

Tesis Doctoral

Estructura y función del dominio D/D de la subunidad regulatoria de la proteína quinasa A de *Saccharomyces cerevisiae*

González Bardeci, Nicolás Diego

2015-03-25

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the doctoral theses collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

González Bardeci, Nicolás Diego. (2015-03-25). Estructura y función del dominio D/D de la subunidad regulatoria de la proteína quinasa A de *Saccharomyces cerevisiae*. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.

Cita tipo Chicago:

González Bardeci, Nicolás Diego. "Estructura y función del dominio D/D de la subunidad regulatoria de la proteína quinasa A de *Saccharomyces cerevisiae*". Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2015-03-25.

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Química Biológica

Estructura y función del dominio D/D de la subunidad regulatoria de la proteína quinasa A de *Saccharomyces cerevisiae*

Tesis presentada para optar por el título de Doctor de la Universidad de Buenos Aires en el área Química Biológica

Autor

Nicolás Diego González Bardeci

Directora de Tesis: Silvia Margarita Moreno de Colonna

Directora Asistente de Tesis: Silvia Graciela Rossi

Consejera de estudios: Silvia Margarita Moreno de Colonna

Lugar de trabajo: Laboratorio de Transducción de Señales, Departamento de Química Biológica e Instituto de Química Biológica de la Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires

Ciudad Autónoma de Buenos Aires, 2015

Fecha de defensa:

Estructura y función del dominio D/D de la subunidad regulatoria de la proteína quinasa A de *Saccharomyces cerevisiae*

Resumen

La proteína quinasa dependiente de cAMP (PKA) es una quinasa de serina y treonina de amplio espectro que fosforila a sus proteínas blanco en respuesta a estímulos específicos. Se encuentra involucrada en la regulación de una gran diversidad de procesos fisiológicos en una amplia variedad de organismos, motivo por el cual es un miembro modelo de estudio de la superfamilia de las Proteínas Quinasa.

En la mayoría de los organismos, en condiciones de baja concentración de cAMP existe como una holoenzima inactiva heterotetramérica conformada por dos subunidades catalíticas (C) y un dímero de subunidades regulatorias (R). En respuesta a estímulos extracelulares específicos, las concentraciones intracelulares del segundo mensajero aumentan; como consecuencia, dos moléculas de cAMP se unen a cada subunidad R, generando un cambio conformacional que disocia la holoenzima en un dímero de subunidades R y dos subunidades C activas.

En mamíferos existen cuatro isoformas de la subunidad R, que presentan una estructura de dominios bien conservada entre los distintos organismos. Ésta consiste de dos dominios de unión a cAMP hacia el extremo C-terminal, y una región responsable de la dimerización hacia el extremo N-terminal. Esta pequeña región de apenas 50 aminoácidos se denomina dominio de dimerización y anclaje (D/D), ya que constituye además una superficie de interacción para una familia de proteínas denominadas AKAPs (A-Kinase Anchoring Proteins). Estas proteínas tienen como función proporcionar a la holoenzima de la PKA la localización subcelular necesaria para garantizar la propagación adecuada de las señales disparadas por cAMP. Todas ellas presentan una hélice anfipática de unos 20 aminoácidos cuya cara no polar interactúa con alta afinidad con una superficie hidrofóbica proporcionada por los dominios D/D.

Se conocen las estructuras de alta resolución de los dominios D/D de mamíferos, tanto en su forma apo como en complejo con péptidos derivados de distintas AKAPs. Sin embargo, hasta la fecha no se han llevado a cabo estudios estructurales de estos dominios en otros organismos.

El objetivo principal de este trabajo consiste en obtener una caracterización estructural y funcional del dominio D/D de Bcy1, la subunidad R de la PKA de *S. cerevisiae*. La motivación principal radica en contribuir con el primer estudio de estos dominios en organismos no mamíferos, a efectos de establecer una comparación con éstos. *S. cerevisiae* es un organismo modelo por excelencia en bioquímica y biología molecular, y resulta muy pertinente para este trabajo dado que en este organismo la PKA es una enzima fundamental involucrada en la regulación de los procesos relacionados con la respuesta a la disponibilidad de nutrientes, respuesta a estrés, desarrollo y entrada en fase estacionaria.

En primer lugar, se realizó un mapeo del dominio D/D de Bcy1 por análisis bioinformático de secuencias y experimentos de entrecruzamiento químico utilizando mutantes de delección de Bcy1 en su extremo N-terminal. En segundo lugar, se procedió al clonado, sobreexpresión en bacterias, y desarrollo de un protocolo de purificación del fragmento recombinante Bcy1 1-50. Posteriormente, se procedió a su caracterización estructural, para lo cual se utilizaron dos abordajes generales: estudiar la estructura en solución y elucidar la estructura cristalina.

Para la primera parte, se determinó el estado oligomérico en solución y se obtuvieron parámetros hidrodinámicos característicos utilizando cromatografía de exclusión molecular (SEC), dispersión estática de la luz (SLS) y dispersión de rayos X a bajo ángulo (SAXS). Por otro lado, se estudió la estructura secundaria utilizando dicroísmo circular (CD). Finalmente, se construyó un modelo de la estructura en solución consistente con los datos experimentales, mediante herramientas de modelado molecular. En segundo lugar, se logró resolver la estructura cristalina utilizando difracción por rayos X (XRD).

Recientemente, en nuestro grupo de investigación se han identificado varias proteínas interactoras de Bcy1. Estas proteínas son consideradas “candidatas” a ser las primeras AKAPs reportadas en levaduras. Para concluir este trabajo, se estudió la interacción entre el fragmento Bcy1 1-50 y un péptido sintético derivado de una de estas proteínas, Ira2. Para ello se utilizaron técnicas de dicroísmo circular y espectroscopía de fluorescencia.

Palabras clave: proteína quinasa A, *Saccharomyces cerevisiae*, Bcy1, proteínas de anclaje de la quinasa A, dicroísmo circular, dispersión de rayos X a bajo ángulo, estructura cristalina.

Structure and function of the D/D domain of the regulatory subunit of protein kinase A in *Saccharomyces cerevisiae*

Abstract

The cAMP dependent protein kinase (PKA) is a broad spectrum serine/threonine kinase which phosphorylates its target proteins in response to specific stimuli. It is involved in the regulation of a great diversity of physiological processes in many organisms. Therefore, it is a model member of study of the Protein Kinase superfamily.

In most organisms, when cAMP levels are low, it exists as an inactive heterotetrameric holoenzyme which is formed by two catalytic (C) subunits and a regulatory (R) subunit dimer. In response to specific extracellular stimuli, the intracellular concentrations of the second messenger rise; as a consequence, two molecules of cAMP bind each R subunit, thus triggering a conformational change that dissociates the holoenzyme in an R subunit dimer and two active C subunits.

In mammals there exist four isoforms of the R subunit, which present a well conserved domain structure among many organisms. It consists of two cAMP binding sites at the C-terminus, and a region responsible for dimerization at the N-terminus. This small region of 50 amino acids is called the dimerization and docking (D/D) domain since it presents a surface for interaction with proteins of the AKAP (A-Kinase Anchoring Proteins) family. These proteins are responsible for the appropriate subcellular localization of the holoenzyme, which is required for the propagation of the signaling events triggered by cAMP. They all present an amphipathic helix of 20 amino acids with a non-polar face that interacts with high affinity with a hydrophobic surface in the D/D domains.

The high resolution structures of mammalian D/D domains are well known, both in their apo forms and in the AKAP peptide complex. However, to date no structural studies of these domains in other organisms have been performed.

The main goal of this work consists in performing a structural and functional characterization of the D/D domain of Bcy1, the R subunit of PKA from *S. cerevisiae*. The motivation for this work consists in contributing with the first structural study of these domains in non-mammalian organisms, in order to compare with mammalian features. *S. cerevisiae* is a prototypic organism in biochemistry and molecular biology, and it is of great interest for this work because PKA is a key enzyme involved in the regulation of nutrient availability, stress response, development and entry into stationary phase.

As a first approach, we performed a mapping of the D/D domain of Bcy1 using sequence analysis and chemical crosslinking experiments with deletion mutants of the N-terminus of Bcy1. Secondly, we cloned, overexpressed, and purified a recombinant fragment, Bcy1 1-50. Finally, we proceeded with its structural characterization, which consisted of two general approaches: to study the solution structure and the crystal structure.

For the solution structure analysis, we determined the oligomeric state and obtained hydrodynamic parameters using size-exclusion chromatography (SEC), static light scattering (SLS), and small-angle X-ray scattering (SAXS). On the other hand, we studied

the secondary structure using circular dichroism (CD). Finally, we generated a model of the solution structure which is consistent with the experimental data, using molecular modelling tools. Finally, we solved the crystal structure using X-ray diffraction (XRD).

Recently, in our group, several proteins that interact with Bcy1 have been identified. These proteins are considered “candidates” to be the first AKAPs reported in yeast. To conclude this work, we studied the interaction between the fragment Bcy1 1-50 and a synthetic peptide derived from one of these proteins, Ira2. To accomplish this goal, we used CD and fluorescence spectroscopy techniques.

Keywords: protein kinase A, *Saccharomyces cerevisiae*, Bcy1, A-Kinase anchoring proteins, circular dichroism, small-angle X-ray scattering, crystal structure.

Colaboraciones

Los experimentos de cromatografía de exclusión molecular, dicroísmo circular, dispersión estática de la luz, y espectroscopía de fluorescencia, fueron realizados en la Fundación Instituto Leloir en colaboración con el Dr. Julio Javier Caramelo.

Las mediciones de dispersión de rayos X a bajo ángulo fueron realizadas en la Universidad de Utah, Salt Lake City, USA, por el Dr. Donald K. Blumenthal, quien también proporcionó asistencia y asesoramiento en el análisis de los datos.

La cristalización y resolución de la estructura cristalina fue realizada en la Unidad de Cristalografía de Proteínas del Institut Pasteur de Montevideo, Uruguay, dirigida por el Dr. Alejandro Buschiazzo. La obtención de cristales se logró durante una pasantía de un mes en esa institución, para lo cual se contó con el asesoramiento y la colaboración de Nicole Larrieux. La estructura fue resuelta posteriormente por el Dr. Felipe Trajtenberg.

Para la construcción de modelos por homología y uso de programas de docking, se contó con el asesoramiento del Dr. Adrián Turjanski, de la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires.

AGRADECIMIENTOS

“ Happiness only real when shared ”

Alexander Supertramp

A **Andrés** y a **Chae**, por todos estos años, y por todos los que faltan. A Andrés por su sentido del humor, y porque siempre vamos a Palermo. A Chae por su amabilidad, que no conoce límites. A la mamá de Andrés.

A **Ceci**. Nada de lo que pueda decirte con este pobre vehículo que son las palabras representa ni remotamente todo lo que quisiera expresar. Por haber hecho absolutamente todo juntos, desde estar naciendo casi al mismo tiempo sin saberlo, pasando por ir al mismo colegio (casi sin saberlo), hasta escribir la tesis de a dos, y bien que lo sabemos. Por la compañía, el aguante incondicional, y por siempre entender. Por todas las demás cosas que no se pueden poner en palabras, pero que por suerte vos ya sabés. Y por supuesto, por ensamblar mi tesis.

A **Coti**, mi hermana del alma, por haber andado juntos estos 7 u 8 años de tan complicado camino. Por ser mi compañera incondicional e irremplazable de laboratorio. Por las penas y las alegrías compartidas. Por estar ahí siempre de una u otra manera, acompañándonos, en silencio muchas veces, con un mate de por medio otras, pero siempre estando.

A **Fiore**, por estar siempre presente cada vez que necesité una ayuda, un consejo, una palabra de aliento, una sopapo, un “tirón de orejas” incluso. Por no dejarme aflojar. Y porque disolver el cAMP siempre va a ser “una boludez”.

A **Flor**, por irrumpir de forma tan perfectamente inesperada en mi vida, y haber sido la mejor compañía que cualquiera podría tener en estas últimas semanas de tortura. Por haberte bancado con santa paciencia las ausencias y las ciclotimias. Y porque sos tan, pero tan increíblemente todo.

A mi **familia** en general, lo que incluye a los animales, que son los dueños de la casa.

Al **Gato Sin Nombre**, por ser el mejor compañero de vida que cualquiera podría querer, pese a haberte portado bastante mal en estas últimas semanas, lo cual no fue de gran ayuda.

A mis **hermanos**, porque aun cuando no nos vemos mucho, siempre están presentes a la distancia, y siempre tienen a mano alguna huevada para decir o mostrar, lo cual nunca viene mal, sobre todo en épocas de reclusión.

A **Iani**, por tantas cosas. Por ser un poco tan como yo, porque entendés tan perfectamente bien todo lo que está ocurriendo dentro de mi cabeza que hasta a veces me asusta un poco. Por estar siempre en las peores, y también en las mejores. Por las películas más malas de Scarlett Johansson y por las películas buenísimas de agujeros negros.

