunto de datos nativos fue utilizado para refinar una estructura perteneciente al conjunto de datos nativos para el mismo cristal que presentaba una mayor resolución (1,5 Å). La estructura refinada presentó el clásico plegamiento descrito para una FU de hemocianina (de tipo -a hasta -g). Se realizó una búsqueda de similitud estructural con la cadena A de la estructura refinada de CCHB-g, utilizando el servidor web Dali (Holm, 2020), el cual arrojó las 3 estructuras de FU disponibles en literatura como las estructuras más similares con un valor de RMSD (del inglés *Root Mean Square Deviation*) menor a 2 Å (**Tabla 6**). La estructura más similar a CCHB-g fue la FU del mismo tipo de la hemocianina tipo 3 del organismo *O. dofleini* (ODH-g) con un RMSD de 1,1 Å sobre 381 residuos con un 53% de identidad de secuencia, seguida por la las FUs de hemocianina tipo 1 correspondientes a RVH2-e y KLH-h. Finalmente, de este mismo análisis se desprende que la estructura cristalográfica de CCHB-g determinada en este trabajo corresponde a la estructura de FU reportada a mayor resolución hasta la fecha (**Tabla 6**).

**Tabla 6. Comparación de parámetros de estructuras similares a CCHB-g.**

FU	Organismo	Resolución (Å)	Nº residuos totales	Nº residuos alineados	RMSD (Å)	Id (%)	PDB
CCHB-g	<i>Concholepas concholepas</i>	1,5	392	-	-	-	-
ODH-g	<i>Octopus dofleini</i>	2,3	382	381	1,1	53	1JS8
RVH2-e	<i>Rapana venosa</i>	3,3	408	387	1,5	41	1LNL
KLH-h	<i>Megathura crenulata</i>	4,0	491	383	1,8	39	3QJO

Respecto al análisis de las glicosilaciones en la estructura de CCHB-g, se puede observar que el residuo Asn123 presenta un árbol de glicosilación Fuc( $\alpha$ 1-6)Man( $\beta$ 1-4)GlcNAc( $\beta$ 1-4)[Fuc( $\alpha$ 1-6)]GlcNAc (**Figura 17, Pág. 52**). En relación con este punto, para la hemocianina KLH, la más estudiada desde el punto de vista bioquímico, se ha descrito que las dos isoformas KLH1 y KLH2 (ambos homodidecámeros) poseen principalmente manosa, galactosa, N-acetilglucosamina, N-acetilgalactosamina y fucosa en distintas proporciones (Stoeva y col., 1999). Adicionalmente, se han encontrado patrones de glicosilación distintivos como Gal( $\beta$ 1-6)Man, el cual no es encontrado frecuentemente en glicoproteínas (Kurokawa y col., 2002). Más aun, se han descrito patrones inmunogénicos en KLH, tales como Fuc( $\alpha$ 1-3)GalNAc( $\beta$ 1-4)[Fuc( $\alpha$ 1-3)]GlcNAc, glicosilación que también está presente en el parásito humano *Schistosoma mansoni* (Wuhrer y col., 2004) y que es relativamente similar a la glicosilación encontrada en la estructura de CCHB-g, ya que también posee dos monómeros de fucosa, pero unidos de forma distinta (enlace  $\alpha$ 1-6 en CCHB y enlace  $\alpha$ 1-3 en KLH) y se reemplaza un monómero de manosa en CCHB por un monómero de N-acetilgalactosamina (GalNAc)

en KLH. Dentro de este mismo punto, un estudio reciente reporta la composición global de N-glicosilación de las hemocianinas de *Concholepas concholepas* (CCH), *Fissurella latimarginata* (FLH) y *Megthura crenulata* (KLH) mediante análisis específicos de identificación por librería de lectinas (Salazar y col., 2019). En dicho estudio, se obtuvo que CCH presenta principalmente monómeros de  $\alpha$ -manosas y complejos con estructuras de GlcNAc, en adición con monómeros de fucosa unidos a estos monosacáridos, las cuales son el mayor componente de las N-glicosilaciones. De la misma forma, CCH presentó N-acetilactosaminas (LacNAc), lactosas y  $\beta$ -galactosas en menor proporción siendo los principales componentes de O-glicosilaciones y N-glicosilaciones mixtas. En el caso de FLH, dado el patrón de unión de lectinas, se sugiere que posee monómeros de manosas terminales, GlcNAc unidas a  $\alpha$ -manosas y GalNAc, en adición a los monosacáridos mencionados anteriormente. La KLH mostró un patrón de unión de lectinas similar a CCH, pero con un contenido mayor de galactosas. En el estudio también se reporta que CCH desglicosilada enzimáticamente muestra una remoción parcial de complejos GlcNAc, LacNAc y galactosas, mientras que la presencia de fucosas no se ve disminuida, al contrario de FLH que también presenta un contenido similar de fucosas, pero que al desglicosilarse, baja notoriamente el contenido de este monosacárido (Salazar y col., 2019). Este último resultado puede explicarse debido a que la enzima utilizada para la N-desglicosilación fue la Péptido:N-glicosidasa F (PNGasa F) la cual está descrita que rompe los enlaces glucosídicos internos en N-glicosilaciones complejas, mixtas o con alto contenido de manosa, excepto cuando el glucano posee un enlace fucosa ( $\alpha$ 1-3) en la GlcNAc unida al residuo de Asn (Freeze & Kranz, 2008), lo que sugiere que CCH tendría una mayor proporción de fucosas unida de esta forma, al contrario de FLH. En

contraposición, la estructura reportada de CCHB-g en este trabajo presenta un monómero de fucosa enlazado de forma ( $\alpha$ 1-6) en la GlcNAc (**Figura 17, Pág. 52**), por lo que es posible que no pueda apreciarse en la estructura cristalográfica un monómero extra de fucosa enlazado ( $\alpha$ 1-3) en la GlcNAc de este árbol de glicosilación, o bien, que simplemente el árbol de glicosilación observado en la FU-g de CCHB no posea una fucosa unida de esta manera.

