



Universidad de Concepción
Dirección de Postgrado
Facultad de Ciencias Biológicas
Programa de Magíster en Bioquímica y Bioinformática

**Determinación de la estructura del complejo Octarrelina
V.1- α Rep mediante cristalización por contradifusión y el
efecto de mutantes de prolina en su formación**



Tesis para optar al grado de Magíster en Bioquímica y
Bioinformática

FABIOLA ANDREA AEDO ALVARADO
CONCEPCIÓN-CHILE
2019

Profesor Guía: Dr. Maximiliano Figueroa Yévenes
Dpto. de Bioquímica y Biología Molecular,
Facultad de Ciencias Biológicas
Universidad de Concepción

Esta tesis ha sido realizada en el Departamento de Bioquímica y Biología Molecular de la Facultad de Ciencias Biológicas, Universidad de Concepción.

Profesor Guía

Dr. Maximiliano Figueroa Yévenes
Facultad de Ciencias Biológicas
Universidad de Concepción

Comisión Evaluadora:

Dr. José Martínez Oyanedel
Facultad de Ciencias Biológicas
Universidad de Concepción



Dr. Víctor Castro Fernandez
Facultad de Ciencias
Universidad de Chile

Director de Programa

Dr. Amparo Uribe Pérez
Facultad de Ciencias Biológicas
Universidad de Concepción

AGRADECIMIENTOS

Sin lugar a dudas, este nuevo paso que decidí dar en el mundo de las ciencias no hubiera sido posible sin la ayuda de mucha gente. A Dios es mi primer agradecimiento por permitirme continuar creciendo cada día. A mis padres quienes han estado para mí como un gran apoyo en todo este proceso. Sin lugar a dudas este grado es tanto suyo como mío. A mis hermanas, quienes a pesar de no comprender mucho aún la vida que decidí seguir, siempre me han levantado y alentado a continuar incesantemente. A mis padrinos quienes siempre que lo he necesitado, me han recibido y aconsejado durante tanto tiempo. A la familia que escogí hace tantos años: Josefa, Natalia, Daniela, Lilybeth y Camila, gracias por las palabras y las risas en mis pocos momentos libres y por no dejar de acompañarme en todo este tiempo.

Siempre agradeceré al laboratorio de biofísica por recibirme: a los Dr. Marta Bunster y José Martínez Oyanedel quienes siempre tuvieron una gran disposición conmigo, y a mi profesor guía el Dr. Maximiliano Figueroa por siempre tener las puertas abiertas para aclarar cualquier duda, por ayudarme con cualquier problema y sobre todo por el apoyo entregado durante la realización de mi tesis. A mis compañeros Ambbar, Jorge, Catalina, Erwin, Diego y Pablo quienes fueron grandes compañeros de laboratorio y grandes apoyos durante mi trabajo.

Mi trabajo no se habría podido completar sin la ayuda del Laboratorio de enzimología, dirigido por la Dr. Amparo Uribe y por sus miembros. A los laboratorios de Cristales Líquidos de la Universidad de Concepción y de biofísica, del centro de Bioinformática y Biología Integrativa de la Universidad Andrés Bello por permitirme realizar experimentos importantes para concluir esta tesis. Finalmente a los profesores, alumnos y personal del departamento de Bioquímica y Biología Molecular por el apoyo directo o indirecto que permiten que pueda presentar este trabajo de tesis.

TABLA DE CONTENIDOS

TABLA DE CONTENIDOS	IV
INDICE DE FIGURAS	VI
INDICE DE TABLAS	VIII
RESUMEN.....	IX
ABSTRACT.....	XI
INTRODUCCIÓN	1
Diseño de Proteínas.....	1
Diseño <i>de novo</i> de proteínas.....	3
Barril TIM	6
Octarrelinas	8
Evolución dirigida.....	12
Octarrelina V.1	13
Octarrelina V.1 Δ 4P y Octarrelina V 4P	17
Propiedades de prolina en la estructura de proteínas	19
Cristalización de proteínas	21
Diagrama de fases de cristalización	22
Métodos de cristalización.....	24
Difusión de interfase libre liquido-liquido	27
Cristalización mediante contra-difusión	30
Ayudantes de cristalización	31
Artificial alpha-helicoidal repeat proteins (α Rep)	34
HIPOTESIS Y OBJETIVO DEL TRABAJO.....	36
MATERIALES	38
METODOLOGÍA	38
Transformación bacteriana mediante choque térmico.....	38
Producción y purificación Octarrelina V.1	39
Producción y purificación α Rep.....	41
Producción complejo Octarrelina V.1- α Rep.....	43
Cristalización de complejo Octarrelina V.1- α Rep	44
Producción y purificación de proteínas Octarrelina V.1 Δ 4P y Octarrelina V4P	45

Identificación de mutantes Octarrelina V.1 Δ 4P y Octarrelina V4P por espectrometría de masas	48
Producción complejos Octarrelinas V4P y V.1 Δ 4P con α Rep	49
Caracterización de las interacciones de unión entre α Rep y Octarrelinas mediante micro calorimetría de titulación isotérmica (ITC).....	49
RESULTADOS	51
Resultados Objetivo I: Producción y purificación de Octarrelina V.1 y α Rep.	51
Purificación Octarrelina V.1	51
Purificación α Rep.....	53
Resultados Objetivo II: Purificación y cristalización del complejo Octarrelina V.1- α Rep	56
Purificación complejo Octarrelina V.1- α Rep.....	56
Cristalización complejo Octarrelina V.1- α Rep.....	57
Resultados Objetivo III: Determinación de estructura cristalina del complejo Octarrelina V.1- α Rep	60
Data cristalográfica y determinación de estructura del complejo	60
Resultados Objetivo IV: Producción y purificación de Octarrelinas mutantes de prolina.....	65
Purificación mutantes Octarrelina V.1 Δ 4P y Octarrelina V4P	65
Análisis de Octarrelina V4P y Octarrelina V.1 Δ 4P	70
Resultados Objetivo V: Estudios de interacción entre α Rep y Octarrelinas mutantes.....	73
Análisis complejos Octarrelinas con α Rep.....	73
Análisis interacciones de unión entre α Rep y Octarrelinas.....	80
DISCUSIÓN	86
Purificación de proteínas.....	86
Cristalización por contradifusión.....	87
Octarrelinas mutantes	89
Complejos Octarrelinas mutantes - α Rep.....	90
CONCLUSIONES	93
BIBLIOGRAFÍA.....	95

INDICE DE FIGURAS

Figura 1: Etapas del diseño de novo de proteínas.....	4
Figura 2: Representación de la estructura característica del plegamiento barril TIM ($\beta\alpha$) ₈	7
Figura 3: Esquema del diseño de novo realizado para la construcción de Octarrelina V.	10
Figura 4: Octarrelina V.1 producida por evolución dirigida.....	13
Figura 5: Alineamiento de secuencias entre Octarrelina V y Octarrelina V.1	14
Figura 6: Estructura por rayos x de Octarrelina V.1 (PDB: 4ZV6).	16
Figura 7: Modificaciones puntuales realizadas para generar las mutantes de Octarrelina.	18
Figura 8: Diagrama de fase para la cristalización de macromoléculas.	23
Figura 9: Esquema de diferentes métodos de cristalización.....	25
Figura 10: Diagrama de fases de la cristalización por difusión de vapor.....	26
Figura 11: Difusión de interfase libre liquido-liquido.....	28
Figura 12: Diagramas de Venn representando diferentes cristalizaciones.	29
Figura 13: Representación de las configuraciones más usadas en contra-difusión.	30
Figura 14: Modelo del motivo α Rep.....	34
Figura 15: Esquema de re-plegamiento de proteínas mediante cromatografía de intercambio iónico.....	47
Figura 16: Análisis por SDS-Page 15% de fracciones de producción y tratamiento por calor de Octarrelina V.1.....	51
Figura 17: Cromatografía de intercambio iónico fracción soluble Octarrelina V.1.....	52
Figura 18: Cromatografía de exclusión molecular Octarrelina V.1.....	53
Figura 19: Cromatografía IMAC α Rep.....	54
Figura 20: Cromatografía de exclusión molecular α Rep.....	55
Figura 21: Cromatografía de exclusión molecular complejo Octarrelina V.1- α Rep.....	56
Figura 22: Análisis SDS-PAGE al 15% del complejo Octarrelina V.1- α Rep concentrado.....	57
Figura 23: Cristalización por contradifusión del complejo Octarrelina V.1- α Rep en 26% PEG 400.....	58
Figura 24: Cristalización por contradifusión de complejo Octarrelina V.1- α Rep en 35% PEG 400.....	59
Figura 25: Sección de mapa de densidad electrónica del complejo Octarrelina V.1- α Rep.....	62
Figura 26: Estructura complejo Octarrelina V.1- α Rep.....	63
Figura 27 : Comparación estructura complejo Octarrelina V.1- α Rep obtenido mediante cristalización por difusión por vapor y contra-difusión.....	64
Figura 28: Cromatografía de intercambio iónico Octarrelina V.1 Δ 4P.....	66
Figura 29: Cromatografía de exclusión molecular Octarrelina V.1 Δ 4P.....	67
Figura 30: Cromatografía de intercambio iónico Octarrelina V4P.....	68

Figura 31: Cromatografía de exclusión molecular Octarrelina iiV4P	69
Figura 32: Muestras proteínas de trabajo purificadas	71
Figura 33: Péptidos obtenidos en MALDI-TOF MS/MS identificados en las secuencias de Octarrelinas mutantes.	72
Figura 34: Cromatografía complejo Octarrelina V.1 Δ 4P- α Rep.....	74
Figura 35: Cromatografía complejo Octarrelina V4P- α Rep.....	75
Figura 36: Comparación perfiles de elución proteínas del complejo Octarrelina V.1- α Rep.....	77
Figura 37: Comparación perfiles de elución proteínas del complejo Octarrelina V4P- α Rep.	78
Figura 38: Comparación perfiles de elución proteínas del complejo Octarrelina V.1 Δ 4P- α Rep.	79
Figura 39: Calorimetría de titulación de α Rep con Octarrelina V.1.	81
Figura 40: <i>Screening</i> de calorimetría de titulación de Octarrelina V.1 a 25 μ M con diferentes concentraciones molares de α Rep	82
Figura 41: Calorimetría de titulación de α Rep con Octarrelina V.1 Δ 4P.....	83
Figura 42: Comparación de calorimetrías de titulación para los complejos Octarrelina V.1 y V.1 Δ 4P con α Rep	84
Figura 43: Estructura de dímero Octarrelina V.1 en complejo con ayudante de cristalización NB1	91



INDICE DE TABLAS

Tabla 1: Resumen de purificación proteínas Octarrelina V.1 y α Rep	55
Tabla 2: Resumen de data colectada por rayos X y estadísticas de refinamiento.	61
Tabla 3: Resumen de purificación proteínas Octarrelina V.1 Δ 4P y Octarrelina V4P	70



RESUMEN

Octarrelinas son proteínas artificiales de más de 200 aminoácidos, diseñadas *de novo* con el fin de adoptar una estructura repetitiva $(\beta\alpha)_8$ conocida como TIM-barrel. Gracias a la técnica de evolución dirigida, una versión de la proteína diseñada *in silico*, Octarrelina V, incremento su solubilidad y estabilidad sin ver modificada su estructura. Esta nueva versión de la proteína artificial, Octarrelina V.1, presenta dentro de las 16 mutaciones introducidas, 4 residuos de prolina en su cadena aminoacídica. La estructura parcialmente resuelta mediante cristalización por difusión de vapor muestra que la proteína diseñada no se plegaba de acuerdo a lo esperado. Sin embargo, la presencia inesperada de los residuos de prolina, despertó el interés para estudiar este residuo y la relación que puede tener con las características finales obtenidas en Octarrelina V.1. Por ello la obtención de un modelo más detallado de Octarrelina V.1 en conjunto con experimentos en mutante de prolinas puede entregar información sobre la influencia de este residuo en las características obtenidas. Para lograr esto, se buscó obtener una estructura más completa del Octarrelina V.1 utilizando el ayudante de cristalización α Rep mediante el método de contradifusión. Además, mutantes de Octarrelina V.1 y V con ausencia y presencia de prolinas, respectivamente, se estudiaron mediante su interacción con α Rep para obtener una mayor información de este residuo.

Las proteínas Octarrelina V.1 y α Rep lograron ser purificadas con altos rendimientos y grados de pureza aceptables para la posterior formación de complejos *in vitro*. La correcta purificación del complejo permitió establecer experimentos de cristalización por contradifusión en capilares. Los cristales obtenidos por esta técnica fueron difractados por rayos X y la data obtenida permitió resolver la estructura a través de reemplazo molecular. Gracias a estos datos, la estructura alcanzada reveló un 90% de los aminoácidos de la secuencia de Octarrelina V.1, aumentando lo anteriormente resuelto mediante cristalización por difusión de vapor.

La producción y purificación de las mutantes de prolina alcanzó bajos rendimientos de producción, procesos en los cuales puede verse involucrado el residuo de prolina. Experimentos de formación de complejos con α Rep para ambas mutantes presentaron una baja unión con este ayudante y experimentos de microcalorimetría de titulación isotérmica que únicamente fueron posibles de realizar para Octarrelina V.1 Δ 4P confirman este comportamiento.

La nueva estructura obtenida puede ser utilizada para estudiar las implicancias de ciertos aminoácidos en el plegamiento de una proteína artificial. Aun cuando Octarrelina V.1 no presenta la estructura terciaria predicha, es un modelo único, sin presión evolutiva, capaz de usarse para caracterizar plegamiento y estructura.



ABSTRACT

Octarrelins are artificial proteins of more than 200 amino acids, designed *de novo* in order to adopt a repetitive structure $(\beta\alpha)_8$ known as TIM-barrel. Thanks to the directed evolution technique, a version of the protein designed *in silico*, Octarrelin V, increases its solubility and stability without seeing its structure modified. This new version of the artificial protein, Octarrelin V.1, presents within the 16 introduced mutations, 4 proline residues in its amino acid chain. The partially resolved structure by vapor diffusion crystallization shows that the designed protein did not fold as expected. However, the unexpected presence of proline residues aroused interest in studying this residue and the relationship it may have with the final characteristics obtained in Octarrelin V.1. Therefore, obtaining a more detailed model of Octarrelin V.1 together with experiments in proline mutants can provide information on the influence of this residue on the characteristics obtained. To achieve this, it was sought to obtain a more complete structure of the Octarrelin V.1 using the α Rep crystallization assistant by means of the counter-diffusion method. In addition, Octarrelin V.1 and V mutants with absence and presence of prolines, respectively, were studied through their interaction with α Rep to obtain more information on this residue.

The Octarrelin V.1 and α Rep proteins were purified with high yields and acceptable degrees of purity for subsequent *in vitro* complex formation. The correct purification of the complex allowed to establish experiments of crystallization by contradiffusion in capillaries. The crystals obtained by this technique were diffracted by X-rays and the data obtained allowed the structure to be resolved through molecular replacement. Thanks to these data, the structure reached revealed 90% of the amino acids of the Octarrelin V.1 sequence, increasing the above solved by vapor diffusion crystallization.

The production and purification of proline mutants achieved low production yields, processes in which proline residue may be involved. Complex formation

experiments with α Rep for both mutants showed low binding with this helper and isothermal titration microcalorimetry experiments that were only possible for Octarrelin V.1 Δ 4P confirm this behavior.

The new structure obtained can be used to study the implications of certain amino acids in the folding of an artificial protein. Although Octarrelin V.1 does not have the predicted tertiary structure, it is a unique model, without evolutionary pressure, capable of being used to characterize folding and structure.



INTRODUCCIÓN

Diseño de Proteínas

Las proteínas son cadenas lineales de aminoácidos de las que organismos y células dependen para mediar procesos fundamentales para la vida como la eliminación de patógenos, señalización, catálisis enzimática o reconocimiento molecular. La obtención de una imagen molecular de las conexiones entre secuencia, estructura y función de las proteínas es, por lo tanto, de sumo interés para comprender las bases moleculares de estos procesos biológicos. Debido a que la organización de los aminoácidos en la cadena polipeptídica determina el plegamiento de una proteína y, con ello, su función biológica, la determinación de las estructuras de proteínas naturales ha permitido un avance en la comprensión de la interacción entre la secuencia, estructura y función. Sin embargo, a pesar del gran conocimiento alcanzado gracias a la determinación estructural de diferentes proteínas, en la actualidad aún no es comprendido del todo como las interacciones del tipo no covalente (fuerzas Van der Waals, puentes de hidrogeno, interacciones hidrofóbicas) pueden conducir al plegamiento de una proteína hasta una estructura nativa relevante biológicamente. El diseño de proteínas provee una gran oportunidad para comprender la relación entre el plegamiento de las proteínas y su estructura.

El diseño de proteínas es un área que busca estudiar de manera experimental como una estructura proteica es alcanzada a través de una secuencia aminoacídica específica (Koga et al., 2012). Además del interés científico, el diseño de proteínas también posee un interés a nivel ingenieril: el diseño de estructuras simples puede conducir en el futuro al diseño de nuevas macromoléculas que puedan resolver problemas químicos o bioquímicos importantes (Carey, 1996; Floudas et al., 2006).

La naturaleza compleja de las proteínas hace que su diseño no sea una tarea trivial. Estas macromoléculas pueden alcanzar largos desde decenas a miles de residuos, haciendo que el número de secuencias posibles sea exponencialmente grande. Por ejemplo, para el diseño de una proteína con 100 residuos variables, existen 20^{100} posibles secuencias. Además de lo anterior, también deben ser consideradas las posibles conformaciones de cadena lateral, lo que incrementa la complejidad en la búsqueda de las secuencias consistentes con una estructura deseada (Muñoz, 2008). Sin embargo, el desarrollo de funciones que describen las fuerzas que determinan la estructura de la proteína y de métodos de optimización que permiten la evaluación de un gran número de secuencias, permitió el progreso del diseño de proteínas (Pokala & Handel, 2001).

El trabajo en esta área comenzó con éxitos limitados, produciendo polipeptidos pequeños con plegamiento compacto aunque con conformaciones desordenadas (Shakhnovich, 1998). Sin embargo, los esfuerzos puestos en diseñar plegamientos con características puntuales han comenzado a tener mayor éxito con el diseño de motivos o proteínas (Correia et al., 2011; Doyle et al., 2015; Farid et al., 2013; King et al., 2012), lográndose incluso avances remarcables como top7 (Kuhlman et al., 2003) o barriles α -helices (Chino et al., 2015; Thomson et al., 2014).

Diferentes vías de diseño de proteínas se han enfatizado a fin de considerar las complejidades antes mencionadas. El rediseño de proteínas existentes busca introducir nuevas propiedades o estudiar el efecto de variaciones en la secuencia y puede ser útil como punto de comienzo para el diseño. Sin embargo permite un trabajo restringido al utilizar proteínas ya conocidas, siendo un enfoque que “se encuentra en el límite entre lo conocido y lo inmediatamente accesible al mundo proteico” (Woolfson et al., 2015). Una mayor exploración de parte del espectro que no es posible ver durante la evolución natural es posible mediante la vía conocida como diseño de proteínas *de novo*.

Diseño *de novo* de proteínas

Se conoce como diseño *de novo* al diseño de proteínas desde sus principios, lo que involucra la determinación previa de la estructura, secuencia y función. Su búsqueda implica el encontrar secuencias compatibles con un cierto plegamiento deseado. Se comienza con un cierto conjunto de supuestos y principios con los cuales se trabaja para diseñar proteínas que se plieguen correctamente. El éxito o fracaso de los diseños resultantes se convierte en conocimiento de importancia para evaluar nuestra comprensión acerca de la estabilidad y el plegamiento de las proteínas, dada la posibilidad que este método ofrece de poder explorar un mayor espacio de secuencias proteicas que las que naturalmente se han optimizado para los procesos biológicos, superando así las limitaciones evolutivas.

Debido a la complejidad de considerar los múltiples factores que contribuyen en las interacciones físicas y químicas que estabilizan las estructuras plegadas de una proteína, el desarrollo del diseño *de novo* se ha encontrado estrechamente ligado al diseño computacional. El desarrollo de métodos computacionales se ha llevado a cabo teniendo en cuenta estas dificultades.

Los primeros esfuerzos realizados el pasado siglo se tornaron infructuosos al no lograrse el pliegue diseñado o con la obtención de estructuras terciarias diferentes a las esperadas (Betz et al., 1993; Lovejoy et al., 1993). Intentos posteriores alcanzaron éxitos en el diseño de núcleos proteicos, lográndose la construcción de núcleos hidrofóbicos con buen plegamiento y estructura estable (Desjarlais & Handel, 1995; Hellinga & Richards, 1994). Progresos en el área metodológica y de aplicación (Koga et al., 2012; Lippow & Tidor, 2007; Pantazes et al., 2011) han permitido la construcción de proteínas de mayor complejidad. Resultados positivos en los últimos años se han alcanzado en el diseño de proteínas con funciones y propiedades deseadas en aplicaciones biotecnológicas y terapéuticas. Entre ellas encontramos avances en el diseño de péptidos y proteínas para aplicaciones terapéuticas como biomarcadores, inhibidores,

andamios de epítopes o péptidos cortos y anticuerpos terapéuticos. También un amplio desarrollo se ha llevado a cabo en el desarrollo de enzimas *de novo* con sitios activos para reacciones de múltiples etapas o enzimas con actividad mejorada. Proteínas y péptidos autoensamblados también han sido realizados *de novo* con miras a aplicaciones en biomateriales (Khoury et al., 2014).

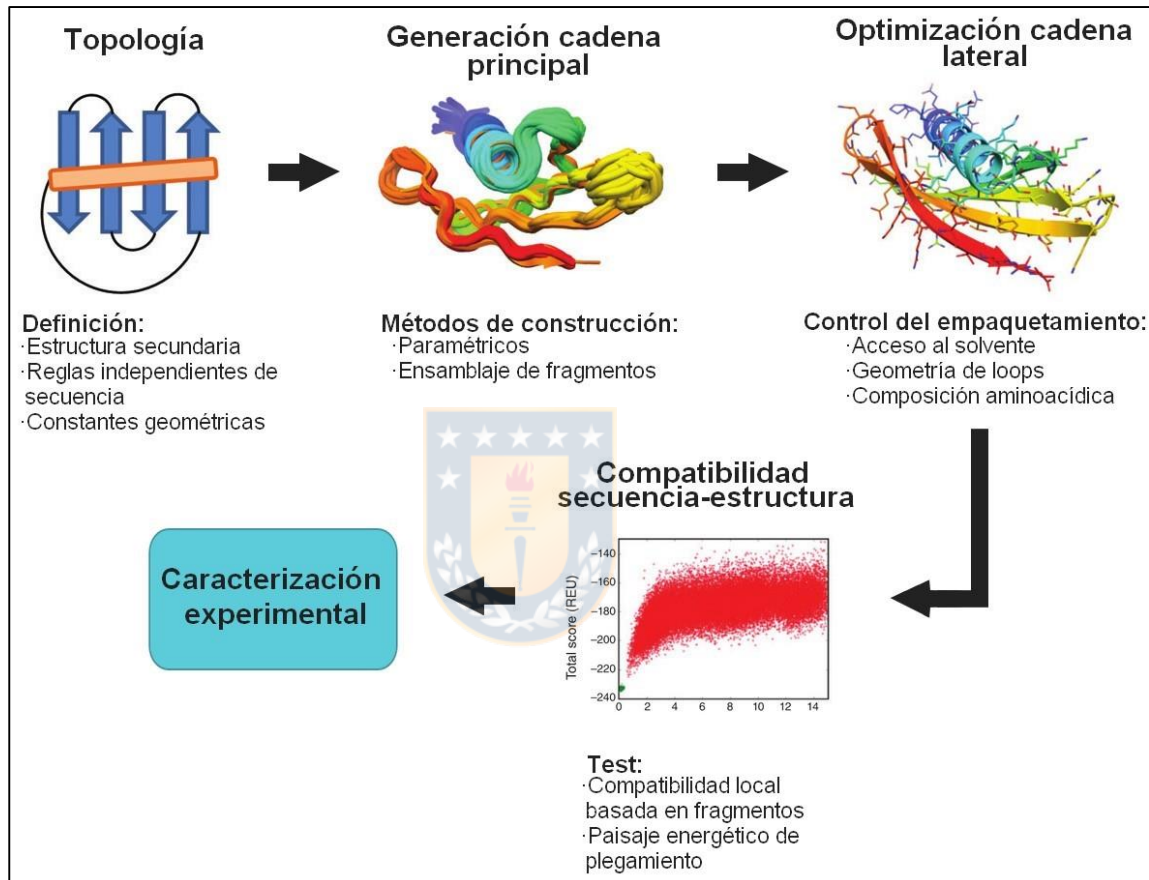


Figura 1: Etapas del diseño *de novo* de proteínas

La definición de la topología corresponde al primer paso, donde características importantes para el diseño deben ser consideradas. De acuerdo al tamaño y plegamiento es elegido el método más adecuado para generar la cadena principal, produciéndose miles de modelos que son filtrados para posteriormente ejecutar los cálculos de diseño de secuencia completa. Los pares de secuencia de la cadena principal mejor clasificados son evaluados por su compatibilidad secuencia-estructura, seleccionándose los que exhiben paisajes de energía en forma de embudo para a continuación, caracterizarse experimentalmente (Modificado de Marcos & Silva, 2018).

En general, el diseño *de novo* de proteínas comienza con la definición de una estructura o topología objetivo para, posteriormente, generar la cadena principal de la proteína que satisfaga los requerimientos del diseñador (elementos de estructura secundaria, conectividad, geometría) (Figura 1). Los principales métodos de construcción de la cadena principal, como los métodos paramétricos o de ensamblaje de fragmentos, se basan en el conocimiento de la información experimental obtenida a partir de estructuras tridimensionales de proteínas conocidas. El siguiente paso involucra la selección de la secuencia de aminoácidos, donde se busca la identidad de los residuos que minimicen la función de energía de la cadena principal realizada previamente. Gran parte de los algoritmos de diseños se basan en funciones de energía donde los pares de conformaciones discretas de los aminoácidos (también conocidos como rotámeros) son calculados con una serie de términos de puntajes descompuestos por pares, lo que a su vez permite calcular previamente las energías del par de rotámeros, mejorando en gran medida la eficiencia del diseño de secuencia. Además de las funciones de energía, energías efectivas que cuantifican las tendencias estructurales pueden incluirse para encontrar soluciones óptimas. Esto puede realizarse preseleccionándose o restringiéndose las identidades de los aminoácidos permitidos en cada posición de la cadena principal o mediante la adición de sesgos a la función de energía para favorecer o desfavorecer ciertas identidades de aminoácidos. Como último paso es necesario la evaluación de la compatibilidad entre la secuencia diseñada y la estructura objetivo. La proteína diseñada debe plegarse en una estructura tridimensional única definida, evitándose o disminuyéndose otras conformaciones diferentes a la objetivo. Para lograr esta especificidad estructural, la estructura diseñada *de novo* debe corresponder a un estado fundamental energético que posea una brecha energética que la separe de otros plegamientos competidores. La mayoría de los estudios de diseños *de novo* evalúan las secuencias diseñadas por su compatibilidad con su estructura, tanto a nivel local como no local, considerando los estados alternativos (Marcos & Silva, 2018; Muñoz, 2008).

Grandes esfuerzos en proteínas *de novo* han estado dirigidos en la construcción de cadenas principales con simetría interna. Esto ayuda a reducir el tamaño de secuencia en la que se debe buscar, además de permitir el uso de un enfoque donde una sola unidad idealizada puede ser repetida varias veces para construir proteínas de mayor tamaño (Doyle et al., 2015; Rämisch et al., 2014). La simetría interna ha sido particularmente observada en estructuras cerradas como toroides α -hélices y barriles TIM (Huang, Boyken, & Baker, 2016). El trabajo con este último plegamiento es de gran interés debido a su amplia presencia en la naturaleza, cumpliendo diversas funcionalidades (Nagano, Orengo, & Thornton, 2002).

Barril TIM

El barril TIM o barril $(\beta\alpha)_8$ es uno de los plegamientos polipeptídicos más comunes en biología. Descrito por primera vez en la estructura de la Triosa Fosfato Isomerasa (Banner et al., 1975), y viniendo de ahí su actual nombre, este plegamiento se presenta como una gran clase enzimática descrita en las bases de datos SCOP y CATH. Las funciones de las enzimas que adoptan este plegamiento incluyen 5 de las principales clasificaciones en los números "Enzyme Commission": Oxidoreductasas, transferasas, liasas, hidrolasas e isomerasas (Nagano, Hutchinson, & Thornton, 1999).

El barril TIM clásico está compuesto por 8 unidades o motivos $\beta\alpha$, ordenados de manera que 8 hebras β paralelas en el interior de la estructura quedan rodeadas por 8 α -hélices en el exterior (Figura 2). Los residuos relacionados con el sitio activo se encuentran localizados en loops de los extremos carboxilo de las hebras β , mientras que los loops ubicados en el otro extremo conectados las hélices con las hebras β poseen una importancia estructural para mantener el plegamiento (Brändén, 1991). Estos elementos estructurales permiten una simetría cuádruple dentro del plegamiento (Nagano, Orengo, & Thornton, 2002).

La localización conservada del sitio activo en enzimas barril $(\beta\alpha)_8$ y similitudes estructurales y de secuencia reportadas (Rice, Goldman, & Steitz, 1990; Scheerlinck et al., 1992; Wilmanns et al., 1991), sugieren que una gran porción de los barriles TIM se encuentran relacionadas por un ancestro común (Nagano, Orengo, & Thornton, 2002).

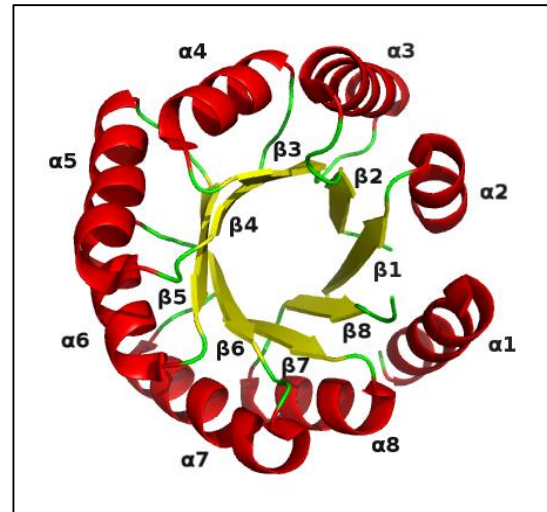
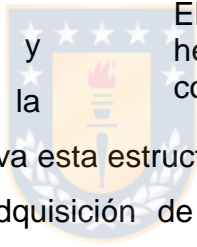


Figura 2: Representación de la estructura característica del plegamiento barril TIM $(\beta\alpha)_8$.

Sus características y múltiples funciones en la



naturaleza han hecho atractiva esta estructura para experimentos de ingeniería, lográndose con éxito la adquisición de nuevas actividades enzimáticas o modificación de especificidades de sustratos (Cheon et al., 2004; Jurgens et al., 2000; Schmidt et al., 2003). Por largo tiempo se ha intentado igualmente el diseño de barriles TIM. Varias estructuras fueron obtenidas utilizando partes de proteínas $(\beta\alpha)_8$ naturales o mediante la utilización de elementos de secuencia repetitivos (Eisenbeis et al., 2012; Fortenberry et al., 2011; Höcker et al., 2009).

El barril se encuentra formado por hebras β (amarillo) y α -helices (rojo) conectadas mediante loops (verde).

El diseño *de novo* de este plegamiento se ha llevado a cabo teniendo especial consideración en las características geométricas y topológicas definidas para este plegamiento (Koga et al., 2012; Lasters et al., 1988; Murzin et al., 1994) y utilizando herramientas computacionales como ROSETTA. En la actualidad, resultados prometedores han sido alcanzados como el diseño en base a una estructura consenso de barril TIM realizada por Nagarajan, Deka, & Rao, (2015); la llevada a cabo utilizando como base a una unidad simétrica $\beta\alpha\beta\alpha$ de Huang et al., 2016 o la producida mediante el uso de técnicas iterativas para la optimización

de los rotámeros (ORBIT) que ha formado parte de las últimas variantes de una serie de proteínas diseñadas para adquirir este plegamiento conocidas como Octarrelinas.

Octarrelinas

El diseño de proteínas ha abierto un mundo de posibilidades, no solo para la obtención de proteínas con interés práctico, sino también para mejorar nuestro entendimiento en la bioquímica de proteínas. Sin embargo, el diseño de proteínas bien plegadas y solubles ha sido difícil debido a su complejidad. No ha sido posible sino hasta hace un poco más de una década lograr avances en la obtención de proteínas bien estructuradas con un solo dominio que excedan los 200 aminoácidos, como los reportados en las variantes Octarrelina V y VI, de un proyecto de larga data que busca la construcción de proteínas con el plegamiento de barril TIM.