A **Lu Álvarez**, por ser la clase de persona absolutamente increíble que sos; por otro lado, por haberte fumado de punta a punta mi tesis sin protestar, y hasta con alegría; por otro lado, por haberme hecho un montón de correcciones que me ayudaron un montón; y no estoy diciendo todo esto sólo por la tesis, dado que, por otro lado, sos una amiga de fierro más allá de todo.

A **Marce**, por preocuparte siempre por mi situación sentimental y por mi bienestar emocional. Por ponerle siempre toda la onda a todo. Por hacerme la mejor decoración de cumpleaños de la historia. Por los “pues no Chae” que nunca dijiste, pero ya está, es como si lo hubieras hecho.

A **Silvia Moreno**. Por su dedicación. Por su paciencia. Por no perder nunca la calma ni el buen humor. Por bancarse mis gastritis, mis alergias, mis cambios de humor, mis altibajos laborales, mis ataques de pánico en los aeropuertos, mis pésimos cálculos de los tiempos. Gracias.

A **Silvia Rossi**. Por recibirme en su laboratorio hace 7 u 8 años cuando, sin conocerme, fui a golpearle la puerta. Por ser un ejemplo de dedicación, esfuerzo, integridad y trabajo. Porque no importa lo ocupada que esté (y en general lo está mucho), siempre tiene un momento para ayudar, y siempre hace sentir que lo que uno está haciendo, sea lo que sea, es lo más importante del mundo. Gracias.

A **Sol Ponte**, por tu alegría que no conoce límites; por estar siempre, pero siempre, pero siempre, aun a la distancia, en las buenas y en las malas.

A mis **viejos**, por hacerme lo que soy, que es lo más importante de todo; sin eso yo no estaría aca.

A **Alejandro Buschiazzo y su grupo de investigación**, que me recibieron como si fuera uno más de ellos. Por el lindo trabajo que hicimos juntos. Y muy en particular, a **Nicole**, que es la persona más buena que existe sobre la faz de la tierra.

Al **Departamento de Química Biológica**, con toda su gente. A las autoridades y a los no docentes. A **Mariela Mezzina** por su buena onda inquebrantable, por estar siempre dispuesta a dar una mano; por los tetrámeros compartidos. A **Lucas Defelipe** por ayudarme a resolver mis ignorancias bioinformáticas.

A **Donald Blumenthal** por su paciencia, por su siempre buena predisposición para ayudar, desinteresadamente, aun cuando le pregunté cosas de mi tesis que nada tenían que ver con nuestro trabajo en conjunto.

A **Euge Tomasini**, por su amistad que data ya de tanto tiempo, cuando nos conocimos en mi primer cuatrimestre de docencia. Por tu apoyo constante a la distancia, que es muy importante para mi.

A **Itu**, por ser tan Itu. Por aparecer cada vez que se abre cualquier envoltorio que contenga algún tipo de producto alimenticio; cosa que a nadie le importa, en el fondo, porque siempre transportás, junto con tus pasos arrastrados, algo interesante para compartir, para contar, para chusmear. Y porque siempre tenés una dosis de alegría y optimismo para infundirle hasta a la cosa más catastrófica.

A **Jose**, por haberme guiado con paciencia, con mucha paciencia, en mis primeros pasos en este extraño submundo en que nos metimos, y porque aun hoy, a la distancia, estás presente, siempre dispuesta a darme algún buen consejo cuando lo necesito.

A **Juancito**, por bancarme en todas, por las juntadas en tu casa, por la energía que infundís. Porque siempre sabés cómo sacarme una sonrisa cuando estoy mal. Y porque siempre tenés un comentario atinado para hacer acerca de las mujeres.

A **Julio Caramelo**, por recibirme siempre con muy buena predisposición en su laboratorio, y ponerle toda la energía del mundo al trabajo que hicimos juntos, aun cuando algunas cosas se nos complicaron.

A **Leyla**, porque a pesar de los intervalos estocásticos de ausencias, los dos sabemos que siempre vamos a estar.

A las **pibas de EL labo**: a Jime (lalalalala! Qué lindo qué lindo qué bien! Chau plaquita que tengas suerte!), a Luli por sus modulaciones incomprensibles y sus derroches de cariños a mansalva, a Enzolón, a Sol, a Adri, a Carla, a Richard, a Pía, a Claribel (por los viajes y las charlas filosóficas). A todos por la compañía de todos los días. A los que pasaron por el labo y me estoy olvidando de nombrar. A Irina por la buena onda de siempre.

A los **pibes de EL colegio**, por estar siempre, desde hace tanto tiempo: a Caño, a Pato (paaaaaatoooooo000000), a Iza (por los problemas gástricos compartidos), a Rodrigo, a Celeste (por ser tan... Celeste), a Dot; y al Ofi (porque de haberle avisado cinco minutos antes, seguramente estaría hoy festejando con nosotros). Porque ya vamos mucho, pero seguimos.

A los **pibes y pibas de LA facu**: a Javi, a Lau, a Maru, a Rolo (por las historias esas que vos contás que ya sabés), a Funes (por haber leído la Introducción metodológica del capítulo de cristalografía... pero no sólo por eso, ya sabés). A Gi por ponerle pilas a no perdernos de vista, aunque hayamos elegido caminos distintos. A Lu B., por haberme bancado en los peores momentos.

A **Molinas**, por su amistad, que si bien nadie sabe muy bien cómo ni cuándo comenzó, ni nunca lo vamos a saber, en el fondo poco importa, porque vale. Por ser más vieja que yo, ahora y para siempre. Por nuestra primera experiencia juntos en docencia, largamente postergada, que definitivamente no fue la mejor (pero al menos compartida fue más llevadera).

A **Paula Angelomé**, porque aunque apenas nos conocemos, sos una máquina de tirar buena onda ininterrumpida por Facebook, y eso levanta un montón.

A **Paula MC**, por animarse a preguntar lo que todos querían saber, delante de todos, para que todos lo escucharan. Por recibirme siempre con ese amor expresivo tan tuyo cada vez que fui a trabajar a tu laboratorio, del estilo "tomá, hacete un café y no jodas". Y porque para mi siempre vas a tener un nombre y dos apellidos.

A **Paulita "into the wild" Portela**, por tu energía, por tu predisposición para ayudar, por tu buena onda de siempre, por los viajes y las anécdotas compartidas.

A **Paulita "ñiñiñiñiñi" Vena**, por la compañía de los sábados a la tarde de cultivos y escrituras, y por la de todos los otros días también. Por los millones de papers que publicamos juntos, y los millones que están por venir.

A **Romi Landa**, porque cómo nos reímos durante todo el cuatrimestre. Por bancársela de frente y defender mi causa perdida contra todo el mundo.

A **Rosario**, por la compañía de todos estos años, y porque siempre está dispuesta a ayudar cuando necesito algo.

A **Sara Aldabe**, por las charlas de pasillo. Por salvarme las papas cuando el horno estaba al mango.

A **Sole Antonel**, por las salidas a la puerta, por las infinitas charlas de pasillo, y un poco también, por qué no decirlo, por ese exceso de optimismo, que si bien a veces me revienta un poco, la realidad es que nunca, nunca viene mal.

A la Universidad de Buenos Aires.

A todos los que me ayudaron y/o acompañaron de una u otra manera en estos años, y que me olvidé de nombrar. Pero ustedes saben.

INDICE

PRÓLOGO	1
INTRODUCCIÓN	5
Transducción de señales y proteínas quinasas	7
La vía de transducción de señales cAMP-PKA	8
Resumen de los principales componentes de la vía de señalización de cAMP-PKA	13
Subunidad C de la PKA	15
Subunidad R de la PKA	16
Microdominios de cAMP y localización subcelular de la PKA	17
Estructura de la subunidad catalítica de la PKA	23
Estructura de la subunidad reguladora de la PKA	29
Dominio de dimerización y anclaje (D/D)	30
Estructura de los complejos D/D-AKB	34
Interacción entre Bcy1 e Ira2	38
Dominios de unión a cAMP	39
Zona bisagra (<i>Linker</i>)	42
Estructura global del dímero de subunidades R	43
Estructura de la holoenzima y aspectos estructurales de la activación de la PKA	45
Objetivos generales	52
Referencias	54
MAPEO, ANÁLISIS BIOINFORMÁTICO Y PRODUCCIÓN DEL DOMINIO D/D DE BCY1	59
Antecedentes	61
Objetivos y abordajes utilizados	61
Resultados y discusión	63
Análisis bioinformático de la región N-terminal de Bcy1	63
Mapeo experimental del dominio D/D de Bcy1	68
Producción del dominio D/D de Bcy1	73
Conclusiones y perspectivas	78
Materiales y métodos	79
Análisis de secuencias	79

Plásmidos y cepas	80
Medios de cultivo y condiciones de crecimiento	82
Clonado de Bcy1 (1-50)	83
Purificación de proteínas	84
Electroforesis	87
Experimentos de entrecruzamiento químico con EGS	88
Espectrometría de masa	89
Referencias	90
ESTRUCTURA EN SOLUCIÓN DEL DOMINIO D/D DE BCY1	93
Antecedentes	95
Objetivos y abordajes utilizados	95
Introducción metodológica	97
Dispersión estática de la luz (SLS)	97
Espectropolarimetría de dicroísmo circular (CD)	98
Determinación de parámetros hidrodinámicos	99
Determinación de la MM en solución utilizando el R_s y el $S_{20,w}$	102
Dispersión de Rayos X a bajo ángulo (SAXS)	102
Resultados y discusión	114
Estado oligomérico de Bcy1 (1-50) en solución	114
Estado oligomérico de BCy1 WT en solución	118
Posible rol fisiológico de la oligomerización de Bcy1	122
Caracterización estructural del tetrámero Bcy1 (1-50)	125
Estructura secundaria en solución	126
Propiedades hidrodinámicas de Bcy1 (1-50)	130
Modelo estructural del tetrámero Bcy1 (1-50)	137
Comparación de parámetros hidrodinámicos de dominios D/D	144
Conclusiones y perspectivas	149
Materiales y métodos	150
Cromatografía de exclusión molecular	150
Experimentos de dispersión estática de la luz	151
Experimentos de entrecruzamiento químico	151
Construcción de la cepa 1115-Bcy1 (1-50)	152
<i>Western blot</i>	153
Ensayos fenotípicos	154

Espectropolarimetría de dicroísmo circular	155
Ultracentrifugación analítica en gradiente de sacarosa	156
Dispersión de Rayos X a bajo ángulo	156
Construcción del modelo atómico del tetrámero	158
Referencias	160
ESTRUCTURA CRISTALINA DEL DOMINIO D/D DE BCY1	163
Antecedentes	165
Objetivos y abordajes utilizados	165
Introducción metodológica	166
Generalidades sobre la técnica	166
Propiedades de los cristales	166
Fundamento del fenómeno de difracción	168
Esquema general de trabajo	170
Obtención de la proteína purificada	171
Cristalización	171
Montado de los cristales y colección de los datos	173
Procesamiento de los datos	175
Resolución de la estructura	176
Construcción del modelo y refinamiento	178
Validación de la estructura	180
Estimadores de la calidad de los datos	180
Estimadores de la calidad del modelo	182
Resultados y discusión	183
Obtención de cristales, colección de los datos y resolución de la estructura	183
Validación de la estructura	186
Topología general de la estructura cristalina	187
Comparación de las estructuras cristalina y en solución	191
Análisis de la estructura del dímero	196
Análisis de la estructura del tetrámero	200
Análisis de las secuencias de los dominios D/D de diversos organismos	206
Conclusiones y perspectivas	215
Materiales y métodos	216
Obtención de cristales	216
Colección de datos	216

Procesamiento de los datos y resolución de la estructura	216
Herramientas bioinformáticas	217
Referencias	221
INTERACCIÓN DEL DOMINIO D/D DE BCY1 CON EL PÉPTIDO IRA2	223
Antecedentes	225
Objetivos y abordajes utilizados	225
Resultados, métodos y discusión	227
Estructura secundaria del péptido Ira2 en solución	227
Interacción entre el péptido Ira2 y las proteínas Bcy1 WT y Bcy1 (1-50)	228
Análisis de las posibles superficies de interacción de Bcy1 (1-50)	233
Conclusiones y perspectivas	237
Referencias	238
CONSIDERACIONES FINALES	239
Conclusiones	241
Perspectivas futuras	244
Referencias	246

PRÓLOGO

“ Tengo el honor de enviarle para su revista un artículo que espero sea usted capaz de comprender mejor que yo. Es una traducción hecha por mi amigo Martin Van Buren Navis (llamado «El brujo de Poughkeepsie») de un manuscrito de extraña apariencia que encontré hace aproximadamente un año dentro de un porrón tapado, flotando en el *Mare Tenebrarum* - mar bien descrito por el geógrafo nubio, pero rara vez visitado en nuestros días, salvo por los trascendentalistas y los buscadores de extravagancias. ”

Edgar Allan Poe

Este trabajo de Tesis Doctoral se enmarca en el contexto de la línea general de investigación del Laboratorio de Transducción de Señales del Departamento de Química Biológica (FCEN-UBA). Desde hace ya muchos años, nuestro grupo de investigación se ha dedicado al estudio de los mecanismos de detección de estímulos extracelulares y de transducción de señales mediados por la proteína quinasa A (PKA).