El último elemento analizado de la estructura de CCHB-g fue el centro activo, consistente en un centro de unión a cobre tipo 3, el cual corresponde a un par de átomos de cobre, cada uno coordinado por 3 residuos de histidina. Este tipo de centros de unión a cobre se encuentra en hemocianinas y tirosinasas (Lewis & Tolman, 2004). Como se mencionó anteriormente, la estructura de CCHB-g reportada en este trabajo se encuentra determinada a una resolución de 1,5 Å, la cual corresponde a la mayor resolución reportada hasta la fecha para una estructura de FU de hemocianina de molusco. La resolución obtenida nos permite analizar en detalle elementos estructurales característicos del centro de unión a cobre tales como el clúster  $\text{Cu}_2\text{O}_2$ , los 6 residuos de histidina que participan en la coordinación y otros elementos conservados en otras estructuras de FU de molusco como un enlace disulfuro y un puente tioéter formado entre una cisteína y una histidina que participa en la coordinación del clúster  $\text{Cu}_2\text{O}_2$  (**Figura 18, Pág. 53**). Más aun, la resolución obtenida permite entregar información detallada acerca del largo de distancia y los ángulos químicos de los enlaces entre los elementos que conforman el clúster  $\text{Cu}_2\text{O}_2$ . Este punto resulta importante, ya que la única información cristalográfica del clúster  $\text{Cu}_2\text{O}_2$  de un centro de unión a cobre tipo 3 de hemocianina, corresponde a la

estructura de la FU tipo -g de hemocianina de pulpo (ODH-g, PDB 1JS8) la cual se encuentra determinada a una resolución de 2,3 Å. Esto tiene implicaciones prácticas, por ejemplo, en el proceso de refinar el clúster  $\text{Cu}_2\text{O}_2$  de la estructura CCHB-g en el software COOT (Emsley & Cowtan, 2004) utilizando las restricciones químicas reportadas en la estructura de ODH-g, se produjeron varios problemas generados por las diferencias principalmente en las distancias y ángulos entre los átomos de oxígeno del clúster (**Tabla 2 y Tabla 3, Pág. 55**), por lo que fue necesario refinar el clúster sin utilizar las restricciones disponibles, sino solamente la densidad electrónica de alta resolución del modelo (1,5 Å).

En resumen, en este trabajo se logró la purificación de una FU de CCH, y posteriormente, la obtención de una estructura cristalográfica a resolución atómica de esta FU permitió obtener su secuencia e identificación como CCHB-g. Además, permitió la identificación de un árbol de N-glicosilaciones, pudiendo compararse con datos de identificación de N-glicosilaciones de CCH obtenidos mediante caracterizaciones bioquímica previas y, finalmente, realizar mediciones detalladas de las interacciones químicas del centro de unión a cobre tipo 3 presente en hemocianinas de molusco.

### 5.3 Modelo de Cryo-EM de CCH

Desde el año 2013, se habla de la “revolución de la resolución” (“resolution revolution”) en referencia al progreso en el análisis de partículas individuales mediante cryo-EM (Kuhlbrandt, 2014). Relacionado a eso, los dos modelos estructurales obtenidos por cryo-EM de hemocianinas de molusco de tipo I: KLH a una resolución de 9 Å (Gatsogiannis & Markl, 2009) y HDH a una resolución de 4,5 Å (Zhang y col.,2013), fueron reportados antes de la “revolución de la resolución”. Como se observa en la **Tabla 7**, hay una clara correlación en la tecnología de los detectores utilizados para la colecta de datos y las resoluciones alcanzadas por los mapas de cryo-EM. En este trabajo se obtuvo un conjunto de datos de cryo-EM de la hemocianina de *C. concholepas* (CCH), el cual fue colectado utilizando el crio-microscopio electrónico de transmisión de última generación Titan Krios, resultando en un mapa a una resolución de 3,2 Å que corresponde a la mayor resolución reportada para un mapa de cryo-EM de hemocianina de molusco tipo I.

**Tabla 7. Comparación de tecnologías de detector asociadas con la resolución obtenida en mapas de cryo-EM para hemocianinas de molusco de tipo I.**

Mapa cryo-EM	Tecnología del detector asociada a la colecta de datos cryo-EM
KLH (Resolución 9,0 Å)	Kodak SO-163. Durante varias décadas y hasta hace no muchos años, la cinta fotográfica se mantuvo como el detector de preferencia para cryo-EM, por lo que era necesario revelar y escanear las cintas. El movimiento de partículas producto de la energía del haz de electrones producía imágenes más borrosas, produciendo a menudo que partículas diferentes estructuralmente se agruparan en una sola reconstrucción. <b>Referencia:</b> Bai y col., 2015.
HDH (Resolución 4,5 Å)	
CCH (Resolución 3,2 Å)	Cámara digital K3. A partir de los 2000s, el ingreso de detectores directos de electrones al campo de cryo-EM permitió detectar directamente la energía depositada por los electrones en los componentes electrónicos fabricados en un chip recubierto por una membrana semiconductor. A partir del 2013, a los detectores directos de electrones, se sumó la tecnología de conteo de electrones únicos, maximizando la relación señal/ruido de las películas digitales de cryo-EM y siendo la principal mejora para la “revolución de la resolución”. <b>Referencia:</b> Sun y col., 2021.