La serie de proteínas octarrelina nace en la década de los 90, con la construcción de una primera versión del péptido artificial mediante la repetición de una unidad peptídica vuelta/hebra β /vuelta/ α -hélice ocho veces. La obtención de la secuencia que construiría esta unidad estructural se realizó mediante la búsqueda de características comunes en proteínas con plegamiento $(\beta\alpha)_8$. Esta primera selección de aminoácidos utilizando criterios como la frecuencia de los residuos, reglas de empaquetamiento α/β o flexibilidad espacial dentro de la estructura secundaria; fue ajustada utilizando el algoritmo de Garnier, también conocido como algoritmo de GOR, para dar finalmente con los aminoácidos escogidos para la secuencia. Se pudo observar en Octarrelina I cierta compactación en su estructura y la capacidad de replegarse de manera cooperativa, con dos estados de transición dependiendo del pH y de la composición del tampón (Goraj et al., 1990). Análisis posteriores indicaron un contenido de estructura secundaria estable mediante dicroísmo circular (CD), con un mayor porcentaje de hebras β que el diseñado. También se muestra que la denaturación con urea es progresiva con el aumento en la concentración de esta

agente, aunque no existe una completa reversibilidad durante la denaturación y renaturación térmica. Además, las tirosinas se encuentran expuestas tanto en la conformación similar a la nativa como en la denaturada. A pesar de que la metodología utilizada en esta versión incremento la posibilidad es establecer residuos relevantes para la estructura secundaria, Octarrelina I presentó finalmente una baja estabilidad en la estructura terciaria similar a la nativa (Beauregard et al., 1991).

Una segunda generación de Octarrelinas, compuesta por Octarrelina II y III, se llevó acabo basada en la secuencia inicial de Octarrelina I, aunque integrando nuevos criterios de diseño. Ambas octarrelinas difieren solo a nivel de hebras β , basando su diseño en características presentes en proteínas TIM naturales, aunque poniendo especial énfasis en la distribución de carga, volumen de los residuos e hidrofobicidad. El análisis tomó en cuenta ideas postuladas en la época como el modelo de “capa” propuesto Lesk, Brändén, & Chothia, 1989. Este proceso reveló que el volumen e hidrofobicidad de los aminoácidos difiere de acuerdo a si los residuos que componen la hebra β apuntan hacia el centro del barril o hacia las α -hélices, siendo estos últimos de mayor tamaño que los anteriores. Manteniendo el concepto de unidades repetitivas y con las anteriores ideas en mente, se diseñaron Octarrelina II, la cual mantiene una simetría óctuple con un solo diseño de secuencia para las hebras β , y Octarrelina III, la cual presenta una simetría cuádruple con un diseño diferente para cada una de las hebras β presente en la unidad. Análisis efectuados como la espectroscopia de infrarrojos con transformada de Fourier (FTIR) o dicroísmo circular (CD) indican un gran contenido de estructura secundaria. En tanto, el experimento de electroforesis en condiciones no denaturantes indican la presencia en la mezcla de formas monoméricas y diméricas para ambas polipéptidos. Experimentos utilizando fluorescencia como ANS o la denaturación con Urea revelan que el estado de Octarrelina III es más organizado y estable que el de Octarrelina II, indicándose así la importancia de la simetría cuádruple de este plegamiento

(Houbrechts et al., 1995), tal como ha sido descrito con posterioridad (Murzin, Lesk, & Chothia, 1994; Nagano, Hutchinson, & Thornton, 1999).

Un cambio radical se presentó en el diseño de la siguiente variante de esta serie. Gracias a los avances en los recursos computacionales, el diseño de una proteína con el plegamiento $(\beta\alpha)_8$ utilizando algoritmos dio origen a Octarrelina V. Para su construcción, fue definido en primer lugar un “backbone” con el uso de parámetros geométricos para que adquiriera una estructura TIM barrel y, posteriormente, la conformación de la cadena lateral (Figura 3).

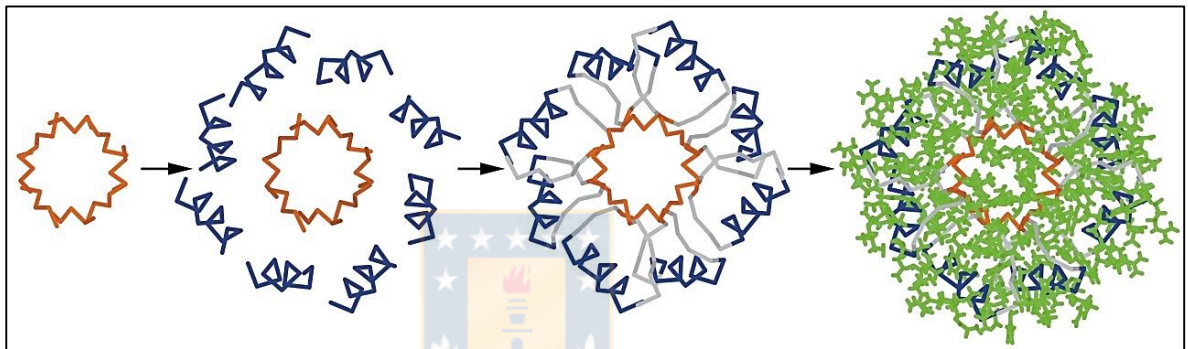


Figura 3: Esquema del diseño *de novo* realizado para la construcción de Octarrelina V.

La construcción de las hebras β se realizó utilizando parámetros geométricos idealizados, añadiendo posteriormente las hélices α y los motivos de conexión. Utilizando el programa ORBIT fue posible definir los aminoácidos compatibles con la cadena predefinida. (Extraído de Offredi et al., 2003)

En el primer paso, se incluyeron parámetros geométricos como distancias y ángulos entre estructuras secundarias para describir la topología del barril con una simetría cuádruple. En tanto, para el siguiente paso, la secuencia de la cadena lateral y su conformación fueron elegidas con ayuda de un algoritmo de selección automatizado. Los residuos se clasificaron según el espacio que ocupaban (núcleo, superficie o posiciones limitantes). Cada una de estas regiones se diseñó por separado utilizando diferentes bibliotecas de rotámeros, seleccionando con especial atención los aminoácidos. Los residuos cisteína y prolina no fueron considerados para el diseño, a fin de evitar la formación de

puentes disulfuro y la isomerización *cis-trans*. El programa ORBIT, mediante la utilización del teorema de eliminación de punto muerto (DEE) (Desmet et al., 1992) identificó los rotámeros que no podían pertenecer a la conformación global de mínima energía. Finalmente, diferentes funciones de energía potencial se utilizaron para puntuar las secuencias obtenidas para dar con la secuencia seleccionada.

El trabajo logró la obtención de una proteína estable y bien empaquetada en solución. Experimentos de CD en UV lejana indican un alto porcentaje de estructuras secundarias, con un 50% de α -hélices. Desplegamiento térmico monitoreado por UV cercana revela una transición cooperativa con una T_m aparente de 65°C. Espectro de resonancia magnética nuclear (NMR) confirmó igualmente la presencia de estructura terciaria en la proteína diseñada. Mediante DLS se confirmó que la máxima concentración posible para la proteína en solución fue de 1mg/mL. Si bien esto fue útil para los diferentes experimentos realizados en Octarrelina V, la baja solubilidad sumado a su presencia en cuerpos de inclusión durante su purificación desde *E. Coli*, no consigue lograr los requerimientos suficientes para experiencias de difracción en rayos X (Offredi et al., 2003).

La última versión actualmente realizada de esta serie corresponde a Octarrelina VI. Esta versión fue realizada utilizando el software Rosetta para el diseño, utilizando como base de las características geométricas observadas en los barriles TIM las coordenadas del “backbone” diseñadas en Octarrelina V. Esta proteína mostró evidencia de tanto de elementos de estructura secundaria como una estable estructura terciaria además de resistencia al desplegamiento mediante calor. Sin embargo, aún se presenta una baja solubilidad en esta versión además de un desplegamiento no cooperativo (Figuroa et al., 2013).

Estos constantes resultados en cuanto a la baja solubilidad, con la proteína presentándose en todas las versiones en cuerpos de inclusión, hizo patente la necesidad de utilizar herramientas de ingeniería de proteínas, como la evolución

dirigida, para mejorar las características de solubilidad y estabilidad que permitieran una determinación más precisa de las proteínas artificiales obtenidas.

Evolución dirigida

La gran e impresionante variedad de características y funciones que las proteínas naturales exhiben, ha creado un gran interés en buscar la forma de utilizar los procesos de diseño evolutivos para conseguir objetivos definidos. La evolución realizada en laboratorios lleva décadas siendo perfeccionada (Cobb et al., 2013), aunque en general en estos experimentos el gen que codifica para una macromolécula de interés, como una proteína, es parcial o completamente sometida a mutaciones aleatorias. La identificación de miembros de este conjunto de mutantes, mediante métodos de selección apropiados, se realiza para la discriminación de las moléculas que posean las propiedades de interés buscadas. Varias metodologías se encuentran disponibles para la realización de estos experimentos en el laboratorio, tanto de forma *in vitro* como *in vivo* (Jäckel et al., 2008). Una de las técnicas más utilizadas enfocada en producir mutaciones puntuales en sitios aleatorios es la PCR propensa a errores, donde la reducción de la fidelidad de la Taq ADN polimerasa provoca la introducción de mutaciones aleatorias durante la síntesis de ADN.

Esta técnica ha permitido introducir cambios en proteínas relacionadas con modificaciones a la especificidad de sustratos (Wymer et al., 2001; Zacco & Gherardi, 1999) o mejoras en la actividad enzimática (Chen & Arnold, 1993), solubilidad (van den Berg et al., 2006) y estabilidad (Komeda et al., 2003; Miyazaki & Arnold, 1999). Debido a esto, esta técnica ha atraído el interés para su aplicación en proteínas diseñadas *de novo*. En el caso de las proteínas de la serie Octarrelina, las características promisorias obtenidas con el primer diseño *in silico* de Octarrelina V, motivaron la utilización de esta técnica para conseguir optimizar sus características de solubilidad y estabilidad. Una nueva variante fue obtenida gracias a esta metodología, la cual fue denominada Octarrelina V.1.

Octarrelina V.1

La variante optimizada de Octarrelina V, fue llevada a cabo el año 2016 por Figueroa y colaboradores (Figueroa, Sleutel, et al., 2016). Para su realización, Octarrelina V fue sometida ciclos consecutivos de PCR propensa a errores a fin de modificar la solubilidad de la proteína en bacterias *E. Coli*. La identificación de las mutantes con modificaciones fue realizada mediante un *screening* basado en GFP, donde el vector con el gen de Octarrelina se clono en fusión con un gen de GFP, pudiéndose seleccionar en cada ciclo las colonias con más proteína soluble en función de la mayor intensidad de GFP que expresaban (Figura 4, A). Después de 8 ciclos de evolución dirigida, la proteína paso de la fracción soluble a la insoluble (Figura 4, B). Octarrelina V.1 presentó finalmente 16 mutaciones localizadas principalmente en las regiones N y C terminales (93% de identidad, figura 5). Uno de los detalles importantes a notar es la aparición de cuatro prolinas dentro de estas mutaciones (L9P, Q16P, L26P, L210P) las cuales se presentaron en una cantidad importante considerando el proceso aleatorio de sustitución de aminoácidos, además de no encontrarse en el diseño original de Octarrelina V, tal como se detalló anteriormente.

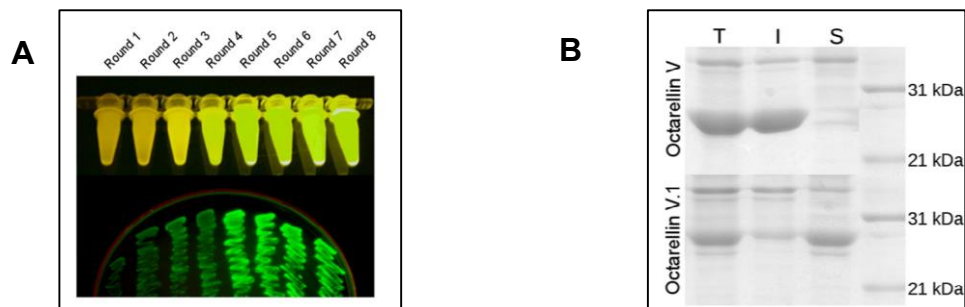


Figura 4: Octarrelina V.1 producida por evolución dirigida.

El seguimiento de la solubilidad en cada ciclo permitió observar las mejores colonias de cada ciclo (A). El resultado final, después de 8 ciclos puede observarse en B (T: Fracción total, I: Fracción insoluble, S: Fracción soluble) (Extraído de Figueroa et al., 2016)

A fin de confirmar que las modificaciones introducidas en la no alteren la estructura, Octarrelina V.1 fue sometida a una serie de caracterizaciones biofísicas. Mediante espectroscopia de CD e infrarroja, se confirmó la presencia de estructura secundaria con un alto contenido de α -hélices. Espectroscopia 2D-NMR validó la presencia de estructura terciaria, aunque con algunas zonas de alta movilidad. Experimentos de denaturación tanto térmica como química con GdmCl presentan una termoestabilidad en la proteína así como un desplegamiento cooperativo de esta. Mediante dispersión de rayos X de ángulo reducido (SAXS) se reveló además que solo una pequeña diferencia entre Octarrelina V y Octarrelina V.1. Todo esto confirma que el proceso de evolución dirigida permitió la obtención de una proteína más soluble, estable y sin alteraciones significantes a nivel de estructura terciaria con Octarrelina V.

Octarrelina V	MAFLIVEGLSEKELKQAVQIANEQGLRAIAFLKQFARNHEKAERFFELLVREGVEAIIIA	60
Octarrelina V.1	MAFLIVKGLSEKDLNIAVQIANEQDPSAIAFLKQFARNHEKAERFFELLVREGVEAIIIA	60
	*****:* ***:*:***** *****	
Octarrelina V	RGVSEREIEQAACLAREKGFALFLAEYERRDRQFDDIIEYFERYGFKAVIVATGLDEK	120
Octarrelina V.1	RGVSEREIEQAACLAREKGFALFLAEYERRDRQFDDIIEYFERYGFKAVIVATGLDEK	120

Octarrelina V	ELKQAAQKIEEKGFKALAFGRIDQENRNINDIFELLQRQGLRAIIAATGLSERELSWAL	180
Octarrelina V.1	ELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQGLRAIIAATGLSERELSWAQ	180
	*****:*****:*****	
Octarrelina V	RAARQYGLDIIIFAYGQFDEQDNQFKHFLELIRRLGAA	217
Octarrelina V.1	RAAQYGLDIIIFANGQFDEQDNRFKHFLELIRRQGAA	217
	:** *****:***** ** **	

Figura 5: Alineamiento de secuencias entre Octarrelina V y Octarrelina V.1

Destacadas en amarillo se indican las mutaciones introducidas posteriores a la evolución dirigida. En tanto, destacadas en rojo se encuentran los residuos de prolina presentes en Octarrelina V.1.

La comparación entre las curvas SAXS de Octarrelinas V y V.1 con una curva SAXS teórica del modelo *de novo* realizado de Octarrelina V dio señales de que ninguna de las dos Octarrelinas se plegaban de acuerdo a lo esperado por el modelo diseñado *de novo*.

Debido a que caracterizaciones clásicas del modelo *de novo* de la secuencia diseñada de Octarrelina V no pudieron llevarse a cabo, la obtención de esta variante optimizada permitió la consideración de caracterizaciones empíricas para la determinación estructural. Estas incluyen técnicas de alta resolución como NMR, criomicroscopía electrónica o cristalización de rayos x (Huang, Boyken, & Baker, 2016; Muñoz, 2008). Sin embargo las primeras dos técnicas requieren tamaños de proteínas incompatibles con la secuencia de cerca de 200 aminoácidos que forma Octarrelina V.1 (Krishnan & Rupp, 2012; Thompson et al., 2016) por lo que la cristalografía fue considerada como la mejor opción, permitida además por las mejoras en solubilidad y estabilidad obtenidas en esta variante. Debido a la movilidad detectada en las primeras caracterizaciones de Octarrelina V.1, *screenings* realizados utilizando la técnica conocida como difusión de vapor por gota colgante de la proteína no lograron una cristalización. A causa de ello, su cristalización tuvo efectuarse mediante co-cristalización, permitiendo la formación de complejos con proteínas ayudantes de cristalización o chaperonas. Una de las chaperonas fue nanobodies (Korotkov et al., 2009; Rasmussen et al., 2011), también conocidos como fragmentos VHH, los cuales corresponden al dominio variable sencillo de fragmentos de anticuerpo derivado de especies *Camelidae*; la otra corresponde a α Rep (Urvoas et al., 2010), una familia de proteínas artificiales basadas en proteínas HEAT repetitivas. Calorimetría isoterma de titulación confirmó la unión estequiométrica 1:1 para ambos complejos.

Mediante la misma técnica de gota colgante, se logró la obtención de cristales para ambos complejos. En el caso de la cristalización del complejo Octarrelina V.1-nanobody, los cristales obtenidos se presentaron en dos diferentes polimorfos, $P2_1$ y $I4_122$, los cuales difractaron a 1,95 y 3,20 Å, respectivamente. En tanto, una forma del complejo Octarrelina- α Rep pudo ser cristalizada en un grupo espacial $P2_1$, la cual difractó a una resolución de 2,22 Å. Las estructuras determinadas de ambos complejos no fueron completas, pero la superposición entre estas presenta el mismo plegamiento, pudiendo así

complementarse para obtener una visión general de la estructura de Octarrelina V.1. Gran parte de los sectores que no se pudieron determinar en ambas estructuras corresponden a zonas de alta movilidad, como el extremo c-terminal de la proteína. El cristal del complejo Octarrelina V.1- α Rep reveló, sin embargo, la porción de una hélice α en este extremo (Figura 6).

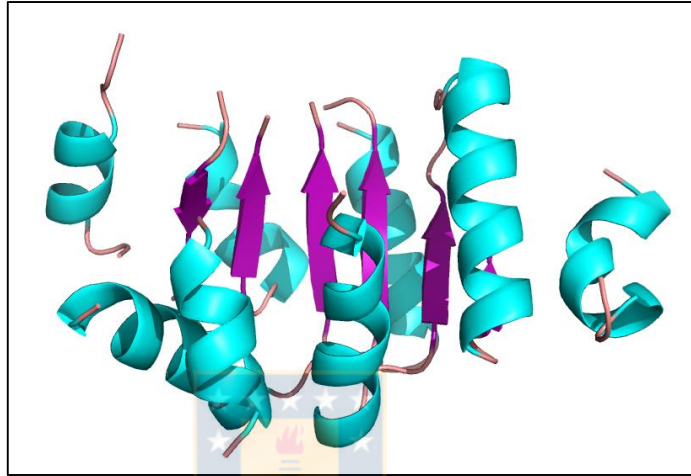


Figura 6: Estructura por rayos x de Octarrelina V.1 (PDB: 4ZV6).

La estructura del cristal obtenida en complejo con α Rep presenta sectores faltantes. Hélice α desde Gly215 a Phe204 (extremo derecho) se puede visualizar únicamente para este complejo

Como se puede apreciar en la figura, Octarrelina V.1 no presenta la arquitectura originalmente diseñada de un barril TIM, aun cuando se presenta un patrón alternante de hebras β y α -hélices organizado en un motivo $\beta\alpha\beta$, lo que hace que su arquitectura se asemeje más a un plegamiento similar al Rossman. Mediante análisis DSSP, fue posible identificar que aproximadamente un 74% de la estructura diseñada a nivel secundario se encuentra presente en la estructura final. Estos resultados exponen que a pesar de encontrarse bien diseñada a nivel secundario, faltan las interacciones que permitan la obtención de una estructura a nivel terciario.

A pesar de que Octarrelina V.1 no presenta el plegamiento esperado de acuerdo al diseño original, la proteína ha mostrado que igualmente puede ser un modelo de estudio de gran interés. Análisis llevados a cabo en el mismo trabajo testearon trece servidores automatizados online de predicción de proteína a fin de predecir la estructura final de Octarrelina V.1 conocida por cristalografía. Ninguno de ellos fue capaz de predecir la estructura de esta proteína: más del 50% de ellos predijeron que su plegamiento era del tipo barril TIM.

Uno de los retos que este modelo ha dejado es conseguir una mejor cristalización del complejo Octarrelina- α Rep para la obtención de una estructura más completa con la cual se pueda llevar a cabo realizar estudios de su estructura tanto a nivel global como local. Otro tema de interés para analizar de esta proteína dice relación con las prolinas obtenidas durante la evolución dirigida y la significancia de su presencia en las características de solubilidad y estabilidad obtenidas por esta proteína.

Octarrelina V.1 Δ 4P y Octarrelina V 4P

Con el fin de conocer un poco más en profundidad el efecto que el aminoácido prolina puede poseer de estabilidad, se generaron dos proteínas mutantes mediante la técnica de mutagenesis sitio-dirigida por PCR. De una manera inversa a la evolución dirigida aplicada, este procedimiento produce modificaciones específicas en el gen de interés in vitro para posteriormente insertar el ADN modificado dentro de una bacteria o célula eucariótica. Dentro de los enfoques para llevarlo a cabo se encuentra el uso de PCR con enzimas termoestables. En esta técnica, partidores mutagénicos que contienen las bases diferentes en comparación a la secuencia molde son sintetizados, para que así cubran la región donde la mutación será realizada. Estos partidores se utilizan posteriormente para llevar a cabo la PCR con la mutación requerida (McLennan et al., 2013; Tao & Raz, 2015).

Mediante este procedimiento, Octarrelina V.1 fue modificada para revertir las cuatro prolinas presentes en su secuencia y sustituirlas por los residuos

presentes en la secuencia original de Octarrelina V. De manera contraria, Octarrelina V fue sometida a los ciclos necesarios de esta técnica a fin de introducir cuatro prolinas en su secuencia en las mismas posiciones de interés (Figura 7). Estudios de estas variantes pueden dar luces de su efecto en las características de Octarrelina V.1. Primeras purificaciones de estas proteínas dieron cuenta que su obtención se debía realizar a partir de la fracción no soluble de bacterias *E. Coli* (Genco, 2017).

A

Octarrelina V.1	1	MAFLIVKGPSEKDLNPAVQIANEQDPSAIAFLKQFARNHEKAERFFELLVREGVEAIIHARGVSEIEQAACLAREKGF	80
Octarrelina V.1 Δ4P	1	MAFLIVKGLSEKDLNPAVQIANEQDPSAIAFLKQFARNHEKAERFFELLVREGVEAIIHARGVSEIEQAACLAREKGF	80
Octarrelina V.1	81	EALAF LAEYERRDRQFDDIIIEYFERYGFKAVIVATGLDEKELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQ	160
Octarrelina V.1 Δ4P	81	EALAF LAEYERRDRQFDDIIIEYFERYGFKAVIVATGLDEKELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQ	160
Octarrelina V.1	161	GLRAIIAATGLSERELSWAQRAAQYGLDIIIFANGQFDEQDNRFKHFLEPIRRQGAA	217
Octarrelina V.1 Δ4P	161	GLRAIIAATGLSERELSWAQRAAQYGLDIIIFANGQFDEQDNRFKHFLEPIRRQGAA	217

B

Octarrelina V	1	MAFLIVEGLSEKELKQAVQIANEQGLRAIAFLKQFARNHEKAERFFELLVREGVEAIIHARGVSEIEQAACLAREKGF	80
Octarrelina V 4P	1	MAFLIVEGPPSEKELKPAVQIANEQGPRAIAFLKQFARNHEKAERFFELLVREGVEAIIHARGVSEIEQAACLAREKGF	80
Octarrelina V	81	EALAF LAEYERRDRQFDDIIIEYFERYGFKAVIVATGLDEKELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQ	160
Octarrelina V 4P	81	EALAF LAEYERRDRQFDDIIIEYFERYGFKAVIVATGLDEKELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQ	160
Octarrelina V	161	GLRAIIAATGLSERELSWALRAARQYGLDIIIFAYGQFDEQDNQFKHFLEPIRRLGAA	217
Octarrelina V 4P	161	GLRAIIAATGLSERELSWALRAARQYGLDIIIFAYGQFDEQDNQFKHFLEPIRRLGAA	217

Figura 7: Modificaciones puntuales realizadas para generar las mutantes de Octarrelina.

Los aminoácidos modificados mediante mutagenesis sitio-dirigida para generar a Octarrelina V.1 Δ4P (A) y a Octarrelina V4P (B) se encuentran señalados en color amarillo.

Propiedades de prolina en la estructura de proteínas

Es sabido que dentro de los 22 aminoácidos naturales encontrados en las proteínas, prolina posee características que la hacen única. La cadena lateral de este aminoácido posee una estructura cíclica especial, un anillo pirrolidina, siendo el único aminoácido que contiene un α -amino secundario (Hermanson, 1997). Esto provoca efectos en la basicidad y polaridad general de este residuo. Otras consecuencias tienen directa relación en su participación dentro de las cadenas polipeptídicas: su presencia impone restricciones importantes en el ángulo diédrico ϕ a valores de -68 o -75 (Yaron & Naider, 1993), lo que le otorga una conformación rígida que reduce la flexibilidad estructural en las regiones polipeptídicas que contienen este aminoácido (Lehninger et al., 2009). En los péptidos en donde este aminoácido participa puede ser encontrado en configuración *cis* además de la forma *trans* encontrada comúnmente en los enlaces peptídicos. La isomeización *cis-trans*, que puede tomar lugar durante el plegamiento, se convierte en un factor limitante durante este proceso (Alderson et al., 2018). Debido a la configuración de este aminoácido, el nitrógeno presente en él carece de un protón, haciendo que este átomo no tenga capacidad de participar en enlaces hidrógenos con sus vecinos. Sin embargo, esto provoca que el átomo de oxígeno del grupo carbonilo que precede a este enlace, tenga una tendencia a aceptar y formar fuertes puentes hidrógeno (Deber & Brodsky, 2001).

Las restricciones estructurales de este residuo también tienen consecuencias en el espacio conformacional, siendo participante principal de estructuras como "loops" y vueltas β . Su presencia en las conformaciones α -helice y hojas β se encuentra principalmente limitada a zonas posteriores a los extremos c-terminales de estas estructuras (MacArthur & Thornton, 1991).

A nivel más general, este aminoácido ha sido relacionado en la mejora de la estabilidad. El trabajo de Suzuki (1989) presentó evidencias de una fuerte relación entre el incremento de la termoestabilidad y la presencia de este residuo en oligo-1,6-glucosidasas de diferentes especies de *Bacillus* del tipo termófilas.

Estudios de sustituciones por prolina en esta enzima para la especie *Bacillus cereus*, una termófila facultativa, pudo corroborar que sustituciones en vueltas β y α -hélices producían mejoras en la estabilidad térmica de esta (K Watanabe et al., 1994; Kunihiro Watanabe et al., 1997). Efectos positivos en estabilidad se han mostrado para sustituciones con este aminoácido en otras proteínas también (Gaseidnes et al., 2003; Hardy et al., 1993; Trevino et al., 2007). Al igual que con la estabilidad, el aminoácido prolina también se ha encontrado conectado a la solubilidad, teniendo funciones importantes para evitar la formación de agregados en proteínas nativas como la cristalino gamma D (Pande et al., 2005) o en la fibronectina tipo III (Steward et al., 2002).

A pesar de lo mencionado anteriormente, también se ha encontrado en otros trabajos que no siempre sus efectos pueden estabilizar una proteína (Allen et al., 1998; Trevino et al., 2007). Debido a las diferencias que se pueden apreciar en sus efectos, es de sumo interés la búsqueda de la relación entre prolina y sus efectos estructurales en la solubilidad y estabilidad.

Como ha sido detallado, de las 16 mutaciones presentes en Octarrelina V.1 que produjeron las mejoras en la solubilidad y estabilidad de la proteína, cuatro de estos aminoácidos corresponden a prolinas que no estaban en la secuencia original. El número de este residuo introducidos es más de lo esperado para una mutación al azar (Figuroa *et al.*, manuscrito en preparación); además, resultados anteriores han mostrado que la sustitución de los cuatro aminoácidos posee efectos en la expresión de la proteína en cuerpos de inclusión (Genco, 2017). Estudiar, por lo tanto, a nivel estructural esta proteína que permita realizar un análisis de la participación de este aminoácido en las características finales obtenidas es de gran interés. Para ello, es necesario revisar las herramientas cristalográficas con las que actualmente se cuenta para poder encontrar alternativas que permitan resolver de mejor manera la estructura de Octarrelina V.1 y con ello poder generar un modelo más claro a partir del cual realizar un análisis estructural más completo que permita una mayor comprensión del efecto de prolinas.

Cristalización de proteínas

El conocimiento estructural de diversas proteínas ha sido indispensable para mejorar la comprensión de sus funciones y mecanismos. Desarrollada por más de 150 años, la cristalización de proteínas ha presentado una enorme expansión en las últimas dos décadas: entre 1999 a la actualidad, más de 120.000 nuevas estructuras se encuentran depositadas en el Protein Data Bank (Berman et al., 2000), siendo un 90% de ellas obtenidas mediante la técnica de difracción de Rayos X (Al 21 de enero de 2019, disponible en <http://www.rcsb.org/pdb/statistics/holdings.do>).

La obtención de estos datos parte desde la preparación de cristales adecuados para ser analizados mediante rayos X lo cual es, sin embargo, una tarea compleja. A pesar de los grandes esfuerzos que se han puesto en desarrollar la teoría, al día de hoy la cristalografía continua siendo en gran parte un proceso de ensayo y error, utilizando como base éxitos que se han producido a lo largo de la historia en otras proteínas. La razón por la cual es dificultoso controlar el proceso de cristalización radica principalmente en la complejidad de las interacciones débiles requeridas para que moléculas de forma irregular con estructuras flexibles en su interior como son las moléculas de proteínas, puedan organizarse en una red cristalina regular y periódica (Rupp, 2010).

Se han descrito múltiples variables que pueden influenciar la cristalización. Dentro de las variables típicas puede incluirse la concentración de la proteína, que debe encontrarse tan alta como razonablemente sea alcanzable para cada proteína que se estudia. La naturaleza de la proteína, ya sea una cristalización de la molécula en su totalidad o proteínas truncadas, con modificaciones químicas o mutaciones, puede generar una diferencia entre una cristalización exitosa y un completo fracaso. La pureza de la proteína, sin lugar a dudas, es un factor determinante, por lo que es recomendado un porcentaje de pureza mayor al 95%, con la presencia de una banda simple en geles de SDS-PAGE (Benvenuti & Mangani, 2007; Rupp, 2010). Parámetros físicos y químicos de la proteína en

solución además de métodos para llevar a cabo la cristalización son variables importantes de considerar y detallar.

Diagrama de fases de cristalización

La formación de un cristal se produce cuando moléculas de proteína son forzadas a salir desde la fase soluble a la insoluble de una manera ordenada, siendo por ello una forma especial de separación de fases a partir de una solución homogénea. A fin de que la separación de fases ocurra, la solución debe encontrarse supersaturada y, dado un evento de nucleación, el exceso de proteína puede separarse de la solución. Si el proceso de nucleación y otros parámetros cinéticos son favorables, se puede retornar a una condición de equilibrio mediante la formación y desarrollo de un cristal (Rupp, 2010).

La interacción entre las variables que influyen la cristalización al producir una precipitación ordenada y la solubilidad de la proteína pueden ser representadas utilizando el diagrama de fases de cristalización (Figura 8). Esta representación del diagrama de fases como un espacio ideal bidimensional dista mucho de lo que realmente ocurre, en donde múltiples factores con diferentes grados de influencia complican esta visión simplificada. A pesar de ello, este esquema es suma utilidad, ya que ilustra las regiones claves para la interpretación de los resultados experimentales.

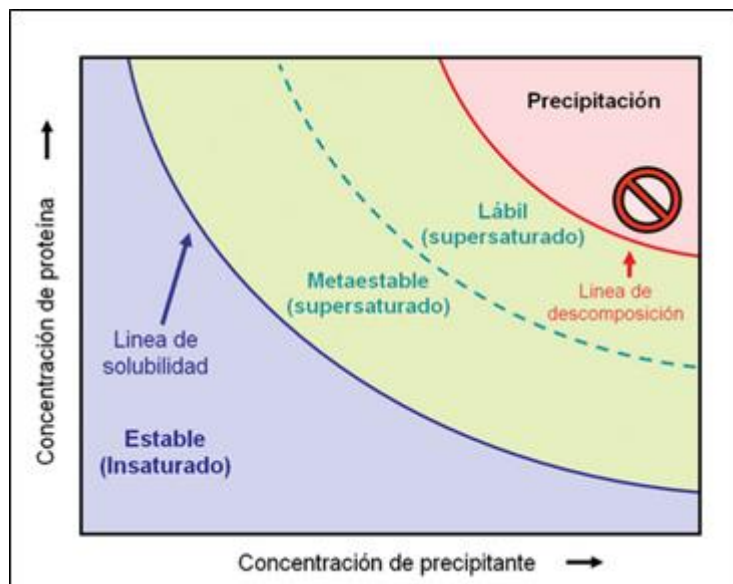


Figura 8: Diagrama de fase para la cristalización de macromoléculas.