La PKA es una quinasa de amplio espectro que fosforila una gran diversidad de sustratos en respuesta a distintos estímulos extracelulares. Está conformada por tipos de subunidades: una subunidad regulatoria y una subunidad catalítica, y existe cierta diversidad en cuanto a la variedad de isoformas de las dos subunidades que presentan los distintos organismos. Como se discutirá en la Introducción, uno de los principales puntos de interés acerca de esta vía de señalización radica en comprender los mecanismos moleculares que determinan la especificidad de la respuesta a distintas señales extracelulares.

El principal modelo de estudio de nuestro laboratorio es la levadura del pan, el vino y la cerveza, *Saccharomyces cerevisiae* (**Figura 1**). Si bien se trata de un organismo modelo central, representa un escalón evolutivo en el que aun quedan muchas cuestiones por resolver en cuanto a la temática de investigación de nuestro grupo. Otros modelos de interés del laboratorio son los hongos dimórficos *Mucor rouxii* y *Mucor circinelloides*; este último resulta de particular interés debido a que presenta una gran cantidad de isoformas de las dos subunidades de la PKA.

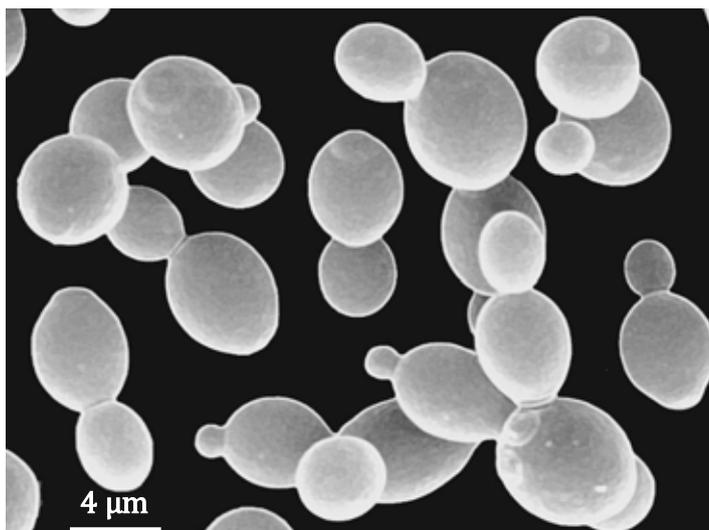
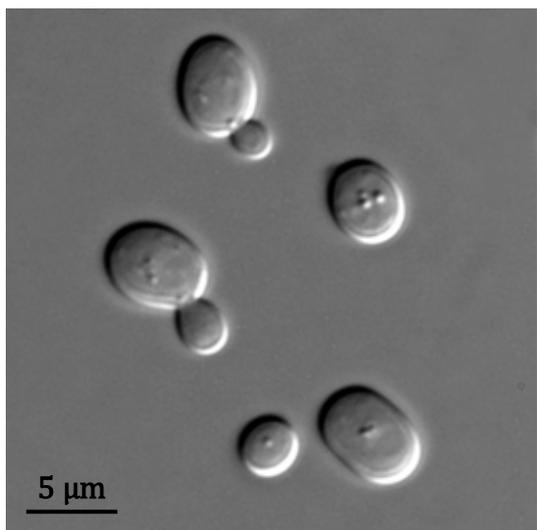


Figura 1: Imágenes de *S. cerevisiae*.

Panel de la izquierda: imagen obtenida por DIC (*differential interference contrast microscopy*). Panel de la derecha: imagen obtenida por SEM (*scanning electron microscopy*).

El reconocimiento específico de sustratos por parte de la PKA es consecuencia de distintos mecanismos que operan de manera independiente: localización subcelular diferencial de las distintas isoformas, participación de los sustratos en la activación de la enzima, relación estructural entre las distintas subunidades, modificaciones post-traduccionales de las mismas, etc. Por lo tanto, el mecanismo de activación dentro de la célula es un proceso muy complejo, lo que requiere para su entendimiento del análisis simultáneo de numerosos aspectos.

Para abordar esta problemática, en nuestro laboratorio se utilizan numerosos enfoques. Algunos de los principales temas abordados que han contribuido a la comprensión de los mecanismos que median la especificidad de la señalización por PKA son: análisis de la regulación de la expresión de las distintas isoformas en *S. cerevisiae* y *M. circinelloides*, consecuencias fenotípicas de las deleciones de los distintos genes, estudio de las propiedades bioquímicas y estructurales de las distintas proteínas, relación estructura-función para las distintas isoformas de los dos tipos de subunidades, etc. En particular, en los últimos años se ha hecho foco en el estudio estructural de la subunidad regulatoria de la PKA de *S. cerevisiae*, Bcy1, que desempeña un papel fundamental en la regulación de la actividad de la PKA mediante distintos mecanismos. Para poder comprender cómo funcionan estos procesos regulatorios, es imprescindible tener un conocimiento acabado acerca de la estructura de esta proteína.

Para contribuir a ese entendimiento, en este trabajo presentamos un estudio estructural exhaustivo del dominio N-terminal de Bcy1. Este dominio resulta crucial en la regulación de la actividad de la PKA, ya que es responsable del estado oligomérico de la enzima y participa también en la determinación de su localización subcelular.

CAPÍTULO 1

INTRODUCCIÓN

“ Afirman los impíos que el disparate es normal en la Biblioteca y que lo razonable (y aun la humilde y pura coherencia) es una casi milagrosa excepción. Hablan (lo sé) de “la Biblioteca febril, cuyos azarosos volúmenes corren el incesante albur de cambiarse en otros y que todo lo afirman, lo niegan y lo confunden como una divinidad que delira”. ”

Jorge Luis Borges, “La biblioteca de Babel”

Transducción de señales y proteínas quinasas

Los mecanismos de transducción de señales son procesos fundamentales para garantizar el funcionamiento y la viabilidad de organismos tanto unicelulares como pluricelulares, dado que se encuentran involucrados en la regulación de procesos biológicos relacionados con el desarrollo, el metabolismo, la proliferación celular, la regulación de la expresión génica, la respuesta a estímulos ambientales, etc. Existen vías de señalización que median la comunicación entre distintas células y que por lo tanto garantizan su funcionamiento coordinado, mientras que otras ocurren a nivel intracelular y son indispensables para la viabilidad de las células individuales. En el interior celular, la información entre las distintas moléculas responsables de la señalización se transmite mediante modificaciones post traduccionales de proteínas que se encuentran minuciosamente reguladas. Una de las modificaciones más comunes en los procesos de señalización es la fosforilación de proteínas, que es controlada por la acción opuesta de quinasas que fosforilan y fosfatasas que defosforilan. El estado de fosforilación de una proteína puede afectar su actividad enzimática, su asociación con otras proteínas, su localización subcelular, etc., por lo que la fosforilación de proteínas proporciona un mecanismo de respuesta universal a estímulos externos, y es probablemente el mecanismo regulatorio más importante en la mayoría de los eucariotas. Como consecuencia, las proteínas quinasas son enzimas clave que se encuentran involucradas en la regulación de numerosos procesos celulares. [1, 2].

La fosforilación de proteínas fue reconocida como mecanismo regulatorio por primera vez por E. Krebs y E. Fischer en 1955 [2, 3]. Estos estudios sentaron las bases para el establecimiento de una de las superfamilias de proteínas más grandes en la mayoría de los eucariotas: las proteínas quinasas eucariotas (EPKs) [3]. La función primaria de las proteínas quinasas es la catálisis de la reacción reversible de transferencia de fosfato a un residuo de serina, treonina o tirosina de la molécula blanco; de esta manera, modifican las propiedades fisicoquímicas de sus sustratos. El principal dador de fosfato en la mayoría de los organismos es el nucleótido adenosina trifosfato (ATP) (**Figura 1.1**). La mayoría de las proteínas quinasas son a su vez fosfoproteínas altamente dinámicas, que oscilan entre numerosos estados conformacionales, y sus grupos fosfato resultan determinantes tanto en su estructura como en su función [2].

En humanos, el quinoma cuenta con alrededor de 500 integrantes [3] lo que representa aproximadamente el 2% del genoma [2, 4]; por otro lado, un tercio del proteoma humano está constituido por proteínas que en algún estadio se encuentran fosforiladas [2]. En plantas, el

4% del genoma codifica para proteínas quinasas [5]. Hasta la fecha se conocen las estructuras de alta resolución de aproximadamente 150 EPKs [3]. Todas estas enzimas comparten el mismo plegamiento, y los subdominios que las componen se encuentran muy conservados [4].

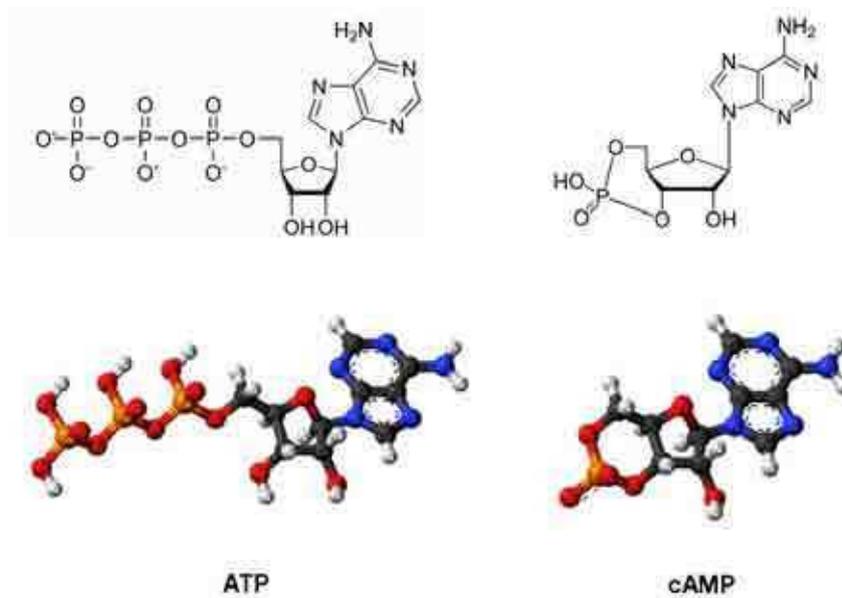


Figura 1.1: Estructura química de nucleósidos fosfato.

En el panel de la izquierda se muestra la estructura de la adenosina trifosfato (ATP), y en el de la derecha la del segundo mensajero 3'-5'-adenosina monofosfato (cAMP). En la parte superior se indica la conectividad entre los átomos, mientras que en la parte inferior se muestra un modelo molecular. Se representan en gris los átomos de carbono, en blanco los de hidrógeno, en azul los de nitrógeno, en rojo los de oxígeno, y en naranja los de fósforo.

La vía de transducción de señales de cAMP-PKA

Una de las vías de transducción de señales mejor caracterizadas en eucariotas es la que involucra al segundo mensajero 3'-5'-adenosina monofosfato (cAMP) (ver **Figura 1.1**). Su descubrimiento por parte de E. Sutherland y T. Rall en el año 1958 como efector intracelular de distintos procesos fisiológicos disparados por señales extracelulares dio origen al concepto de segundo mensajero [3].

En esta vía de señalización (**Figura 1.2**), la unión de un ligando extracelular específico a receptores acoplados a proteínas G (GPCRs) genera un cambio conformacional en la proteína G heterotrimérica asociada (conformada por las subunidades α , β y γ). Como consecuencia, el

complejo se disocia, liberando la subunidad $G\alpha_s$ (estimuladora), que a su vez interactúa con la enzima adenilato ciclasa (AC), activándola. Esta enzima es responsable de la síntesis de cAMP a partir de ATP, por lo que la señal extracelular desencadena en última instancia un aumento de la concentración intracelular del segundo mensajero, que resulta a su vez en la activación de diferentes moléculas efectoras. La duración de la señal mediada por cAMP es regulada, entre otros factores, por la variación temporal de la concentración intracelular del segundo mensajero, que resulta del balance entre la velocidad de su síntesis por la AC y la de su degradación por parte de enzimas denominadas fosfodiesterasas de cAMP (PDEs). [1, 6].