La influencia del uso de mejores tecnologías en detectores también se evidencia en la cantidad de partículas de proteína que pueden aislarse a partir de una colecta de datos de cryo-EM. En el caso de este trabajo, para la colecta de datos de cryo-EM de CCH se utilizó una cámara K3 en modo conteo de electrones a super-resolución. El modo conteo de electrones permite una mayor velocidad de fotogramas y un recuento de pixeles más alto, mientras que la super-resolución toma el hecho de que un electrón es detectado simultáneamente por múltiples pixeles adyacentes, lo que permite dividir en dos un cuadrante del píxel físico 1,1 Å, quedando en 0,55 Å en modo super-resolución, permitiendo entonces estimar la ubicación del electrón incidente con mayor precisión (Sun y col., 2021). En resultados concretos, la implementación de esta tecnología se traduce en 558.149 partículas iniciales extraídas por colecta de datos de cryo-EM para CCH y 130.952 partículas finales

utilizadas para la reconstrucción del modelo 3D refinado, comparadas con las 6.100 partículas iniciales extraídas para KLH, con 4.094 partículas finales para la reconstrucción del modelo (Gatsogiannis & Markl, 2009) y 65.000 partículas iniciales extraídas para HDH, con 28.641 partículas finales para la reconstrucción del modelo (Zhang y col.,2013).

En conjunto con la mejora tecnológica de las cámaras digitales, la mejora en los algoritmos de los softwares de procesamiento de imágenes cryo-EM y la reconstrucción de modelos también han contribuido para convertir la cryo-EM a alta resolución en una herramienta de rutina. En este trabajo, se utilizó el *software* cryoSPARC el cual presenta la combinación de dos algoritmos para reconstrucción de modelos *ab initio*: descenso de gradiente estocástico (SGD) y optimización de máxima probabilidad de ramificación y límite, los cuales reducen considerablemente el gasto computacional, esto junto con la utilización de una interfaz gráfica amigable con el usuario permiten reducir considerablemente el gasto en *hardware* y el nivel de habilidad computacional para poder realizar el tratamiento de los datos de cryo-EM. El uso del software cryoSPARC permitió realizar todo el procesamiento de datos de cryo-EM: tratamiento de las imágenes de cryo-EM (corrección de movimiento, estimación de CTF), la reconstrucción *ab initio* (selección de partículas, clasificación 2D, reconstrucción de modelos) y el refinamiento de modelos (refinamiento homogéneo/heterogéneo y determinación de resolución). En comparación, los modelos de mapas de KLH y HLH fueron generados utilizando diferentes softwares para tratamiento de imágenes, reconstrucción de modelos y estimación de resolución (CTFFIND3, IMAGIC, EMAN), más aún, el *software* EMAN necesita de un modelo 3D preliminar, como el modelo de KLH que necesitó un modelo 3D de la hemocianina de



*Haliotis tuberculata* (HTH) reportada anteriormente a 11 Å (modelo no disponible en ninguna base de datos, Meissner y col., 2007) lo cual genera modelos sesgados y en este caso específicamente, un modelo con errores en la distribución de FUs dentro del decámero de KLH (Gatsogiannis y col., 2009). En el caso del modelo de HDH, el *software* EMAN1 necesitó de varias rondas de reconstrucción de modelos 3D preliminares para generar un modelo 3D inicial con el cual poder reconstruir el modelo final de HDH (Zhang y col., 2013), mientras que en el caso de CCH, el utilizar cryoSPARC con la combinación de algoritmos, permitió reconstruir un mapa 3D muy similar al modelo final, disminuyendo considerablemente los tiempos de cálculo y el uso de recursos computacionales.

Una vez obtenido el mapa 3D de CCH, se realizó un modelo de *docking* rígido utilizando el modelo 3D de HDH que corresponde a un homodidecámero. Al colorear el mapa 3D de CCH por proximidad con las cadenas del modelo de HDH, se observa que el modelo calza bastante bien con el mapa, y que solo queda una parte sin colorear correspondiente a las regiones de ubicación de los dominios extra *cupredoxin-like* de las FU-h correspondientes al collar interno del didecámero (**Figura 25, Pág. 65**). Este experimento resulta fundamental, ya que muestra que el mapa de CCH que corresponde a un heterodidecámero posee la misma simetría de subunidades y distribución de FUs dentro de las subunidades que el modelo de HDH que corresponde a un homodidecámero. De la misma forma, una posibilidad de organización de subunidades dentro del heterodidecámero, es que cada decámero esté formado solo por subunidades CCHA o CCHB, sin embargo, se ha visto a través de microscopía electrónica de tinción negativa

que la asociación entre subunidades CCHB forman polímeros tubulares, mientras que la asociación entre subunidades CCHA no forman polímeros, sino que presentan formas más bien irregulares, similares a agregados (González y col., 2017). Estos resultados descartan la opción de decámeros solo formados por CCHA o CCHB, concluyéndose que los únicos dos modelos posibles de organización de subunidades CCHA y CCHB para el heterodidecámero de CCH son los presentados en la **Figura 27, Pág. 68**.