El diagrama se encuentra dividido entre en una región de insaturación y una región de superpaturación mediante la línea que denota la máxima solubilidad a específicas concentraciones de un precipitante. La región de supersaturación, a su vez se divide en regiones poco determinadas denominadas metaestables y lábiles. En estas dos regiones se producen los fenómenos de nucleación y crecimiento de los cristales de proteínas (Modificado de Rupp, 2010).

Como se puede apreciar en la figura 8, el diagrama define zonas con características termodinámicas definidas donde las condiciones fisicoquímicas en la solución determinan el estado de la proteína. La región con solución insaturada corresponde a la proteína en solución en una sola fase. La línea o curva de solubilidad separa esta zona de la región supersaturada, donde se presenta tanto proteína única como proteína saturada en equilibrio termodinámico. Dentro de esta región, la zona de precipitación corresponde una solución con alto grado de supersaturación que hace inviable la agregación ordenada. Esta zona se separa de otras dos zonas menos definidas conocidas como lábil y metaestable, las cuales son relevantes en el proceso de cristalización de proteínas. Si las condiciones de la solución conducen a la proteína a la zona lábil de mayor sobresaturación, la proteína puede experimentar una espontánea y homogénea nucleación. Una vez formado el núcleo, el crecimiento continúa a

un nivel menor de sobresaturación en la zona metaestable. El crecimiento cristalino finalmente se detiene debido a que el equilibrio es finalmente alcanzado (Luft et al., 2011; McPherson & Gavira, 2014; Rupp, 2010).

La alteración de la solubilidad a fin de alcanzar la supersaturación, puede realizarse de diferentes formas entre las que se incluye: alterar a la proteína mediante cambios de pH, alterando la actividad química del agua mediante el uso de sales o alterando la naturaleza de las interacciones entre las moléculas de proteína y el solvente mediante la adición de polímeros o iones (McPherson & Gavira, 2014).

Métodos de cristalización

La cristalización de proteínas puede realizarse en algunos sistemas diseñados que permiten la alteración de la solubilidad de la proteína. A pesar de la existencia de diferentes estrategias, las características generales de todas ellas son la misma: la solución con proteína es combinada con agentes precipitantes en diferentes proporciones, en sistemas más o menos cerrados, para posteriormente dejar que el sistema alcance el equilibrio hasta donde la cinética lo permita. La figura 9 presenta algunos de los métodos utilizados para la cristalización de proteínas. Sin embargo, solo dos de ellos (difusión de vapor y difusión de interfase libre) serán considerados por sus aplicaciones y éxitos que son de interés.

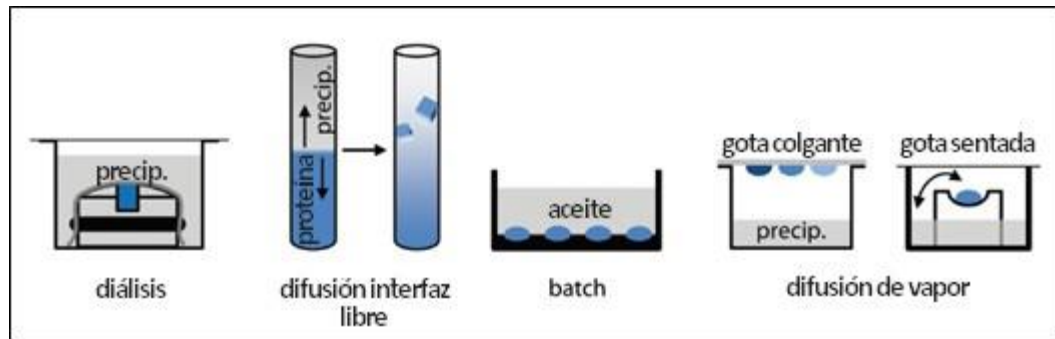


Figura 9: Esquema de diferentes métodos de cristalización

En todos los métodos presentados, la solución de proteína se encuentra de color azul mientras que el agente precipitante se presenta de color gris. En diálisis, una membrana separa a la proteína del agente precipitante. La difusión de interfaz libre se realiza en un diseño restringido. En la cristalización en batch, la proteína se mezcla con el agente precipitante y son colocados bajo aceite. La técnica de difusión de vapor posee dos variantes: gota sentada y gota colgante. (Imagen extraída de Klostermeier & Rudolph, 2018)

Difusión por vapor

En esta metodología la gota de cristalización busca alcanzar el equilibrio a través de fase de vapor contra un reservorio de gran volumen que contiene una alta concentración de precipitante. Este intercambio permite en la gota un aumento de la concentración tanto de los agentes precipitantes como de la proteína, pudiéndose conducir así la solución hasta un estado de supersaturación que le permita la cristalización.

En un caso ideal como se detalla en la figura 10, la mezcla entre el stock de proteína a una concentración p , se mezcla con la solución de cristalización a una concentración c en cantidades iguales. La gota resultante se encuentra en una posición $p/2$ y $c/2$, punto de partida para el experimento. Una vez que se sella el sistema, el vapor difunde entre la gota y el reservorio, alcanzándose la zona de nucleación. Con los núcleos formados se puede continuar con el crecimiento de los cristales. Alcanzado el equilibrio con la solución de proteína saturada, el crecimiento de los cristales termina, alcanzando su tamaño final.

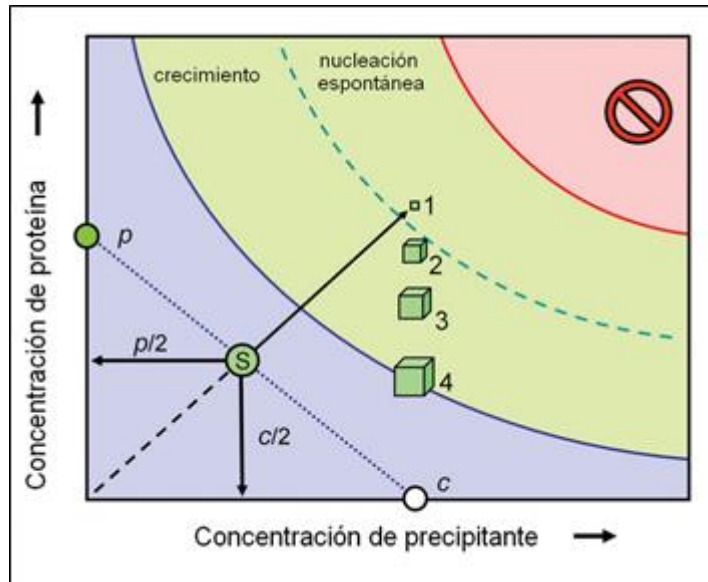


Figura 10: Diagrama de fases de la cristalización por difusión de vapor

La mezcla de las soluciones de proteína y agente precipitante da como resultado una gota que comienza el proceso de cristalización en el punto S. Esto lleva al incremento proporcional de ambos elementos a lo largo del diagrama. La vía que recorre es mostrada desde el origen. Una vez que el núcleo es formado en la zona de nucleación, el cristal puede comenzar a crecer (1 a 3) hasta alcanzar el equilibrio (4). (Modificada de Rupp, 2010)

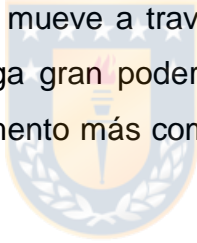
Diferentes variaciones de esta técnica son utilizadas: en la difusión por gota sentada, la gota se encuentra ubicada en una plataforma elevada. Es una técnica útil cuando soluciones de proteína poseen poca tensión superficial, manteniendo la gota a salvo en una pequeña depresión. Algunos inconvenientes incluyen el tiempo de demora entre el montaje de la primera gota hasta que todas las otras gotas sean montadas y se pueda cerrar la placa completa o la posibilidad de que los cristales formados puedan adherirse a la superficie plástica.

Una de las variantes que es más comúnmente utilizada en los laboratorios corresponde a la difusión de vapor por gota colgante. La gota es preparada en una tapa de vidrio deslizante mediante la mezcla de la solución de la proteína con la solución de cristalización que se encuentra en el reservorio. La tapa de vidrio es girada y posicionada como tapa sobre el reservorio, el cual tiene sus bordes previamente engrasados para permitir la adherencia de la tapa. Así, el vapor de

agua puede difundir entre la gota colgada y el reservorio que contiene una mayor concentración de agente precipitante (Drenth & Mesters, 2007; Rupp, 2010; Wlodawer et al., 2017).

Difusión de interfase libre líquido-líquido

La cristalización por difusión de interfase popularizada en la época de los 70 (Salemme, 1972), se basa en que la solución de proteína y la que contiene el agente precipitante son puestas en contacto en un capilar estrecho sin una mezcla previa, permitiendo que los componentes se equilibren uno contra otro. Una gran diferencia que es apreciable en comparación a las técnicas previamente descritas radica en que los perfiles de concentración de proteína y de precipitante varían en diferentes formas, pasando por diferentes localizaciones en el sistema (Figura 11). Así una “onda” de alta concentración de solución de cristalización se mueve a través de una cámara de proteína. Esta característica hace que tenga gran poder al permitir una mejor cobertura del espacio, haciendo el experimento más complejo y completo (Drenth & Mesters, 2007; Rupp, 2010).



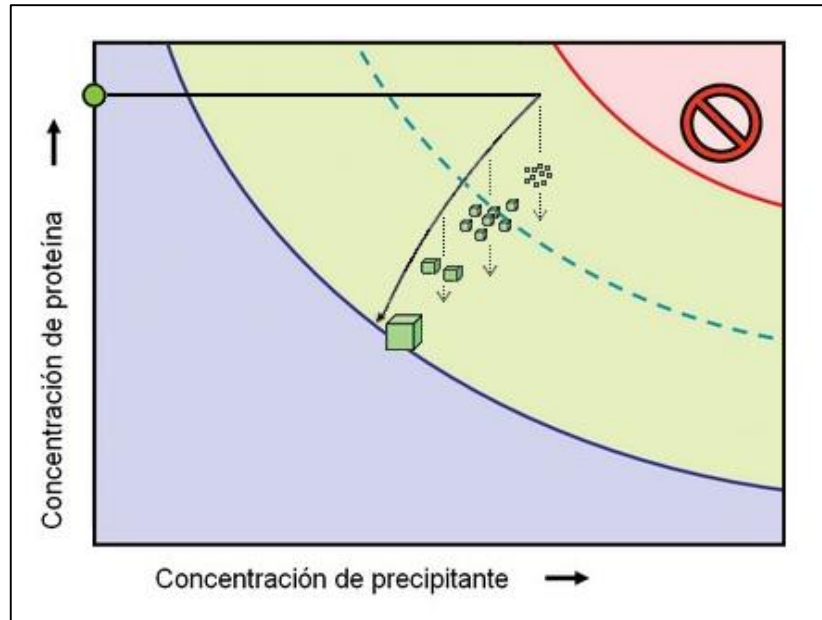


Figura 11: Difusión de interfase libre líquido-líquido.

A medida que el precipitante se difunde en la solución de proteína la interacción de la precipitación y el transporte de masa generan una onda de sobresaturación que viaja a lo largo del capilar a medida que aumenta el tiempo. (Modificada de Rupp, 2010).

Diferentes variantes existen para estos experimentos, entre los que se encuentran la difusión en interfase libre en chips de litografía suave multicapa o la contra-difusión en capilares. Esta última técnica se ha convertido en una técnica simple, económica y práctica para la cristalización de proteínas, permitiendo su uso para la búsqueda de condiciones de supersaturación y también para la optimización de la cristalización para difracción con rayos X (Ng et al., 2003).

A pesar de que no existen diferencias a nivel químico, las diferencias en las cinéticas y puntos finales de los experimentos varían enormemente entre diferentes técnicas de cristalización. Si las condiciones de partida que se utilizan entre las técnicas de batch y de difusión de vapor, pueden generarse similares condiciones resultantes (Baldock et al., 1996) como se aprecia en la Figura 12. Además, ambos experimentos se pueden considerar como ciegos y el éxito de

encontrar un cristal depende del número de experiencias que se lleven a cabo con diferentes condiciones de inicio. A diferencia de ellos, las técnicas de difusión de interfase presentan una mayor posibilidad de éxito (Figura 12), ya que además de permitir la exploración de varias condiciones simultáneamente, estas técnicas fueron diseñadas para alcanzar la sobresaturación crítica para la nucleación muy lentamente, buscando un solo evento de nucleación que pueda generar cristales utilizables (Juan Ma García-Ruiz, 2003; Rupp, 2010).

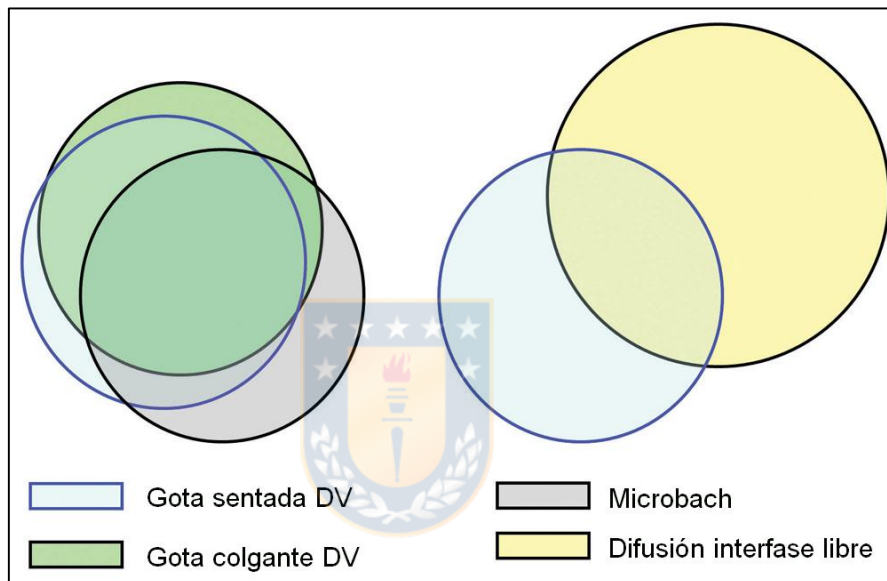


Figura 12: Diagramas de Venn representando diferentes cristalizaciones.

Los círculos representan diferentes técnicas de acuerdo a la leyenda, indicándose un igual éxito en los diferentes métodos con círculos con los mismos diámetros. Las técnicas de difusión de vapor, gota sentada y colgante, con las mismas condiciones iniciales de reactivos y tamaños de gota, presentan una superposición bastante grande. En tanto los experimentos de microbatch poseen algunas condiciones en común con ambas técnicas. En el segundo diagrama se representa una superposición entre las técnicas de difusión de interfase libre, con un potencial alto grado de éxito, y la técnica de gota sentada (Modificada de Rupp, 2010).

Cristalización mediante contra-difusión

En esta variante de la difusión de interfase libre, la solución de proteína interactúa con el agente precipitante, conectándose entre sí directamente u opcionalmente, separándose mediante una cámara intermedia que funciona como barrera física. Al momento de entrar en contacto estas soluciones, las

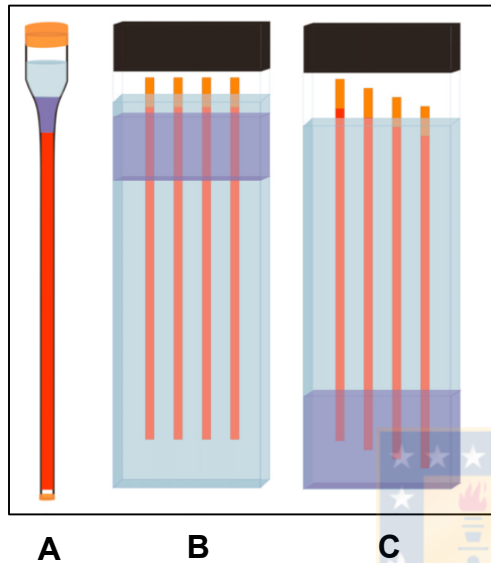


Figura 13: Representación de las configuraciones más usadas en contra-difusión.

(A) Configuración de triple capa (3L), (B) Caja de cristalización Granada Box Domino, (C) Método de acupuntura en gel en caja de cristalización Granada Box Domino. En todos los esquemas se presenta el gel de agarosa (violeta), la solución de proteína (rojo) y la solución precipitante (celeste). (Imagen extraída de Otálora, Gavira, Ng, & García-Ruiz, 2009)

moléculas del agente precipitante difunden rápidamente a través de la solución de proteína, provocando la formación de una fase amorfa en la interfase del líquido-líquido. Alrededor de los precipitados microcristalinos que se forman en esta fase, la concentración de proteína disminuye, pero el fenómeno de difusión continúa. Las moléculas de precipitante difunden más allá del precipitado, generándose un gradiente gradual de supersaturación, lo que permite la nucleación de la proteína y el consiguiente crecimiento del cristal en regiones de menor saturación (Otálora et al., 2009; Rupp, 2010). Es así como la formación de cristales por esta técnica se da mediante la interrelación no lineal entre transporte de masa, nucleación

de cristales y crecimiento (Carotenuto et al., 2002).

Existen diferentes configuraciones para la aplicación de este sistema, aunque todos se caracterizan por realizarse en una geometría restringida mediante el uso de capilares con diámetros de fracciones de milímetros. En la

configuración de tres cámaras o tres capas (3L) (Figura 13, A), la solución de cristalización y la proteína se encuentran en cámaras separadas por una cámara de gel (generalmente de agarosa). Esta capa intermedia ayuda a retrasar el proceso de difusión. En el sistema diseñado por Granada (Figura 13, B), el extremo inferior del capilar se mantiene abierto para permitir el fenómeno de contradifusión, permitiendo un retardo del evento de nucleación el cual es útil en experimentos en microgravedad (García-Ruiz et al., 2002). En el método de acupuntura en gel (Figura 13, C), el capilar lleno de proteína es puesto sobre un gel y la solución con el agente precipitante es vertida sobre el gel. Esto hace posible la modificación de precipitante en durante el experimento (Drenth & Mesters, 2007).

Debido a las características de este método, es posible llevar a cabo experimentos tanto de *screening* como de optimización de las condiciones de cristalización de una proteína. Estudios con proteínas modelo (insulina, lisozima y taumatina) realizando *screening* de matriz dispersa logró la obtención de en promedio el doble de hits que en la técnica de difusión por vapor, además se tuvo éxito en generación de cristales adicionando crioprotectores y átomos pesados para dispersión de rayos x permitiendo la determinación de fases de manera *ab initio* (Ng et al., 2003). Otros *screening* han sido llevados con éxito para esta técnica (González-Ramírez et al., 2017).

Ayudantes de cristalización

A pesar de los esfuerzos puestos en las técnicas que permitan la cristalización, existen ciertos tipos de proteína como los que poseen multidominios con uniones flexibles, proteínas de membrana, estados conformacionales transientes, moléculas de RNA de alta flexibilidad o proteínas intrínsecamente desordenadas; que presentan dificultades al intentarse su cristalización. Esfuerzos puestos en la estabilización de estados únicos han incluido el uso de inhibidores para enzimas o en aislamiento de dominios en proteínas multi dominios, preparación de complejos biológicos macromoleculares

estables, mediante mutaciones que reducen o incrementan la estabilidad de la molécula, o mediante el uso de proteínas de unión específicas tanto naturales como no naturales (Bukowska & Grütter, 2013). Estas proteínas son conocidas como ayudantes de cristalización (Klostermeier & Rudolph, 2018) y más antiguamente como chaperonas de cristalización (Koide, 2009). Se caracterizan por ser estables y unirse a proteínas que necesiten ser cristalizadas con una alta afinidad en áreas hidrofóbicas o regiones de loops, lo cual permite la reducción de la hidrofobicidad y/o la entropía de la muestra, facilitando los contactos entre las moléculas en la red cristalina (Klostermeier & Rudolph, 2018).

La tecnología de anticuerpos monoclonales (mAbs) se ha establecido por décadas como una opción importante a la hora de ser necesario reactivos con alta especificidad de unión. Por más de una década, fragmentos de anticuerpos han sido utilizados para ayudar en cristalizaciones. El uso de fragmentos Fab ha tenido grandes éxitos en el campo de las proteínas de membrana (Fang et al., 2009; Krishnamurthy & Gouaux, 2012; Tsukazaki et al., 2008) o en proteínas con alto grado de desorden como el caso de Tat, proteína activadora transcripcional del virus HIV-1 (Serrière et al., 2011). El caso ya mencionado de nanobodies, fragmentos de cadena simple de anticuerpos de camélidos, que presenta un largo historial de éxitos (Chaikuad et al., 2014; Domanska et al., 2011; Löw et al., 2013; Ring et al., 2013; Wu et al., 2011) al que se suma la estructura obtenida en Octarrelina V.1.

A pesar de la efectividad claramente demostrada para su uso en cristalización, el uso de fragmentos de anticuerpos aún en la actualidad presenta limitaciones en los costos de producción o patentamiento existentes para esta tecnología (Koide, 2009).

La introducción de proteínas artificiales con sitios de unión con especificidad deseada, también conocidas como proteínas andamio, han supuesto un importante avance en la ingeniería de proteínas, permitiendo la obtención de proteínas con capacidad de unión diferentes a los anticuerpos (Binz

& Plückthun, 2005; Skerra, 2007). Estos andamios se definen como bases proteicas que poseen tolerancia a la alteración de aminoácidos o a inserción de secuencias, lo que le confiere funciones variables relacionadas con la unión a targets específicos. Sus usos no solo se limitan al área de la cristalización, incluyéndose también aplicaciones en áreas terapéuticas o diagnóstico (Binz et al., 2005).

Los andamios basados en proteínas de unión son generalmente producidos a través de evolución dirigida, creándose grandes librerías combinatorias donde la secuencia de aminoácidos de un parche en la superficie del andamio, se encuentra diversificada (Gilbreth & Koide, 2012). Dentro de los grupos más importantes de este tipo de chaperonas se encuentra los monobodies, derivados de la fibronectina tipo III FN3 (también conocidos como dominios FN3 o adnectinas); affibodies, derivados del dominio de unión a inmunoglobulina de la proteína A; DARPin, proteínas diseñadas a partir de módulos repetitivos "Ankyrin"; o anticalinas, derivados de la proteína de unión a billina de las lipocalinas y de la lipocalina 2 humana. Un grupo creciente de estructuras resueltas con estos andamios se puede observar en el PDB. Gilbreth & Koide, 2012 presentan el detalle de algunas de ellas.

Las proteínas repetitivas, como la ya mencionada DARPin, se conocen por estar compuestas de repeticiones de un motivo proteico con favorables propiedades biofísicas (Kajava, 2001) Estas arquitecturas proteicas naturales se han estudiado con gran interés, generando andamios versátiles para el reconocimiento molecular. Algunos ejemplos más de este tipo de proteínas generadas incluyen las basadas en la proteína repetitiva tetratricopeptido (TRP) (Jackrel et al., 2010; Main et al., 2005), las proteínas con plegamiento repetitivo armadillo (Parmeggiani et al., 2008; Varadamsetty, et al., 2012) o las proteínas con motivos repetitivos HEAT, como las ya nombradas previamente α Rep.

Artificial alpha-helicoidal repeat proteins (α Rep)

El motivo repetitivo HEAT se ha encontrado como un motivo común tanto en procariontes como eucariontes (Andrade et al., 2001). Su nombre deriva de las proteínas en las cuales fue identificada primeramente: huntingtina (relacionado con la enfermedad de Huntington), el factor de elongación 3, la subunidad PR65/A de la proteína fosfatasa A y en TOR (diana de rapamicina) (Andrade & Bork, 1995). Este motivo se ha encontrado relacionado biológicamente a diferentes interacciones proteína-proteína (Andrade, Perez-Iratxeta, et al., 2001).

Las repeticiones HEAT se encuentran compuestas de α -hélices antiparalelas, las cuales se apilan en un arreglo consecutivo (Groves & Barford, 1999). La yuxtaposición de las hélices resulta en superficies extendidas donde es posible la unión a blancos (Figura 14). La curvatura y flexibilidad local del solenoide formado puede permitir que las repeticiones del motivo envuelvan a la proteína asociada. Se ha encontrado una alta divergencia entre las diferentes secuencias de las proteínas con motivo HEAT.

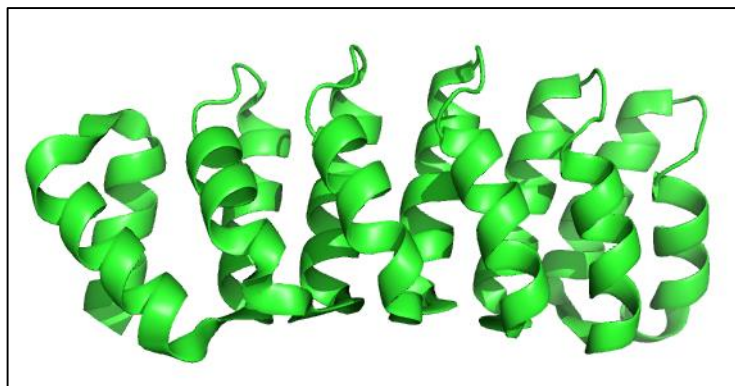


Figura 14: Modelo del motivo α Rep.

Un motivo repetitivo HEAT corresponde a dos α -hélices consecutivas. Para esta estructura el modelo se encuentra compuesto de cuatro repeticiones del motivo. Motivos en los extremos N y C terminal, denominados N-cap y C-cap, poseen secuencias específicas para estos extremos por lo que no se incluyen como parte de las repeticiones. (Imagen elaborada a partir de PDB 3LTM)

En el año 2010 Urvoas y colaboradores diseñaron una nueva proteína artificial basada en esta estructura (Urvoas et al., 2010). Para ello, un motivo consenso fue definido en base a clases específicas de proteínas HEAT, determinándose zonas de alta conservación, con un rol estructural, y zonas altamente variables, las cuales fueron posteriormente utilizadas como sitios de aleatorización en el diseño. Una librería con hasta 10 repeticiones del motivo fue creada y proteínas seleccionadas de esta se encontraron bien expresadas, plegadas y con alta estabilidad, lo cual permitió la purificación de estas en formas monoméricas y diméricas para proteínas con ciertos números de repetición interna, las cuales se encuentran corroboradas mediante ensayos biofísicos y cristalografía.

Complejos con diferentes proteínas blancos han sido igualmente purificados, caracterizados y cristalizados (Guellouz et al., 2013) lo que da cuenta de su utilidad como un ayudante para la cristalización de proteínas.

Debido a que Octarrelina V.1 presentó zonas resueltas diferentes y de mayor interés que Nanobody (Figueroa, Sleutel, et al., 2016), es que se tomó a este ayudante para realizar un proceso de re-cristalización mediante la técnica de contra-difusión que permita crear una estructura mejor resuelta para su estudio. Una nueva estructura resuelta abriría un nuevo camino para poder explorar las implicancias de ciertos aminoácidos no solo en el diseño *de novo* de proteínas, sino que igualmente en el plegamiento de proteínas. El presente trabajo también permitirá la utilización del ayudante de cristalización α Rep como una herramienta más allá de la cristalización: su uso en diferentes técnicas utilizadas para el estudio de complejos nos puede abrir la posibilidad de obtener más información sobre los efectos que aminoácido de prolina pueda tener durante la formación del complejo entre el ayudante de cristalización α Rep y las diferentes mutantes de Octarrelina.

HIPOTESIS Y OBJETIVO DEL TRABAJO

Resultado de un proceso de evolución dirigida, las características de solubilidad tanto al interior de *E. coli* como *in vitro*, y estabilidad de la proteína artificial Octarrelina V fueron mejoradas, mediante una nueva variante denominada Octarrelina V.1. La presencia inesperada de prolina dentro de las mutaciones introducidas hace pensar que este aminoácido puede tener una influencia significativa en las características finales obtenidas por Octarrelina V.1. A partir de esto, se hace necesario realizar estudios adicionales para encontrar el grado de relación existente prolina y las características. Un modelo estructural completo donde todas las prolina se encuentren presentes para su estudio además de experimentos en proteínas con modificaciones de este aminoácido pueden ser de gran relevancia al generar mayor información sobre la influencia de este residuo en proteínas diseñadas *de novo*.

Se plantea que:

“La obtención de un modelo más detallado de Octarrelina V.1 en conjunto con experimentos en mutante de prolina puede entregarnos información sobre la influencia de este residuo en las características de Octarrelina V.1”

A fin de demostrar la hipótesis, se plantea el como objetivo general:

“Determinación de una estructura más completa de Octarrelina V.1 junto con la examinación de mutantes de prolina”

Los objetivos específicos del presente proyecto incluyen:

- I. Producción y purificación de Octarrelina V.1 y α Rep
- II. Purificación y cristalización del complejo Octarrelina V.1- α Rep
- III. Determinación de estructura cristalina del complejo Octarrelina V.1- α Rep

- IV. Producción y purificación de Octarrelinas mutantes de prolina
- V. Estudios de interacción entre α Rep y Octarrelinas mutantes



MATERIALES

Cepa y plásmidos

Escherichia coli, cepa BL21 (DE3) fue transformada con los plásmidos que codifican para α Rep y Octarrelina V.1. Se propagó en medio Luria-Bertani (LB) sólido o líquido adicionado ampicilina para alcanzar una concentración de 100 μ g/mL dependiendo los volúmenes de trabajo.

α Rep se encuentra subclonado en el vector de expresión pQE81 (Qiagen). Este vector de expresión adiciona una secuencia de 6-His en el extremo N-terminal y se encuentra bajo el control del promotor T5. El gen que codifica para α Rep se encuentra localizado entre los sitios de restricción BamHI y HindIII. Octarrelina V.1 y las mutantes Octarrelina V.1 Δ 4P y Octarrelina V4P, en tanto, se encuentran subclonadas en el vector pJB122, plásmido derivado del vector pNS3785 el cual a su vez proviene del vector comercial pTrcHisA (Invitrogen). El gen de Octarrelina se encuentra bajo el control del promotor Tcr. El constructo de estos genes se encuentra, en tanto, entre los sitios de restricción NcoI and BamHI. Ambos plásmidos son inducibles por IPTG.

METODOLOGÍA

Transformación bacteriana mediante choque térmico

La amplificación de plásmidos fue realizada utilizando *E. coli* competentes almacenadas a -80 °C, las cuales fueron descongeladas en hielo. 50 μ L de estas bacterias se mezclaron con 2 μ L de ADN plasmidial. Posteriormente se incubó la mezcla en hielo por 30 minutos. Luego de ello se realizó un choque térmico a 42 °C durante 90 segundos trasladando el tubo a una baño termo-regulado seco. Las bacterias fueron nuevamente puestas en hielo durante 2 minutos, siendo posteriormente añadido 200 μ L de medio LB sin antibiótico e incubadas en agitación a 37°C durante 30 a 60 minutos. Las bacterias transformadas fueron

trasladadas a una placa de agar LB con ampicilina (100µg/mL) y se incubaron durante 16 h a 37 °C.

Para un almacenamiento a largo plazo de los plásmidos de trabajo, una colonia obtenida en esta placa se transfirió a un tubo con 5 mL de medio LB con ampicilina 100 µg/mL y se dejó en incubación durante 16 horas a 37 °C. 500 µL del cultivo bacteriano se trasladaron a tubos de tapa rosca, donde se mezcló gentilmente con 500 µL de glicerol al 50% estéril. Los tubos se almacenaron a -80 °C hasta su uso.

Producción y purificación Octarrelina V.1

La recuperación de las bacterias *E.coli* a -80°C con plásmido codificante de Octarrelina V.1 se realizó tomando con una pipeta Pasteur parte del cultivo bacteriano congelado y sembrándolo en una placa LB con ampicilina (100 µg/mL). La placa se llevó a una estufa a 37°C donde se incubó por toda la noche. Una colonia de la placa fue seleccionada y se inoculó en 200 mL de medio LB con 100 µg/mL de ampicilina. El matraz se trasladó a un incubador con agitación y temperatura controlada donde se agitó el cultivo a 150 rpm durante un máximo de 16 horas. Posterior a ello, 30 mL del pre-inóculo fueron trasladados a matraces con 1L de medio LB con ampicilina 100 µg/mL. El cultivo se agitó a 180 rpm y 37 °C hasta que se alcanzó una OD₆₀₀ de 0.6, donde la expresión de la proteína se indujo mediante la adición de isopropil β-D-tiogaláctosido (IPTG) a 1 mM, incubando posteriormente las células por 4 horas más. El cultivo se recuperó mediante centrifugación a 5.000 x g por 10 minutos a 4°C, descartándose el sobrenadante.

El pellet bacteriano obtenido se resuspendió en 10 mL por litro de cultivo en tampón Tris(hidroximetil)aminometano/Ácido Clorhídrico (Tris-HCl) 100 mM pH 8.0 y ácido etilendiaminotetraacético (EDTA) 2 mM. Las bacterias resuspendidas se trasladaron a tubos centrífugos de 50mL, los cuales se llevaron a una caja con hielo. El pellet bacteriano se lisó mediante sonda de sonicación

(Vibra-Cell™ Ultrasonic Processors, Sonics) utilizando un pulso de 80% de amplitud por 30 segundos seguido de 15 segundos de descanso en un tiempo total de lisis de 3 minutos. El lisado obtenido se centrifugó por 45 minutos a 20.000 x g y 4°C, donde se recuperó y almacenó el sobrenadante en refrigerador a 4°C.