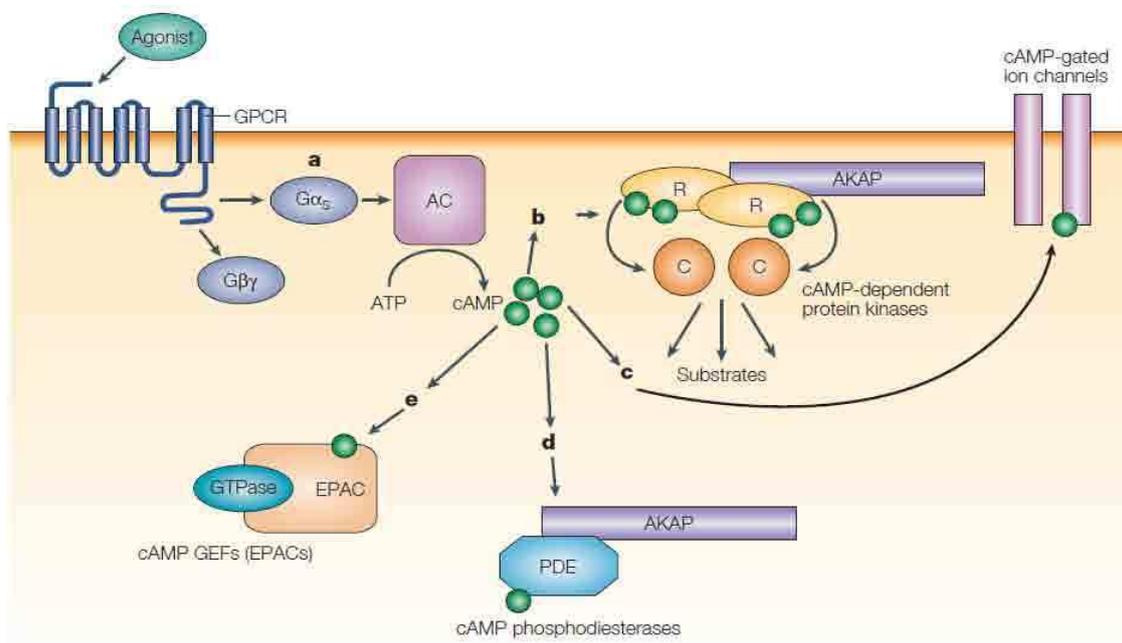


Figura 1.2: La vía de señalización de cAMP-PKA en mamíferos.

La estimulación del receptor acoplado a proteína G (GPCR) por parte de un agonista extracelular dispara la cascada de señalización por cAMP. (a) La subunidad $G\alpha_s$ activa la adenilato ciclasa (AC), responsable de la síntesis de cAMP, lo que a su vez ejerce su efecto sobre diferentes blancos celulares: la proteína quinasa A (b), los canales iónicos disparados por cAMP (c), y el factor intercambiador de guanina activado por cAMP (EPAC) (e). Por otro lado, el cAMP es degradado por la acción de fosfodiesterasas (PDEs) (d). R: subunidad reguladora de la PKA; C: subunidad catalítica de la PKA; AKAP: proteína de anclaje de la PKA. (Figura tomada de [6]).

El principal blanco intracelular del cAMP es la proteína quinasa dependiente de cAMP o proteína quinasa A (PKA) (ver **Figura 1.2**), que fue descrita por primera vez por D. Walsh, J.

Perkins y E. Krebs en el año 1968 [7]. Esta fue la primera molécula efectora de cAMP identificada, y desde entonces ha sido la más estudiada. Posteriormente, se han identificado otras moléculas que unen cAMP en diversos organismos: la proteína activadora de genes por catabolito de *Escherichia coli* [8], el receptor quimiotáctico de *Dictyostelium discoideum* [9], los canales iónicos disparados por nucleótido cíclico [10] (ver **Figura 1.2**), y el factor intercambiador de guanina activado por cAMP (EPAC) [11] (ver **Figura 1.2**). Los dominios de unión a nucleótidos cíclicos (CNBs), y en particular los dominios de unión a cAMP (CBDs), están presentes en todos los *phyla* y se encuentran muy conservados en secuencia y estructura [12].

Desde entonces, la PKA se ha convertido en un importante modelo de estudio no sólo en las vías de transducción de señales mediadas por cAMP, sino también en un miembro prototípico de la superfamilia de las EPKs, tanto desde un punto de vista funcional como estructural, por numerosas razones. Por un lado, es un sistema que en una primera aproximación resulta relativamente sencillo de estudiar, dado que consta de dos tipos de subunidades con funciones bien definidas: una subunidad regulatoria (R) y una subunidad catalítica (C), que pueden disociarse por acción de cAMP y consecuentemente ser estudiadas por separado. Ambas proteínas son relativamente pequeñas, dado que su masa molecular es de entre 40 y 60 kDa, y pueden expresarse en bacterias conservando su funcionalidad y ser purificadas con alta eficiencia en grandes cantidades. La mayoría de los estadios del ciclo catalítico de la subunidad C han sido capturados en estado cristalino, y por lo tanto se tiene un amplio conocimiento acerca de su mecanismo de acción. [2, 4]. Por otro lado, se expresa en todos los tipos celulares de mamíferos, y está presente en una amplia diversidad de organismos, incluyendo insectos, hongos, y hasta patógenos unicelulares como *Plasmodium falciparum* y *Trypanosoma* [12].

En metazoos, la PKA se encuentra involucrada en una amplia variedad de procesos celulares: en el núcleo, participa en la regulación de la expresión génica mediante la fosforilación directa de factores de transcripción; también desempeña un papel importante en la regulación del ciclo celular, ya que participa del armado y desarmado de la envoltura nuclear durante la división celular, de la dinámica de microtúbulos, y de la condensación y decondensación de la cromatina. Por otro lado, regula una amplia variedad de procesos metabólicos mediante la fosforilación de enzimas clave de distintas vías. Dado que se expresa en todos los tipos celulares, su participación en procesos tan fundamentales hace que la PKA

sea una enzima clave en estos organismos, involucrada en la regulación de numerosas funciones básicas como la función cardiovascular, la biosíntesis de esteroides, la función reproductiva, la función inmunológica, etc. [13].

En *Saccharomyces cerevisiae*, organismo que es modelo de estudio de este trabajo, la PKA también participa del control del ciclo celular, del metabolismo, y de la regulación de la expresión génica. Se encuentra involucrada en eventos de transducción de señales asociados a procesos celulares fundamentales, como por ejemplo el crecimiento fermentativo, la detección de la disponibilidad y del tipo de nutrientes, la entrada en fase estacionaria, la respuesta a distintos tipos de estrés, etc. [14-17].

En este organismo, el mecanismo de activación de la vía de la PKA es sustancialmente más complejo [17], dado que se requiere la acción concomitante de dos vías que convergen en la activación de la AC. La primera consta de un sensor de glucosa extracelular que funciona de manera análoga a la vía de mamíferos, e involucra al GPCR Gpr1 y a la subunidad $G\alpha$ activadora de la AC, Gpa2. La segunda consiste en un sensor de glucosa intracelular, que funciona a través del intercambio de guanosina difosfato (GDP) por guanosina trifosfato (GTP) que ocurre en la proteína Ras en respuesta al aumento de los niveles intracelulares de glucosa 6-fosfato, que se genera por acción de la enzima de entrada a la glucólisis, la hexoquinasa (Hxk) (**Figura 1.3**).

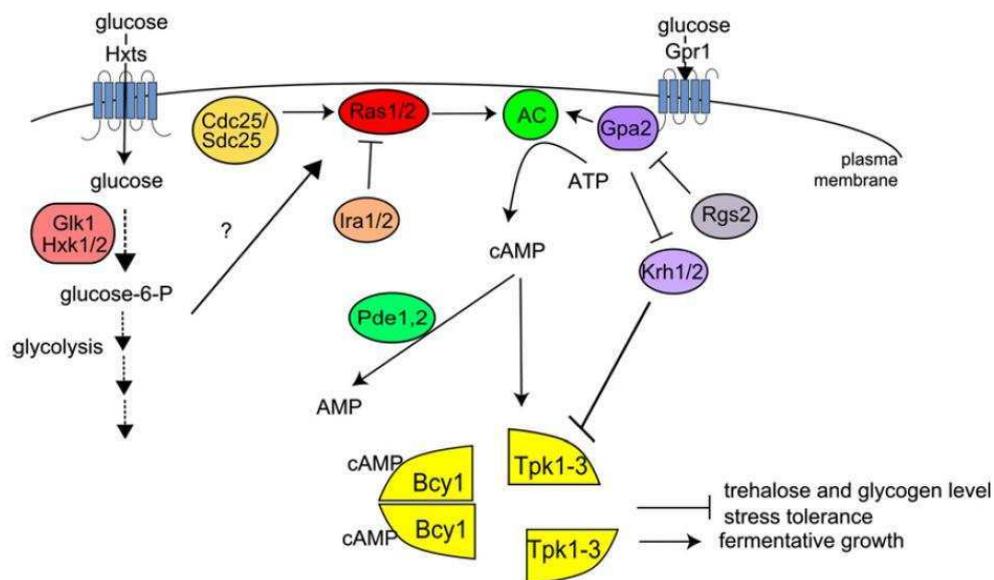


Figura 1.3: La vía de señalización de cAMP-PKA en *Saccharomyces cerevisiae* mediada por glucosa.

La activación de la AC en *S. cerevisiae* requiere de la convergencia de dos vías de señalización. La glucosa ingresa a la célula mediante transportadores de hexosas (Hxts) y es posteriormente transformada en glucosa 6-fosfato por la hexoquinasa (Hxk). La proteína Ras funciona como sensor de glucosa 6-fosfato intracelular mediante un mecanismo que aun no se conoce, pero en su forma activa (unida a GTP) es indispensable para la activación de la AC. Las proteínas Cdc25 y Sdc25 funcionan como factores intercambiadores de guanina (GEFs) que promueven el intercambio de GDP por GTP activando Ras, mientras que las proteínas Ira son proteínas activadoras de la actividad GTPasa de Ras (GAPs) que por lo tanto la inhiben. Por otro lado, el sensor de glucosa extracelular funciona mediante el GPCR de glucosa Gpr1, que al unir el ligando extracelular funciona como GEF de la subunidad $G\alpha$ de la proteína G, Gpa2, que en su estado unido a GTP activa la AC. Bcy1 y Tpk1-3 son las subunidades regulatoria y catalítica de la PKA, respectivamente, y Pde1-2 son las fosfodiesterasas, responsables de la degradación del cAMP intracelular. (Figura tomada de [17]).

La PKA es una quinasa de serina y treonina de amplio espectro que fosforila a sus moléculas blanco en respuesta a estímulos extracelulares específicos. Todas las proteínas sustrato presentan una serina o una treonina que se fosforilan en su grupo oxidrilo terminal, en alguno de los siguientes contextos: Arg/Lys-Arg/Lys-X-Ser/Thr, en donde X puede ser cualquier aminoácido [1]. La PKA pertenece a la familia de las AGC quinasas dentro de las EPKs, cuya denominación proviene de las quinasas A, G, y C, un grupo de quinasas de serina y treonina que se caracterizan por contener: (i) dominio/s N-terminal/es que determinan total o parcialmente su localización, (ii) un dominio catalítico, (iii) un dominio regulatorio C-terminal con un sitio consenso hidrofóbico [18].

Como ya se señaló, la característica más distintiva de la PKA es la existencia de dos tipos de subunidades con roles bien definidos. En la mayoría de los organismos, la forma inactiva de la PKA, que existe a bajas concentraciones intracelulares de cAMP, está representada por una holoenzima heterotetramérica formada por dos subunidades catalíticas (C) unidas a un dímero de subunidades regulatorias (R). En este estado, el sitio inhibitorio (IS) de la subunidad R ocupa el sitio activo de la subunidad C, inhibiendo su actividad. Al aumentar los niveles intracelulares de cAMP, una molécula del segundo mensajero se une con alta afinidad a cada uno de los dos dominios de unión a cAMP (CBD) presentes en cada subunidad R. De esta manera, tiene lugar un cambio conformacional que genera una disminución de la afinidad

entre las subunidades R y C, disociando consecuentemente el complejo en un dímero de subunidades R unidas a cAMP y dos subunidades C libres, catalíticamente activas (**Figura 1.4**; ver **Figura 1.2** y **Figura 1.3**). [1, 2].

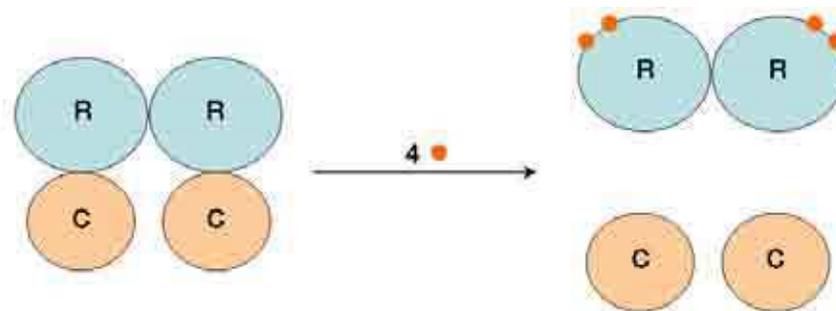


Figura 1.4: Mecanismo clásico de activación de la PKA.