Posteriormente, una vez planteados ambos modelos posibles de organización heterodidecámerica, se procedió a completar los dímeros de subunidades CCHA y CCHB utilizando para ello las secuencias de las FU-d, -g y -h de CCH de las cuales conocemos la secuencia y para las secuencias de FU para cuales no conocemos la secuencia, FU-a, -b, -c, -e y -f se utilizaron las secuencias de FU de cada subunidad correspondiente a la hemocianina del organismo *Rapana venosa* (RVH), dado que según las secuencias de CCH que tenemos disponibles, son equivalentes a las secuencias de subunidades de RVH: RVH1 equivalente a CCHB y RVH2 equivalente a CCHA. Después de generar ambos modelos de organización de subunidades (**Figura 28, Pág. 71**), estos fueron alineados al mapa de cryo-EM de CCH mediante *docking* rígido y se procedió a evaluar ambos modelos utilizando la herramienta refinamiento en espacio real de Phenix, comparando los valores de coeficientes de correlación (CC) para determinar el grado de ajuste los modelos a la densidad del mapa de cryo-EM de CCH. Los valores de CC obtenidos indicaron que el modelo de distribución con las subunidades CCHA hacia ambos extremos del cilindro del didecámero y las subunidades CCHB ubicadas hacia el interior del cilindro del didecámero es el que mejor se ajusta al mapa de cryo-EM (**Tabla 5, Pág. 72**), lo que

también se confirma mediante el análisis visual de ajuste de la densidad electrónica circundante en residuos no conservados de la posición de FU CCHA-d en ambos modelos.

El carácter inmunogénico y adyuvante de las hemocianinas de mamíferos se ha atribuido a su gran tamaño (4-8 MDa), lo intrincado de su estructura cuaternaria y la presencia de glicosilaciones en su superficie. Bajo la misma línea, se ha evaluado el requerimiento del contenido de glicosilaciones tanto para la mantención de la estructura cuaternaria de hemocianinas de molusco como para sus actividades inmunológicas en mamíferos (Salazar y col., 2019). Al respecto, los resultados obtenidos en este trabajo constituyen un aporte para el desarrollo de las aplicaciones biomédicas de las hemocianinas pues entregan información estructural empírica acerca una de las unidades funcionales que componen la hemocianina del molusco *Concholepas concholepas* (CCH), en conjunto con un patrón de glicosilación el cual contiene fucosa, azúcar encontrado solo en N-glicosilaciones de mamíferos, insectos y plantas. Esto podría abrir un marco de investigación acerca de la relevancia de estas glicosilaciones en la generación de una mayor respuesta inmune en mamíferos, ya que se ha descrito que la hemocianina de *Fissurella latimarginata* (FLH) posee una mayor proporción de fucosas comparada con CCH y con la hemocianina de “keyhole limpet” (KLH) (Salazar y col., 2019) y produce una respuesta más inmunogénica en conjunto con una actividad antitumoral más potente que CCH y KLH en modelos murinos de melanoma (Arancibia y col., 2014). Por otra parte, en este trabajo se propone un modelo de distribución de las distintas subunidades en el heterodidecámero de CCH, propiedad estructural única descrita para esta hemocianina, con lo que dicho modelo permitirá que una vez obtenida la secuencia aminoacídica completa de CCH, se pueda

generar un modelo tridimensional que podría contribuir a dilucidar por qué esta hemocianina posee mayor estabilidad y solubilidad comparada con hemocianinas homólogas.

## 6. CONCLUSIONES

- 1.- La utilización de los *tags* de solubilización SUMO y TrxA no permitieron la producción de FUs de CCH de secuencia conocida de manera recombinante en *E. coli*, probablemente debido a la ausencia de glicosilaciones, las cuales serían fundamentales para el plegamiento de las FU.
- 2.- La estructura cristalográfica de la FU CCHB-g obtenida a resolución atómica (1,5 Å) permitió completar la secuencia de esta FU, analizar detalladamente el patrón de las glicosilaciones (presencia de monómeros de fucosa) y analizar en detalle las distancias y ángulos de las interacciones químicas del centro de unión a cobre.
- 3.- La construcción de un dendrograma con las secuencias de FU provenientes de la hemocianina de *Rapana venosa* (RVH) permitió agrupar e identificar el tipo de FU proveniente de las secuencias disponibles de CCH.
- 4.- El mapa de cryo-EM de CCH obtenido a una resolución cercana a la atómica (3,2 Å) permitió ubicar los modelos de FU creados a partir de la estructura cristalográfica de CCHB-g, generando modelos de organización tridimensional de las subunidades CCHA y CCHB dentro del heterodidecámero de CCH.
- 5.- El análisis de los modelos de organización tridimensional de subunidades de CCH, sugiere fuertemente que las subunidades CCHA se ubican hacia los extremos del heterodidecámero de CCH, mientras que las subunidades CCHB se ubican hacia el interior de la estructura, constituyendo el primer modelo de organización tridimensional de un heterodidecámero de hemocianina de tipo I reportado hasta la fecha.