El sobrenadante obtenido se incubó a 70-75 °C por 45 minutos para realizar una precipitación de las proteínas bacterianas de la mezcla y separarlas de Octarrelina V.1, que presenta termostabilidad. Luego, esto se llevó a centrifugación a 25.177 x g por 30 minutos a 4 °C y se recuperó el sobrenadante, el cual se filtró con filtro de jeringa de 0,45µm y trasladó a un nuevo tubo para ser almacenado a 4°C por 12 horas.

La purificación de Octarrelina V.1 se llevó a cabo por cromatografía de intercambio iónico utilizando para ello el equipo Akta Prime GE Healthcare mediante el uso de la columna HiTrap Q XL (GE Healthcare). La columna se equilibró en primer lugar con 3 volúmenes de columna de tampón Tris-HCl 100mM pH 8.0, EDTA 2mM y Cloruro de sodio (NaCl) 1M (Tampón B); seguido de 5 volúmenes de tampón Tris-HCl 100mM pH 8.0 y EDTA 2mM (Tampón A), ambas realizadas a un flujo de 5 mL/min. La elución de la proteína se realizó utilizando un gradiente lineal de 0 a 100% de Tampón B en un largo de 20 volúmenes de columna a un flujo de 5mL/min, siguiendo la elución de la proteína por un detector UV a una absorbancia de 280 nm.

Las muestras que presentaron absorbancia a 280 nm se analizaron mediante la técnica de electroforesis en gel de poliacrilamida con dodecilsulfato sódico (SDS-PAGE) (Laemmli, 1970). Se prepararon geles al 15% de acuerdo a la cantidad de muestras a analizar y se cargó un total de 5 µL de marcador de peso molecular Maestrogen y 20 µg de proteína total en cada carril. La electroforesis se llevó a cabo corriendo a 70V en la zona del gel concentrador y a 100V en la zona del gel separador. Concluida la electroforesis, los geles se incubaron durante 20 minutos en una solución de metanol:agua:ácido acético en

una relación 3:6:2 y posteriormente se tiñeron en una solución de azul de coomasie coloidal G250 durante toda la noche a temperatura ambiente y agitación. Finalmente los geles se destiñeron utilizando agua destilada y en agitación.

Las fracciones con presencia de Octarrelina V.1 (25KDa) se seleccionaron, mezclaron y concentraron utilizando para ello unidades de filtración centrífuga Amicon® Ultra (Merck) de 15mL y 10000 NMWL. 2 mL de muestra concentrada se filtraron con filtro de 0,45 µm y se llevó a un segundo paso de purificación por cromatografía de exclusión molecular utilizando para ello la columna Hiload™ 16/60 Superdex 75 (GE Healthcare) en el equipo Akta Prime equilibrando previamente con 2 volúmenes de columna de tampón fosfato de potasio (K₂HPO₄ /KH₂PO₄) 50 mM pH 8.0, NaCl 150 mM. La cromatografía se ejecutó en este mismo tampón a un flujo de 0,5 mL/min, seleccionándose las fracciones absorbancia a 280nm, las cuales se analizaron mediante SDS-PAGE en geles de 15% de la misma forma detallada anteriormente. Las fracciones seleccionadas se mezclaron y su concentración se midió utilizando el ensayo de proteínas Bio-Rad. La muestra de proteína fue dosificada en tubos centrífugos de 1,5 mL y almacenada a -80 °C hasta su uso.

Producción y purificación αRep

Se recuperaron las bacterias *E.coli* a -80°C con plásmido codificante de αRep tomando con una pipeta pasteur parte de la bacteria congelada y sembrándola en una placa LB con ampicilina (100µg/mL). La placa se trasladó a una estufa a 37°C donde se incubó por toda la noche. Una colonia de la siembra obtenida se seleccionó e inoculó en 100 mL de medio LB con 100µg/mL de ampicilina. El matraz se trasladó a un incubador con agitación y temperatura controlada donde se creció el cultivo a 150 rpm durante un máximo de 16 horas. Posterior a ello, 30mL del pre-inóculo fueron trasladados a matraces con 1L de medio LB con ampicilina 100µg/mL. El cultivo se agitó a 180 rpm y 37°C hasta que se alcanzó una OD₆₀₀ de 0.6, induciéndose la expresión de la proteína

mediante la adición de IPTG 1mM, incubando posteriormente las células por 4 horas más. El cultivo se recuperó mediante centrifugación a 5000x *g* por 10 minutos a 4°C, descartándose el sobrenadante.

El pellet bacteriano obtenido se resuspendió en 10 mL por litro de cultivo en tampón Fosfato de potasio 50 mM pH 8.0, NaCl 150 mM. Las bacterias resuspendidas se trasladaron a tubos centrífugos de 50 mL y en frío se realizó la lisis por sonda de sonicación (Vibra-Cell™ Ultrasonic Processors, Sonics) utilizando un pulso 80% de amplitud por 30 segundos seguido de 15 segundos de descanso en un tiempo total de lisis de 3 minutos. El lisado obtenido se centrifugó por 45 minutos a 25.177 x *g* y 4 °C, trasladándose posteriormente el sobrenadante, filtrándolo con filtro de jeringa de 0,45um y almacenándose a 4°C por toda la noche.

Al día siguiente, la purificación se realizó en el equipo Akta prime mediante cromatografía IMAC con la columna IMAC HP HiTrap (GE Healthcare). Esta se lavó con 3 volúmenes de columna de tampón fosfato de potasio 50 mM pH 8.0, NaCl 150 mM y imidazol 200 mM (tampón B) y posteriormente con 5 volúmenes de K₂HPO₄ /KH₂PO₄ 50 mM pH 8.0 y NaCl 150 mM (tampón A) a un flujo de 5 mL/min. La elución de αRep se realizó mediante un gradiente de 0 a 100% del tampón B en un total de 20 volúmenes de columna a un flujo de 5mL/min. Las fracciones con absorbancia a 280nm se analizaron mediante SDS-PAGE en geles de 15%.

De acuerdo al análisis de SDS-PAGE, se seleccionó las fracciones con presencia de proteína a los 30 KDa y la mezcla se concentró hasta los 2 mL utilizando unidades de filtración centrifuga Amicon® Ultra (Merck) de 15 mL y 10000 NMWL. Este volumen se filtró a 0,45μm y se ingresó al equipo Akta prime para un segundo paso cromatografico por exclusión molecular por medio de la columna Hiloat™ 16/60 Superdex 75 (GE Healthcare), la cual se equilibró con anticipación con tampón fosfato de potasio 50 mM pH 8.0 y NaCl 150 mM y que se utilizó para las posterior cromatografía a un flujo de 0,5 mL/min. La elución de

proteínas desde la columna se detectó mediante UV a una absorbancia de 280 nm y las fracciones con absorbancia se analizaron mediante SDS-PAGE en geles de 15%.

Las fracciones seleccionadas se mezclaron y la concentración de la muestra fue calculada utilizando el ensayo de proteínas Bio-Rad. La muestra de α Rep se dosificó en tubos centrifugos de 1,5mL y almacenó a -80°C hasta su uso.

Producción complejo Octarrelina V.1- α Rep

Las fracciones de las proteínas purificadas Octarrelina V.1 y α Rep almacenadas a -80°C fueron descongeladas. Ambas alícuotas se mezclaron en una relación molar 2:1 de OctarrelinaV.1- α Rep y se incubaron a temperatura ambiente por 15 minutos. La mezcla posteriormente se filtró a $0,45\ \mu\text{m}$.

La purificación del complejo formado se realizó mediante cromatografía de exclusión molecular en el equipo FPLC Merck Hitachi con bomba L-6210 Intelligent Pump, detector L-4250 UV-VIS e integrador D 2500. La columna Hiloal™ 16/60 Superdex 75 (GE Healthcare) se conectó en su salida a un recolector de fracciones Super Fraction Collector SF-2120, Advantec. La columna se equilibró previamente con 2 volúmenes de columna de tampón fosfato de potasio 50 mM pH 8.0 y NaCl 150 mM para luego ingresar la muestra a un flujo constante de $0,5\ \text{mL}/\text{min}$, manteniéndose a lo largo de toda la cromatografía.

La presencia del complejo se analizó por SDS-PAGE en gel al 15%. La selección de las fracciones se llevó a cabo en función de la presencia de 2 bandas (aproximadamente a 25 y 30 KDa) y baja presencia de bandas contaminantes. La mezcla de las fracciones seleccionadas se concentró mediante sistema de filtración con centrifuga (Amicon® ultra-0.5 30K, Merck Millipore) a $4.000\ \times\ g$ por 10 minutos. El complejo fue cuantificado mediante medición de absorbancia a 280 nm y su pureza medida mediante densitometría mediante el programa ImageJ para ser finalmente almacenado a -80°C .

Cristalización de complejo Octarrelina V.1- α Rep

La obtención de cristales del complejo se realizó mediante la técnica de cristalización por contradifusión en capilares utilizando el sistema “granada box”. Para la aplicación de esta técnica, capilares de 0,3 mm de diámetro interno y 50 mm de largo, se llenaron por capilaridad del stock de complejo proteico previamente descongelado, utilizándose aproximadamente un volumen de 4.2 μ L de proteína a una concentración de 11,9 mg/mL. Los capilares se sumergieron en diferentes cajas plásticas selladas con agarosa 0,5% donde se depositaron dos condiciones de cristalización. Una de ellas contenía 26% PEG 400, 200 mM MgCl₂, HEPES 100 mM pH 7,0; en tanto la segunda fue llenada con 35% PEG 400, 200 mM MgCl₂, HEPES 100 mM pH 7,0. Todos los estudios cristalográficos se llevaron a cabo a 17°C.

El seguimiento del proceso de cristalización se realizó mediante su visualización semanal en el microscopio Ivens SE2200. La captura de imágenes con filtro polarizador se efectuó en el microscopio Olympus BX41 acoplado a una cámara MicroPublisher 5.0 & 3.3 RTV (QImaging) junto con un adaptador de cámara U-TV0.5XC-3 (Olympus) conectado al programa Q-capture Pro 6.0 en el laboratorio de Cristales Líquidos de la Facultad de Química de la Universidad de Concepción. La medición de los tamaños de los cristales formados se realizó utilizando el programa ImageJ.

Los capilares con presencia de cristales se enviaron mediante correo hasta el European Synchrotron Radiation Facility (ESRF) en Grenoble, Francia, donde se realizó la difracción por rayos X y colección de la data cristalográfica en la línea BM30A, dedicada a la investigación de proteínas.

La estructura del complejo Octarrelina V.1- α Rep fue determinada mediante reemplazo molecular utilizando la estructura del complejo Octarrelina V.1- α Rep obtenida por gota colgante (PDB: 4ZV6) como templado en software *Phaser-MR* (McCoy et al., 2007). El modelo obtenido fue finalizado después de unos ciclos

de reconstrucción manual utilizando Coot (Emsley et al., 2010) y refinamiento utilizando *phenix.refine* (Afonine et al., 2012). Las coordenadas atómicas y factores estructurales se encuentran depositados en el Protein Data Bank bajo el código de acceso 6BZX. Las imágenes estructurales fueron realizadas con el software Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC)

Producción y purificación de proteínas Octarrelina V.1 Δ 4P y Octarrelina V4P

Las bacterias *E.coli* almacenadas a -80°C con cada uno de los plásmidos de las proteínas se recuperó tomando con una pipeta Pasteur parte de las bacterias y sembrándolas en una placa LB con ampicilina ($100\ \mu\text{g}/\text{mL}$). La placa se llevó a una estufa a 37°C donde se incubó por toda la noche. Una colonia de la siembra obtenida se seleccionó e inoculó en 100 mL de medio LB con $100\ \mu\text{g}/\text{mL}$ de ampicilina. El matraz con la colonia se trasladó a un agitador incubador donde se creció el cultivo a 150 rpm y 37°C durante un máximo de 16 horas. Posterior a ello, 30mL del pre-inóculo fueron trasladados a matraces con 1L de medio LB + ampicilina $100\ \mu\text{g}/\text{mL}$. El cultivo se agitó a 180 rpm y 37°C hasta que se alcanzó una OD_{600} de 0.7, induciéndose la expresión de la proteína mediante la adición de IPTG 1 mM, incubando posteriormente las células por 5 horas más. El cultivo se recuperó mediante centrifugación a $5000\ \times\ g$ por 10 minutos a 4°C , descartándose el sobrenadante.

El pellet bacteriano obtenido se resuspendió en 10 mL por litro de cultivo en tampón Tris-HCl 50 mM pH 8.5, el que posteriormente se trasladó a tubos centrífugos de 50 mL. La lisis del pellet se realizó por sonda de sonicación (Vibra-Cell™ Ultrasonic Processors, Sonics) utilizando un pulso de 80% de amplitud por 30 segundos seguido de 15 segundos de descanso en un tiempo total de lisis de 3 minutos, todo efectuado en hielo.

A fin de remover el contenido de ADN de los cuerpos de inclusión donde

se encontraba la muestra, se estandarizó el siguiente protocolo detallado a continuación: el pellet bacteriano lisado se equilibró a temperatura ambiente para posteriormente añadir 100 μ L de tampón de reacción (100 mM Tris-HCl pH 7.5, 25 mM $MgCl_2$, 10 mM $CaCl_2$) por mL de extracto obtenido, agitando a continuación los tubos. Luego de ello se añadió 5-50 unidades de DNasa I por mL de extracto obtenido, mezclando por inversión. Los tubos se trasladaron a una estufa a 37°C por 60 minutos. Posteriormente se añadió 20 μ L de Lisozima (10 mg/mL) por mL de extracto con el fin de poder eliminar residuos de pared celular. Los tubos se agitaron en vortex por 1 minuto.

El pellet obtenido se centrifugó a 16.000 x *g* por 45 minutos a 4°C, descartándose el sobrenadante. Luego de ello, el pellet se resuspendió en 100 mL de tampón Tris-HCl 50 mM pH 8.5 y se llevó a sonicación aplicando un pulso de 56% por 60 segundos seguido de 30 segundos de descanso en un tiempo total de lisis de 5 minutos. La muestra fue centrifugada por 10 minutos a 16000x*g*, manteniéndose el pellet. Nuevamente el pellet fue resuspendido en 100 mL de tampón Tris-HCl 50 mM pH 8.5. Este ciclo de sonicación y centrifugación se repitió 3 veces. Posterior a esto, el pellet se centrifugó a 16.000 x *g* por 45 minutos, descartando el sobrenadante. El pellet se resuspendió en 200 mL tampón Tris-HCl 50 mM pH 8.5 y Urea 8 M, y luego se centrifugó a 16.000 x *g* por 45 minutos manteniendo el sobrenadante y filtrándolo finalmente en filtro de 0,45 μ m en sistema de bomba de vacío. Esta solución caotrópica con urea permite la solubilización de los cuerpos de inclusión en los cuales se encuentran las proteínas Octarrelina V4P y V.1 Δ 4P.

El re-plegamiento y purificación de las proteínas se llevó a cabo mediante cromatografía de intercambio iónico. Para llevar a cabo el proceso en este tipo de columnas la proteína, denaturada mediante urea, debe unirse reversiblemente a la resina de intercambio iónico. Posterior a ello, un gradiente es utilizado para remover el agente denaturante, lo que mantiene unida la proteína a la columna además de inducir el re-plegamiento y formación de puentes disulfuro. Mediante un gradiente de sal, la proteína re-plegada puede ser posteriormente eluida

(Freydell et al., 2010). Estos pasos pueden ser observados en el esquema presentado en la figura 15.

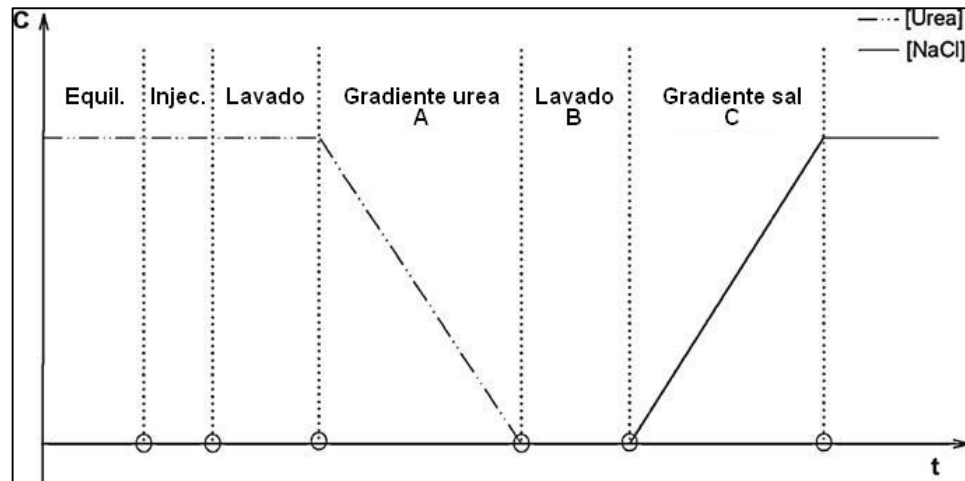


Figura 15: Esquema de re-plegamiento de proteínas mediante cromatografía de intercambio iónico

Línea de puntos: concentración de Urea; línea sólida: concentración de cloruro de sodio. Especificados con letras mayúscula se indican los pasos más relevante de la cromatografía: decrecimiento de la concentración de urea (A), lavado de columna con tampón sin urea (B) y elución de la proteína mediante gradiente salino (C). (Modificada de Freydell et al., 2010)

Para llevar a cabo el re-plegamiento de las proteínas mediante cromatografía de intercambio iónico en el equipo Akta prime con la columna HiTrap Q XL (GE Healthcare), el procedimiento se realizó como se detalla a continuación: la columna se lavó con 3 volúmenes de columna de Tris-HCl 50mM pH 8.5 y Urea 8M a un flujo de 5mL/min. Luego de la introducción de la muestra a la columna, se llevó a cabo un gradiente de 0 a 100% de tampón Tris-HCl 50mM pH 8.5 por 10 volúmenes de columna, seguido de un lavado por 5 volúmenes de columna con este tampón. Finalmente la elución de las proteínas se realizó mediante gradiente de 0 a 100% de tampón Tris-HCl 50 mM pH 8.5 y NaCl 1M por 10 volúmenes de columna. Las muestras obtenidas en este último gradiente se analizaron mediante SDS-PAGE en geles de 15% y por absorbancia a 260 y 280nm a fin de descartar muestras con una relación de absorbancia 260/280 mayor a 1.

Las muestras seleccionadas de acuerdo a su peso cercano a los 25KDa para ambas mutantes se mezclaron y concentraron utilizando para ello unidades de filtración centrifuga Amicon® Ultra (Merck) de 15 mL y 10000 NMWL. 2 mL de muestra concentrada se filtraron a 0,45 µm y se llevó a un segundo paso de purificación por cromatografía de exclusión molecular utilizando para ello la columna Hiload™ 16/60 Superdex 75 (GE Healthcare) en el equipo Akta Prime equilibrando previamente con 2 volúmenes de columna de tampón fosfato de potasio 50 mM pH 8.0 y NaCl 150mM. La cromatografía se ejecutó en este mismo tampón a un flujo de 0,5 mL/min, seleccionándose las fracciones de interés de acuerdo a su peso molecular, las cuales se analizaron mediante SDS-PAGE en geles al 15% de la misma forma detallada anteriormente. Las fracciones se mantuvieron a 4 °C. Al momento de ser utilizadas, las fracciones seleccionadas se mezclaron y concentraron en unidades de filtración centrifuga Amicon® Ultra (Merck) de 15 mL y 10000 NMWL hasta alcanzar la concentración deseada en función de los experimentos de formación de complejos o de micro calorimetría de titulación isotérmica, los cuales serán detallados a posteriormente, utilizando para la medición el ensayo de proteínas Bio-Rad.

Identificación de mutantes Octarrelina V.1Δ4P y Octarrelina V4P por espectrometría de masas

Las proteínas mutantes se identificaron por mapeo peptídico utilizando *matrix-assisted laser desorption/ionization time-of-flight mass spectrometry* (MALDI-TOF MS/MS) a partir de bandas de gel de proteínas teñido con azul de coomasie coloidal G250. Para ello, un gel SDS-PAGE al 15% (1,5 mm de espesor) fue cargado con 5 µg de proteína purificada de cada mutante. Una vez desteñido y lavado el gel de acuerdo al protocolo normal, las bandas individuales fueron cortadas y trasladadas por separado a un tubo centrifugo de 1,5 mL. Estas muestras fueron enviadas por correo hasta la unidad de bioquímica y proteómica analíticas del Instituto Pasteur de Montevideo, Uruguay, donde se llevaron a cabo los análisis.

Producción complejos Octarrelinas V4P y V.1Δ4P con αRep

Alícuotas de fracciones de cada una de las mutantes seleccionadas y concentradas se tomaron junto con concentrado purificado almacenado de αRep a -80°C previamente descongelado, en consideración con el volumen requerido para la cromatografía manteniendo una relación molar 2:1 de OctarrelinaV4P-V.1Δ4P-αRep. Se mezclaron ambas alícuotas y se incubaron a temperatura ambiente por 15 minutos para luego ser filtrada por jeringa en filtro de 0,45 μm.

La purificación del complejo formado se realizó posteriormente mediante cromatografía de exclusión molecular en el equipo FPLC Merck Hitachi con bomba L-6210 Intelligent Pump, detector L-4250 UV-VIS e integrador D 2500. La columna Hiloal™ 16/60 Superdex 75 (GE Healthcare) se equilibró previamente con 2 volúmenes de columna de tampón fosfato de potasio 50 mM pH 8.0 y NaCl 150 mM para luego ingresar la muestra a un flujo constante de 0,5 mL/min, manteniéndose este flujo a lo largo de toda la cromatografía.

Caracterización de las interacciones de unión entre αRep y Octarrelinas mediante micro calorimetría de titulación isotérmica (ITC)

Se estudió la unión entre αRep y Octarrelina V.1 o Octarrelina V.1 Δ4P mediante micro calorimetría de titulación isotérmica en colaboración con el laboratorio de biofísica, del centro de Bioinformática y Biología Integrativa de la Universidad Andrés Bello, Santiago, Chile.

En la titulación, utilizando el equipo NanoITC (TA Instrument), la muestra purificada de αRep (entre 2,48 a 6,98 mg/mL) fue colocada en la jeringa del equipo, y se añadió a la muestra de Octarrelina V.1 u Octarrelina V.1Δ4P ubicada en la celda a concentraciones entre 0,60 a 1,02 mg/mL . En cada experimento se llevó a cabo un total de 20 inyecciones controladas por computador de un total de 100 μL de αRep que tituló 500 μL de Octarrelina. La solución en la celda de

titulación fue agitada a 200 rpm durante todos los experimentos y la temperatura fue mantenida en los 25°C. La celda de referencia del microcalorímetro se encontraba con de agua ultrapura (Winkler). Un blanco del tampón de trabajo (fosfato de potasio 50 mM pH 8.0 y NaCl 150mM) se realizó a fin de obtener el calor de dilución el cual se sustrajo del calor de interacción obtenido a fin de obtener el valor de interacción real. La evaluación de la data fue realizada mediante el programa NanoAnalyze Data Analysis 3.7.5 (TA Instrument) utilizando el modelo “independiente” provisto por el software.



RESULTADOS

Resultados Objetivo I: Producción y purificación de Octarrelina V.1 y α Rep

Purificación Octarrelina V.1

La producción de Octarrelina para su posterior utilización realizada en *E. Coli* permitió la obtención de cantidades considerables de proteína que pudo ser llevada posteriormente a los siguientes pasos de purificación. El proceso de eliminación de proteínas por calor, gracias a la termoestabilidad de esta proteína, permite la eliminación de una porción importante de las proteínas intracelulares presente en la bacteria (Figura 16), sin embargo aun con este procedimiento es necesario los pasos por columnas de intercambio iónico y exclusión molecular.

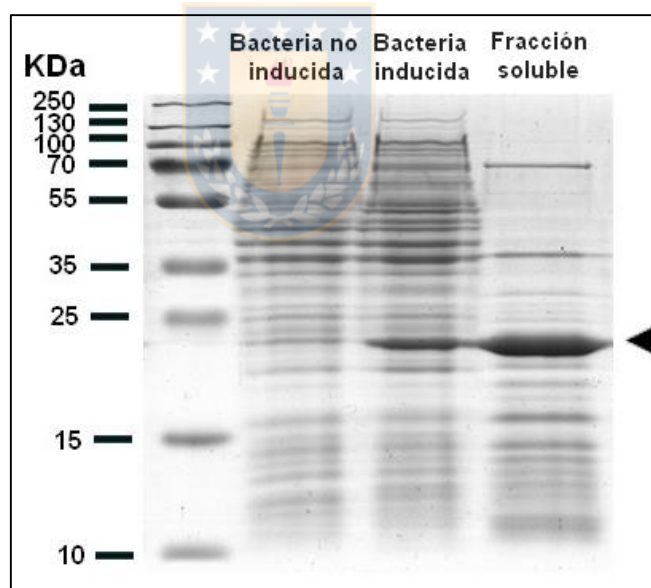


Figura 16: Análisis por SDS-Page 15% de fracciones de producción y tratamiento por calor de Octarrelina V.1

El gel fue cargado con 50 μ g de proteína total para bacterias no inducidas e inducidas, además de 100 μ g de sobrenadante tratado térmicamente (Fracción soluble). Flecha en negro indica el peso esperado para la proteína (aproximadamente 25KDa).

La cromatografía de intercambio iónico del tipo aniónico utilizada permitió

la captura de gran parte de la proteína de interés, y mediante la utilización de un gradiente de NaCl al 1M, se pudo efectuar la liberación de Octarrelina V.1 de la columna en la primera fase del gradiente (Figura 17).

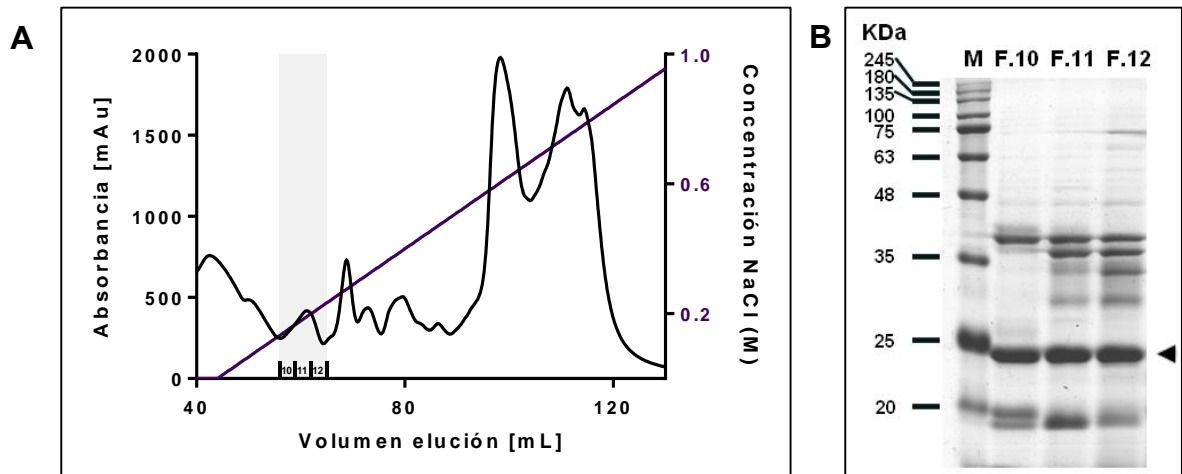


Figura 17: Cromatografía de intercambio iónico fracción soluble Octarrelina V.1

A) Cromatograma de la cromatografía realizada con un gradiente de NaCl 1M. La proteína fue detectada mediante absorbancia de 280nm. El peak marcado en gris corresponde a la fracciones con presencia de Octarrelina. **B)** Gel SDS-Page 15% con fracciones de interés. Flecha en negro corresponde al peso esperado para Octarrelina V.1 (24KDa). Fracciones 10 a 12 fueron cargadas con 4 ug de proteína total.

Aún con este paso de purificación y tal como se puede apreciar en la Figura 16,B, las fracciones seleccionadas de esta cromatografía aún presentan contaminantes que pueden ser interferentes importantes. Gracias a la cromatografía de exclusión molecular, fue posible obtener fracciones de Octarrelina V.1 con una baja cantidad de contaminantes (Figura 18). Las Fracciones seleccionadas pudieron ser finalmente concentradas y almacenadas hasta su uso.

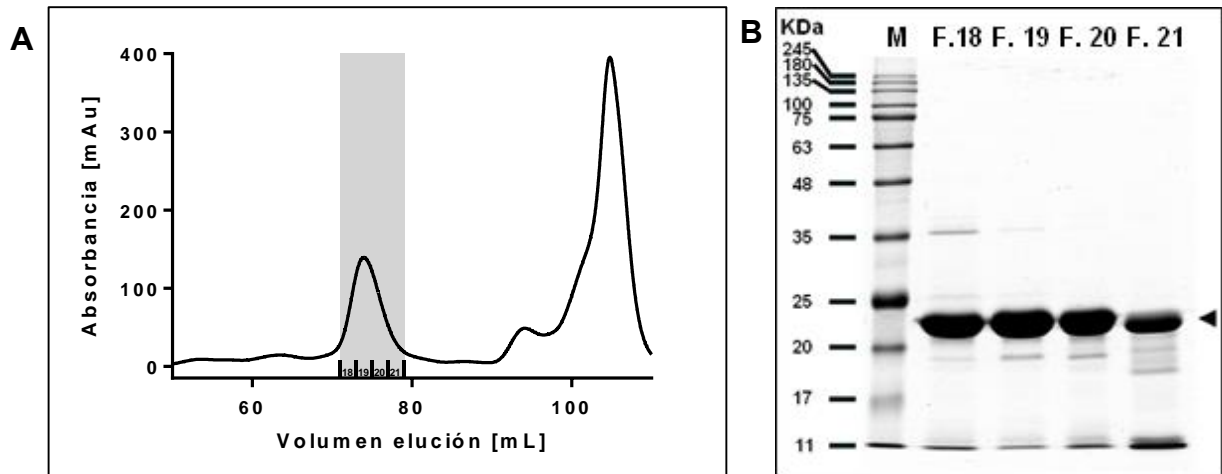


Figura 18: Cromatografía de exclusión molecular Octarrelina V.1

A) Cromatograma de la cromatografía de exclusión molecular. La proteína fue detectada a una absorbancia de 280 nm. Peak en gris corresponde a fracciones con presencia de proteína. **B)** Análisis de fracciones en SDS-PAGE al 15. Flecha en negro corresponde al peso esperado para Octarrelina V.1 (24 KDa). Fracciones de 18 a 21 se cargaron con 20 μ g de proteína total.

Purificación α Rep

El plásmido codificante para α Rep, adiciona un tag de 6-His en el extremo N terminal de la proteína, lo que permite su purificación mediante cromatografía de afinidad por metal inmovilizado (IMAC). Esta técnica permite, por lo tanto, la separación de α Rep de parte de las proteínas celulares de *E. Coli* por la capacidad que los residuos de histidina poseen para quelar metales divalentes, como níquel, y así unirse más fuertemente a la columna, además de permitir una gran recuperación de la proteína, como se puede apreciar por la mayor cantidad de fracciones seleccionadas (Figura 19).

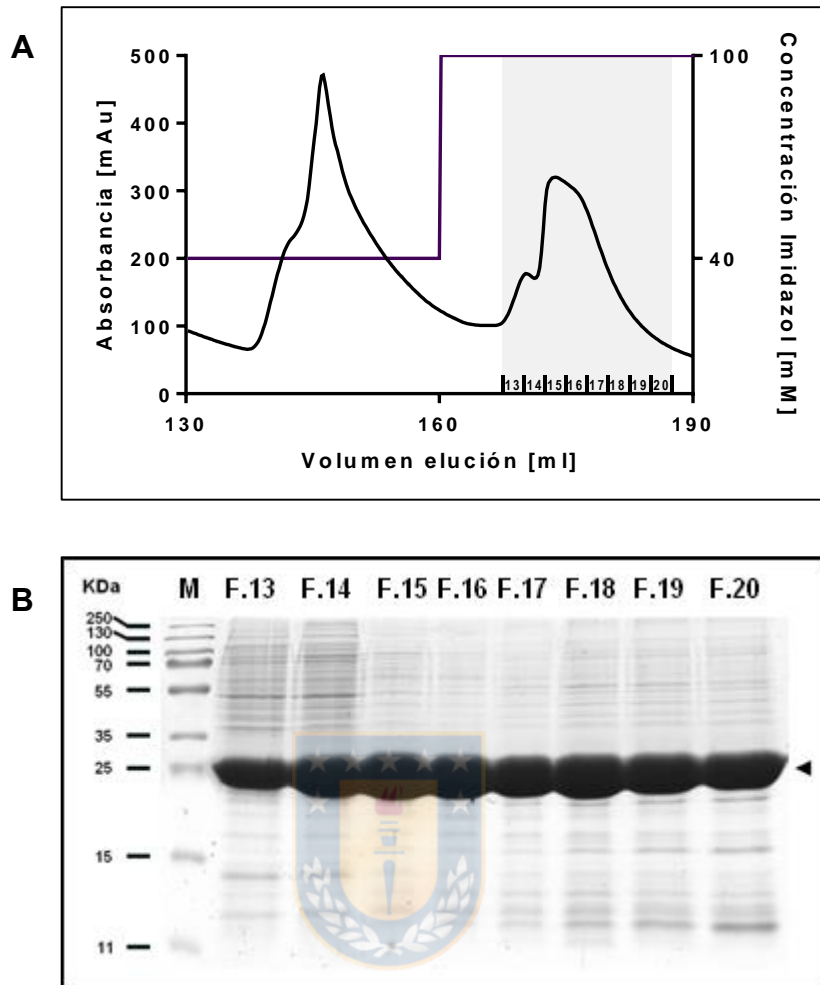


Figura 19: Cromatografía IMAC α Rep

A) Cromatograma de cromatografía realizada con un gradiente de Imidazol 200 mM. Se destaca en gris el peak de fracciones con presencia de α Rep. **B)** Gel SDS-PAGE al 15% de fracciones de interés. Marcado con flecha negra se encuentra el peso molecular esperado para la proteína α Rep (30KDa). Fracciones 13 a 20 se cargaron con 30 μ g de proteína total.