Como consecuencia del aumento de los niveles intracelulares de cAMP (representado como puntos rojos), dos moléculas del segundo mensajero se unen a cada subunidad R de la holoenzima, generando un cambio conformacional que provoca su disociación en un dímero de subunidades R y dos subunidades C activas.

Resumen de los principales componentes de la vía de señalización de cAMP-PKA

Receptores acoplados a proteína G (GPCRs): son los blancos primarios de las señales extracelulares que disparan la vía de señalización por cAMP. En mamíferos existen más de 30 familias de estos receptores [3], y por lo tanto los GPCRs representan probablemente la superfamilia de proteínas más grande en estos organismos. En *S. cerevisiae*, el GPCR que funciona como sensor extracelular de glucosa es Gpr1.

Subunidad $G\alpha_s$ de la proteína G: en su estado inactivo, se encuentra unida a guanosina difosfato (GDP) y forma parte de un complejo heterotrimérico junto con las subunidades $G\beta$ y $G\gamma$, que se encuentra asociado a los GPCRs. Como consecuencia de la unión de un ligando extracelular al receptor, el complejo se activa aumentando la velocidad de intercambio de GDP por GTP en la subunidad $G\alpha_s$, que en su forma activa (unida a GTP) se libera y estimula la actividad de la enzima adenilato ciclasa (AC). La actividad GTPasa intrínseca de esta subunidad es uno de los mecanismos de terminación de la señal. Esta proteína se expresa en

todos los tejidos y, dependiendo de en cuál se encuentre, está asociada a distintas isoformas de la AC. [1]. En *S. cerevisiae*, Gpa2 es la subunidad $G\alpha_s$ de la vía de señalización mediada por glucosa.

Adenilato ciclasas (ACs): son las enzimas responsables de la síntesis de cAMP a partir de ATP. En mamíferos, existen nueve isoformas de membrana y una isoforma citoplasmática. Todas las formas de membrana presentan una actividad basal que es aumentada como consecuencia de la unión de la subunidad $G\alpha_s$ de la proteína G, y reducida por la subunidad $G\alpha_i$ (inhibitoria). La expresión de cada isoforma es específica del tipo de tejido, e incluso depende del estadio de diferenciación celular. [13].

Fosfodiesterasas (PDEs): son enzimas responsables de la hidrólisis de los nucleótidos cíclicos (en particular, cAMP y cGMP), y desempeñan un papel fundamental en la regulación de los niveles intracelulares de estos segundos mensajeros, en particular en lo que respecta a la terminación de la señal. Por otro lado, contribuyen al establecimiento de gradientes locales de cAMP al ser dirigidas a compartimientos subcelulares específicos, proporcionando especificidad a las distintas señales disparadas por cAMP, dado que regulan la disponibilidad del segundo mensajero en las cercanías de sus moléculas blanco (ver siguiente sección). Las PDEs son una amplia superfamilia de enzimas; en mamíferos existen alrededor de 50 de estas proteínas con distintas especificidades de sustrato y características alostéricas. Existen PDEs específicas de cAMP, específicas de cGMP, y duales. [13].

Proteínas Ras: En *S. cerevisiae*, las proteínas Ras1 y Ras2 representan un componente indispensable de la vía de señalización por cAMP-PKA. Forman parte de la maquinaria de detección de los niveles intracelulares de glucosa 6-fosfato, que al aumentar disparan mediante un mecanismo desconocido la activación de Ras, que es indispensable para la activación de la AC. [17].

Proteínas Ira: En *S. cerevisiae*, las proteínas Ira1 e Ira2 funcionan como inhibidoras de las proteínas Ras, mediante su actividad estimuladora de la actividad GTPasa intrínseca de éstas. La hidrólisis de GTP genera proteínas Ras unidas a GDP, que son inactivas desde el punto de vista de la activación de la AC. [17].

Proteínas de anclaje de la PKA (AKAPs): son proteínas de andamiaje fundamentales en la regulación de la señal mediada por cAMP en mamíferos, dado que son responsables de proporcionar una localización subcelular adecuada a la PKA (entre otros componentes). Constituyen una familia de proteínas estructuralmente muy diversa, que se clasifican en base a su funcionalidad. Existen hasta el momento unos 50 miembros caracterizados, que se expresan diferencialmente en los distintos tejidos, y se encuentran vinculadas a diversas funciones y procesos fisiológicos. [1, 13].

Proteínas fosfatasas (PPs): son enzimas que catalizan la reacción inversa a la PKA, es decir, la remoción del grupo fosfato de las proteínas blanco, y participan por lo tanto en la terminación de la señal disparada por cAMP.

Proteína quinasa A (PKA): es la proteína quinasa activable por el segundo mensajero cAMP. Se encuentra conformada por dos tipos de subunidades, que se describen a continuación.

Subunidad C de la PKA: es la subunidad responsable de la reacción de transferencia de fosfato. Se activa al disociarse de la subunidad R como consecuencia de la unión de cAMP a la holoenzima heterotetramérica inactiva.

En mamíferos existen tres isoformas: C α , C β y C γ , codificadas por tres genes distintos. Entre C α y C β las secuencias de aminoácidos difieren aproximadamente en un 7%. La isoforma C α se expresa constitutivamente en todos los tejidos, mientras que la expresión de C β es dependiente del tejido; la presencia de C γ sólo se ha demostrado en testículos. Esta isoforma presenta una identidad del 79% y 83% con C α y con C β respectivamente. [1, 13].

En *S. cerevisiae* existen tres genes que codifican para la subunidad C, TPK1, TPK2 y TPK3 [19]. Las secuencias de las tres proteínas tienen una identidad de alrededor del 50% con las subunidades C de mamíferos. La identidad entre Tpk1 y Tpk3 es del 86%, entre Tpk2 y Tpk1 de 75% y entre Tpk2 y Tpk3 de 76%. Los genes no son esenciales, dado que la mutación de uno e incluso de dos de ellos no altera la viabilidad; sin embargo, la triple mutante *tpk1 tpk2 tpk3* es inviable, por lo que se requiere de al menos uno de ellos para proporcionar la mínima actividad de PKA necesaria para la supervivencia celular [19]. Por lo tanto, los tres genes son parcialmente redundantes, y efectivamente comparten la mayor parte de los sustratos [20];

de hecho, *tpk1* y *tpk2* presentan la misma actividad específica para muchos de sus sustratos [21].

Subunidad R de la PKA: es la principal proteína regulatoria de la actividad de la subunidad C. Además, es el principal blanco intracelular del cAMP, y funciona como molécula adaptadora que vincula la actividad catalítica de la PKA con las AKAPs.

En mamíferos existen cuatro isoformas (RI α , RI β , RII α y RII β) que presentan una alta identidad de secuencia en los dominios de unión a cAMP (CBDs), mientras que el grado de identidad varía en el resto de la molécula. Como se discutirá más adelante, las cuatro isoformas desempeñan distintos roles, presentan distinto patrón de expresión en los diferentes tejidos, y difieren en numerosas características estructurales; además, son funcionalmente no redundantes. [1, 5, 13].

En la gran mayoría de los hongos (incluyendo a *S. cerevisiae*), existe solamente un gen que codifica para la subunidad R. En *S. cerevisiae* este gen se denomina BCY1 (bypass of cyclic AMP dependence) [22, 23]; la proteína Bcy1 (que es el objeto de estudio de este trabajo) presenta alrededor de 40% de identidad de secuencia con las subunidades R de mamíferos. El gen no es esencial, dado que la mutante *bcy1* es viable; sin embargo, presenta un fenotipo alterado, muy característico de una actividad de PKA constitutivamente elevada, e insensible a la regulación por cAMP. Algunos de los fenotipos más característicos son: la mutante *bcy1* es incapaz de alcanzar la fase estacionaria en cultivo (lo que evidencia el rol que desempeña la PKA en la regulación del ciclo celular); por otro lado, no tiene la capacidad de crecer en fuentes de carbono no fermentables; finalmente, como consecuencia de la actividad de PKA constitutiva, no acumula el polisacárido de reserva glucógeno, presenta baja tolerancia al estrés térmico, y es incapaz de esporular [20, 23, 24].

Como comentario final de esta sección, es importante destacar la diversidad de isoformas que presentan los distintos componentes de esta vía de señalización. Considerando que una amplia variedad de señales extracelulares (que requieren en última instancia respuestas intracelulares específicas) convergen en esta vía de señalización, siendo el resultado último en todos los casos la síntesis de cAMP, una interrogante clave ha sido cómo es capaz el mismo segundo mensajero de regular específicamente tan amplia diversidad de procesos celulares. Para lograrlo, se requieren niveles adicionales de regulación para garantizar que cada señal

extracelular genere selectivamente el efecto adecuado, mediante la fosforilación de sustratos específicos [13]. Esto se logra esencialmente mediante dos mecanismos: la expresión específica en cada tejido de las distintas isoformas de los componentes involucrados en la vía de señalización (principalmente, en lo que respecta a los GPCRs, las AKAPs, y las subunidades R de PKA), y la regulación de la localización subcelular de los distintos componentes, que son posteriormente ensamblados en “focos” discretos de señalización [3]. En las siguientes secciones se explican en profundidad estos mecanismos regulatorios, y se describen en mayor detalle los componentes de la vía de señalización de cAMP-PKA más relevantes para este trabajo.

Microdominios de cAMP y localización subcelular de la PKA

Inicialmente, la visión del cAMP como segundo mensajero era la de una molécula capaz de difundir libremente a través de la célula, alcanzando mediante este mecanismo todas las regiones subcelulares; sin embargo, posteriormente se observó mediante técnicas de microscopía avanzada la existencia de gradientes locales en lugar de una distribución uniforme, lo que indica que el cAMP se acumula en regiones discretas en el interior celular. Más aun, se ha observado que la naturaleza de estos focos de cAMP es altamente dinámica, lo cual en ocasiones se manifiesta en forma de pulsos [6]. Estas regiones discretas de elevada concentración relativa de cAMP, o microdominios, se localizan en distintas regiones de la célula, y son fundamentales en la señalización por PKA. Las concentraciones locales de cAMP son reguladas por la acción opuesta de las ACs que lo sintetizan, y las PDEs que lo degradan [1, 13] (**Figura 1.5**).

En las últimas décadas, uno de los hallazgos más relevantes en biología celular fue la elucidación del rol central que desempeñan las interacciones proteína-proteína en las vías de señalización. El principal ejemplo de este fenómeno son las proteínas de andamiaje y proteínas adaptadoras que funcionan como núcleos de ensamblado de grandes complejos macromoleculares. Estas proteínas sirven de plataforma para la integración y diseminación de múltiples señales; por otro lado, al localizar una determinada enzima en un compartimiento subcelular específico en un momento determinado, contribuyen a la regulación espacio-temporal de la señal y a su especificidad, dado que restringen a la enzima a un subconjunto fisiológicamente relevante de todos sus posibles blancos intracelulares (en particular, los que

se requieren para desencadenar la respuesta al estímulo específico que disparó esa señal) [25].

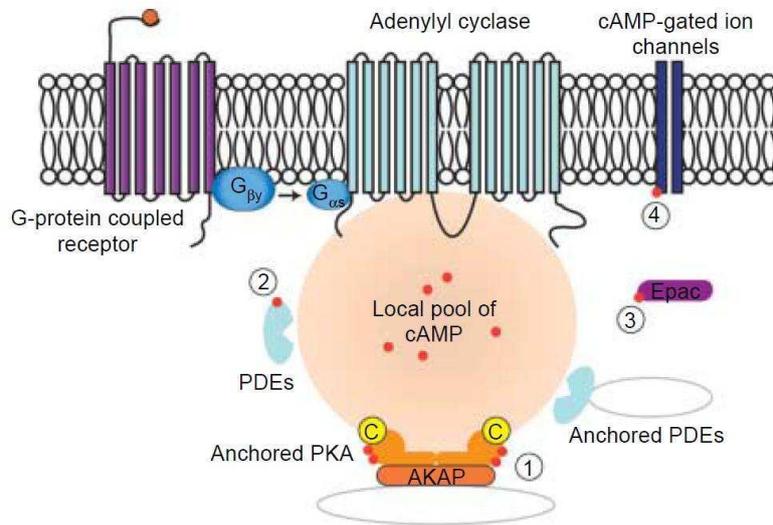


Figura 1.5: Microdominios de cAMP.