## 7. BIBLIOGRAFÍA

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J, Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., & Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* 66, 213–221.

Arancibia, S., Espinoza, C., Salazar, F., Del Campo, M., Tampe, R., Zhong, T.Y., De Ioannes, P., Moltedo, B., Ferreira, J, Lavelle, E.C, Manubens, A., De Ioannes A.E, & Becker, M.I. (2014). A novel immunomodulatory hemocyanin from the limpet *Fissurella latimarginata* promotes potent anti-tumor activity in melanoma. *PLoS One*, 9(1), e87240.

Bai, X., McMullan, G., & Scheres, S.H.W. (2015). How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences*. Elsevier Ltd, 40(1), 49–57.

Beck, L., Van-Lüme, D.S., de Souza, J.R., Morais, C.N., Melo, W G., Xavier, E., Barbosa, C.S., Aroucha, M.L., Domingues, A.L., Favre, T., Pieri, O., Abath, F.G., & Montenegro, S.M. (2004). Evaluation of tests based on the antibody response to keyhole limpet haemocyanin and soluble egg antigen to differentiate acute and chronic human *schistosomiasis mansoni*. *Memórias do Instituto Oswaldo Cruz*, 99, 97-98.

Becker, M. I., Fuentes, A., Del Campo, M., Manubens, A., Nova, E., Oliva, H., Faunes, F., Valenzuela, M.A., Campos-Vallette, M., Aliaga, A., Ferreira, J., De Ioannes, A.E., De Ioannes, P., & Moltedo, B. (2009). Immunodominant role of CCHA subunit of *Concholepas* hemocyanin is associated with unique biochemical properties. *International immunopharmacology*, 9(3), 330-339.

Boisguérin, V. (2006). Recombinant expression of molluscan hemocyanin (KLH) substructures in a prokaryotic system: *E. coli* (Doctoral dissertation, Mainz, Univ., Diss., 2006).

Boisset, N., & Mouche, F. (2000). *Sepia officinalis* hemocyanin: a refined 3D structure from field emission gun cryoelectron microscopy. *J Mol Biol* 296(2):459–472

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, 72(1–2), 248–254.

Butt, T.R., Edavettal, S.C., Hall, J.P., & Mattern, M.R. (2005). SUMO fusion technology for difficult-to-express proteins. *Protein expression and purification*, 43(1), 1-9.

- Chojnowski, G., Simpkin, A. J., Leonardo, D. A., Seifert-Davila, W., Vivas-Ruiz, D. E., Keegan, R. M., & Rigden, D. J. (2022). findMySequence: a neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCrJ*, 9(1).
- Cuff, M.E., Miller, K.I., van Holde, K.E., & Hendrickson, W.A. (1998). Crystal structure of a functional unit from Octopus hemocyanin. *J Mol Biol* 278(4):855–870
- Dauter, M., & Dauter, Z. (2007). Phase determination using halide ions. In *Macromolecular Crystallography Protocols* (pp. 149-158). Humana Press.
- De Ioannes, P., Moltedo, B., Oliva, H., Pacheco, R., Faunes, F., De Ioannes, A.E., & Becker, M.I. (2004). Hemocyanin of the molluscan *Concholepas concholepas* exhibits an unusual heterodecameric array of subunits. *Journal of Biological Chemistry*, 279(25), 26134-26142.
- De Pourcq, K., De Schutter, K., & Callewaert, N. (2010). Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Applied microbiology and biotechnology*, 87(5), 1617-1631.
- Emsley, P., & Crispin, M. (2018). Structural analysis of glycoproteins: building N-linked glycans with Coot. *Acta Crystallographica Section D: Structural Biology*, 74(4), 256-263.
- Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr.* 60, 2126–2132.
- Evans, P.R., & Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr.* 69, 1204–1214.
- Freeze, H. H., & Kranz, C. (2008). Endoglycosidase and glycoamidase release of N-linked glycans. *Current protocols in immunology*, 83(1), 8-15.
- Gai, Z., Matsuno, A., Kato, K., Kato, S., Khan, M., Shimizu, T., Yoshioka, T., Kato, Y., Kishimura, H., Kanno, G., Miyabe, Y., Terada, T., Tanaka, Y., & Yao, M. (2015). Crystal structure of the 3.8 MDa respiratory supermolecule hemocyanin at 3.0 Å resolution. *Structure* 23:2204–2212
- Gatsogiannis, C., Hofnagel, O., Markl, J., & Raunser, S. (2015). Structure of mega-hemocyanin reveals protein origami in snails. *Structure* 23:93–103
- Gatsogiannis, C., & Markl, J. (2009). Keyhole limpet hemocyanin: 9-A CryoEM structure and molecular model of the KLH1 didecamer reveal the interfaces and intricate topology of the 160 functional units. *J Mol Biol* 385(3):963–983

Gatsogiannis, C., Moeller, A., Depoix, F., Meissner, U., & Markl, J. (2007). *Nautilus pompilius* Hemocyanin: 9 Å Cryo-EM structure and molecular model reveal the subunit pathway and the interfaces between the 70 functional units. *J Mol Biol* 374(2):465–486

Geyer, H., Wuhrer, M., Resemann, A., & Geyer, R. (2005). Identification and characterization of keyhole limpet hemocyanin N-glycans mediating cross-reactivity with *Schistosoma mansoni*. *J Biol Chem* 280: 40731–40748.