Las fracciones obtenidas de esta etapa poseen una alta presencia de proteína con bandas contaminantes, las cuales fueron en gran medida eliminadas durante la posterior cromatografía por exclusión molecular (Figura 20), permitiendo obtener fracciones que se mezclaron y concentraron para su uso final.

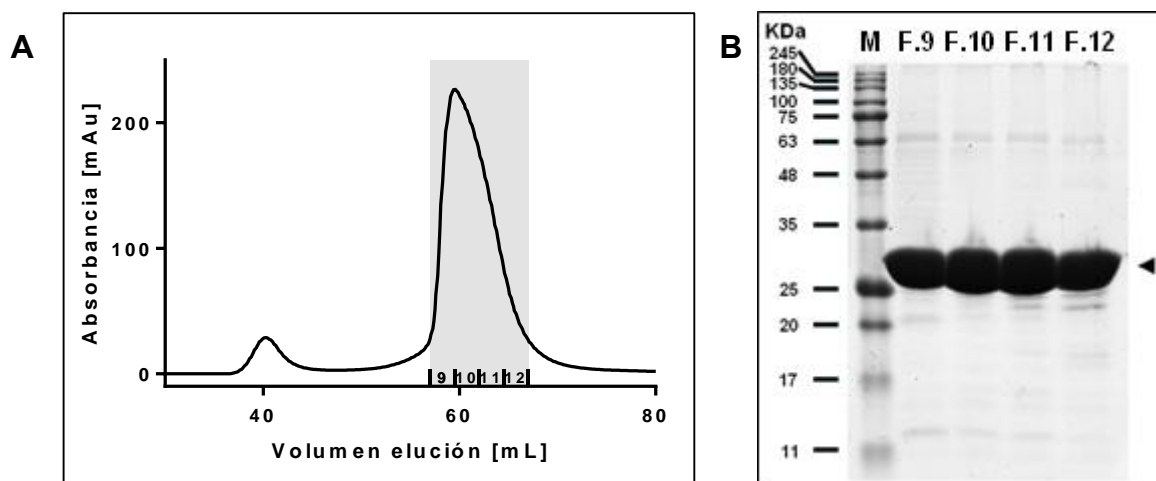


Figura 20: Cromatografía de exclusión molecular α Rep.

A) Cromatograma de cromatografía de exclusión molecular. El peak en gris corresponde a las fracciones con presencia de α Rep detectado mediante absorbancia a 280 nm. **B)** Análisis de fracciones mediante SDS-Page al 15%. Flecha en negro indica el peso molecular esperado para α Rep (30 KDa). Fracciones 9 a la 12 fueron cargadas con 30 μ g de proteína total.

Los volúmenes y concentraciones de las proteínas obtenidas de las purificaciones se encuentran detalladas en la Tabla 1

Tabla 1: Resumen de purificación proteínas Octarrelina V.1 y α Rep

Proteína	Volumen por litro de cultivo de <i>E.coli</i>	Concentración
Octarrelina V.1	2 mL	0,69 mg/mL
α Rep	3 mL	4, 884 mg/mL

Resultados Objetivo II: Purificación y cristalización del complejo Octarrelina V.1- α Rep

Purificación complejo Octarrelina V.1- α Rep

Para la formación de complejos estables de ambas proteínas, se tomaron stock almacenados a -80°C de Octarrelina V.1 a una concentración de 1,65 mg/mL y α Rep a 0,60 mg/mL con los cuales se realizó la producción *in vitro* del complejo mediante incubación de ambas proteínas, con un leve exceso de Octarrelina, y una posterior purificación del complejo obtenido por un paso cromatográfico por exclusión molecular, lográndose la obtención de fracciones con presencia del complejo Octarrelina V.1- α Rep (Figura 21).

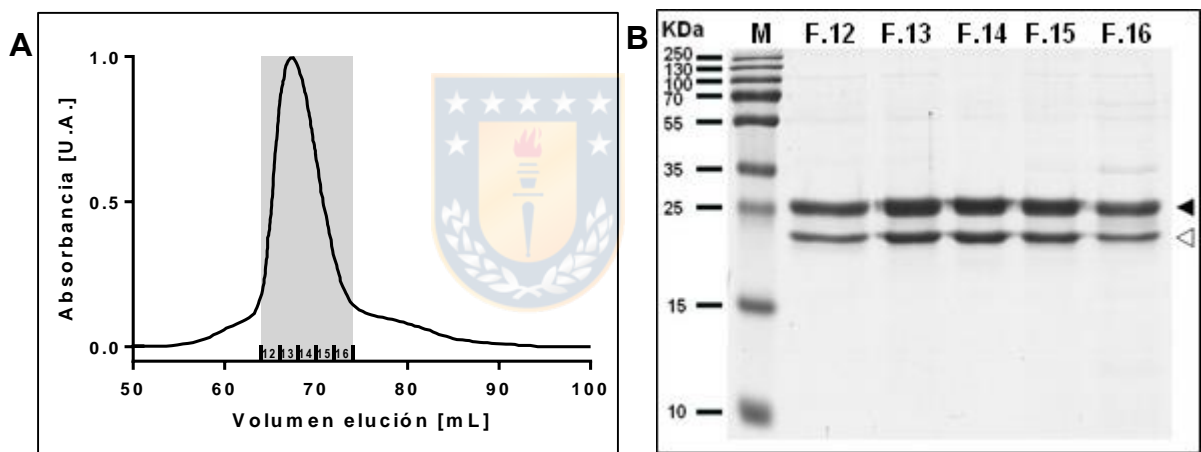


Figura 21: Cromatografía de exclusión molecular complejo Octarrelina V.1- α Rep

A) Cromatograma de la purificación realizada (U.A: Unidades Arbitrarias). Peak en gris corresponden a fracciones con presencia de complejo, detectado mediante absorbancia a 280 nm. **B)** SDS-Page realizado para el análisis de las fracciones al 15%. Flecha negra indica peso esperado para proteína α Rep (30KDa), mientras que flecha blanca indica el peso esperado para Octarrelina V.1 (24KDa). Fracciones 12 a 17 fueron cargadas con 0,7 μg de proteína total.

Las fracciones con pocos contaminantes fueron finalmente mezcladas y concentradas para su uso en cristalización (Figura 22). Fue obtenido finalmente un volumen de 10ul de complejo a una concentración de 11,34 mg/mL con un

porcentaje de pureza de aproximadamente un 90% calculado mediante densitometría.

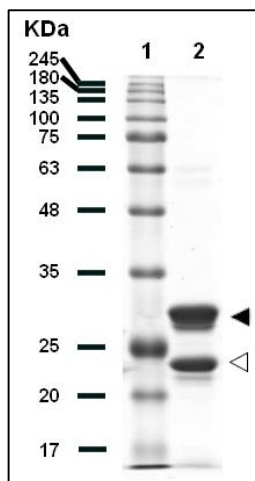


Figura 22: Análisis SDS-PAGE al 15% del complejo Octarrelina V.1- α Rep concentrado.

Carril 1 presenta patrón de peso molecular mientras que carril 2 presenta muestra de complejo cargada con $1\mu\text{g}$ de proteína total. Flecha en negro indica el peso molecular esperado para α Rep (30KDa), flecha en blanco indica peso molecular de Octarrelina V.1 (24KDa)

Cristalización complejo Octarrelina V.1- α Rep

El montaje fue realizado con el concentrado obtenido en la purificación por cromatografía de exclusión molecular del complejo, el cual fue introducido en capilares de 0,3 mm de diámetro y sumergido en tampones previamente estudiados para producir su cristalización. Para el caso del montaje realizado en tampón HEPES 100 mM pH 7,0, 26% PEG 400, 200 mM MgCl_2 , se logró apreciar en solo 4 días la formación de precipitado proteico en la base del capilar y el núcleo de lo que semanas después se convirtió en uno de los cristales finalmente formados en el capilar (Figura 23A). El cristal finalmente formado alcanzó unas dimensiones aproximadas de 380 μm de largo por 103 μm de ancho.

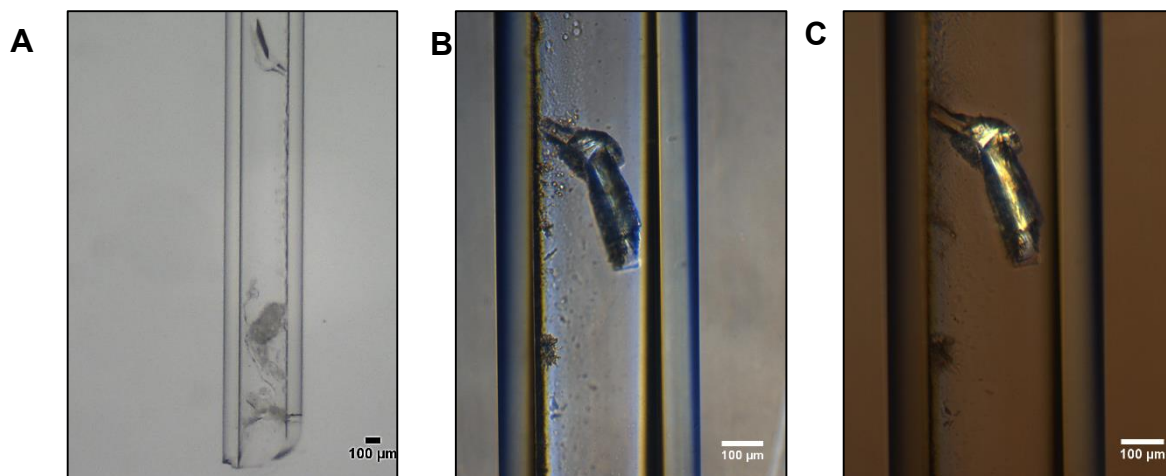


Figura 23: Cristalización por contradifusión del complejo Octarrelina V.1- α Rep en 26% PEG 400.

A) Capilar en la primera semana de su montaje. **B)** Capilar con cristal formado a la semana 3 observado con filtro polarizador. **C)** Cristal en la semana 12 con filtro polarizador.

En el caso de la condición con tampón HEPES 100 mM pH 7,0, 35% PEG 400, 200 mM $MgCl_2$ se logró apreciar a los 4 días, al igual que en la condición anterior, una formación importante de núcleos a lo largo del capilar (Figura 24A), los cuales dieron paso a cristales con una formación planar y más desorganizada que la condición anterior, aunque con tamaños más considerables. El mejor cristal de este capilar formado alcanzó aproximadamente los 357 μm de largo x 340 μm de ancho (Figura 24C).

El sistema de capilares permitió en ambos casos la mantención de los cristales formados en ellos por más de 10 semanas como se puede apreciar en las Figuras 23 y 24, lo que aseguró su estabilidad sin la adición de ningún otro reactivo hasta su envío al sincrotrón en Francia para su debida exposición a rayos X y obtención del patrón de difracción de estos.

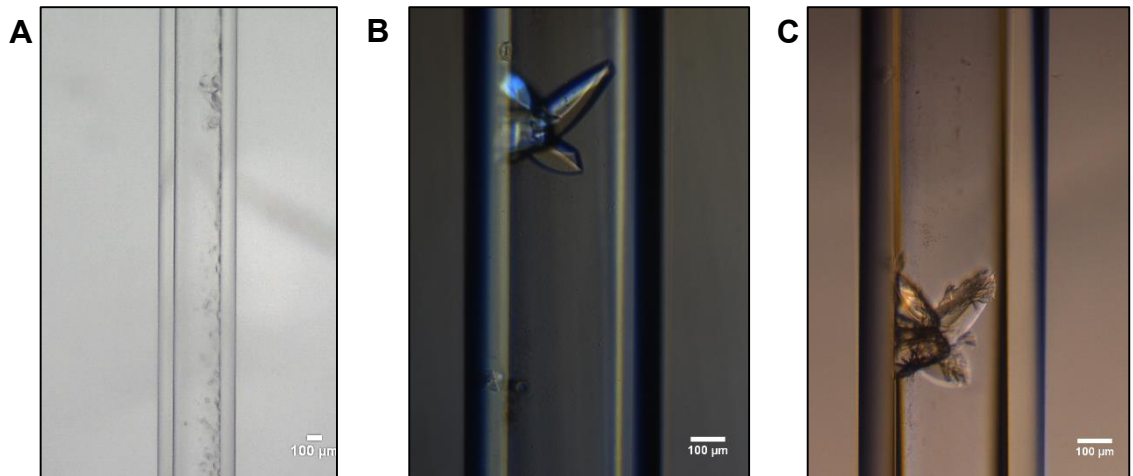


Figura 24: Cristalización por contradifusión de complejo Octarrelina V.1- α Rep en 35% PEG 400.

A) Capilar en la primera semana de su montaje. **B)** Capilar con cristal formado a la semana 3 observado con filtro polarizador. **C)** Cristal en la semana 12 con filtro polarizador.



Resultados Objetivo III: Determinación de estructura cristalina del complejo Octarrelina V.1- α Rep

Data cristalográfica y determinación de estructura del complejo

La difracción y colección de data de los cristales enviados a Francia se llevó a cabo con la colaboración del Dr. Carlos Contreras Martel del "Bacterial Pathogenesis Group" en el "Institut de Biologie Structurale", en las instalaciones del "European Synchrotron Radiation Facility" (ESRF) en Grenoble, Francia.

La data obtenida presenta una formación del cristal en el grupo espacial P 1 21 1 a una resolución de 3.1 Å. En tanto, la resolución de la estructura fue realizada mediante reemplazo molecular utilizando la estructura del complejo Octarrelina- α Rep cristalizada mediante la técnica de gota colgante (PDB:4ZV6).

La tabla 2 presenta un resumen de la data colectada y del proceso de refinamiento realizado a esta. La estructura obtenida se encuentra depositada en el Protein Data Bank bajo el identificador 6BZX.

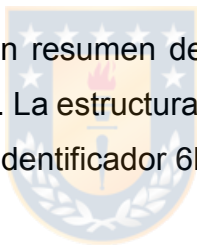


Tabla 2: Resumen de data colectada por rayos X y estadísticas de refinamiento.

Números en paréntesis corresponden a valores para la capa de mayor resolución.

<i>Colección de data</i>		Octarrelina V.1- α Rep
Longitud de onda (Å)		0.979742
Rango de resolución		38.94 - 3.107 (3.22 - 3.11)
Grupo espacial		P 1 21 1
Dimensiones de la celda		
	a,b,c (Å)	70.562, 41.108, 81.409
	α, β, γ (°)	90, 106.567, 90
N° reflexiones observadas/únicas		45861/8115
Llenado de datos (%)		96.75 (87.23)
I/ σ (I)		1,91 (3,12)
R _{sym}		9.8 (104.2)
<i>Refinamiento</i>		
Reflexiones usadas en el refinamiento		8093 (738)
Reflexiones usadas para el R-free		862 (95)
R-work / R-free		0.2453 / 0.3030
N° de átomos		3716
	Macromoléculas	3715
	Ligandos/ion	1
	Proteínas	475
Derivaciones RMS		
	Distancias de enlace (Å)	0.002
	Ángulos de enlace (°)	0.53
	Promedio factores B	125.08
	Macromoléculas	125.10
	Ligandos/ion	81.31

Los datos obtenidos lograron aumentar las zonas visibles dentro del mapa de densidad electrónica de regiones que habían sido poco claras en su predecesora de gota colgante como se ejemplifica en la Figura 25.

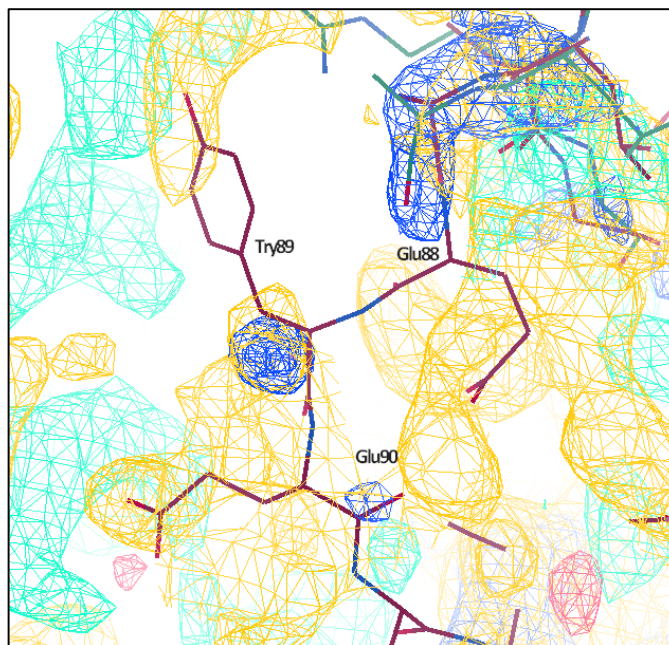


Figura 25: Sección de mapa de densidad electrónica del complejo Octarrelina V.1- α Rep

Región donde se presenta nueva densidad del cristal obtenido por el contra-difusión por capilares. En este caso, una nueva región presentada de la proteína Octarrelina que comprende entre los aminoácidos Glu88-Glu90 es presentada. En azul se muestra el mapa de densidad electrónica del cristal obtenido por gota colgante (PDB: 4ZV6), mientras que en amarillo se presenta el mapa del cristal producido por el contrario-difusión (PDB: 6BZX).

La estructura obtenida presentó una correcta formación del complejo con un solo complejo por celda unidad (Figura 26A). Como se comentó anteriormente, la estructura alcanzada por medio de los datos de densidad electrónica aumentó los aminoácidos visibles, lográndose visualizar un 90% de los aminoácidos que componen la estructura de Octarrelina V.1, además de la visualización de una prolina más que la estructura antes reportada, correspondiente a Pro 9 (Figura 26B). Las nuevas zonas incluidas corresponden principalmente a loops con alta movilidad como Gly8-Asp13, Asn22-Pro26, Ala87-Asp97, Lys120-Lys123, Lys135-Asn147 (Figura 27). Alguno de los sectores que aún quedan inconclusos incluyen los segmentos entre Asp13-Asn22, Gly62-Glu65, Ala114-Ala120, Ala193-Asp198 y Gln214-Ala217.

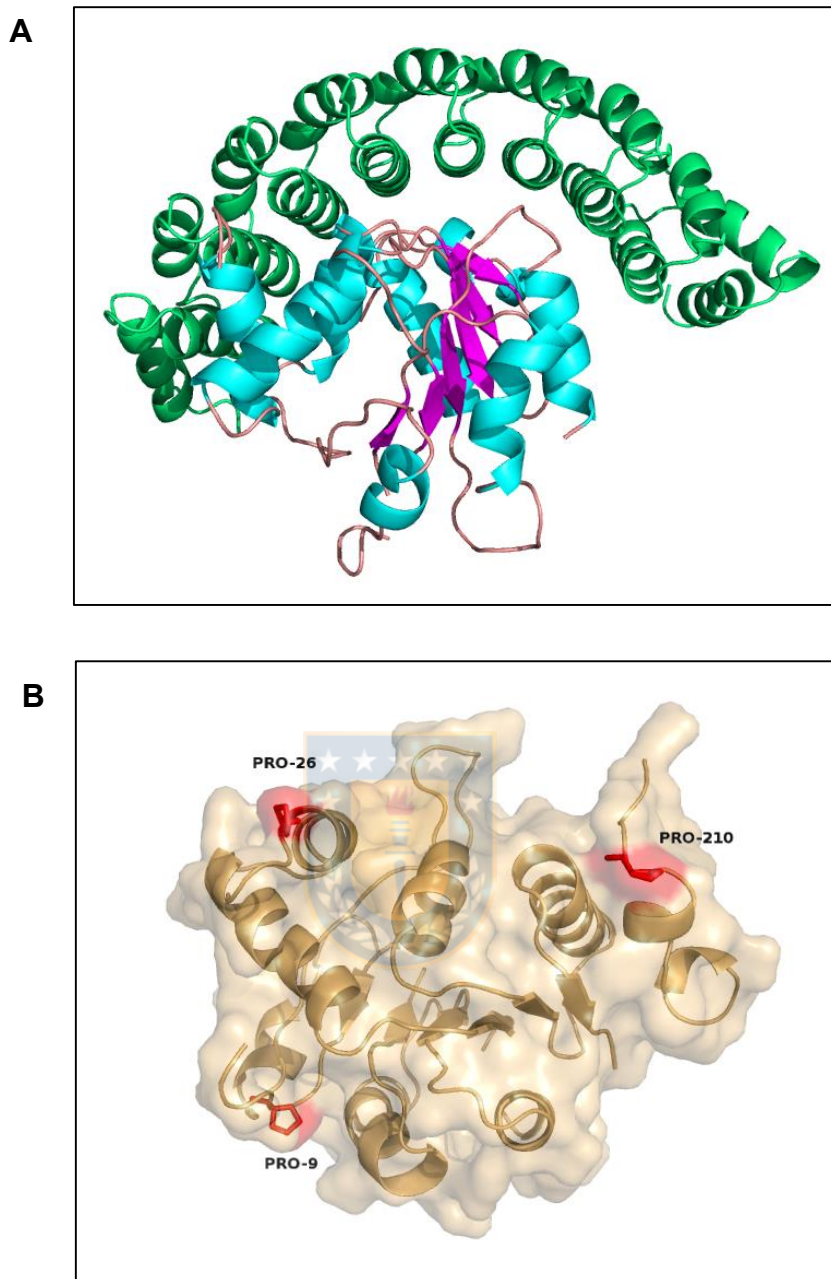


Figura 26: Estructura complejo Octarrelina V.1- α Rep.

En **A** se presenta el complejo en visualización *cartoon*. α Rep se presenta en color verde, mientras que Octarrelina V.1 se presenta en celeste (α -hélices), magenta (laminas β) y rosa (loops). En **B** se presenta la estructura de Octarrelina V.1 en vista de superficie. Residuos de prolina se encuentran destacados en rojo.

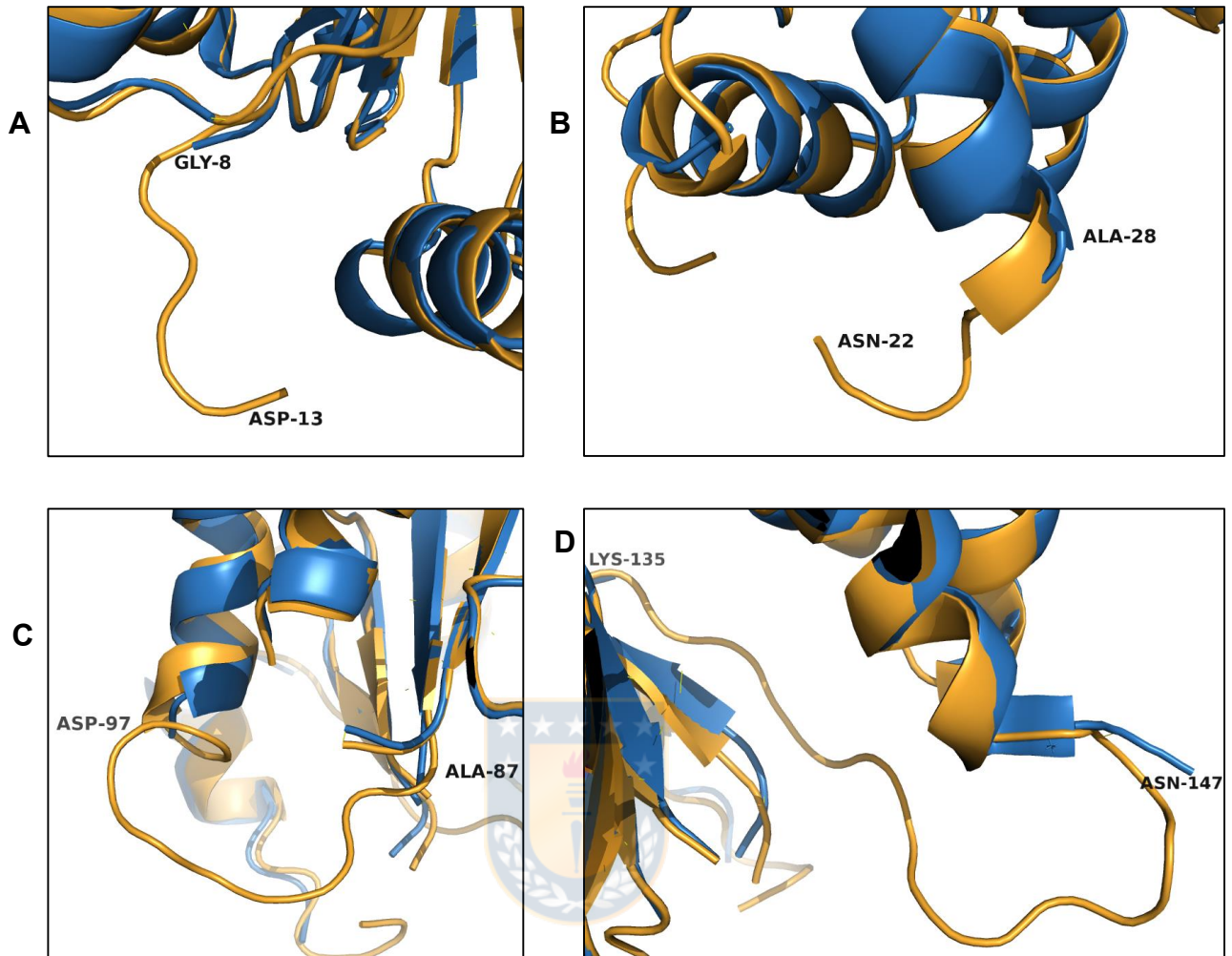


Figura 27 : Comparación estructura complejo Octarrelina V.1- α Rep obtenido mediante cristalización por difusión por vapor y contra-difusión

Zonas del complejo obtenido, identificador PDB: 6BZX (amarillo) se compararon con el complejo PDB: 4ZV6 (azul). (A) Gly8-Asp13, (B) Asn22-Ala28, (C) Ala87-Asp97, (D) Lys135-Asn147.

Resultados Objetivo IV: Producción y purificación de Octarrelinas mutantes de prolina

Purificación mutantes Octarrelina V.1 Δ 4P y Octarrelina V4P

Para la obtención de las mutantes, se aplicó como paso previo al proceso de purificación, una metodología en base a tratamiento enzimático y sonicación a fin de reducir el contenido de ADN presente en las muestras de cuerpos de inclusión donde se encontraban las proteínas mutantes.

Una vez realizado con este procedimiento, fue posible llevar a cabo la cromatografía de intercambio iónico para la purificación y el re-plegamiento proteico, donde en una sola etapa se consiguió ambas cosas para estas proteínas. Una vez realizada la disolución de los cuerpos de inclusión en Urea, el proceso de re-plegamiento se llevó a cabo en la columna cromatográfica de intercambio iónico, durante el primer lavado, al disminuir la presencia del agente caotrópico. La proteína retenida en la columna fue entonces liberada mediante la utilización de un gradiente de NaCl 1M. Para el caso de Octarrelina V.1 Δ 4P las fracciones seleccionadas correspondieron al primer peak obtenido (Figura 28), el cual presentó fracciones con una relación 260/280 igual o menor a 1. Posterior a esta cromatografía, se llevó a cabo una cromatografía de exclusión molecular, donde todas las fracciones con presencia de proteína y bajo contenido de ADN se seleccionaron y almacenaron hasta su uso (Figura 29).

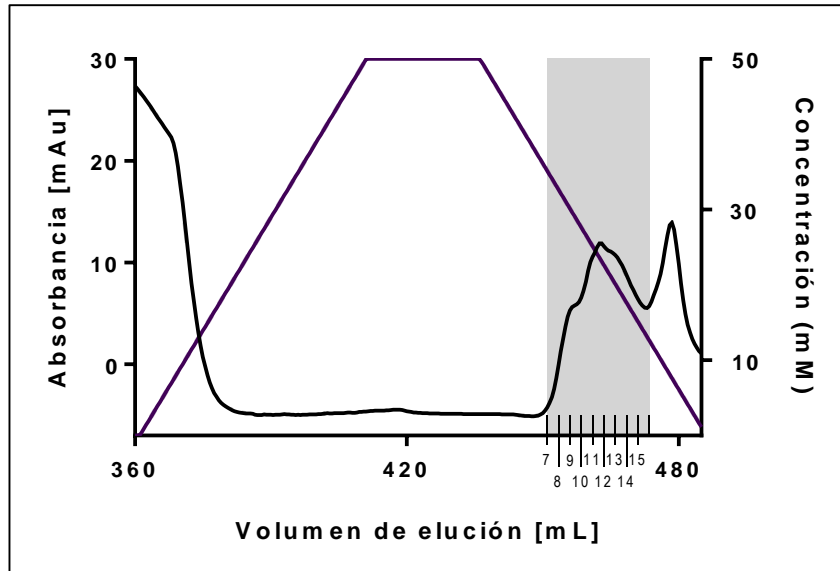
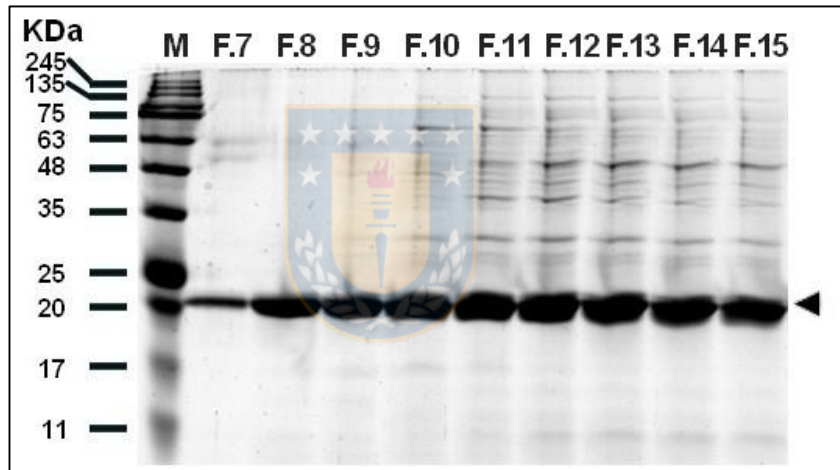
A**B**

Figura 28: Cromatografía de intercambio iónico Octarrelina V.1Δ4P.

A) Cromatograma de cromatografía de re-plegamiento. Línea morada muestra gradientes realizados y lavado con buffer Tris-HCl 50mM pH 8.5 a fin de reducir la concentración de buffer con Urea para el re-plegamiento y posteriormente aumentar la concentración de NaCl para la elución. Peak en gris corresponden a fracciones con presencia de proteína, detectado mediante absorbancia a 280 nm.

B) Gel SDS-Page al 15% con fracciones seleccionadas. Flecha en negro indica peso esperado para Octarrelina V.1Δ4P (24,9 KDa). Fracción 7 fue cargada con aproximadamente 1μg de proteína total, fracción 8 con 3 μg, fracción 9 con 4 μg, fracción 10-12 con 5 μg y fracción 13-14 con 4 μg aproximadamente de proteína total.