La unión de un ligando extracelular a los GPCRs activa a las ACs cercanas y genera regiones confinadas de elevada concentración de cAMP. La concentración local y la distribución de los gradientes es controlada por las PDEs que lo degradan (2). Algunos GPCRs específicos son confinados a determinadas regiones de la membrana celular mediante su asociación con organelas y componentes del citoesqueleto. Adicionalmente, las AKAPs (1) dirigen a isoformas de la PKA específicas hacia los gradientes. Otros componentes de la vía de señalización por cAMP como EPAC (3) y los canales iónicos disparados por cAMP (4) también operan en el entorno de los microdominios. (Figura tomada de [1]).

Una de las familias de proteínas de andamiaje más relevantes son las proteínas de anclaje de la PKA (AKAPs). Como ya se señaló, las AKAPs son una familia secuencial y estructuralmente muy diversa, que se definen por lo tanto en base a su funcionalidad, que está determinada por tres características centrales: 1) todas las AKAPs contienen un dominio de unión a la subunidad R dimérica de la PKA (dominio AKB); 2) todas las AKAPs contienen un único dominio de anclaje dirigido a alguna localización subcelular específica, como ser membranas, organelas u otro tipo de estructuras; 3) la gran mayoría de las AKAPs contienen además dominios adicionales de interacción con otros componentes de la misma vía o bien componentes de otras vías, lo que les permite establecer grandes complejos macromoleculares de señalización [13]. La expresión específica en los distintos tejidos de las

diferentes isoformas de los GPCRs, las AKAPS, y la PKA, junto con su ensamblado en localizaciones subcelulares discretas, garantizan la regulación específica de las diferentes funciones celulares.

Las primera AKAP identificada fue una proteína asociada a microtúbulos (MAP2), que fue detectada como contaminante asociado a las purificaciones de las subunidades R [6]. Posteriormente, mediante numerosos abordajes, esta familia se ha ampliado hasta contener hoy alrededor de 50 miembros. Algunas AKAPs se expresan ubicuamente, mientras que otras presentan una función específica en tipos celulares especializados. Hasta la fecha, solo se han identificado AKAPs en un número reducido de organismos, que incluye mamíferos [6], *Drosophila* [26], y *Caenorhabditis elegans* [27].

Actualmente, es ampliamente aceptado que la interacción de las AKAPs con la PKA ocurre a través del dímero de subunidades R de la holoenzima; en particular, como se discutirá más adelante, todas las subunidades R contienen en su extremo N-terminal una secuencia de aproximadamente 50 aminoácidos que es responsable de la dimerización y de la interacción con las AKAPs (dominio D/D). Luego de la disociación por cAMP de la PKA anclada, la localización de la subunidad C no está restringida por la interacción con la subunidad R, y puede por lo tanto fosforilar los sustratos que se encuentran cercanos en el espacio. Por otro lado, como ya se señaló, las AKAPs no solo sitúan a la PKA en regiones específicas cercanas a sus sustratos, sino que también interactúan con las PDEs, los GPCRs, las ACs, y las proteínas sustrato de la PKA, entre otros componentes. De esta manera, la regulación de la vía del cAMP ocurre a diferentes niveles: la expresión de los distintos componentes, la generación localizada del cAMP, la regulación dinámica de la localización de los efectores, y la terminación de la señal.

En *S. cerevisiae*, la localización de las dos subunidades de la PKA es dinámica y dependiente de la disponibilidad de nutrientes. En lo que respecta a las subunidades C, durante el crecimiento exponencial en glucosa Tpk2 se localiza en el núcleo, mientras que Tpk1 y Tpk3 se encuentran distribuidas entre el núcleo y el citoplasma. En células creciendo en fuentes de carbono no fermentables o en fase estacionaria, las subunidades C de la PKA se localizan en el citoplasma, encontrándose Tpk2 y Tpk3, asociadas a *P-bodies* o *EGP-bodies* [28].

En cuanto a Bcy1, en células creciendo en glucosa se encuentra localizada principalmente en el núcleo; por otro lado, en células que se encuentran creciendo en fuentes de carbono no fermentables o en fase estacionaria, se localiza en el citoplasma [29] pero nunca se encuentra asociada a *P-bodies* o *EGP-bodies* [28], que son estructuras que se generan en el citoplasma bajo ciertas condiciones de estrés. La localización de la fracción nuclear de Tpk1 en condiciones de fase exponencial en glucosa depende de los niveles de cAMP, por lo que se postula que esa localización es consecuencia de la interacción con Bcy1 para formar una holoenzima nuclear [29]. Este comportamiento es opuesto al que se observa en células de mamíferos: la subunidad RII (y consecuentemente la subunidad C asociada) se encuentra típicamente anclada a organelas mediante AKAPs, y excluida del núcleo; cuando se adiciona cAMP, se observa que la subunidad C se libera de esos complejos e ingresa al núcleo por difusión [30].

El patrón de localización descrito para Bcy1 se encuentra comprometido si se delecionan los primeros 48 aminoácidos del extremo N-terminal de la molécula, y totalmente alterado si se delecionan los primeros 124 aminoácidos de ese extremo [29]; como ya se señaló, en esta región de la molécula se encuentra el dominio responsable de la interacción con AKAPs y consecuentemente de la localización subcelular de la subunidad R de mamíferos. Si bien se ha estipulado que el dominio D/D de *S. cerevisiae* no resulta similar en secuencia al de mamíferos [29, 31], y hasta la fecha no se han reportado proteínas que puedan definirse como AKAPs en este organismo, la idea de que en *S. cerevisiae* la localización subcelular de la subunidad R se encuentra mediada por proteínas de anclaje resulta una interesante hipótesis de trabajo [29].

Por otro lado, se ha reportado que Bcy1 se fosforila en su extremo N-terminal (1-124) en dos segmentos enriquecidos en residuos de serina, denominados *cluster I* y *cluster II*. El *cluster I* se ubica en el dominio D/D (Ser3, Ser4, Ser9), mientras que el *cluster II* se ubica en una región intrínsecamente desestructurada ubicada hacia el extremo C-terminal del dominio D/D, denominada *linker I* (Ser74, Ser77, Ser79, Ser81, Ser83 y Ser84). La fosforilación en estos dos *clusters* desempeña un papel fundamental en la determinación de la localización subcelular de Bcy1 ya que las distintas isoformas presentan una localización diferencial: las formas fosforiladas se encuentran principalmente en el citoplasma mientras que la fracción que permanece sin fosforilar se localiza en el núcleo. El estado de fosforilación de estos dos *clusters* de serina y la consecuente localización subcelular de Bcy1 se encuentra regulado por la fuente de carbono, de manera dependiente de la expresión del gen YAK1. [32].

Yak1 es una proteína quinasa cuya actividad es antagónica a la de la PKA [33], y su expresión es dependiente de los factores de transcripción Msn2 y Msn4, que son regulados negativamente por la PKA [34]. Esto sugiere que la PKA podría regular su propia localización subcelular: en presencia de glucosa (que es una situación de altos niveles de cAMP y por lo tanto alta actividad PKA), la PKA (que se encuentra localizada en el núcleo bajo estas condiciones) inhibe la expresión de Yak1 mediante los factores de transcripción Msn2 y Msn4, y por lo tanto Bcy1 se mantiene en su forma no fosforilada, nuclear; en fuentes de carbono no fermentables (cuando la actividad PKA se encuentra atenuada), se dispara la expresión de Yak1, lo que a su vez repercute en un aumento de la fosforilación de Bcy1, disparando su traslocación al citoplasma [32] (**Figura 1.6**).

En este mismo trabajo, se identificaron por el sistema de doble híbrido dos proteínas, Zds1 y Zds2, que interactúan con la región N-terminal (1-124) de Bcy1; estas dos proteínas se encuentran involucradas en la regulación del ciclo celular [35]. Se observó que la localización citoplasmática de Bcy1 en células creciendo en fuente de carbono no fermentable requiere de la presencia de estas proteínas [32]. Más aun, parece existir un efecto diferencial en cuanto al rol que desempeñan los dos *clusters* de serinas fosforilables en la localización mediada por Zds, dado que la localización citoplasmática mediada por Zds resulta dependiente de la fosforilación del *cluster* II pero no de la del *cluster* I. Consecuentemente, se postula en este trabajo que la fosforilación del *cluster* II de serinas incrementaría la afinidad de Bcy1 por Zds reteniendo a Bcy1 fosforilada en el citoplasma, por lo que esta proteína podría desempeñar funciones similares a las de las AKAPs, pese a que no presenta las características de esta familia. Por otro lado, dado que la localización citoplasmática relacionada con la fosforilación del *cluster* I no requiere de Zds, se propone la participación de otra proteína o factor aun no identificado que interactúe con esta región. [32].

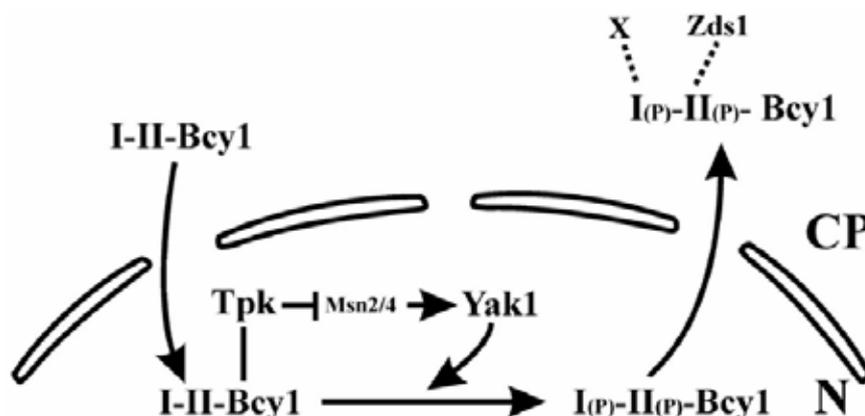


Figura 1.6: Regulación de la localización subcelular de Bcy1 mediada por Yak1 y Zds1.

Las formas no fosforiladas de Bcy1 (I-II-Bcy1) se localizan preferencialmente en el núcleo (N), mientras que las formas fosforiladas (I_(P)-II_(P)-Bcy1) son traslocadas al citoplasma (CP). En condiciones de alta concentración de cAMP (por ejemplo, crecimiento en fase exponencial en glucosa) la PKA se encuentra activa y localizada principalmente en el núcleo, donde inhibe a los factores de transcripción Msn2 y Msn4 evitando la expresión de la quinasa Yak1. Cuando las concentraciones de cAMP son bajas (por ejemplo, crecimiento en fase exponencial en fuente de carbono no fermentable), la actividad de PKA es baja, lo que dispara la expresión de Yak1, que fosforila a su vez a Bcy1 en sus dos *clusters* de serinas (I y II). Como consecuencia de esta fosforilación, Bcy1 es retenida en el citoplasma mediante la interacción del cluster II en su estado fosforilado con Zds1, y la del cluster I en su estado fosforilado con un factor desconocido (X). (Figura tomada de [32]).

Como ya se señaló, hasta la fecha no se han reportado proteínas que puedan definirse como AKAPs en *S. cerevisiae* en base a los criterios previamente establecidos. Sin embargo, recientemente ciertas proteínas interactoras de Bcy1, algunas de las cuales podrían presentar esas características, han sido reportadas por nuestro grupo de investigación [36]. En ese trabajo, se demostró que las proteínas Ira2 (que, como ya se mencionó, es una proteína inhibitoria de Ras), Hsp60 (una chaperonina mitocondrial), Myo2 (miosina, una proteína motora asociada al citoesqueleto) y Ptp1 (una fosfatasa específica de fosfotirosina) interactúan *in vitro* con Bcy1. Adicionalmente, se encontraron los péptidos de esas proteínas que son responsables de la interacción, y se elucidaron algunos de los determinantes que la rigen (que se discutirán más adelante).

La primera característica de estas interacciones que postula a algunas de estas proteínas como “candidatas” a AKAPs de levaduras es que la misma es dependiente del extremo N-terminal (1-85) de Bcy1, donde se encuentra localizado el dominio D/D que, como ya se señaló, en mamíferos es responsable de la interacción con las AKAPs. La candidata más fuerte de ese grupo resulta ser Ira2, que además de interactuar *in vitro* con el extremo N-terminal de Bcy1, colocaliza con ésta *in vivo* en fracciones de membrana. Por otro lado, dos péptidos de Ira2 que contienen los residuos Ser1745 (en un contexto de PKA fuerte) y Ser1018, resultaron ser fosforilados por PKA *in vitro*, por lo que Ira2 es sustrato de la PKA. Finalmente, las dos proteínas participan de la misma vía de señalización (ver **Figura 1.3**). [36].