González, A., Nova, E., Del Campo, M., Manubens, A., De Ioannes, A., Ferreira, J., & Becker, M. I. (2017). The oxygen-binding properties of hemocyanin from the mollusk *Concholepas concholepas*. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1865(12), 1746-1757.

Guerrero, F., Ciragan, A., & Iwai, H. (2015). Tandem SUMO fusion vectors for improving soluble protein expression and purification. *Protein expression and purification*, 116, 42-49.

Guo, M., Hang, H., Zhu, T., Zhuang, Y., Chu, J., & Zhang, S. (2008). Effect of glycosylation on biochemical characterization of recombinant phytase expressed in *Pichia pastoris*. *Enzyme and Microbial Technology*, 42(4), 340-345.

Han, M., Wang, X., Ding, H., Jin, M., Yu, L., Wang, J., & Yu, X. (2014). The role of N-glycosylation sites in the activity, stability, and expression of the recombinant elastase expressed by *Pichia pastoris*. *Enzyme and microbial technology*, 54, 32-37.

Harris, J.R., & Markl, J. (1999). Keyhole limpet hemocyanin (KLH): a biomedical review. *Micron* 30:597–623.

Holm, L. (2020) Using Dali for protein structure comparison. *Methods Mol. Biol.* 2112, 29-42.

van Holde, K.E. & Miller, K.I. (1995). ‘Hemocyanins’, *Advances in protein chemistry*. United States, 47, pp. 1–81.

Jaenicke, E., Büchler, K., Markl, J., Decker, H., & Barends, T.R.M. (2010). Cupredoxin-like domains in haemocyanins. *Biochem J* 426(3):373–378.

Jaenicke, E., Büchler, K., Decker, H., Markl, J., & Schroder, G.F. (2011). The refined structure of functional unit h of keyhole limpet hemocyanin (KLH1-h) reveals disulfide bridges. *IUBMB Life* 63(3):183–187.

Kabsch, W. (2010). XDS. *Acta Crystallogr.* 66, 125–132.

Kato, S., Matsui, T., & Tanaka, Y. (2020). Molluscan Hemocyanins. *Vertebrate and Invertebrate Respiratory Proteins, Lipoproteins and other Body Fluid Proteins*, 195-218.



- Kuhlbrandt, W. (2014). The Resolution Revolution. *Science*, 343(6178), 1443–1444.
- Kurokawa, T., Wuhrer, M., Lochnit, G., Geyer, H., Markl, J., & Geyer, R. (2002). Hemocyanin from the keyhole limpet *Megathura crenulata* (KLH) carries a novel type of N-glycans with Gal ( $\beta$ 1–6) Man-motifs. *European journal of biochemistry*, 269(22), 5459-5473.
- LaVallie, E.R., DiBlasio, E.A., Kovacic, S., Grant, K.L., Schendel, P.F., & McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/technology*, 11(2), 187-193.
- Lewis, E. A., & Tolman, W. B. (2004). Reactivity of dioxygen– copper systems. *Chemical reviews*, 104(2), 1047-1076.
- Manubens, A., Salazar, F., Haussmann, D., Figueroa, J., Del Campo, M., Pinto, J.M., Huaquín, L., Venegas, A., & Becker, M.I. (2010). *Concholepas* hemocyanin biosynthesis takes place in the hepatopancreas, with hemocytes being involved in its metabolism. *Cell and tissue research*, 342(3), 423-435.
- Matsuno, A., Gai, Z., Tanaka, M., Kato, K., Kato, S., Katoh, T., Shimizu, T., Yoshioka, T., Kishimura, H., Tanaka, Y., & Yao, M. (2015) Crystallization and preliminary X-ray crystallographic study of a 3.8-MDa respiratory supermolecule hemocyanin. *J Struct Biol* 190:379–382.
- Meissner, U., Gatsogiannis, C., Moeller, A., Depoix, F., Harris, J.R., & Markl, J. (2007). Comparative 11 Å structure of two molluscan hemocyanins from 3D cryo-electron microscopy. *Micron*, 38(7), 754-765.
- Melnikov, I., Polovinkin, V., Kovalev, K., Gushchin, I., Shevtsov, M., Shevchenko, V., Mishin, A., Alekseev, A., Rodriguez-Valera, F., Borshchevskiy, V., Cherezov, V., Leonard, G.A., Gordeliy, V., & Popov, A. (2017). Fast iodide-SAD phasing for high-throughput membrane protein structure determination. *Science advances*, 3(5), e1602952.
- Miller, L.A., Fagerstone, K.A., & Eckery, D.C. (2013). Twenty years of immunocontraceptive research: lessons learned. *Journal of Zoo and Wildlife Medicine*, 44(4s).
- Molledo, B., Faunes, F., Haussmann, D., De Ioannes, P., De Ioannes, A.E., Puente, J., & Becker, M.I. (2006). Immunotherapeutic effect of *Concholepas* hemocyanin in the murine bladder cancer model: evidence for conserved antitumor properties among hemocyanins. *The Journal of urology*, 176(6), 2690-2695.
- Nordberg, J., & Arnér, E.S. (2001). Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. *Free radical biology and medicine*, 31(11), 1287-1312.