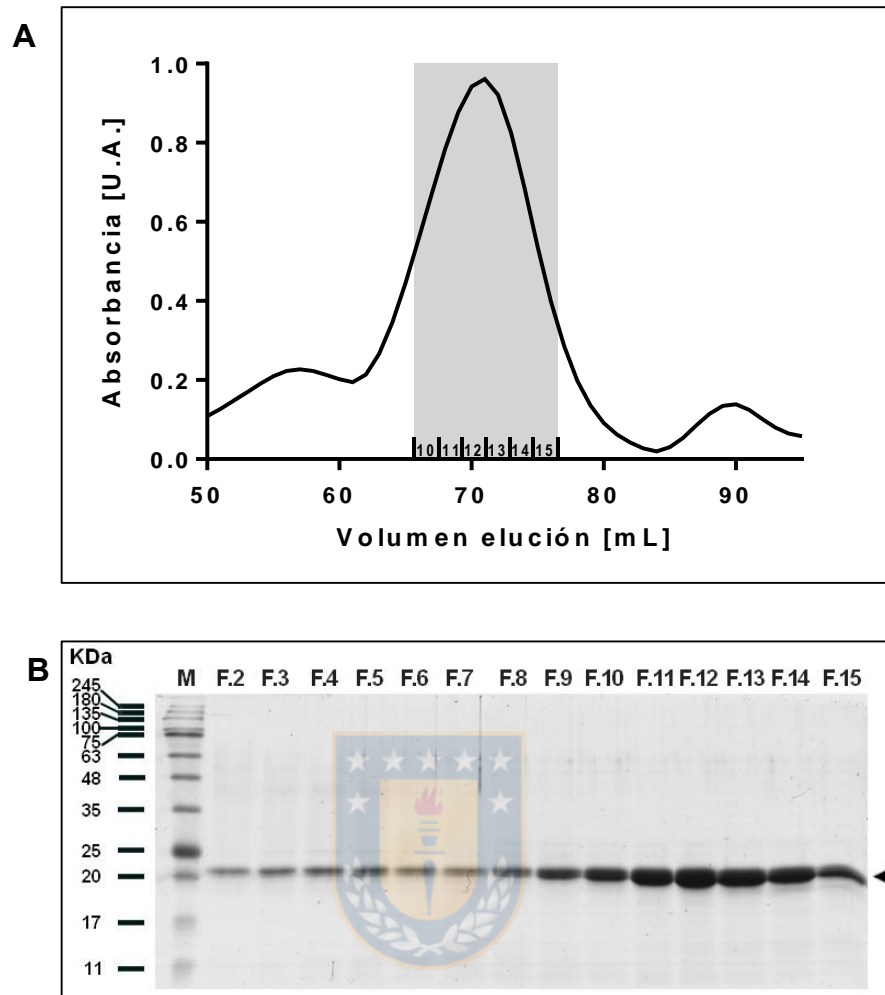


Figura 29: Cromatografía de exclusión molecular Octarrelina V.1Δ4P.

A) Cromatograma de purificación por exclusión molecular de Octarrelina V.1Δ4P (U.A: Unidades Arbitrarias). Peak en gris corresponden a fracciones de interés con presencia de proteína, detectada mediante absorbancia a 280 nm. **B)** SDS-Page al 15% de fracciones analizadas de peak de interés. Flecha en negro presenta peso esperado para Octarrelina V.1Δ4P (24,9 KDa). Fracciones de interés corresponden a últimos carriles del gel. Fracción 10 se cargó con aproximadamente 1 μg de proteína total, fracción 11 con 3 μg, fracción 12 con 6 μg, fracción 13 y 14 con 9 μg y fracción 15 con 7 μg aproximadamente de proteína total.

Para el caso de Octarrelina V4P, al igual que con la cromatografía de intercambio iónico de Octarrelina V.1Δ4P, el primer peak corresponden a las fracciones utilizables para esta cromatografía y con una menor cantidad de ADN (Figura 30).

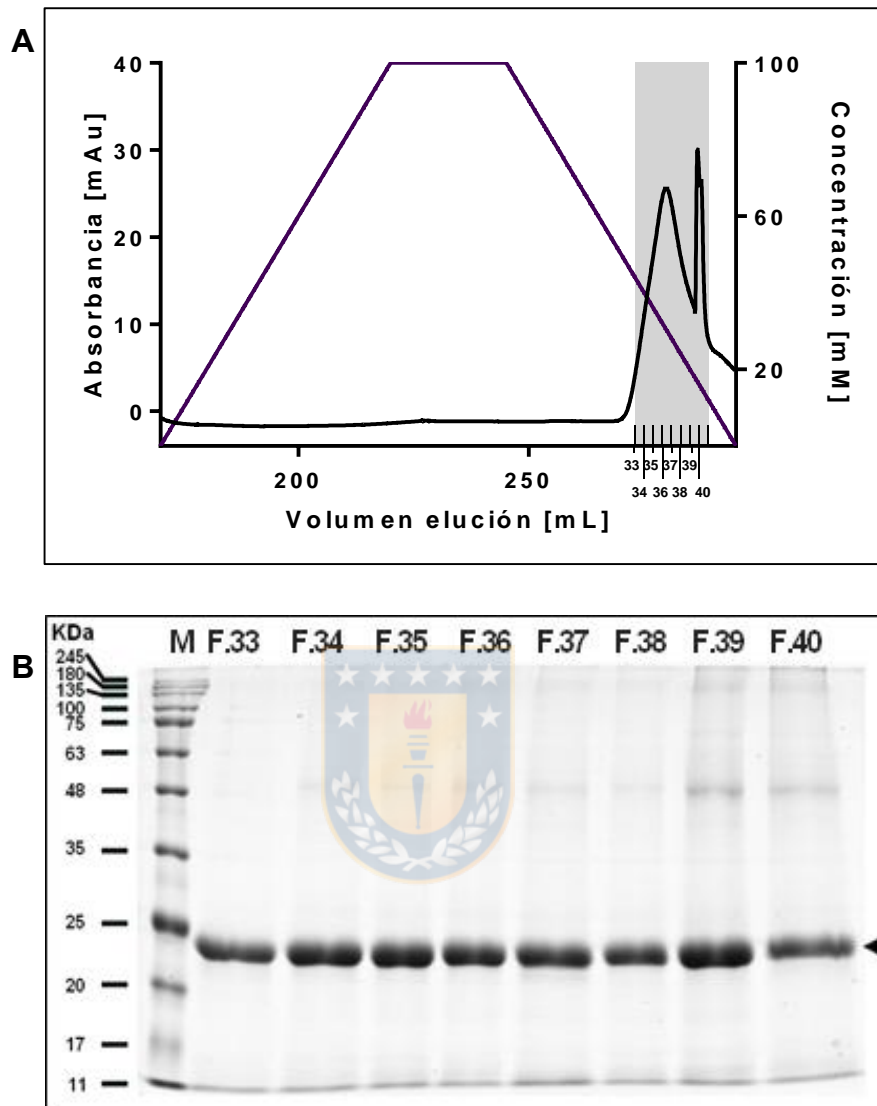


Figura 30: Cromatografía de intercambio iónico Octarrelina V4P

A) Cromatograma de re-plegamiento de proteína. Línea morada muestra gradientes realizados y lavado con buffer Tris-HCl 50mM pH 8.5 a fin de reducir la concentración de buffer con Urea para el re-plegamiento y posteriormente aumentar la concentración de NaCl para la elución. Peaks en gris corresponden a fracciones con presencia de proteína, detectado mediante absorbancia a 280 nm. **B)** SDS-PAGE al 15% de fracciones seleccionadas. Flecha en negro indica peso esperado para proteína (24,9KDa). Fracciones 33 a 40 fueron cargadas con 20 µg de proteína total.

Las fracciones seleccionadas fueron llevadas al paso de cromatografía de exclusión molecular que permitió definir fracciones con presencia de proteínas y baja cantidad de ADN (Figura 31). Sin embargo la producción de esta proteína no superó los 0,5 mL de proteína purificada por litro de cultivo inicial.

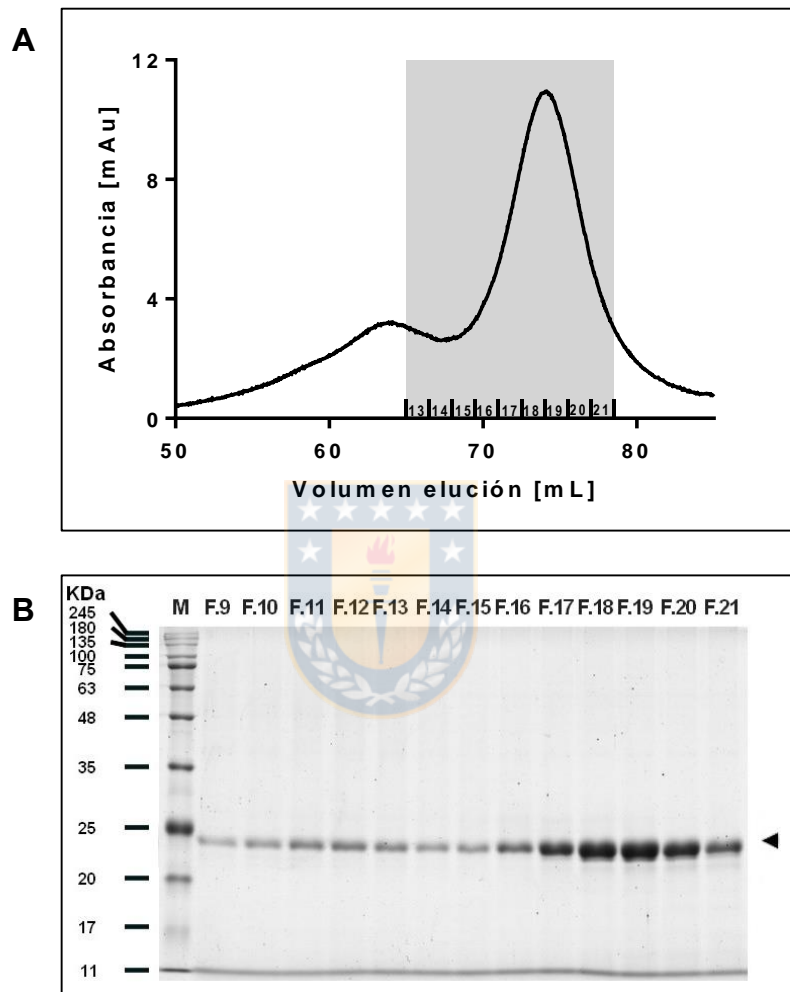


Figura 31: Cromatografía de exclusión molecular Octarrelina iiV4P

A) Cromatograma de purificación por exclusión molecular de Octarrelina V4P. Peak en gris corresponden a fracciones de interés con presencia de proteína, detectada mediante absorbancia a 280 nm. **B)** Análisis SDS-PAGE al 15% de fracciones con presencia de Octarrelina V4P. Flecha en negro presenta peso esperado para Octarrelina V4P (24,9 KDa). Fracciones seleccionadas fueron cargadas con las siguientes cantidades aproximadas de proteína total: Fracción 13, 3 μ g; fracción 14, 2 μ g; fracción 15, 3 μ g, fracción 16, 2 μ g; fracción 17, 1 μ g, fracción 18, 5 μ g; fracción 19, 6 μ g; fracción 20, 4 μ g y fracción 21, 3 μ g.

Los volúmenes y concentraciones de la purificación de las Octarrelinas mutantes se encuentran presentadas en la tabla 3.

Tabla 3: Resumen de purificación proteínas Octarrelina V.1 Δ 4P y Octarrelina V4P

Proteína	Volumen por litro de cultivo de <i>E.coli</i>	Concentración
Octarrelina V.1 Δ 4P	0,5 mL	0,59 mg/mL
Octarrelina V4P	0,33 mL	0,46 mg/mL

Análisis de Octarrelina V4P y Octarrelina V.1 Δ 4P

Los procesos de purificación realizados para las distintas proteínas permitieron obtener muestras de las proteínas de interés para los experimentos. Sin embargo hay una mayor presencia de contaminantes en los concentrados finales de las proteínas mutantes a diferencia de la proteína Octarrelina V.1 (Figura 32), existiendo diferencia de entre un 9 a 20% de pureza entre las mutantes contra esta proteína. Debido a la similitud de pesos moleculares entre Octarrelina V.1 (24.833 Da) y Octarrelinas V.1 Δ 4P y V4P (24.912 y 24.937 Da respectivamente) también es difícil la simple identificación por geles SDS-Page de ambas proteínas, por lo que se realizó un análisis MS/MS del tipo MALDI para la identificación de las mutantes utilizando un gel de similares características al presentado en la figura 32, pero con una carga de 5 μ g de proteína.

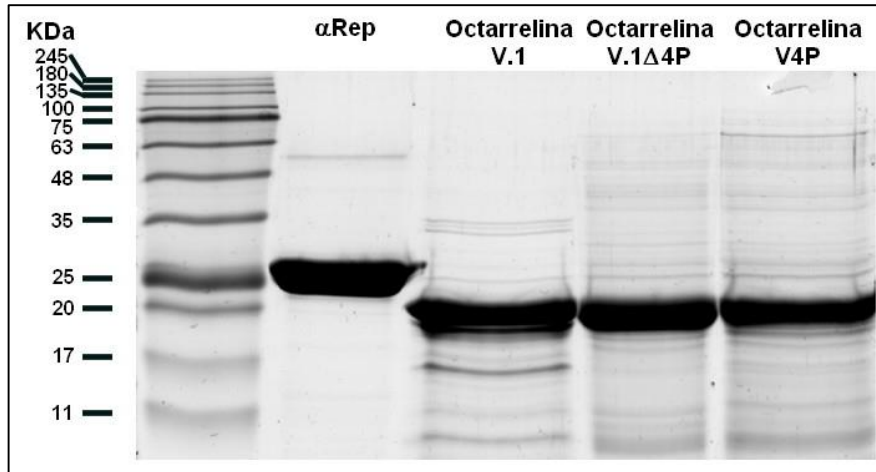


Figura 32: Muestras proteínas de trabajo purificadas

Gel SDS-Page al 15% presenta concentrados finales obtenidos de cada una de las proteínas de trabajo. 15 µg de proteína total fueron cargados en cada uno de los carriles para las muestras de αRep, Octarrelina V.1, Octarrelina V.1Δ4P y Octarrelina V4P.

Las bandas de las mutantes fueron cortadas y enviadas para analizar al Instituto Pasteur de Montevideo en Uruguay. Los resultados obtenidos indican una correcta identificación de cada de una de las mutantes. En el caso de Octarrelina V4P se obtuvo una cobertura de los fragmentos cortados con tripsina de un 78%, mientras que para Octarrelina V.1Δ4P se obtuvo un 72% (Figura 33), confirmándose así la correcta purificación de cada una de las mutantes.

A

```

MAFLIVEGPSEKELKPAVQIANEQGPRAIAFLKQFARNHEKAERFFELLVREGVEAIIIA 60
RGVSEREIEQAAKLAREKGFALAF LAEYERRDRQFDDIIEYFERYGFKAVIVATGLDEK 120
ELKQAAQKIEEKGFKALAF LGRIDQENRNINDIFELLQRQGLRAIIAATGLSERELSWAL 180
RAARQYGLDIIFAYGQFDEQDNQFKHFLEPIRR LGAA 217
  
```

B

```

MAFLIVKGLSEKDLNQAVQIANEQDLSAIAFLKQFARNHEKAERFFELLVREGVEAIIIA 60
RGVSEREIEQAAKLAREKGFALAF LAEYERRDRQFDDIIEYFERYGFKAVIVATGLDEK 120
ELKQAAQKIEEKGFKALAFSGRIDQENHNINDIFELLQRQGLRAIIAATGLSERELSWAQ 180
RAAQYGLDIIFANGQFDEQDNRFKHFLELIR RQAA 217
  
```

Figura 33: Péptidos obtenidos en MALDI-TOF MS/MS identificados en las secuencias de Octarrelinas mutantes.

Péptidos correspondientes a Octarrelina V4P (A) y Octarrelina V.1Δ4P (B).



Resultados Objetivo V: Estudios de interacción entre α Rep y Octarrelinas mutantes

Análisis complejos Octarrelinas con α Rep

Una vez realizada la formación *in vitro* de los complejos, se procedió a su purificación en la columna exclusión molecular. Para el caso del complejo formado con el complejo formado de α Rep con Octarrelina V.1 Δ 4P presenta una curva de bastante amplitud de base y con máximo de peak a los 64,9 mL (Figura 34, A). El gel realizado de esta cromatografía (Figura 34,B) muestra que en las fracciones analizadas del peak se presenta tanto α Rep como Octarrelina V.1 Δ 4P, aunque es notoria la mayor presencia de concentración de la primera proteína.

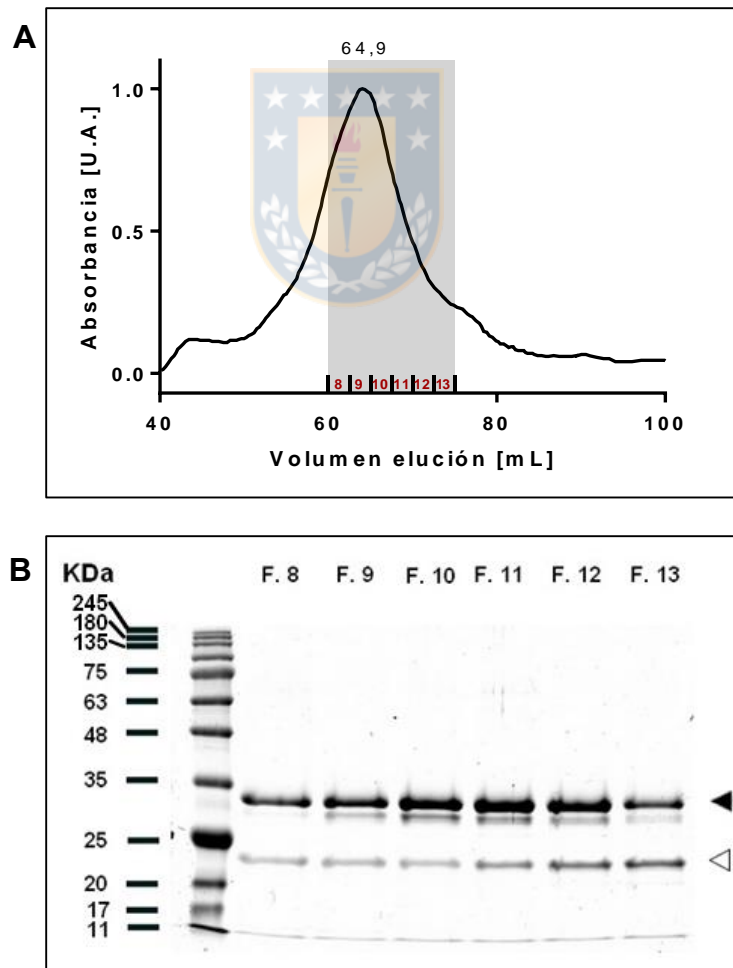


Figura 34: Cromatografía complejo Octarrelina V.1Δ4P-αRep

A) Cromatograma de cromatografía de exclusión molecular Octarrelina V.1Δ4P-αRep. (U.A: Unidades Arbitrarias). Fue cargado 2,5 mL de complejo con una cantidad de proteína total de aproximadamente 4.0 mg. La elución de proteína fue seguida mediante absorbancia a 280 nm. La zona en gris corresponde a la zona analizada mediante gel. **B)** SDS-Page al 15% de fracciones de interés. Fracciones 8 a 13 fueron cargadas con 0,98 μg de proteína total. Flecha en negro indica peso esperado para proteína αRep (30KDa), mientras que flecha en blanco indica peso esperado para Octarrelina V.1Δ4P (24KDa)

A diferencia del complejo anterior, en la formación del complejo αRep con Octarrelina V4P se puede apreciar la presencia de 2 peaks solapados con máximos a los 66,2 mL y 77,4 mL (Figura 35, A). Mientras que el primer peak presentado en el cromatograma se presume la presencia de complejo, el peak que se aprecia a los 77,4mL se encuentra relacionado con proteína Octarrelina V4P que no interaccionó con el ayudante (V_E Octarrelina V4P: 74,1 mL) El análisis mediante SDS-Page del primer peak corrobora lo visto en el cromatograma, presentándose tanto proteína αRep como Octarrelina V4P.



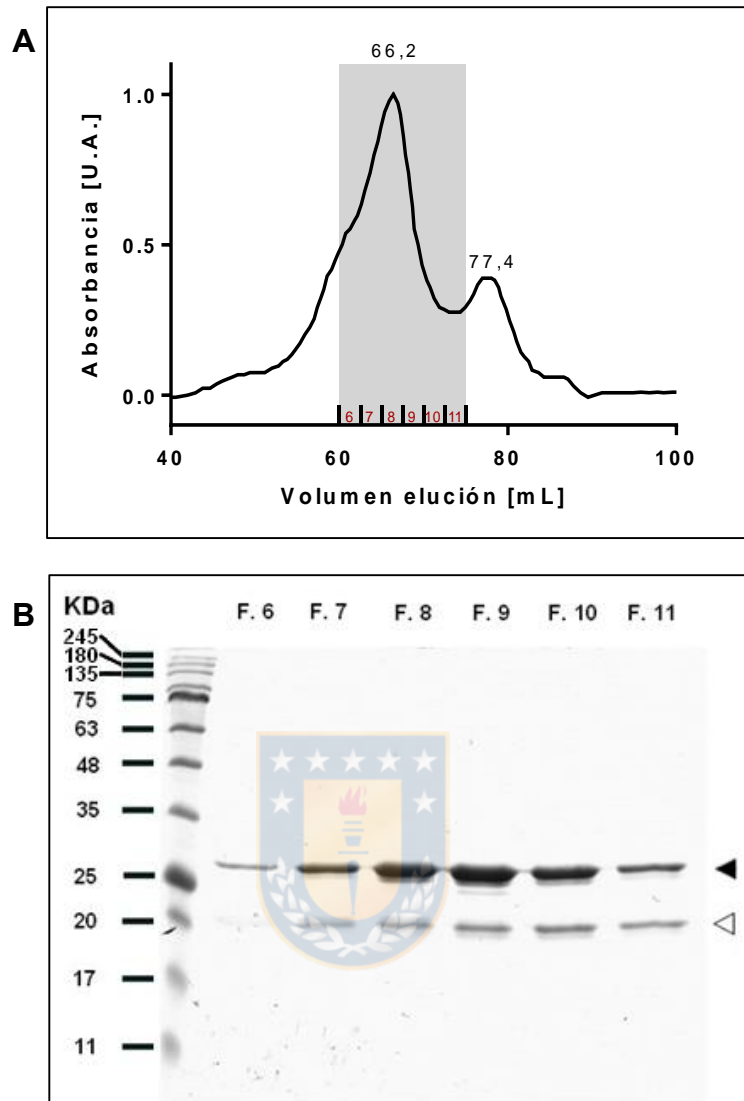


Figura 35: Cromatografía complejo Octarrelina V4P- α Rep

A) Cromatograma de cromatografía de exclusión molecular Octarrelina V4P- α Rep. (U.A: Unidades Arbitrarias). Fue cargado 2,5 mL de complejo con aproximadamente 3,0 mg de proteína total. La elución de proteína fue seguida mediante absorbancia a 280 nm. La zona en gris corresponde a la zona analizada mediante gel. **B)** SDS-Page al 15% de fracciones de interés. Fracciones 6 a 11 fueron cargadas con 0,75 μ g de proteína total. Flecha en negro indica peso esperado para proteína α Rep (30KDa), mientras que flecha en blanco indica peso esperado para Octarrelina V4P (24KDa)

Los peaks obtenidos en los complejos con las proteínas mutantes presentan diferencias destacables en comparación con el complejo formado entre Octarrelina V.1 y α Rep. La figura 36 presenta los perfiles de elución en la columna Hiload™ 16/60 Superdex 75 para Octarrelina V.1 y su complejo con α Rep, donde los pesos moleculares por columna fueron calculados. Es importante señalar en primer lugar que α Rep a pesar de unirse como monómero con octarrelina V.1 como se puede confirmar en los resultados anteriores, se espera sea purificado en una forma dimérica, tal como se encuentra descrito en bibliografía. A pesar de que el tamaño del monómero es de 32kDa, fue visible durante su purificación que el volumen de elución de esta proteína fue a los 61,9 mL, correspondiéndose a un tamaño de proteína calculado por columna cercano a los 100KDa (Figura 36). El peak de octarrelina V.1 presenta un volumen de elución de 74mL, lo que se encuentra cerca del peso molecular esperado de la proteína de 24,9 KDa, siendo el valor calculado por columna de 30KDa. El complejo de ambas proteínas formado y correctamente purificado se presenta a los 68,0 mL, con un peso molecular calculado de 55KDa que es el esperado teórico para este complejo. Es posible apreciar en la figura 36 que α Rep se encuentra unido prácticamente en su totalidad con Octarrelina V.1, lo que queda demostrado con la ausencia de peak del dímero de α Rep en la curva del complejo.

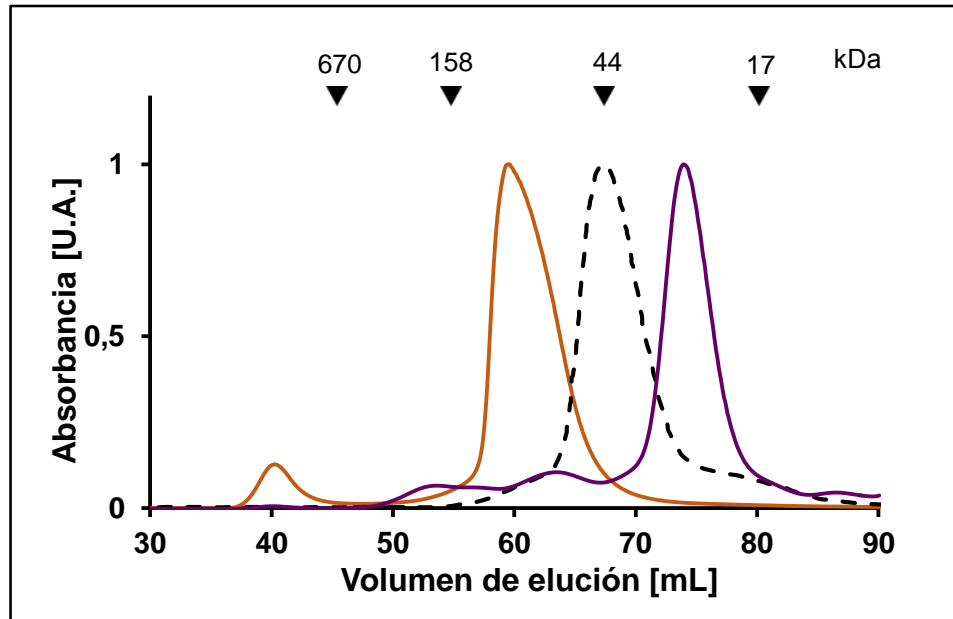


Figura 36: Comparación perfiles de elución proteínas del complejo Octarrelina V.1- α Rep.

Cromatogramas obtenidos para las purificaciones en cromatografía de exclusión molecular de α Rep (línea naranja) y Octarrelina V.1 (línea morada) se compararon con cromatograma obtenido del complejo Octarrelina V.1- α Rep (línea punteada negra). Data se encuentra normalizada para su comparación (A.N.: Absorbancia normalizada). Estándares marcados corresponden a Tiroglobulina (670kDa), γ -globulina (158kDa), Ovoalbúmina (44kDa), Mioglobina (17kDa) y Vitamina B12 (1,350kDa).

Para el caso del complejo con Octarrelina V4P, es posible corroborar mediante la comparación de los diferentes perfiles que el peak de 77,4mL correspondería a proteína OctarrelinaV4P que no se encuentra ligada al complejo, con un tamaño molecular calculado por columna de 20,2KDa (Figura 37). La existencia de complejo con α Rep en poca cantidad puede visualizarse por el desplazamiento del peak del complejo desde los 61,9 mL del homodímero α Rep a los 66,4 mL (figura 37, línea punteada), siendo aún visible en el peak tanto la presencia de homodimero α Rep como complejo con Octarrelina V4P como se pudo visualizar en el gel SDS-Page de la figura 35,B. El peso molecular calculado para este máximo corresponde a 60,7 KDa.

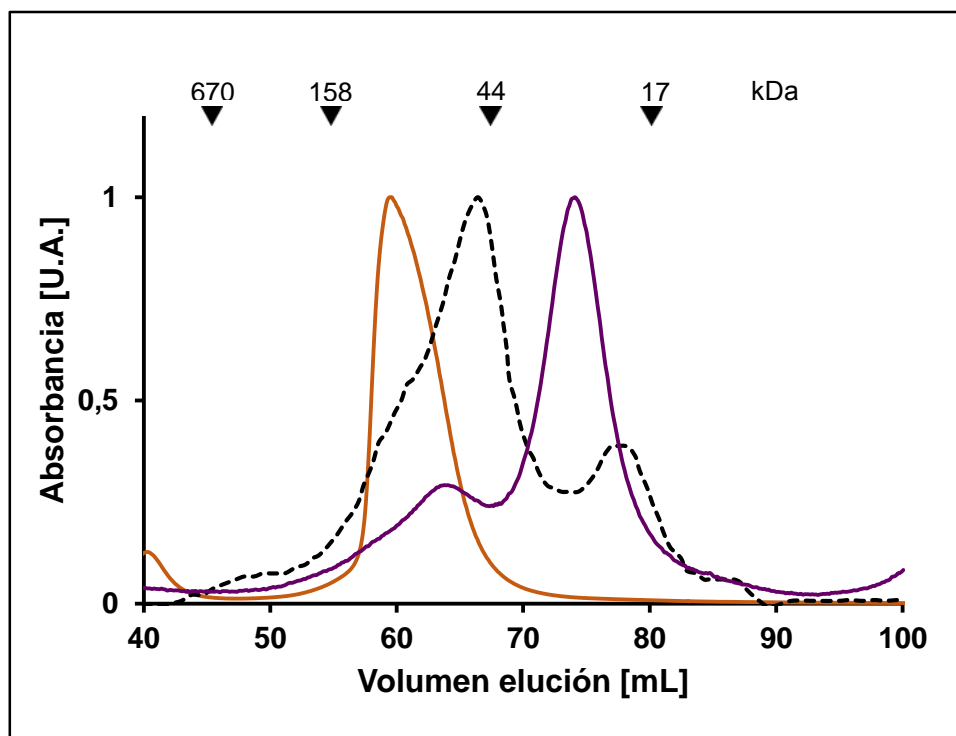


Figura 37: Comparación perfiles de elución proteínas del complejo Octarrelina V4P- α Rep.

Cromatogramas de las purificaciones en cromatografía de exclusión molecular de α Rep (línea naranja) y Octarrelina V4P (línea morada) se compararon con cromatograma obtenido de la purificación del complejo Octarrelina V4P- α Rep (línea punteada negra). Data se encuentra normalizada para su comparación (A.N.: Absorbancia normalizada). Estándares marcados corresponden a Tiroglobulina (670kDa), γ -globulina (158kDa), Ovoalbúmina (44kDa), Mioglobina (17kDa) y Vitamina B12 (1,350kDa).

Para el caso del complejo Octarrelina V.1 Δ 4P – α Rep (Figura 38), no es visible un pick de proteína libre como en el caso de Octarrelina V4P, aunque el ancho del peak y la forma presentada en la parte final del peak del complejo hacen pensar que igualmente quedó parte de Octarrelina libre aunque en menor cantidad que el complejo anterior (V_E Octarrelina V.1d4P: 70,44, peso molecular estimado por columna: 43 KDa). Es apreciable también que el desplazamiento del peak desde los 61,9mL de volumen de elución del homodímero de α Rep a los 64,94mL del complejo que también hubo una aparente formación del complejo. Por lo tanto el peak de complejo presente en la figura 38 se encuentra compuesto

por Octarrelina V.1Δ4P libre, homodímero αRep y complejo Octarrelina V.1Δ4P-αRep, lo que puede dar luces al igual que con octarrelina V4P, de una menor afinidad de las mutantes con αRep. El peso estimado por columna del peak máximo del complejo Octarrelina V.1Δ4P- αRep es de 74,6 KDa.

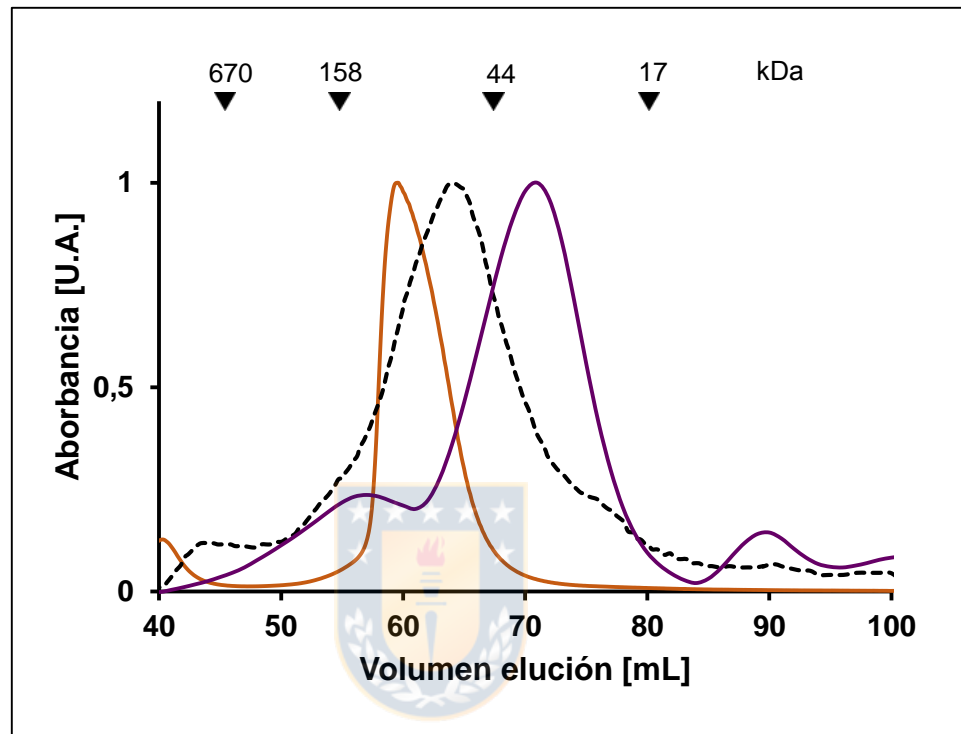


Figura 38: Comparación perfiles de elución proteínas del complejo Octarrelina V.1Δ4P-αRep.