Estructura de la subunidad catalítica de la PKA

Como ya se mencionó, en mamíferos existen 3 isoformas para la subunidad catalítica: $C\alpha$, $C\beta$ y $C\gamma$. Dado que la primera es la única que se expresa constitutivamente en todos los tejidos, ha sido históricamente la más estudiada. En el año 1991 se obtuvo la estructura cristalina de un complejo ternario formado por la subunidad $C\alpha$, ATP, y un fragmento correspondiente a los aminoácidos 5 a 24 del inhibidor de PKA (PKI), que es un péptido que regula la actividad y la localización de la subunidad C [37, 38]. Esta fue la primera estructura de alta resolución de una EPK, y permitió definir los elementos fundamentales que se encuentran conservados en todos los miembros de la superfamilia. Desde entonces, numerosas estructuras de diversas EPKs han sido resueltas, confirmando la universalidad de esos determinantes [12].

La estructura consta de un núcleo catalítico bilobular muy conservado en toda la superfamilia de las EPKs, cuyos residuos se agrupan alrededor del bolsillo correspondiente al sitio activo que se encuentra entre los dos lóbulos, y contribuyen a la unión de ATP o bien a la reacción de transferencia de fosfato (**Figura 1.7**). Este núcleo se encuentra flanqueado por un motivo predominantemente helicoidal de unos 40 aminoácidos hacia el extremo N-terminal, y por un largo segmento extendido de unos 50 aminoácidos hacia el extremo C-terminal, intrínsecamente desordenado en ausencia de ligando; estos segmentos desempeñan funciones regulatorias [2, 12].

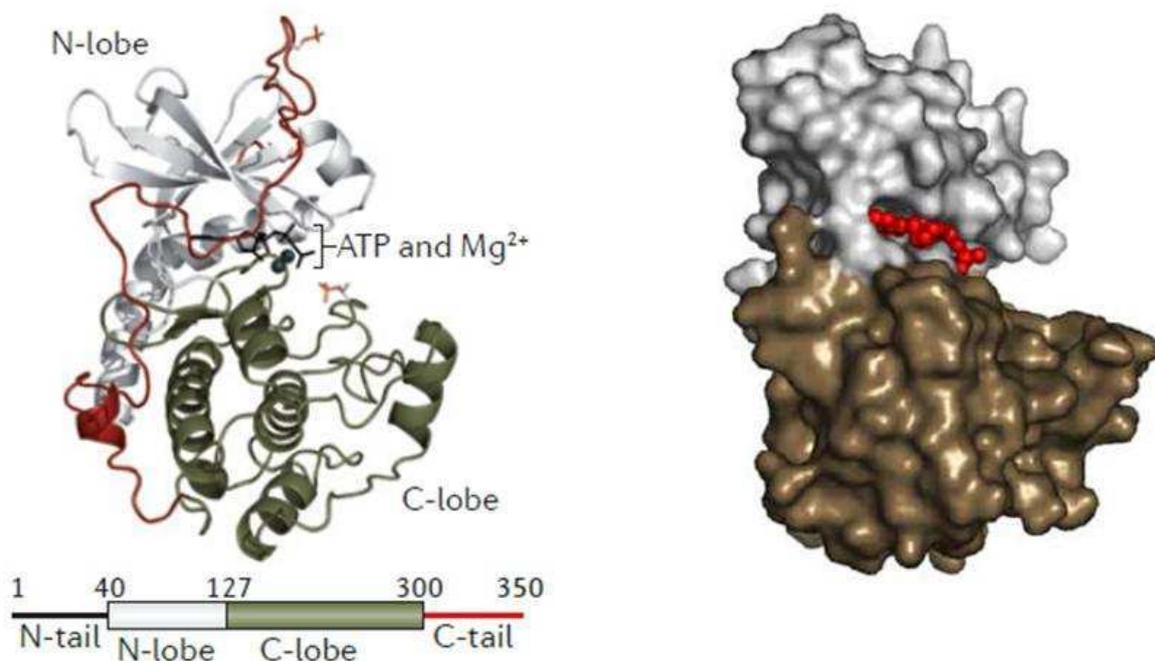


Figura 1.7: Estructura global de la subunidad catalítica de la PKA.

En el panel de la izquierda se muestra el plegamiento de la enzima, donde pueden apreciarse los distintos elementos de estructura secundaria así como también los lóbulos menor (*N-lobe*) y mayor (*C-lobe*). El esquema en la parte inferior indica la extensión y el código de colores utilizado para cada una de las regiones. También se muestra en el bolsillo catalítico la posición del ATP y del Mg^{2+} , cofactor de la enzima. En el panel de la derecha se muestra una representación de la superficie del núcleo catalítico solamente (residuos 40-300), con el lóbulo menor (40-127) en blanco, y el lóbulo mayor (127-300) en marrón. En el centro se aprecia la posición que ocupa el ATP. (La figura de la izquierda fue tomada de [12], y la de la derecha de [4]).

El núcleo bilobular consta de dos dominios que presentan distintas funcionalidades y características dinámicas. El lóbulo menor (lóbulo N-terminal, o lóbulo N), más dinámico y maleable, contribuye principalmente al sitio de unión a ATP, mientras que el lóbulo mayor (lóbulo C-terminal, o lóbulo C), más rígido, presenta una superficie expuesta para la interacción con el péptido/sustrato y contiene la mayor parte de la maquinaria catalítica [2]. Los dos lóbulos forman una hendidura profunda en la cual se une el ATP.

El lóbulo menor está formado por cinco hebras β dispuestas de forma antiparalela, y una α -hélice denominada hélice αC (**Figura 1.7**, panel de la izquierda). Los principales motivos conservados en este lóbulo son: el *loop* rico en glicina, la hoja β -3, y la hélice αC . El *loop* rico en glicina (**Figura 1.8**) es una característica muy distintiva de toda la superfamilia de las EPKs, ya que constituye un elemento fundamental del sitio de unión a ATP, siendo su principal función proporcionar el posicionamiento adecuado que requiere el grupo γ -fosfato del ATP para la reacción de transferencia de fosfato. Este *loop* es el elemento más flexible de todo el núcleo catalítico. Los residuos de glicina permiten un acercamiento muy estrecho a los fosfatos del ATP, y como consecuencia de esta unión el *loop* sufre un cambio conformacional profundo [2, 5].

Por otro lado, el lóbulo mayor está formado por cinco α -hélices (**Figura 1.7**, panel de la derecha) y contiene los principales elementos de la maquinaria catalítica, que se encuentran ubicados en *loops* que conectan las distintas α -hélices: el *loop* catalítico, el *loop* de posicionamiento de magnesio, el *loop* de activación, y el *loop* de posicionamiento del sustrato (*loop* P+1) (**Figura 1.8**). La complejidad de estos elementos es realmente llamativa, dado que casi todos los residuos desempeñan una función específica [2]. Dado que la subunidad C no es

objeto central de estudio de este trabajo, se mencionará muy brevemente la función de cada uno de ellos.

El *loop* catalítico (residuos 166-171) es un elemento bastante rígido, y contiene residuos muy conservados que se encuentran involucrados en la reacción de transferencia de fosfato, siendo Asp 166 el residuo que actúa como nucleófilo. El *loop* de posicionamiento de magnesio (residuos 184-187) también contribuye a la catálisis mediante el residuo Asp 184, que coordina al ion Mg^{2+} (cofactor de todas las quinasas), el cual desempeña la función de coordinar los grupos fosfato β y γ del ATP, volviendo el enlace que los conecta más lábil. El *loop* de activación (residuos 195-197) contiene un residuo clave, Thr 197, que es fosforilado autocatalíticamente y resulta crucial para el establecimiento de una red de interacciones que conectan a este grupo fosfato con el *loop* catalítico y el *loop* de posicionamiento de magnesio. Finalmente, el *loop* P+1 (residuos 198-205) conforma un sitio de unión para los residuos de la secuencia consenso del sustrato; en particular, contiene una cavidad hidrofóbica que puede acomodar al residuo P+1 (es decir, el residuo contiguo hacia el C-terminal del residuo de serina o treonina fosforilable) del sustrato.

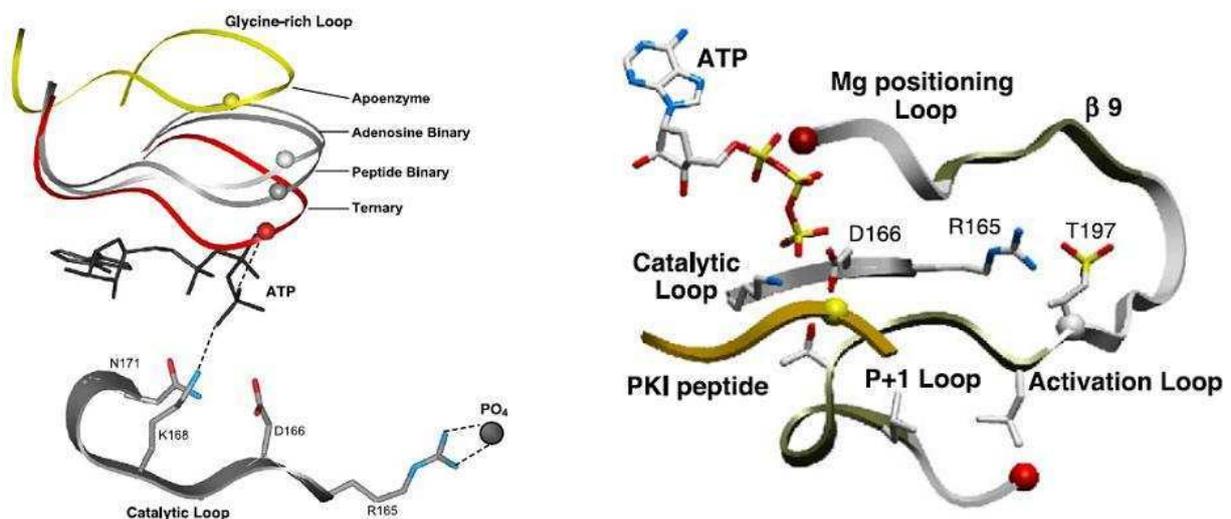


Figura 1.8: Elementos característicos de la subunidad catalítica de la PKA.

En el panel de la izquierda se ilustran los cambios conformacionales que sufre el *loop* rico en glicina en los distintos estados de la subunidad C: estado abierto, representado por la apoenzima (amarillo), complejos binarios (o con sustrato, o con ATP) (gris y blanco, respectivamente), y estado cerrado, representado por el complejo ternario con sustrato y ATP (rojo). Por otro lado, se muestra también la posición del ATP y la del *loop* catalítico; se indican los principales residuos involucrados en la catálisis; la Arg 165 coordina el grupo fosfato de la Thr 197 fosforilada del *loop* de activación. En el panel de la

derecha se muestran el segmento de activación del lóbulo mayor (residuos 184-208), formado por: el *loop* catalítico, el *loop* de posicionamiento de magnesio, el *loop* de activación, y el *loop* de posicionamiento del sustrato (*loop* P+1). El sustrato se muestra en amarillo (*PKI peptide*) con su residuo de serina fosforilable remarcado como una esfera. Se muestra también el residuo Asp 166 para resaltar el rol que desempeña en la reacción de transferencia del grupo γ -fosfato del ATP, y la Thr 197 fosforilada en el *loop* de activación para mostrar el rol que desempeña en el mantenimiento de la estructura mediante la interacción con Arg 165 del *loop* catalítico. (Figura tomada de [5])

La subunidad C se ha cristalizado con una amplia variedad de péptidos correspondientes a distintos sustratos e inhibidores, y también con diferentes análogos de nucleótidos. Estas estructuras permitieron obtener información acerca de cómo los diferentes elementos se unen al núcleo catalítico, así como también acerca de los distintos estados conformacionales que adopta la enzima en las distintas etapas de su ciclo catalítico (**Figura 1.8**, panel de la izquierda). Concretamente, se ha observado que en ausencia de sustratos, esencialmente no existen contactos entre los lóbulos del núcleo catalítico, y la enzima se encuentra por lo tanto en una conformación abierta; como consecuencia de la unión de los sustratos, se adquiere una conformación cerrada. Estas observaciones permitieron establecer el concepto de la plasticidad conformacional de las subunidad C, que es una característica de todas las EPKs, y hace referencia a que la posición relativa de ambos lóbulos es variable como consecuencia de la apertura y el cierre del sitio activo a medida que la enzima atraviesa su ciclo catalítico [2].

Por otro lado, la comparación de numerosas proteínas quinasas en sus conformaciones activa e inactiva, permitió definir cómo los residuos hidrofóbicos del núcleo de la quinasa contribuyen al ensamblado de la proteína y a su arquitectura general. Este núcleo estable pero dinámico está constituido por tres elementos esenciales, que se encuentran conservados en todas las EPKs, y es bastante distinto a los núcleos hidrofóbicos de la mayoría de las proteínas globulares.: una hélice hidrofóbica proporcionada por el lóbulo mayor (hélice αF), y dos “columnas vertebrales” (*spines*) también hidrofóbicas que están conformadas por residuos no contiguos de ambos lóbulos. Estos dos elementos proporcionan una conexión firme pero flexible entre los dos lóbulos, y se encuentran anclados a la hélice αF . [3].