Ogawa, S., Shimizu, T., Ohki, H., Araya, T., Okuno, T., & Miyairi, K. (2009). Expression, purification, and analyses of glycosylation and disulfide bonds of *Stereum purpureum* endopolygalacturonase I in *Pichia pastoris*. *Protein expression and purification*, 65(1), 15-22.

Oliva, H., Moltedo, B., De Ioannes, P., Faunes, F., De Ioannes, A. E., & Becker, M.I. (2002). Monoclonal antibodies to molluscan hemocyanin from *Concholepas concholepas* demonstrate common and specific epitopes among subunits. *Hybridoma and hybridomics*, 21(5), 365-374.

Perbandt, M., Guthöhrlein, E.W., Rypniewski, W., Idakieva, K., Stoeva, S., Voelter, W., Genov, N., & Betzel, C. (2003). The structure of a functional unit from the wall of a gastropod hemocyanin offers a possible mechanism for cooperativity. *Biochemistry* 42(21):6341–6346.

Peroutka III, R. J., Orcutt, S. J., Strickler, J. E., & Butt, T. R. (2011). SUMO fusion technology for enhanced protein expression and purification in prokaryotes and eukaryotes. *Heterologous Gene Expression in E. coli*, 15-30.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. & Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.*, 30(1):70-82.

Punjani, A., Rubinstein, J.L., Fleet, D.J., & Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, 14(3), 290-296.

Reyes, D., Salazar, L., Espinoza, E., Pereda, C., Castellón, E., Valdevenito, R., Huidobro, C., Becker, M.I., Lladser, A., López, M.N., & Salazar-Onfray, F. (2013). Tumour cell lysate-loaded dendritic cell vaccine induces biochemical and memory immune response in castration-resistant prostate cancer patients. *British journal of cancer*, 109(6), 1488-1497.

Rosano, G.L., & Ceccarelli, E.A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in microbiology*, 5, 172.

Rosenthal, P.B., & Henderson, R. (2003). Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology*, 333(4), 721–745.

Salazar, M.L., Jiménez, J.M., Villar, J., Rivera, M., Báez, M., Manubens, A., & Becker, M.I. (2019). N-Glycosylation of mollusk hemocyanins contributes to their structural stability and immunomodulatory properties in mammals. *Journal of Biological Chemistry*, 294(51), 19546-19564.

- Sheldrick, G.M. (2008). A short history of SHELX. *Acta Crystallographica Section A: Foundations of Crystallography*, 64(1), 112-122.
- Siddiqui, N.I., Idakieva, K., Demarsin, B., Doumanova, L., Compernelle, F., & Gielens, C. (2007). Involvement of glycan chains in the antigenicity of *Rapana thomasiana* hemocyanin. *Biochem Biophys Res Commun* 361:705–711.
- Sun, M., Azumaya, C.M., Tse, E., Bulkley, D.P., Harrington, M.B., Gilbert, G., Frost, A., Southworth, D., Verba, K.A., Cheng, Y., & Agard, D.A. (2021). Practical considerations for using K3 cameras in CDS mode for high-resolution and high-throughput single particle cryo-EM. *Journal of Structural Biology*, 213(3), 107745.
- Stoeva, S., Schütz, J., Gebauer, W., Hundsdörfer, T., Manz, C., Markl, J., & Voelter, W. (1999). Primary structure and unusual carbohydrate moiety of functional unit 2-c of keyhole limpet hemocyanin (KLH). *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, 1435(1-2), 94-109.
- Tanaka, Y., Kato, S., Stabrin, M., Raunse, S., Matsui, T., & Christos, G. (2019). Cryo-EM reveals the asymmetric assembly of squid hemocyanin. *IUCrJ* 6:426–437.
- Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics*, 54(1), 5-6.
- Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Vishal, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall III, W.B., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S., & Richardson, D.C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* 27: 293-315.
- Wuhrer, M., Robijn, M. L., Koeleman, C. A., Balog, C. I., Geyer, R., Deelder, A. M., & Hokke, C. H. (2004). A novel Gal (beta1-4) Gal (beta1-4) Fuc (alpha1-6)-core modification attached to the proximal N-acetylglucosamine of keyhole limpet haemocyanin (KLH) N-glycans. *Biochemical Journal*, 378(2), 625-632.
- Zhang, Q., Dai, X., Cong, Y., Zhang, J., Chen, D.H., Dougherty, M.T., Wang, J., Ludtke, S.J., Schmid, M.F., & Chiu, W. (2013). Cryo-EM structure of a molluscan hemocyanin suggests its allosteric mechanism. *Structure* 21(4): 604–613.

## 8. ANEXOS

### Anexo 8.1. Péptidos de FU silvestre purificada obtenidos por espectrometría de masas.

#### A) Péptidos identificados utilizando la secuencia CCHX.

Muestra	Secuencia	Cobertura	N° Péptidos	Abundancia
FU purificada en gel	CCHX	27%	19	208
	AIDAFNYDQYGYQYDHLHFHGLTVDK			53
	AIDAFNYDQYGYQYDHLHFHGLTVDKLDEK			28
	GLDYNTANCHIQDLR			7
	MSADVVFDCNAQGTGNFAGTFAILGGPLE MPWNFDR			1
	AVNGMQLDPNLEPPSVTFVPGKK			2
	HRGLDYNTANCHIQDLR			2
	EQDRVFLNFMLR			15
	DPEHGEESEFFR			13
	APRPQLFKDPEHGEESEFFR			4
	VLFPTEEGIAGPIR			1
	QFAIWQELQK			30
	QALLAFEQR			25
	LRPDSNYTIPIR			3
	KPLEPFNR			9
	EQDRVFLNFMLR			3
	LDEKLEK			1
	ANNPVLVTR			8
	APRPQLFK			1
	AENITTR			1
	VFLNFMLR			1