Cromatogramas obtenidos para las purificaciones en cromatografía de exclusión molecular de αRep (línea naranja) y Octarrelina V.1Δ4P (línea morada) se compararon con cromatograma obtenido del complejo Octarrelina V.1Δ4P-αRep (línea punteada negra). Data se encuentra normalizada para su comparación (A.N.: Absorbancia normalizada). Estándares marcados corresponden a Tiroglobulina (670kDa), γ-globulina (158kDa), Ovoalbúmina (44kDa), Mioglobina (17kDa) y Vitamina B12 (1,350kDa).

Análisis interacciones de unión entre α Rep y Octarrelinas

Se estudió mediante micro calorimetría de titulación isotérmica (ITC), la interacción entre α Rep y Octarrelina V.1 u Octarrelina V.1 Δ 4P a fin de poder evaluar las diferencias que pudieran existir. En primer lugar se llevó a cabo un experimento control de unión entre Octarrelina V.1, de acuerdo al experimento de unión previamente realizado para este compuesto (Figuroa et al, 2016). Para ello, muestra de Octarrelina V.1 fue concentrada hasta alcanzar los 0,6 mg/mL a fin de preparar una muestra de 25 μ M. En tanto α Rep fue concentrada hasta los 6,98 mg/mL a fin de cargar una muestra de 216 μ M.

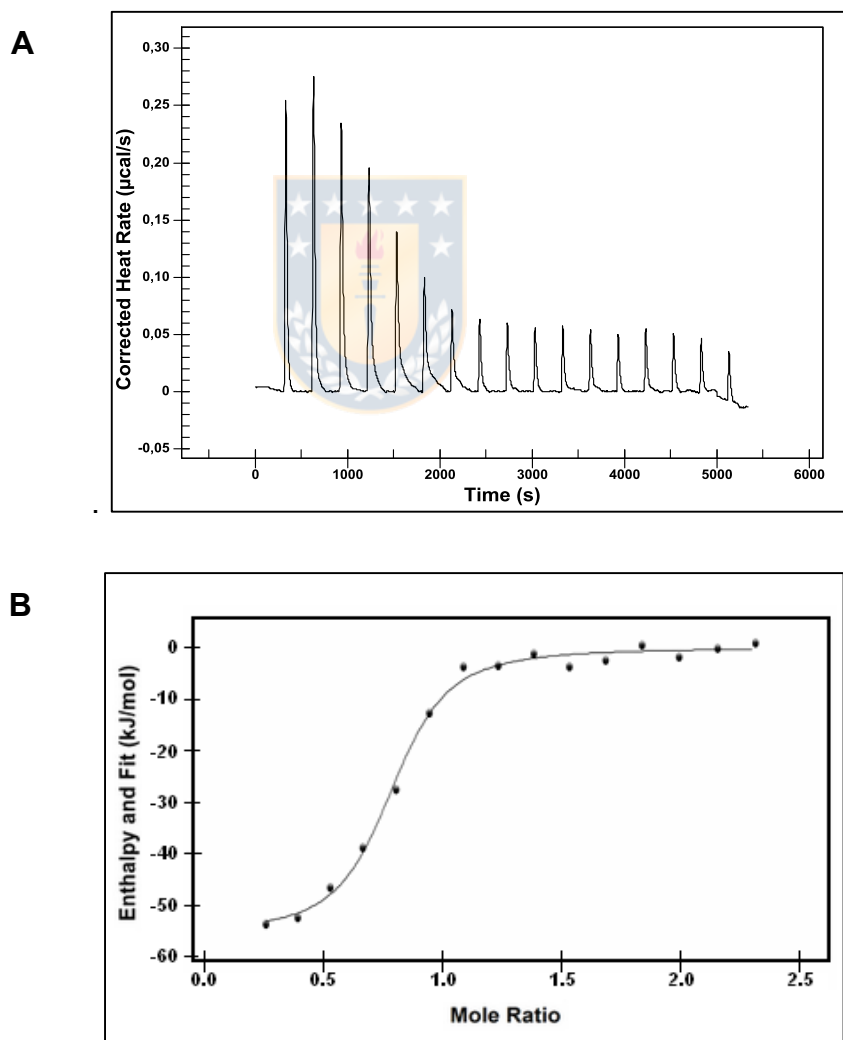


Figura 39: Calorimetría de titulación de α Rep con Octarrelina V.1.

Proteína α Rep a $216\mu\text{M}$ fue inyectada un total de 20 veces a muestra de Octarrelina V.1 en la celda a $25\mu\text{M}$. La data cruda presentada en **A**, fue integrada a fin de obtener la isoterma de unión presentada en panel **B**. La K_d obtenida fue de $0,456\mu\text{M}$ con una estequiometría 1:1 ($n: 0,737$)

La data reportada en este experimento (Figura 39) fue similar en el valor de K_d previamente reportado ($0,45\mu\text{M}$) aunque presentó una variación con respecto al valor de n ($n: 1,16$), a pesar de que se confirma igualmente la unión estequiométrica de 1:1 del complejo.

Debido a las bajas concentraciones y volúmenes obtenidos para octarrelina V4P, no fue posible llevar a cabo experimentos de ITC para esta mutante, por lo que solo fue estudiada la interacción con Octarrelina V.1 Δ 4P. Por otro lado, teniendo en cuenta la información previa respecto a una posible disminución de la afinidad de las octarrelinas mutantes con respecto a α Rep, se llevaron a cabo *screening* de diferentes concentraciones de α Rep con Octarrelina V.1 a una concentración de $25\mu\text{M}$, a fin de poder detectar el límite hasta el cual era posible lograr una detección de señal mediante calorimetría (Figura 40).

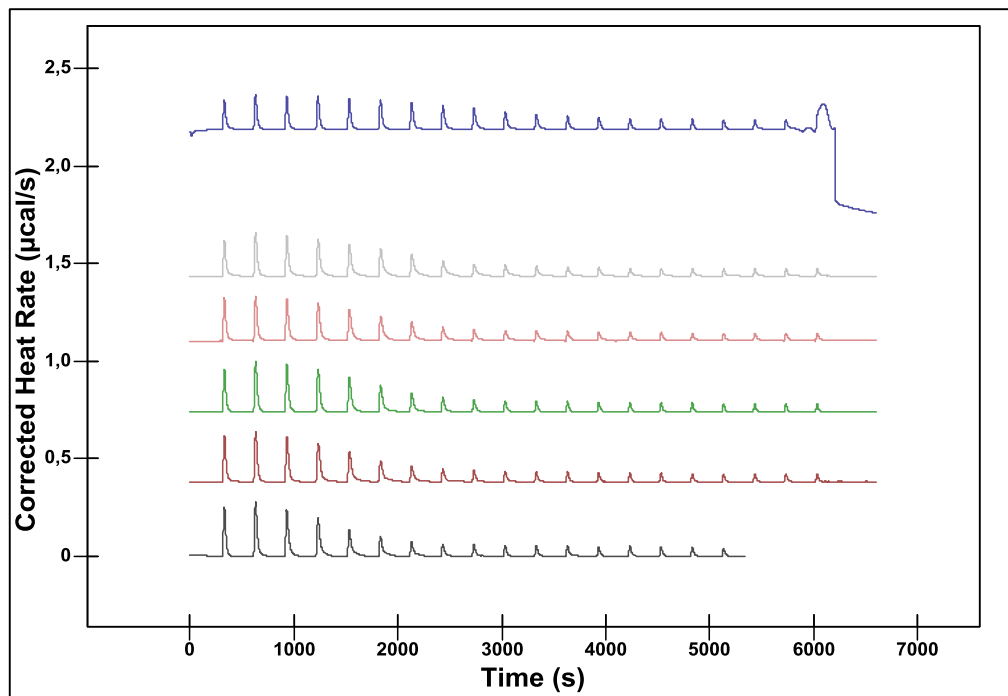


Figura 40: *Screening* de calorimetría de titulación de Octarrelina V.1 a 25 μM con diferentes concentraciones molares de αRep

Stock de αRep fue diluido a fin de obtener diferentes concentraciones molares que permitieran estudiar el límite de detección de señal de calor para la interacción con Octarrelina V.1. De abajo hacia arriba: línea negra, 200 μM ; línea roja, 190 μM ; línea verde, 170 μM ; línea rosa, 160 μM ; línea gris, 150 μM ; línea azul, 100 μM .

Como se puede apreciar en la figura 40, es posible estudiar la interacción de Octarrelina con αRep hasta los 100 μM . Con esto definido, se llevó a cabo un experimento con una concentración de 123 μM de αRep (3,97 mg/mL) con 25 μM de Octarrelina V.1 Δ4P (0,622 mg/mL), donde se puede apreciar una rápida dilución de αRep en la muestra de Octarrelina V.1 Δ4P (Figura 41A). La isoterma de unión nos muestra un alto valor de K_d de 3,13 μM , además de una estequiometría de unión extremadamente baja (n : 0,164). Los valores de K_d del complejo con Octarrelina V.1 para los valores de 100 μM y 150 μM se encuentran entre los 0,588 μM y los 0,548 μM respectivamente, mientras que los de n van desde los 0,657 a 0,714. Esto nos entrega evidencia de una menor afinidad de Octarrelina V.1 Δ4P con respecto a Octarrelina V.1

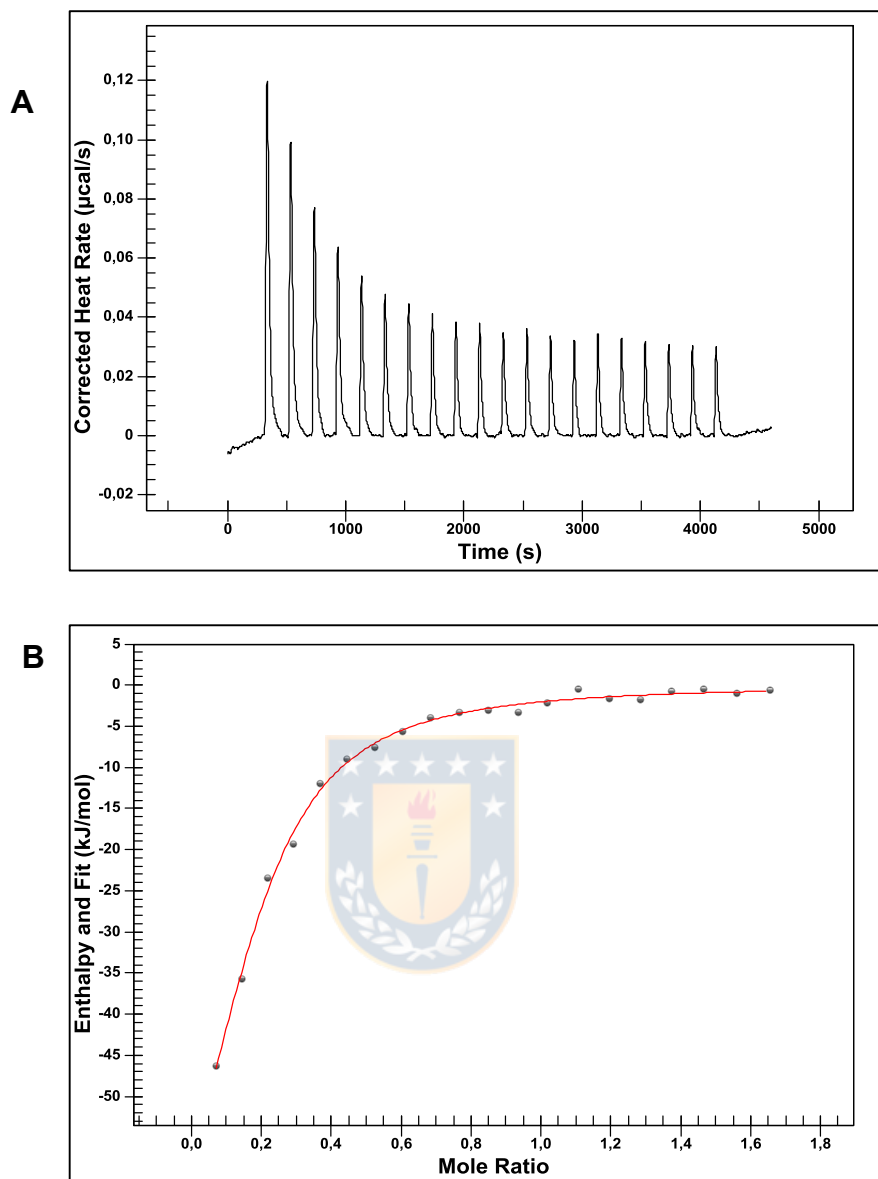


Figura 41: Calorimetría de titulación de αRep con Octarrelina V.1 Δ 4P.

Proteína αRep a $123\mu\text{M}$ fue inyectada un total de 20 veces a muestra de Octarrelina V.1 Δ 4P en la celda a $25\mu\text{M}$. La data cruda presentada en **A**, fue integrada a fin de obtener la isoterma de unión presentada en panel **B**. La K_d obtenida fue de $3,13\mu\text{M}$ con una estequiometría de unión de $n: 0,164$.

A fin de confirmar esto, se decidió aumentar la concentración molar de la mutante Octarrelina V.1 Δ 4P y Octarrelina V.1 a fin de evaluar si existía alguna diferencia entre ambas a concentraciones más similares, sin diluir en extremo la proteína αRep para los experimentos.

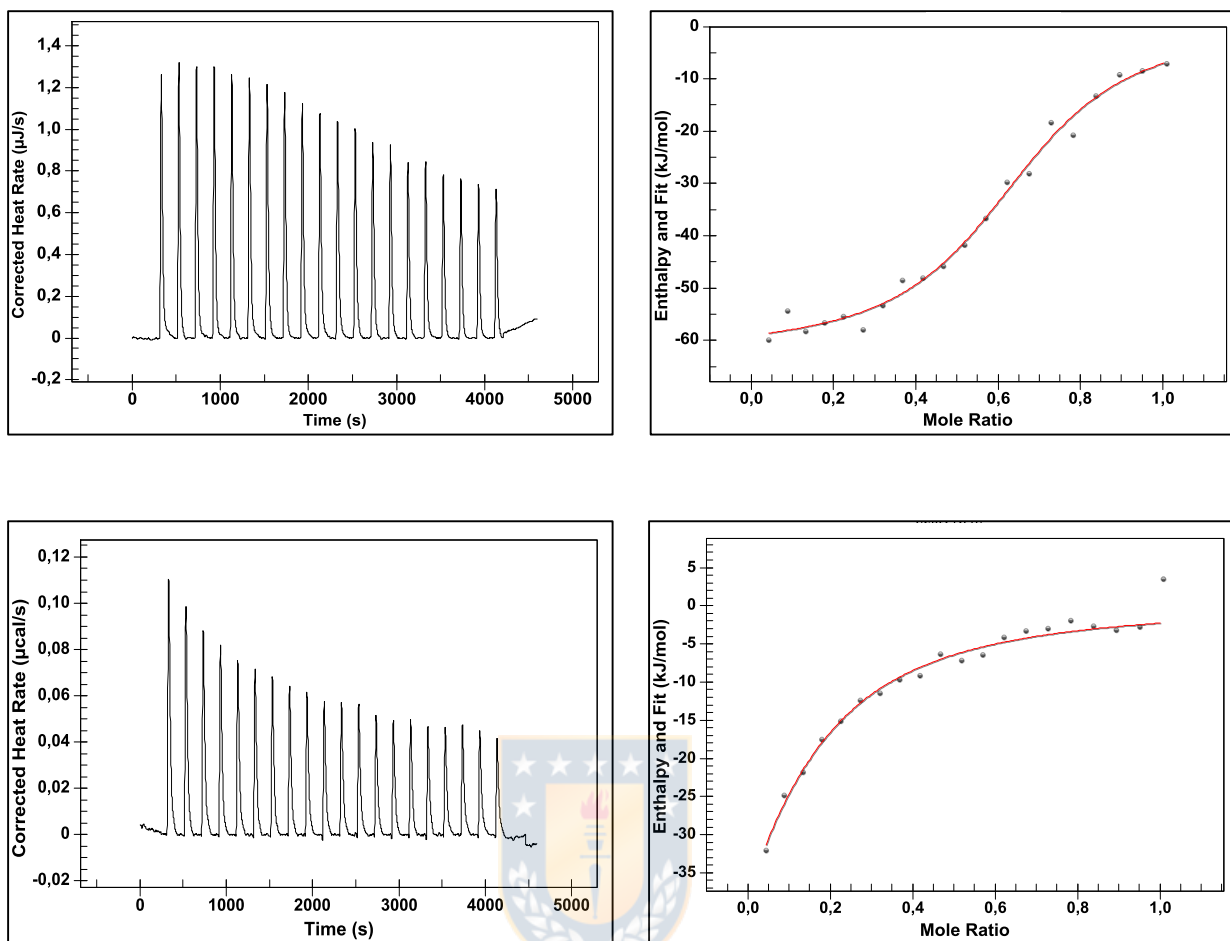


Figura 42: Comparación de calorimetrías de titulación para los complejos Octarrelina V.1 y V.1Δ4P con α Rep

Octarrelina V.1(A) y Octarrelina V.1Δ4P (B) a $41\mu\text{M}$ ($1,02\text{ mg/mL}$) fueron titulados con α Rep a $120\mu\text{M}$ ($3,97\text{ mg/mL}$) en un total de 20 inyecciones. Para cada una de las reacciones fue integrada la data cruda a partir de la cual fue obtenida la isoterma de unión en cada caso. Los valores de K_d y n para Octarrelina V.1 fueron de $1,272\mu\text{M}$ y $0,651$ respectivamente, mientras que estos valores para Octarrelina V.1Δ4P fueron de $11,56\mu\text{M}$ y $0,1$ para K_d y n .

Es apreciable en la figura 42 una diferencia importante en las isotermas obtenidas para los complejos de α Rep con Octarrelina V.1 y V.1Δ4P. En el caso de Octarrelina V.1 es apreciable una reacción casi completa con α Rep, lo que nos reporta un valor de K_d más alto que los apreciados en los experimentos anteriores, pero manteniéndose la estequiometría 1:1 del complejo. Sin embargo Octarrelina V.1Δ4P presenta un valor de estequiometría de unión similar al obtenido a $25\mu\text{M}$ y con una K_d 10 veces más alta que la de Octarrelina V.1, lo

que confirma que con solo una diferencia de 4 aminoácidos, Octarrelina V.1 Δ 4P posee una baja afinidad con α Rep a diferencia de Octarrelina V.1.



DISCUSIÓN

Purificación de proteínas

El proceso de cristalización de proteínas se ve gobernado por múltiples factores. Uno de los principales, y sobre el cual es posible tener una mayor control, es el nivel de pureza de la proteína a cristalizar (McPherson & Gavira, 2014). El uso de técnicas de purificación apropiadas tanto para Octarrelina V.1 como para α Rep, las cuales consideraron los elementos ventajosos para cada una de las proteínas, como la termo-estabilidad en el caso de la primera, y la presencia de un His-tag para la segunda, permitió la purificación de ambas proteínas para realizar la formación de complejos *in vitro* con la obtención de un stock de proteína de un 90% de pureza (Figura 22). A pesar de que la literatura recomienda para el uso de en procesos de cristalización u optimización un 95% de pureza (Rupp, 2010), procesos de cristalización se pueden llevar a cabo con un 90% de pureza, tal como ocurrió en este caso, debido a que el proceso de cristalización puede ser considerado por sí mismo un proceso de purificación (Benvenuti & Mangani, 2007). Un ejemplo de esto puede ser visto en el trabajo de Zang et al., 2011 donde la cristalización de IgG4 monoclonal puede remover en más de un 90% los contaminantes presentes en la solución stock de la proteína. En algunos casos, masas heterogéneas de la solución pueden servir como centros de nucleación en donde los cristales pueden crecer (McPherson et al., 1996). Incluso, algunas proteínas recombinantes pueden ser completamente purificadas desde el extracto crudo de la fracción soluble bacteriana post inducción, como ocurre con la enzima glucosa isomerasa (Visuri, 1984).

Cristalización por contradifusión

La técnica de cristalización utilizada en este trabajo para el complejo α Rep-Octarrelina V.1 también significó una mejora sustancial a la hora de poder determinar una estructura más completa. La estructura originalmente resuelta para este complejo con los diferentes ayudantes de cristalización, Nanobody y α Rep, logró en su conjunto resolver las posiciones de tan solo un 70% de los aminoácidos de la estructura final de Octarrelina V.1 (Figueroa, Sleutel, et al., 2016). En contraste, un 90% de los aminoácidos que constituyen esta proteína fueron visualizados en esta oportunidad solo utilizando α Rep. Elementos que representan una gran dificultad como las zonas de loops de alta movilidad pudieron ser resueltas gracias a esta técnica. El interés en el uso de este ayudante por sobre nanobody se debió a la presencia de una hélice entre Phe204 hasta Gly215 en el C-terminal de la estructura obtenida por gota colgante. Sin embargo en esta nueva estructura, a pesar de confirmarse la presencia de este loop, no se obtuvo una significativa mejora en la resolución de este extremo.

Parte de las mejoras pueden relacionarse al proceso de cristalización por contradifusión. En general el proceso de cristalización se divide principalmente en las etapas de nucleación y crecimiento del cristal. Se ha postulado que el proceso de nucleación, base para comenzar el crecimiento de cristales, se puede llevar a cabo en dos etapas: en primer lugar un clúster de líquido denso debe formarse para que posteriormente un núcleo cristalino pueda formarse al interior de este clúster (Vekilov, 2010). La velocidad de nucleación puede ser controlada variando las propiedades (viscosidad, tamaño y número) de los clusters que se deben formar en el líquido denso (Vekilov, 2012). En el caso de la contradifusión la formación, el número y tamaño de estos clúster se va controlando posterior al encuentro de las fases del agente precipitante con la fase de la proteína. Estableciéndose primeramente el estado de supersaturación y formación de precipitados, las moléculas del precipitado continúan avanzando más allá permitiendo la nucleación de la proteína. Los puntos de nucleación formados se

convierten en un frente que comienza a avanzar combinándose con el fenómeno de la precipitación. Con el paso del tiempo nuevas regiones se ven envueltas en la nucleación y el crecimiento avanza. Debido a que la relación proteína-precipitante es diferente en varias partes del experimento, diferentes condiciones serán analizadas y se optimizarán automáticamente, por lo que existen cristales que crecerán en (o cerca de) las condiciones óptimas de manera ordenada (Carotenuto et al., 2002; Juan Manuel García-Ruiz et al., 2016; Otálora et al., 2009).

Por otra parte, se ha reportado el uso de la técnica de contradifusión para realizar la optimización de las condiciones de cristalización (González-Ramírez et al., 2017). Convencionalmente, en las técnicas de difusión de vapor o batch, diferentes concentraciones de proteína y agente precipitante se ajustaban de manera sistemática. Esto puede ser reducido aun solo ensayo gracias a la contradifusión. Éxitos han sido obtenidos tanto para enzimas (Drebes et al., 2016; Khusaini et al., 2011; Martínez-Rodríguez et al., 2006) como para complejos (Biertümpfel et al., 2002; Ruiz-Martínez et al., 2009), produciéndose, al igual que en este trabajo, grandes cristales y con mejor calidad para poder ser difractados mediante la técnica de rayos X. Otro de los beneficios importantes que esta técnica ofrece es que es un procedimiento simple de aplicar y práctico para transportar, lo que ofrece sobretodo grandes beneficios en la producción de cristales en países como el nuestro alejados de sincrotrones. Los costos y riesgos del transporte convencional de cristales producidos por otros métodos, como difusión de vapor, que implica colecta, congelado y transporte en sistemas fríos con un Dewar (Rupp, 2010), son totalmente eliminados y reemplazados por un simple envío postal hasta el sincrotrón donde se posea disponibilidad de difractar. Este proceso de manipulación y transporte no altera la calidad de los cristales. La adición de crioprotectores mediante el sistema de contradifusión también es fácilmente aplicable, así como la incorporación de átomos pesados para determinación de fases de modo experimental (Ng et al., 2003).

Octarrelinas mutantes

La producción de Octarrelina Mutantes presentó bajos rendimientos en comparación a Octarrelina V.1. En el caso de Octarrelina V4P, Offredi y colaboradores ya habían presentado dificultad de producir Octarrelina V en grandes cantidades y las mutaciones de prolinas L9P, Q16P, L26P y L210P no significaron mejoras en los aspectos de solubilidad y estabilidad por sí solas. Aunque de manera contraria, el reemplazo de las cuatro prolinas en Octarrelina V.1Δ4P ya había mostrado que alteraba su expresión soluble en *E. Coli* (Genco, 2017). Es claro que todas las mutaciones en conjunto adicionadas para llegar a Octarrelina V.1 presentan una gran importancia en el resultado final. Un bajo contenido de prolina en proteínas nativas puede relacionarse con la solubilidad (Schein, 1989). En trabajos que han realizado mejoras de solubilidad y estabilidad utilizando este aminoácido también hablan del efecto acumulativo que las modificaciones introducidas que prolina, junto con otros residuos, poseen para generar las mejoras (Allen et al., 1998; Broom et al., 2017), por lo que el efecto relacionado con otros residuos debe ser considerado.

Procedimientos metodológicos también pueden ser mejorados a fin de obtener una mayor cantidad de proteína. Por ejemplo, mejoras de condiciones durante el cultivo como la temperatura (de Groot & Ventura, 2006; Strandberg & Enfors, 1991), o modificaciones en el proceso de solubilización de los cuerpos de inclusión (Simpson, 2010; Singh & Panda, 2005; Vallejo & Rinas, 2004), además de modificaciones a las técnicas añadidas para la eliminación del contenido de ADN de las muestras que también pueden haber contribuido en la cantidad de proteína finalmente recuperada.

De acuerdo a lo observado durante las purificaciones de Octarrelinas mutantes, es posible plantear que además de la existencia de la forma nativa plegada existen otras formas de la proteína en un plegamiento o grado de compactación diferente. En el caso de las cromatografías de intercambio iónico, todos los peaks obtenidos en las cromatografías mediante el gradiente salino

fueron igualmente analizados y se observó la presencia de proteína en los pesos moleculares esperados para todos ellos (resultados no mostrados). Estas fracciones se descartaron únicamente por el alto contenido de ADN que presentaban (relación Abs 260/280 mayor a 1). Las posteriores cromatografías SEC también dan fuerza a esta idea, pues la presencia de dos máximos y/o hombros en la zona esperada para la elución de Octarrelina darían cuenta de esta observación. La formación de especies plegadas incorrectamente y de agregados suele estar presente en el proceso de re-plegamiento a partir de cuerpos de inclusión. Uno de los elementos que puede haber influenciado en este caso corresponde al método utilizado de dilución de la concentración de agente denaturante, lo que permite la solubilización de los agregados pero no promueve específicamente un plegamiento apropiado (De Bernardez Clark, 1998). Otro factor que puede haber afectado el proceso en el caso de Octarrelina V4P, corresponde a la influencia que prolina puede poseer durante el proceso de re-plegamiento. Es sabido que en proteínas desplegadas existe una mezcla de isómeros *cis-trans* donde aproximadamente un 80% de prolinas se encuentran en conformación *trans*, y un 20% en conformación *cis* (Osváth & Gruebele, 2003; Salahuddin, 1984). Al momento de replegarse pueden, por lo tanto, encontrarse prolinas en una conformación diferente a la conformación nativa, por lo que tienen que esperar a que el proceso de isomerización ocurra (Burgess & Deutscher, 2009) por lo que pueden caer en estados intermedios de plegamiento.

Complejos Octarrelinas mutantes - α Rep

Tanto los experimentos de formación de complejos como la posterior calorimetría isotérmica de titulación dan cuenta de una aparente baja unión entre α Rep y las Octarrelinas mutantes. Para el caso de Octarrelina V4P, en la cromatografía SEC el desplazamiento del peak de los 62,9mL que conforman el homodímero α Rep hacia un volumen de elución de 67mL puede dar luces de la formación de un complejo aunque con menor cantidad de α Rep, ya que parte de la solución restante de la proteína sin interaccionar se encuentra presente en un peak a los

78 mL de volumen de elución. En el caso de Octarrelina V.1 Δ 4P se ve una mayor cantidad de proteína mutante interaccionando con α Rep, pero aún se observa tanto α Rep como Octarrelina V.1 Δ 4P sin interactuar. Durante los experimentos de microcalorimetría, para caracterizar estas interacciones es claramente visible la diferencia de afinidad existente entre α Rep y esta mutante en comparación con Octarrelina V.1. Una de las razones que puede haber causado esta disminución de la afinidad de Octarrelina V.1 Δ 4P con α Rep puede deberse a la presencia de un fenómeno de dimerización de esta proteína. Los pesos moleculares calculados para Octarrelina V.1 y V4P corresponden a 30,1 y 30,4 KDa respectivamente, valores cercanos a los pesos moleculares esperados para estas proteínas (24,8 y 24,9 KDa respectivamente). Sin embargo el peso molecular calculado por columna para Octarrelina V.1 Δ 4P es de 43,0 KDa siendo el peso molecular de la proteína de 24,9 KDa, lo que se encuentra más cerca del tamaño de un posible dímero de esta proteína. Un antecedente más que puede sumarse a la idea se presenta es que cristales de Octarrelina V.1 con el ayudante Nanobody han dado cuenta de una posible formación de un dímero de esta proteína por lo extremos c-terminal (Figuroa et al, 2016). El PDB de esta

estructura no está disponible aun, sin embargo la Figura 43 muestra un modelo de lo observado en este cristal. La posibilidad de la existencia de formas diméricas en ambas mutantes, especialmente de Octarrelina V.1 Δ 4P es posible. Así, puede darse que la mezcla de dímeros de α Rep y los dímeros de Octarrelina no existían



Figura 43: Estructura de dímero Octarrelina V.1 en complejo con ayudante de cristalización NB1

Estructura resuelta a una resolución de 2,9 Å en un grupo espacial I 4₁ 2 2.

las condiciones termodinámicas para lograr la ruptura de los complejos y formación de nuevos complejos Octarrelina- α Rep. Para poder analizar este planteamiento, experimentos de SAXS o DLS pueden ser llevados a cabo para demostrar la presencia en solución de dímeros de mutantes. La búsqueda de versiones de α Rep compatibles dentro de la librería (Urvoas et al., 2010) pueden abrir la posibilidad de realizar nuevos estudios más detallados de la mutantes de prolina, acomplejándola y estabilizándola con esta nueva proteína. Otros ayudantes de cristalización como nanobody, que presenta una mayor afinidad por Octarrelina V.1 que α Rep (Figuroa et al., 2016), también pueden ser explorados como opción para estudiar a Octarrelina V4P y V.1 Δ 4P.

Sin duda alguna, el diseño de proteínas artificiales seguirá siendo un desafío mediante el cual podemos probar nuestro conocimiento respecto al plegamiento y estructura de proteínas. Esta tesis ha sentado las bases para poder estudiar en un futuro cercano las implicancias de ciertos aminoácidos en el plegamiento de una proteína artificial, entregando una nueva estructura proteica artificial de 217 amino ácidos. Esta proteína, que sin bien puede ser considerada como un resultado negativo al no plegarse como se esperaba, entrega un modelo único, sin presión evolutiva, capaz de usarse para caracterizar plegamiento y estructura.

CONCLUSIONES

El diseño *de novo* de proteínas plantea una nueva posibilidad de obtener productos útiles a nivel industrial, pero además, mejorar la comprensión de la bioquímica de las proteínas. La serie de proteínas de más de 200 aminoácidos conocidas como Octarrelinas, buscan alcanzar el plegamiento barril TIM, el cual presenta una variedad de funciones posibles de explotar. Octarrelina V, primera de la serie Octarrelinas diseñada *in silico*, no pudo ser determinada estructuralmente. Mediante el proceso de evolución dirigida se introdujeron mutaciones que permitieron la obtención de una variable más estable y soluble conocida como Octarrelina V.1. Determinada en parte estructuralmente, se pudo ver que no alcanzaba el plegamiento esperado. Sin embargo su uso como modelo de estudio puede entregar datos de interés. La presencia de prolina inesperadas en su secuencia puede entregar información respecto a este aminoácido y las características de solubilidad y estabilidad en proteínas.