**B) Péptidos identificados utilizando la secuencia RVH1.**

Muestra	Secuencia	Cobertura	N° Péptidos	Abundancia
FU purificada en gel	RVH1	11%	25	125
	GLDYNTANCHIQDLR			7
	SLQADSSADGFQSLASFHAQPLCPYPEATK			2
	KPLQPFAQTSVTNPDPVTR			1
	LGLRLDSAADFLEVK			18
	HRGLDYNTANCHIQDLR			2
	LFVEQVENALLDHGSAVSPYWDWTEDFD HLPK			3
	DVETLTDEEVYLR			1
	FSILGGEYEMPFTFDRLYK			1
	QFAIWQELQK			30
	AIDAFNYDQYGYQYDHLRFHGLSVQELEEK			1
	DINTLTAEMKNLR			1
	DPEHGEESEFFYR			6
	QALLAFEQR			25
	KPLEPFNRAGNPVLVTR			1
	LRPDSNYTIPIR			3
	KPLEPFNR			9
	GKVSMSMLDYSAFDPIFFLHHANMDR			1
	RVQAGTGR			5
	EAGEFYLLGDEFEMPWEYDRLFK			1
	EHAHPQTVFDYEDLGHYHDDFK			1
	QMEVALSWEGAK			1
	VFAGFMLHGIQK			1
	QMEVALSWEGAK			1
	APRPQLFK			1
	AENITTR			1
	VFLNFMLR			1

**Anexo 8.2. Protocolo de tratamiento de datos de cryo-EM de CCH utilizando el software CryoSPARC.**

**1.- Imports - Import movies:**

**Raw pixel size (Å): 0.55**

**Accelerating voltage (kV): 300**

**Spherical aberration (mm): 0.01**

**Total exposure dose (e/Å<sup>2</sup>): 30**

**2.- Motion Correction - Patch motion correction (multi):**

**Output F-crop factor: 1/2**

**3.- CTF Estimation: Patch CTF Estimation (multi)**

**4.- Exposure Curation - Manually Curate Exposures**

\* Se filtran las partículas con una estimación de CTF de 8 Å. De 2.060 películas quedaron 1.793.

**5.- Particle Picking – Blob Picker:**

**Minimum particle diameter (Å): 250**

**Maximum particle diameter (Å): 400**

**Number of pics to process: 100**

**6.- Extract from micrographs:**

**Extract box size (pix): 600**

**Fourier crop to box size (pix): 200**

**7.-Particle Curation – 2D Classification:**

**Number of 2D classes: 50**

\* Selección de 13.928 partículas iniciales correspondientes a las 50 clases.

**8.- Particle Curation – Select 2D Classes**

\* Selección de 15 clases por selección visual (13 orientadas en vista lateral y 2 orientadas en vista frontal). Se seleccionan 1.604 partículas de acuerdo con las 15 clases elegidas.

**9.- Particle Picking - Template Picker:**

**Particle diameter (Å): 400**

\* De 1.793 películas se seleccionan 558.149 partículas.

**10.- Extract from micrographs**

**Extract box size (pix): 600**

**Fourier crop to box size (pix): 200**

**11.- Particle Curation – 2D classification:**

**Number of 2D classes: 150**

\* Selección de 432.482 partículas correspondientes a las 150 clases.

**12.- Particle Curation – Select 2D Classes**

\* Selección de 26 clases por selección visual correspondiendo con 213.006 partículas.

**13.- Particle Curation – 2D classification:**

**Number of 2D classes: 150**

\* Las 213.006 partículas seleccionadas anteriormente se clasifican en 150 clases.

**14.- Particle Curation – Select 2D Classes**

\* Selección de 32 clases por selección visual correspondiendo con 199.045 partículas.

**15.- 3D Reconstruction – Ab-Initio Reconstruction:**

**Number of Ab-initio classes: 3**

\* Se utilizan las partículas seleccionadas anteriormente para realizar reconstrucción de 3 modelos de volúmenes *Ab-initio*.

**16.- 3D Refinement – Heterogeneous Refinement**

**Symmetry: D5**

\* Se utilizan las mismas partículas seleccionadas anteriormente y el primer modelo de volumen reconstruido anteriormente.

**17.- Extract from micrographs**

**Extract box size (pix): 600**

\* Se extraen 130.952 partículas finales correspondientes al modelo refinado anteriormente.

**18.- 3D Refinement – Heterogeneous Refinement**

**Symmetry: D5**

\* Se utilizan las 130.952 partículas extraídas en el paso anterior y el modelo de volumen reconstruido en el paso 16. En este punto se calcula una resolución de mapa de cryo-EM de 3,18 Å calculada con el estándar dorado de correlación de la capa de Fourier (GSFSC), basándose en el criterio de corte de 0,143.

\* Los pasos seguidos se encuentran enumerados y se incluyen comentarios principalmente enfocados en los resultados obtenidos y en ciertos criterios de trabajo.