En el presente trabajo, el proceso de purificación realizado logró la obtención de las proteínas Octarrelina V.1 y α Rep con importantes rendimientos y baja presencia de contaminantes. Esto pudo llevar a la formación *in vitro* del complejo Octarrelina V.1- α Rep y a su posterior purificación, obteniéndose proteína con un 90% de pureza, una cantidad aceptable para poder conducir los experimentos de cristalización por contradifusión en capilares. Se logró la obtención de cristales, los cuales fueron difractados mediante rayos X, obteniéndose data de difracción que pudo ser utilizada para la generación de una estructura de Octarrelina V.1 completada en un 90% de acuerdo a su secuencia y con la presencia de 3 de las 4 prolina de interés a estudiar, pudiéndose alcanzar así un mejor modelo de estudio para esta proteína a través de una técnica sencilla y fácil de aplicar.

La producción y purificación de las mutantes de prolina, en tanto, generó mutantes correctamente identificadas aunque con bajos rendimientos en su producción. Experimentos de formación de complejos con α Rep para ambas

mutantes presentaron una baja unión con esta y experimentos de microcalorimetría de titulación isotérmica que únicamente fueron posibles de realizar para Octarrelina V.1Δ4P confirma este comportamiento. Prolina puede verse involucrado en los fenómenos de solubilidad y estabilidad modificados en estas mutantes, aunque debe considerarse el efecto de los otros aminoácidos involucrados. La baja afinidad de las mutantes con αRep, en tanto, puede encontrarse más relacionada a una formación de dímeros de proteínas que impiden la formación de los complejos αRep/Octarrelina mutantes.



BIBLIOGRAFÍA

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., & Adams, P. D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallographica Section D Biological Crystallography*, 68(4), 352–367.

Alderson, T. R., Lee, J. H., Charlier, C., Ying, J., & Bax, A. (2018). Propensity for cis-Proline Formation in Unfolded Proteins. *ChemBioChem*, 19(1), 37–42.

Allen, M. J., Coutinho, P. M., & Ford, C. F. (1998). Stabilization of *Aspergillus awamori* glucoamylase by proline substitution and combining stabilizing mutations. *Protein engineering*, 11(9), 783–788.

Andrade, M. A., & Bork, P. (1995, octubre). HEAT repeats in the Huntington's disease protein. *Nature Genetics*, Vol. 11, pp. 115–116.

Andrade, M. A., Perez-Iratxeta, C., & Ponting, C. P. (2001). Protein Repeats: Structures, Functions, and Evolution. *Journal of Structural Biology*, 134(2–3), 117–131.

Andrade, M. A., Petosa, C., O'Donoghue, S. I., Müller, C. W., & Bork, P. (2001). Comparison of ARM and HEAT protein repeats. *Journal of Molecular Biology*, 309(1), 1–18.

Baldock, P., Mills, V., & Stewart, P. S. (1996). A comparison of microbatch and vapour diffusion for initial screening of crystallization conditions. *Journal of Crystal Growth*, 168(1–4), 170–174.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D., & Waley, S. G. (1975). Structure of chicken muscle triose phosphate

isomerase determined crystallographically at 2.5Å resolution: Using amino acid sequence data. *Nature*, 255(5510), 609–614.

Beauregard, M., Goraj, K., Goffin, V., Heremans, K., Goormaghtigh, E., Ruyschaert, J. M., & Martial, J. A. (1991). Spectroscopic investigation of structure in octarellin (a de novo protein designed to adopt the alpha/beta-barrel packing). *Protein engineering*, 4(7), 745–749.

Benvenuti, M., & Mangani, S. (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature protocols*, 2(7), 1633–1651.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic acids research*, 28(1), 235–242.

Betz, S. F., Raleigh, D. P., & DeGrado, W. F. (1993). De novo protein design: from molten globules to native-like states. *Current Opinion in Structural Biology*, 3(4), 601–610.

Biertümpfel, C., Basquin, J., Suck, D., & Sauter, C. (2002). Crystallization of biological macromolecules using agarose gel. *Acta Crystallographica Section D: Biological Crystallography*, 58(10 II), 1657–1659.

Binz, H. K., Amstutz, P., & Plückthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. *Nature Biotechnology*, 23(10), 1257–1268.

Binz, H. K., & Plückthun, A. (2005). Engineered proteins as specific binding reagents. *Current Opinion in Biotechnology*, 16(4), 459–469.

Brändén, C.-I. (1991). The TIM barrel—the most frequently occurring folding motif in proteins. *Current Opinion in Structural Biology*, 1(6), 978–983.

Broom, A., Jacobi, Z., Trainor, K., & Meiering, E. M. (2017). Computational tools help improve protein stability but with a solubility tradeoff.

The Journal of biological chemistry, 292(35), 14349–14361.

Bukowska, M. A., & Grütter, M. G. (2013). New concepts and aids to facilitate crystallization. *Current Opinion in Structural Biology*, 23, 409–416.

Burgess, R. R., & Deutscher, M. P. (2009). *Guide to Protein Purification* (2^a ed.). San Diego, United State of America: Elsevier Science & Technology Books.

Carey, P. R. (1996). *Protein Engineering and Design* (2^a ed.). San Diego, United State of America: Elsevier Science.

Carotenuto, L., Piccolo, C., Castagnolo, D., Lappa, M., Tortora, A., & García-Ruiz, J. M. (2002). Experimental observations and numerical modelling of diffusion-driven crystallisation processes. *Acta Crystallographica Section D Biological Crystallography*, 58(10), 1628–1632.

Chaikuad, A., Keates, T., Vincke, C., Kaufholz, M., Zenn, M., Zimmermann, B., Gutiérrez, C., Zhang, R.-G., Hatzos-Skintges, C., Joachimiak, A., Muyldermans, S., Herberg, F. W., Knapp, S., & Müller, S. (2014). Structure of cyclin G-associated kinase (GAK) trapped in different conformations using nanobodies. *The Biochemical journal*, 459(1), 59–69.

Chen, K., & Arnold, F. H. (1993). Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proceedings of the National Academy of Sciences*, 90(12), 5618–5622.

Cheon, Y. H., Park, H. S., Kim, J. H., Kim, Y., & Kim, H. S. (2004). Manipulation of the active site loops of D-hydantoinase, a (β/α)₈-barrel protein, for modulation of the substrate specificity. *Biochemistry*, 43(23), 7413–7420.

Chino, M., Maglio, O., Nastri, F., Pavone, V., DeGrado, W. F., & Lombardi, A. (2015). Artificial Diiron Enzymes with a De Novo Designed Four-Helix Bundle

Structure. *European Journal of Inorganic Chemistry*, 2015(21), 3371–3390.

Cobb, R. E., Chao, R., & Zhao, H. (2013). Directed Evolution: Past, Present and Future. *AIChE journal. American Institute of Chemical Engineers*, 59(5), 1432–1440.

Correia, B., Ban, Y.-E. A., Friend, D., Ellingson, K., Xu, H., Boni, E., Bradley-Hewitt, T., Bruhn-Johannsen, J., Stamatatos, L., Strong, R., & Schief, W. (2011). Computational Protein Design Using Flexible Backbone Remodeling and Resurfacing: Case Studies in Structure-Based Antigen Design. *Journal of Molecular Biology*, 405(1), 284–297.

De Bernardez Clark, E. (1998). Refolding of recombinant proteins. *Current Opinion in Biotechnology*, 9(2), 157–163.

de Groot, N. S., & Ventura, S. (2006). Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Letters*, 580(27), 6471–6476.

Deber, C. M. & Brodsky, B. (2001). Proline Residues in Proteins. In *Encyclopedia of Life Sciences*. pp 1-7. London, UK: Nature Publishing Co.

Desjarlais, J. R., & Handel, T. M. (1995). De novo design of the hydrophobic cores of proteins. *Protein science : a publication of the Protein Society*, 4(10), 2006–2018.

Desmet, J., Maeyer, M. De, Hazes, B., & Lasters, I. (1992). The dead-end elimination theorem and its use in protein side-chain positioning. *Nature*, 356(6369), 539–542.

Domanska, K., Vanderhaegen, S., Srinivasan, V., Pardon, E., Dupeux, F., Marquez, J. A., Giorgetti, S., Stoppini, M., Wyns, L., Bellotti, V., & Steyaert, J. (2011). Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic 2-microglobulin variant. *Proceedings of the National Academy of Sciences*, 108(4), 1314–1319.

Doyle, L., Hallinan, J., Bolduc, J., Parmeggiani, F., Baker, D., Stoddard, B. L., & Bradley, P. (2015). Rational design of α -helical tandem repeat proteins with closed architectures. *Nature*, 528(7583), 585–588.

Drebes, J., Künz, M., Windshügel, B., Kikhney, A. G., Müller, I. B., Eberle, R. J., Oberthür, D., Cang, H., Svergun, D. I., Perbandt, M., Betzel, C., & Wrenger, C. (2016). Structure of ThiM from Vitamin B1 biosynthetic pathway of *Staphylococcus aureus* - Insights into a novel pro-drug approach addressing MRSA infections. *Scientific Reports*, 6(1), 22871.

Drenth, J., & Mesters, J. (2007). Principles of protein X-ray crystallography: Third edition. In *Principles of Protein X-Ray Crystallography: Third Edition* (Vol. 9780387333).

Eisenbeis, S., Proffitt, W., Coles, M., Truffault, V., Shanmugaratnam, S., Meiler, J., & Höcker, B. (2012). Potential of fragment recombination for rational design of proteins. *Journal of the American Chemical Society*, 134(9), 4019–4022.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D Biological Crystallography*, 66(4), 486–501.

Fang, Y., Jayaram, H., Shane, T., Kolmakova-Partensky, L., Wu, F., Williams, C., Xiong, Y., & Miller, C. (2009). Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature*, 460(7258), 1040–1043.

Farid, T. A., Kodali, G., Solomon, L. A., Lichtenstein, B. R., Sheehan, M. M., Fry, B. A., Bialas, C., Ennist, N. M., Siedlecki, J. A., Zhao, Z., Stetz, M. A., Valentine, K. G., Anderson, J. L. R., Wand, A. J., Discher, B. M., Moser, C. C., & Dutton, P. L. (2013). Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nature chemical biology*, 9(12), 826–833.

Figuroa, M., Oliveira, N., Lejeune, A., Kaufmann, K. W., Dorr, B. M.,

Matagne, A., Martial, J. A., Meiler, J., Van de Weerd, C. (2013). Octarellin VI: Using Rosetta to Design a Putative Artificial (β/α)₈ Protein. *PLoS ONE*, 8(8), e71858.

Figuroa, M., Sleutel, M., Vandevenne, M., Parvizi, G., Attout, S., Jacquin, O., Vandenameele, J., Fischer, A. W., Damblon, C., Goormaghtigh, E., Valerio-Lepiniec, M., Urvoas, A., Durand, D., Pardon, E., Steyaert, J., Minard, P., Maes, D., Meiler, J., Matagne, A., Van de Weerd, C. (2016). The unexpected structure of the designed protein octarellin V.1 forms a challenge for protein structure prediction tools. *Journal of Structural Biology*, 195(1), 19–30.

Figuroa, M., Vandenameele, J., Goormaghtigh, E., Valerio-Lepiniec, M., Minard, P., Matagne, A., & Weerd, C. Van De. (2016). Biophysical characterization data of the artificial protein Octarellin V.1 and binding test with its X-ray helpers. *Data in Brief*, 8, 1221–1226.

Floudas, C. A., Fung, H. K., McAllister, S. R., Mönnigmann, M., & Rajgaria, R. (2006). Advances in protein structure prediction and de novo protein design: A review. *Chemical Engineering Science*, 61(3), 966–988.

Fortenberry, C., Bowman, E. A., Proffitt, W., Dorr, B., Combs, S., Harp, J., Mizoue, L., & Meiler, J. (2011). Exploring symmetry as an avenue to the computational design of large protein domains. *Journal of the American Chemical Society*, 133(45), 18026–18029.

Freydell, E. J., van der Wielen, L., Eppink, M., & Ottens, M. (2010). Ion-exchange chromatographic protein refolding. *Journal of Chromatography A*, 1217(46), 7265–7274.

Garcia-Ruiz, J. M., Gonzalez-Ramirez, L. A., Gavira, J. A., & Otálora, F. (2002). Granada Crystallisation Box: a new device for protein crystallisation by counter-diffusion techniques. *Acta Crystallographica Section D Biological Crystallography*, 58(10), 1638–1642.

García-Ruiz, Juan Ma. (2003). Counterdiffusion Methods for Macromolecular Crystallization. *Methods in Enzymology*, 368, 130–154.

García-Ruiz, Juan Manuel, Otálora, F., & García-Caballero, A. (2016). The role of mass transport in protein crystallization. *Acta crystallographica. Section F, Structural biology communications*, 72(Pt 2), 96–104.

Gaseidnes, S., Synstad, B., Jia, X., Kjellesvik, H., Vriend, G., & Eijsink, V. G. H. (2003). Stabilization of a chitinase from *Serratia marcescens* by Gly-->Ala and Xxx-->Pro mutations. *Protein Engineering Design and Selection*, 16(11), 841–846.

Genco, A. (2017). Protocol standardization to purify Octarellin V mutants and biophysical characterization. Recuperado de <https://etd.adm.unipi.it/t/etd-07022017-164600/>

Gilbreth, R. N., & Koide, S. (2012). Structural insights for engineering binding proteins based on non-antibody scaffolds. *Current Opinion in Structural Biology*, 22(4), 413–420.

González-Ramírez, L. A., Ruiz-Martínez, C. R., Estremera-Andújar, R. A., Nieves-Marrero, C. A., García-Caballero, A., Gavira, J. A., López-Garriga, J., & García-Ruiz, J. M. (2017). Efficient screening methodology for protein crystallization based on the counter-diffusion technique. *Crystal Growth and Design*, 17(12), 6780–6786.

Goraj, K., Renard, A., & Martial, J. A. (1990). Synthesis, purification and initial structural characterization of octarellin, a de novo polypeptide modelled on the α/β -barrel Proteins. *Protein Engineering, Design and Selection*, 3(4), 259–266.

Groves, M. R., & Barford, D. (1999). Topological characteristics of helical repeat protein. *Current Opinion in Structural Biology*, 9(3), 383–389.

Guellouz, A., Valerio-Lepiniec, M., Urvoas, A., Chevrel, A., Graille, M., Fourati-Kammoun, Z., Desmadril, M., van Tilbeurgh, H., & Minard, P. (2013). Selection of Specific Protein Binders for Pre-Defined Targets from an Optimized Library of Artificial Helicoidal Repeat Proteins (alphaRep). *PLoS ONE*, 8(8).

Hardy, F., Vriend, G., Veltman, O. R., van der Vinne, B., Venema, G., & Eijssink, V. G. H. (1993). Stabilization of *Bacillus stearothermophilus* neutral protease by introduction of prolines. *FEBS Letters*, 317(1–2), 89–92.

Hellinga, H. W., & Richards, F. M. (1994). Optimal sequence selection in proteins of known structure by simulated evolution. *Proc Natl Acad Sci USA*, 91(13), 5803–5807

Hermanson, G. T. (1997). Bioconjugate techniques. *General Pharmacology: The Vascular System*, 28(1), 167.

Höcker, B., Lochner, A., Seitz, T., Claren, J., & Sterner, R. (2009). High-Resolution Crystal Structure of an Artificial (Ba)₈-Barrel Protein Designed. *Biochemistry*, 48, 1145–1147.

Houbrechts, A., Moreau, B., Abagyan, R., Mainfroid, V., Préaux, G., Lamproye, A., Poncin, A., Goormaghtigh, E., Ruyschaert, J. M., & Martial, J. A. (1995). Second-generation octarellins: two new de novo (beta/alpha)₈ polypeptides designed for investigating the influence of beta-residue packing on the alpha/beta-barrel structure stability. *Protein engineering*, 8(3), 249–259.

Huang, P.-S. S., Feldmeier, K., Parmeggiani, F., Velasco, D. A. F., Hocker, B., Baker, D., Fernandez Velasco, D. A., Höcker, B., Baker, D., Velasco, D. A. F., Höcker, B., & Baker, D. (2016). De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nature Chemical Biology*, 12(1), 29–34.

Huang, P. S., Boyken, S. E., & Baker, D. (2016). The coming of age of de novo protein design. *Nature*, Vol. 537, pp. 320–327.

Jäckel, C., Kast, P., & Hilvert, D. (2008). Protein Design by Directed Evolution. *Annual Review of Biophysics*, 37(1), 153–173.

Jackrel, M. E., Cortajarena, A. L., Liu, T. Y., & Regan, L. (2010). Screening Libraries To Identify Proteins with Desired Binding Activities Using a Split-GFP Reassembly Assay. *ACS Chemical Biology*, 5(6), 553–562.

Jurgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M., & Sterner, R. (2000). Directed evolution of a (beta alpha)₈-barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proceedings of the National Academy of Sciences*, 97(18), 9925–9930.

Kajava, A. V. (2001). Review: Proteins with Repeated Sequence—Structural Prediction and Modeling. *Journal of Structural Biology*, 134(2–3), 132–144.

Khoury, G. A., Smadbeck, J., Kieslich, C. A., & Floudas, C. A. (2014). Protein folding and de novo protein design for biotechnological applications. *Trends in Biotechnology*, 32(2), 99–109.

Khusaini, M. S., Rahman, R. N. Z. R. A., Mohamad Ali, M. S., Leow, T. C., Basri, M., & Salleh, A. B. (2011). Crystallization and preliminary X-ray crystallographic analysis of a thermostable organic solvent-tolerant lipase from *Bacillus* sp. strain 42. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 67(Pt 3), 401–403.

King, N. P., Sheffler, W., Sawaya, M. R., Vollmar, B. S., Sumida, J. P., Andre, I., Gonen, T., Yeates, T. O., & Baker, D. (2012). Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science*, 336(6085), 1171–1174.

Klostermeier, D., & Rudolph, M. G. (2017). *Biophysical Chemistry* (2^a ed.). Boca Raton, United State of America: CRC Press, Taylor & Francis Group.

Koga, N., Tatsumi-Koga, R., Liu, G., Xiao, R., Acton, T. B., Montelione, G. T., & Baker, D. (2012). Principles for designing ideal protein structures. *Nature* 2012 491:7423, 491(7423), 222.

Koide, S. (2009, agosto). Engineering of recombinant crystallization chaperones. *Current Opinion in Structural Biology*, Vol. 19, pp. 449–457.

Komeda, H., Ishikawa, N., & Asano, Y. (2003). Enhancement of the thermostability and catalytic activity of d-stereospecific amino-acid amidase from *Ochrobactrum anthropi* SV3 by directed evolution. *Journal of Molecular Catalysis B: Enzymatic*, 21(4–6), 283–290.

Korotkov, K. V., Pardon, E., Steyaert, J., & Hol, W. G. J. (2009). Crystal Structure of the N-Terminal Domain of the Secretin GspD from ETEC Determined with the Assistance of a Nanobody. *Structure*, 17(2), 255–265.

Krishnamurthy, H., & Gouaux, E. (2012). X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature*, 481(7382), 469–474.

Krishnan, V., & Rupp, B. (2012). Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy. *En eLS*. <https://doi.org/10.1002/9780470015902.a0002716.pub2>

Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., & Baker, D. (2003). Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science*, 302(5649), 1364–1368.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680–685.

Lasters, I., Wodak, S. J., Alard, P., & van Cutsem, E. (1988). Structural principles of parallel beta-barrels in proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 85(10), 3338–3342.

Lehninger, A. L. (2007). *Principios de bioquímica Lehninger* (5ª ed.). Barcelona, España: Ediciones Omega, S.A.

Lesk, A. M., Brändén, C.-I., & Chothia, C. (1989). Structural principles of α/β barrel proteins: The packing of the interior of the sheet. *Proteins: Structure, Function, and Genetics*, 5(2), 139–148.

Lippow, S. M., & Tidor, B. (2007). Progress in computational protein design. *Current Opinion in Biotechnology*, 18(4), 305–311.

Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., & Eisenberg, D. (1993). Crystal structure of a synthetic triple-stranded alpha-helical bundle. *Science (New York, N.Y.)*, 259(5099), 1288–1293.

Löw, C., Yau, Y. H., Pardon, E., Jegerschöld, C., Wåhlin, L., Quistgaard, E. M., Moberg, P., Geifman-Shochat, S., Steyaert, J., & Nordlund, P. (2013). Nanobody Mediated Crystallization of an Archeal Mechanosensitive Channel. *PLoS ONE*, 8(10), e77984.

Luft, J. R., Wolfley, J. R., & Snell, E. H. (2011). What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Crystal Growth and Design*, 11(3), 651–663.

Main, E. R. G., Stott, K., Jackson, S. E., Regan, L., & Baldwin, R. L. (2005). Local and long-range stability in tandemly arrayed tetratricopeptide repeats. *PNAS*, 102(16), 5721–5726.

Marcos, E., & Silva, D. A. (2018). *Essentials of de novo protein design: Methods and applications*. Wiley Interdisciplinary Reviews: Computational Molecular Science, 8(6), 1–19.

Martínez-Rodríguez, S., González-Ramírez, L. A., Clemente-Jiménez, J. M., Rodríguez-Vico, F., Las Heras-Vázquez, F. J., Gavira, J. A., & García-Ruíz, J. M. (2006). Crystallization and preliminary crystallographic studies of the

recombinant dihydropyrimidinase from *Sinorhizobium meliloti* CECT4114. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 62(Pt 12), 1223–1226.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). *Phaser* crystallographic software. *Journal of Applied Crystallography*, 40(4), 658–674.

McLennan, A., Turner, P., Bates, A., & White, M. (2012). *Molecular Biology* (5^a ed.). New York, United States of America: Garland Science.

McPherson, A., & Gavira, J. A. Introduction to protein crystallization. , 70 *Acta Crystallographica Section F: Structural Biology Communications* (2014).

McPherson, A., Malkin, A. J., Kuznetsov, Y. G., & Koszelak, S. (1996). Incorporation of impurities into macromolecular crystals. *Journal of Crystal Growth*, 168(1–4), 74–92.

Miyazaki, K., & Arnold, F. H. (1999). Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. *Journal of molecular evolution*, 49(6), 716–720.

Muñoz, V. (2008). *Protein Folding, Misfolding and Aggregation: Classical Themes and Novel Approaches*. Cambridge, UK: Royal Society of Chemistry.

Murzin, A. G., Lesk, A. M., & Chothia, C. (1994). Principles determining the structure of β -sheet barrels in proteins I. A theoretical analysis. *Journal of Molecular Biology*, 236(5), 1369–1381.

Nagano, N, Hutchinson, E. G., & Thornton, J. M. (1999). Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels. *Protein science : a publication of the Protein Society*, 8(10), 2072–2084.

Nagano, Nozomi, Orengo, C. A., & Thornton, J. M. (2002). One Fold with

Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *Journal of Molecular Biology*, 321(5), 741–765.

Nagarajan, D., Deka, G., & Rao, M. (2015). Design of symmetric TIM barrel proteins from first principles. *BMC biochemistry*, 16, 18.

Ng, J. D., Gavira, J. A., & García-Ruiz, J. M. (2003). Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *Journal of Structural Biology*, 142(1), 218–231.

Offredi, F., Dubail, F., Kischel, P., Sarinski, K., Stern, a S., van de Weerd, C., Hoch, J. C., Prospero, C., Francois, J. M., Mayo, S. L., & Martial, J. a. (2003). *De novo* Backbone and Sequence Design of an Idealized α/β -barrel Protein: Evidence of Stable Tertiary Structure. *Journal Of Molecular Biology*, 325(1), 163–174.

Osváth, S., & Gruebele, M. (2003). Proline can have opposite effects on fast and slow protein folding phases. *Biophysical journal*, 85(2), 1215–1222.

Otálora, F., Gavira, J. A., Ng, J. D., & García-Ruiz, J. M. (2009). Counterdiffusion methods applied to protein crystallization. *Progress in Biophysics and Molecular Biology*, 101(1–3), 26–37.

Pande, A., Annunziata, O., Asherie, N., Ogun, O., Benedek, G. B., & Pande, J. (2005). Decrease in Protein Solubility and Cataract Formation Caused by the Pro23 to Thr Mutation in Human γ D-Crystallin,. *Biochemistry*, 44(7), 2491–2500.

Pantazes, R. J., Grisewood, M. J., & Maranas, C. D. (2011). Recent advances in computational protein design. *Current Opinion in Structural Biology*, 21(4), 467–472.

Parmeggiani, F., Pellarin, R., Larsen, A. P., Varadamsetty, G., Stumpp, M.

T., Zerbe, O., Caflich, A., & Plückthun, A. (2008). Designed Armadillo Repeat Proteins as General Peptide-Binding Scaffolds: Consensus Design and Computational Optimization of the Hydrophobic Core. *Journal of Molecular Biology*, 376(5), 1282–1304.

Pokala, N., & Handel, T. M. (2001). Review: Protein design - Where we were, where we are, where we're going. *Journal of Structural Biology*, 134(2–3), 269–281.

Rämisch, S., Weininger, U., Martinsson, J., Akke, M., & André, I. (2014). Computational design of a leucine-rich repeat protein with a predefined geometry. *Proceedings of the National Academy of Sciences of the United States of America*, 111(50), 17875–17880.

Rasmussen, S. G. F., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the β 2adrenoceptor. *Nature*, 469(7329), 175–181.

Rice, P. A., Goldman, A., & Steitz, T. A. (1990). A helix-turn-strand structural motif common in α - β proteins. *Proteins: Structure, Function, and Genetics*, 8(4), 334–340.

Ring, A. M., Manglik, A., Kruse, A. C., Enos, M. D., Weis, W. I., Garcia, K. C., & Kobilka, B. K. (2013). Adrenaline-activated structure of β 2-adrenoceptor stabilized by an engineered nanobody. *Nature*, 502(7472), 575–579.

Ruiz-Martínez, C. R., Nieves-Marrero, C. A., Estremera-Andújar, R. A., Gavira, J. A., González-Ramírez, L. A., López-Garriga, J., & García-Ruiz, J. M. (2009). Crystallization and diffraction patterns of the oxy and cyano forms of the *Lucina pectinata* haemoglobins complex. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 65(Pt 1), 25–28.

Rupp, B. (2010). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. New York, United States of America: Garland Science.

Salahuddin, A. (1984). Proline peptide isomerization and protein folding. *Journal of Biosciences*, 6(4), 349–355.

Salemme, F. R. (1972). A free interface diffusion technique for the crystallization of proteins for X-ray crystallography. *Archives of Biochemistry and Biophysics*, 151(2), 533–539.

Scheerlinck, J.-P. Y., Lasters, I., Claessens, M., De Maeyer, M., Pio, F., Delhaise, P., & Wodak, S. J. (1992). Recurrent $\alpha\beta$ loop structures in TIM barrel motifs show a distinct pattern of conserved structural features. *Proteins: Structure, Function, and Bioinformatics*, 12(4), 299–313.

Schein, C. H. (1989). Production of Soluble Recombinant Proteins in Bacteria. *Nature Biotechnology*, 7(11), 1141–1149.

Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., & Gerlt, J. A. (2003). Evolutionary Potential of $(\beta/\alpha)_8$ -Barrels: Functional Promiscuity Produced by Single Substitutions in the Enolase Superfamily. *Biochemistry*, 42(28), 8387–8393.

Serrière, J., Dugua, J. M., Bossus, M., Verrier, B., Haser, R., Gouet, P., & Guillon, C. (2011). Fab'-induced folding of antigenic N-terminal peptides from intrinsically disordered HIV-1 Tat revealed by X-ray crystallography. *Journal of Molecular Biology*, 405(1), 33–42.

Shakhnovich, E. I. (1998). Protein design: A perspective from simple tractable models. *Folding and Design*, 3(3), 45–58.

Simpson, R. J. (2010). Solubilization of Escherichia coli Recombinant Proteins from Inclusion Bodies. *Cold Spring Harbor Protocols*, 2010(9),

pdb.prot5485-pdb.prot5485.

Singh, S. M., & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*, 99(4), 303–310.

Skerra, A. (2007). Alternative non-antibody scaffolds for molecular recognition. *Current Opinion in Biotechnology*, 18(4), 295–304.

Steward, A., Adhya, S., & Clarke, J. (2002). Sequence Conservation in Ig-like Domains: The Role of Highly Conserved Proline Residues in the Fibronectin Type III Superfamily. *Journal of Molecular Biology*, 318(4), 935–940.

Strandberg, L., & Enfors, S. O. (1991). Factors influencing inclusion body formation in the production of a fused protein in *Escherichia coli*. *Applied and environmental microbiology*, 57(6), 1669–1674.

Suzuki, Y. (1989). A general principle of increasing protein thermostability. *Proc. Japan Acad.*, 65(6), 146–148.

Tao, A., & Raz, E. (2015). *Allergy Bioinformatics*. New York, United States of America: Springer Netherlands.

Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P., & Ranson, N. A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*, 100, 3–15.

Thomson, A. R., Wood, C. W., Burton, A. J., Bartlett, G. J., Sessions, R. B., Brady, R. L., & Woolfson, D. N. (2014). Computational design of water-soluble α -helical barrels. *Science*, 346(6208), 485–488.

Trevino, S. R., Scholtz, J. M., & Pace, C. N. (2007). Amino Acid Contribution to Protein Solubility: Asp, Glu, and Ser Contribute more Favorably than the other Hydrophilic Amino Acids in RNase Sa. *Journal of Molecular Biology*, 366(2), 449–460.

Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., Perederina, A., Sugita, Y., Vassilyev, D. G., Ito, K., & Nureki, O. (2008). Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature*, 455(7215), 988–991.

Urvoas, A., Guellouz, A., Valerio-Lepiniec, M., Graille, M., Durand, D., Desravines, D. C., van Tilbeurgh, H., Desmadril, M., & Minard, P. (2010). Design, Production and Molecular Structure of a New Family of Artificial Alpha-helical Repeat Proteins (α Rep) Based on Thermostable HEAT-like Repeats. *Journal of Molecular Biology*, 404(2), 307–327.

Vallejo, L., & Rinas, U. (2004). Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microbial Cell Factories*, 3(1), 11.

van den Berg, S., Löfdahl, P.-Å., Härd, T., & Berglund, H. (2006). Improved solubility of TEV protease by directed evolution. *Journal of Biotechnology*, 121(3), 291–298.

Varadamsetty, G., Tremmel, D., Hansen, S., Parmeggiani, F., & Plückthun, A. (2012). Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High Specificity. *Journal of Molecular Biology*, 424(1–2), 68–87.

Vekilov, P. G. (2010). Nucleation. *Crystal Growth and Design*, 10(12), 5007–5019.

Vekilov, P. G. (2012). Phase diagrams and kinetics of phase transitions in protein solutions. *Journal of Physics Condensed Matter*, 24(19).

Visuri, K. (1984). Stable glucose isomerase concentrate and a process for the preparation thereof. US4699882A.

Watanabe, K., Masuda, T., Ohashi, H., Mihara, H., & Suzuki, Y. (1994).

Multiple proline substitutions cumulatively thermostabilize *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule. *European journal of biochemistry*, 226(2), 277–283.

Watanabe, Kunihiko, Hata, Y., Kizaki, H., Katsube, Y., & Suzuki, Y. (1997). The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *Journal of Molecular Biology*, 269(1), 142–153.

Wilmanns, M., Hyde, C. C., Davies, D. R., Kirschner, K., & Jansonius, J. N. (1991). Structural conservation in parallel .beta./alpha.-barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry*, 30(38), 9161–9169.

Wlodawer, A., Dauter, Z., & Jaskolski, M. (2018). *Protein Crystallography: Methods and Protocols*. New York, United States of America: Springer New York.

Woolfson, D. N., Bartlett, G. J., Burton, A. J., Heal, J. W., Niitsu, A., Thomson, A. R., & Wood, C. W. (2015). De novo protein design: How do we expand into the universe of possible protein structures? *Current Opinion in Structural Biology*, 33, 16–26.

Wu, M., Park, Y. J., Pardon, E., Turley, S., Hayhurst, A., Deng, J., Steyaert, J., & Hol, W. G. J. (2011). Structures of a key interaction protein from the *Trypanosoma brucei* editosome in complex with single domain antibodies. *Journal of Structural Biology*, 174(1), 124–136.

Wymer, N., Buchanan, L. V., Henderson, D., Mehta, N., Botting, C. H., Pocivavsek, L., Fierke, C. A., Toone, E. J., & Naismith, J. H. (2001). Directed Evolution of a New Catalytic Site in 2-Keto-3-Deoxy-6-Phosphogluconate Aldolase from *Escherichia coli*. *Structure*, 9(1), 1–9.

Yaron, A., & Naider, F. (1993). Proline-Dependent Structural and

Biological Properties of Peptides and Proteins. *Crit. Rev. Biochem. Mol. Biol.*, 28, 31–81.

Zaccolo, M., & Gherardi, E. (1999). The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 β -lactamase. *Journal of Molecular Biology*, 285(2), 775–783.

Zang, Y., Kammerer, B., Eisenkolb, M., Lohr, K., & Kiefer, H. (2011). Towards Protein Crystallization as a Process Step in Downstream Processing of Therapeutic Antibodies: Screening and Optimization at Microbatch Scale. *PLoS ONE*, 6(9), e25282.