Todas las quinasas contienen una *spine* conformada por dos residuos del lóbulo menor y dos del lóbulo mayor, cuya conformación depende del estado de activación de la subunidad C: este elemento está quebrado en las conformaciones inactivas, pero se ensamblado en las activas; por este motivo, se denomina *R-spine* (regulatoria). Típicamente, se establece como

consecuencia de la fosforilación del *loop* de activación, de forma tal que un residuo aromático (Phe185 en la PKA) se acomoda para completar la *R-spine*, y deja un residuo de aspártico posicionado para interactuar con uno de los iones Mg^{2+} unidos al ATP. El elemento se completa con Tyr164 del *loop* catalítico, que interactúa con Phe185 por interacciones tipo *stacking*, y dos residuos hidrofóbicos del lóbulo menor: Leu95 en la hélice αC , y Leu106 en la hebra $\beta 5$ (Figura 1.9). El ensamblado de la *R-spine* de esta forma dinámica es una característica fundamental de la plasticidad conformacional de las EPKs. [3].

El otro elemento, que se ubica de forma paralela a la *R-spine*, se completa por la inserción dinámica del anillo de adenina del ATP. Este elemento se denomina *C-spine* (catalítica) (Figura 1.9). De esta manera, el ATP posiciona los dos lóbulos para que los residuos catalíticos queden perfectamente alineados para la catálisis. [3].

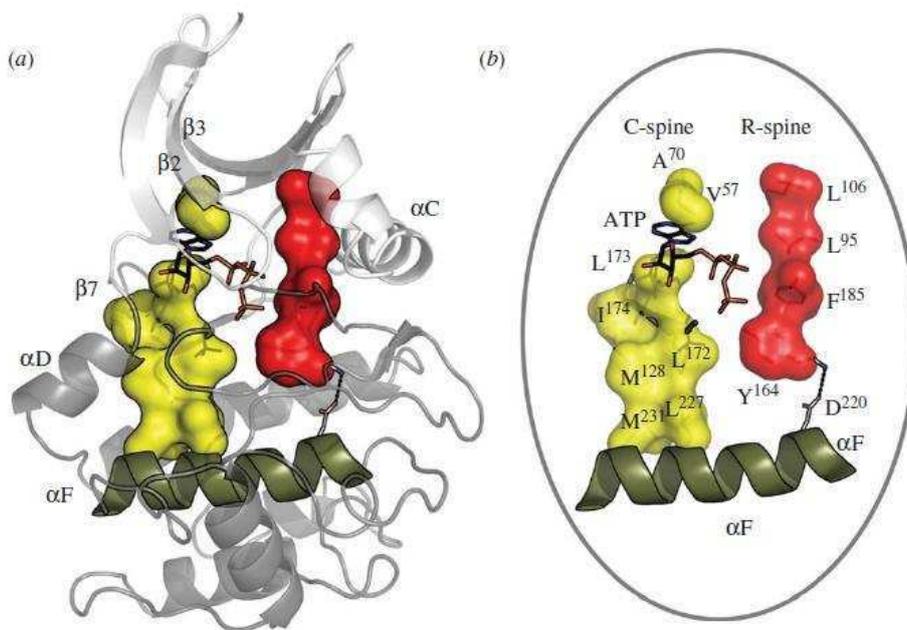


Figura 1.9: Spines hidrofóbicas que definen la arquitectura interna de la subunidad C.

En el panel de la izquierda se muestran las dos *spines* hidrofóbicas que atraviesan los dos lóbulos del núcleo de la quinasa, en la conformación activa: en rojo la *R-spine*, y en amarillo la *C-spine*. Se observa cómo ambas se encuentran ancladas a la hélice αF . En el panel de la derecha se aprecian los diferentes residuos que las constituyen: la *R-spine* contiene cuatro residuos de diferentes subdominios de la proteína, y se encuentra anclada a la hélice αF mediante el residuo muy conservado Asp220. Por otro lado, se aprecia cómo la *C-spine* se encuentra completada por el ATP. (Figura tomada de [3]).

Las proteínas quinasas en general no son catalizadores eficientes *per se*, sino que han evolucionado para funcionar como interruptores moleculares minuciosamente regulados, controlados típicamente por fosforilación. Como ya se indicó, además del núcleo bilobular conservado, la subunidad C de la PKA contiene numerosos elementos regulatorios, entre ellos dos segmentos flanqueando el núcleo a ambos lados, y dos sitios de fosforilación [12].

Los elementos N y C terminales que rodean al núcleo bilobular (ver **Figura 1.7**, panel de la derecha) funcionan como elementos regulatorios en *cis*. Por ejemplo, si bien la mayor parte de los residuos de unión a ATP son residuos del núcleo bilobular, existen dos residuos hidrofóbicos en el segmento C-terminal que forman parte del motivo Phe-Asp-Asp-Tyr (motivo FDDY), que son una parte esencial del sitio de unión; este segmento se encuentra conservado en todas las AGC quinasas. Por otro lado, el segmento N-terminal es único de la subunidad C de la PKA de mamíferos, y consiste en una larga α -hélice que se une a la superficie de ambos lóbulos del núcleo catalítico. Esta hélice contiene un sitio de miristoilación en su extremo N-terminal, es importante para la catálisis, y constituye el sitio de unión para la proteína interactora de la quinasa A 1 (AKIP1), que se encuentra involucrada en la traslocación de la PKA entre el núcleo y el citoplasma.

En cuanto a los sitios de fosforilación, el primero consiste en la Thr 197 del *loop* de activación; este residuo es fosforilado en *trans* de manera post-traducciona, ya sea por otra subunidad C, o bien por la proteína quinasa dependiente de 3-fosfoinosítido 1 (PDK1). Esta fosforilación es esencial para la actividad de la quinasa, dado que estabiliza la subunidad C en una posición que permite el acomodamiento del sustrato en una conformación extendida, en las cercanías del grupo γ -fosfato del ATP, ambos requisitos indispensables para la catálisis. La segunda fosforilación clave tiene lugar cerca del extremo C-terminal (Ser 338), y es consecuencia de una autofosforilación en *cis* que ocurre co-traduccionalmente, cuando la proteína aun está unida al ribosoma. A diferencia de lo que ocurre con la mayoría de las otras quinasas, una vez que la subunidad C se encuentra fosforilada en estas dos posiciones, los fosfatos son muy resistentes a la acción de las fosfatasas; por lo tanto, la actividad catalítica sólo depende de la liberación de la subunidad C de la holoenzima, es decir, de los niveles de cAMP [12].

En lo que respecta a *S. cerevisiae*, se conoce solamente la estructura cristalina de la mutante de delección $\Delta(1-80)$ de la isoforma Tpk1 en ausencia de ligandos [39]; esta es muy similar a la de la subunidad C α de mamíferos, lo que indica que la estructura de estas

proteínas se encuentra muy conservada aun entre organismos poco relacionados evolutivamente.

Estructura de la subunidad regulatoria de la PKA

Como ya se mencionó, en mamíferos existen cuatro isoformas de la subunidad R de la PKA: RI α , RI β , RII α , y RII β , que difieren en numerosas características: afinidad por cAMP, afinidad por la subunidad C, AKAPs con las que interactúan (y consecuentemente localización subcelular), aspectos estructurales (que serán discutidos en esta sección) y patrón de expresión en los diferentes tejidos. En *S. cerevisiae* existe una sola isoforma, Bcy1. Sus principales funciones son: regular la actividad de la subunidad C mediante su efecto inhibitorio, servir como los principales blancos intracelulares de cAMP, y proporcionar la localización subcelular adecuada a la enzima de la PKA mediante la interacción con las AKAPs.

Pese a estas numerosas diferencias, todas estas isoformas comparten la misma organización de dominios, que además se encuentra conservada en numerosos organismos [4] (**Figura 1.10**). Hacia el extremo N-terminal existe un dominio responsable del estado dimérico de la subunidad R, que sirve además de plataforma de interacción para las AKAPs que, como ya se señaló, son un componente clave de la maquinaria de transducción de señales por cAMP en mamíferos; éste se denomina dominio de dimerización y anclaje (D/D). Hacia el extremo C-terminal existen dos dominios de unión a cAMP (CBD) ubicados en *tandem*. Estas dos regiones bien definidas tienen estructuras características, y se encuentran conectadas por una región intrínsecamente desordenada denominada *linker*, que incluye un sitio inhibitorio (IS) de la subunidad C. Este sitio emula la secuencia consenso de fosforilación de los sustratos de la PKA, que en la holoenzima inactiva ocupa el bolsillo catalítico de la subunidad C. La secuencia del *linker* es altamente repetitiva y muy poco conservada, y se han predicho sitios de fosforilación y de interacción proteína-proteína en esta región [40]. La región hacia el N-terminal del IS se conoce como *linker I*, mientras que la región hacia el C-terminal del IS se denomina *linker II*. El IS consta solamente de cuatro residuos: Arg-Arg-X-Ser/Ala, donde X puede ser cualquier residuo; en la cuarta posición, las subunidades RI presentan una alanina mientras que las subunidades RII presentan una serina, por lo que son fosforiladas por la subunidad C, mientras que las subunidades RI no; esto constituye diferencia entre las dos isoformas de R.

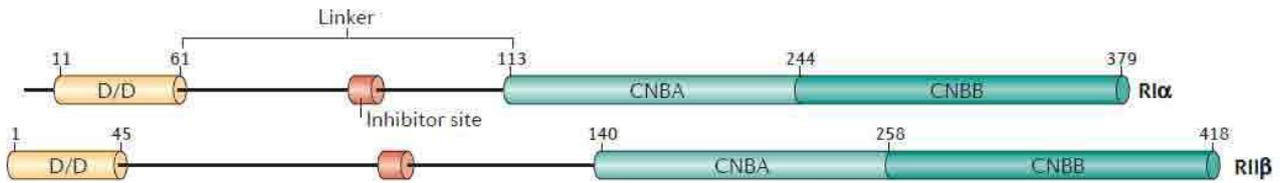


Figura 1.10: Arquitectura de la subunidad reguladora de la PKA.

Organización general de las subunidades RI y RII, tomando como ejemplos RI α y RII β (RI β es análoga a RI α , y RII α es análoga a RII β). Las distintas regiones son: D/D (dominio de dimerización y anclaje); *Inhibitor site* (sitio inhibitorio de la subunidad C); CNBA/CNBB (dominios de unión a cAMP). Los números indican la extensión de cada uno de los dominios. (Figura tomada de [12]).

Durante muchos años, no se pudo conocer la estructura de alta resolución de la subunidad R completa, debido a que el desorden intrínseco del *linker* atenta contra la cristalización; por lo tanto, inicialmente se estudiaron por separado las estructuras de los dominios D/D y de los CBD. Recientemente, sin embargo, se ha logrado resolver una estructura del homodímero de R completo [41]. Dado que el foco de este trabajo es la subunidad R de la PKA de *S. cerevisiae*, en las siguientes secciones se describirá con mayor detalle la estructura de cada uno de estos dominios.

Dominio de dimerización y anclaje (D/D)

Como ya se señaló, la función del dominio D/D es proporcionar a la subunidad R el estado dimérico en solución y servir de superficie de interacción con las AKAPs para garantizar la localización subcelular adecuada de la holoenzima de la PKA. Es importante mencionar que existen organismos en los cuales la subunidad R no presenta dominio D/D; algunos ejemplos son: *Dictyostelium discoideum* [42] y *Paramecium* [43]; en estos organismos, las subunidades R son monoméricas, y las holoenzimas son dímeros RC.

La primera estructura de alta resolución de un dominio D/D fue obtenida en el año 1999 por resonancia magnética nuclear (NMR), para la subunidad RII α [44]. Posteriormente, se resolvió por cristalografía de rayos X (DRX) la estructura de la misma construcción en su forma apo y en complejo con un péptido derivado de una AKAP [45]. Con el tiempo, también se obtuvieron las estructuras en solución [46] y cristalina [47] del dominio D/D de la subunidad RI α , así como de distintos complejos entre dominios D/D y péptidos derivados de AKAPs [47-49].

Todas las estructuras consisten en un dímero en el que las dos cadenas polipeptídicas presentan un motivo hélice-giro-hélice cada una (**Figura 1.11**). Ambas cadenas se encuentran orientadas de forma antiparalela una respecto de la otra, formando un plegamiento de tipo manojo de cuatro α -hélices en forma de X (*X-type four helix bundle*). La estructura está sostenida principalmente por residuos alifáticos de cadena larga, que constituyen un núcleo hidrofóbico de dimerización. Estos residuos se encuentran muy conservados entre todas las isoformas de la subunidad R.

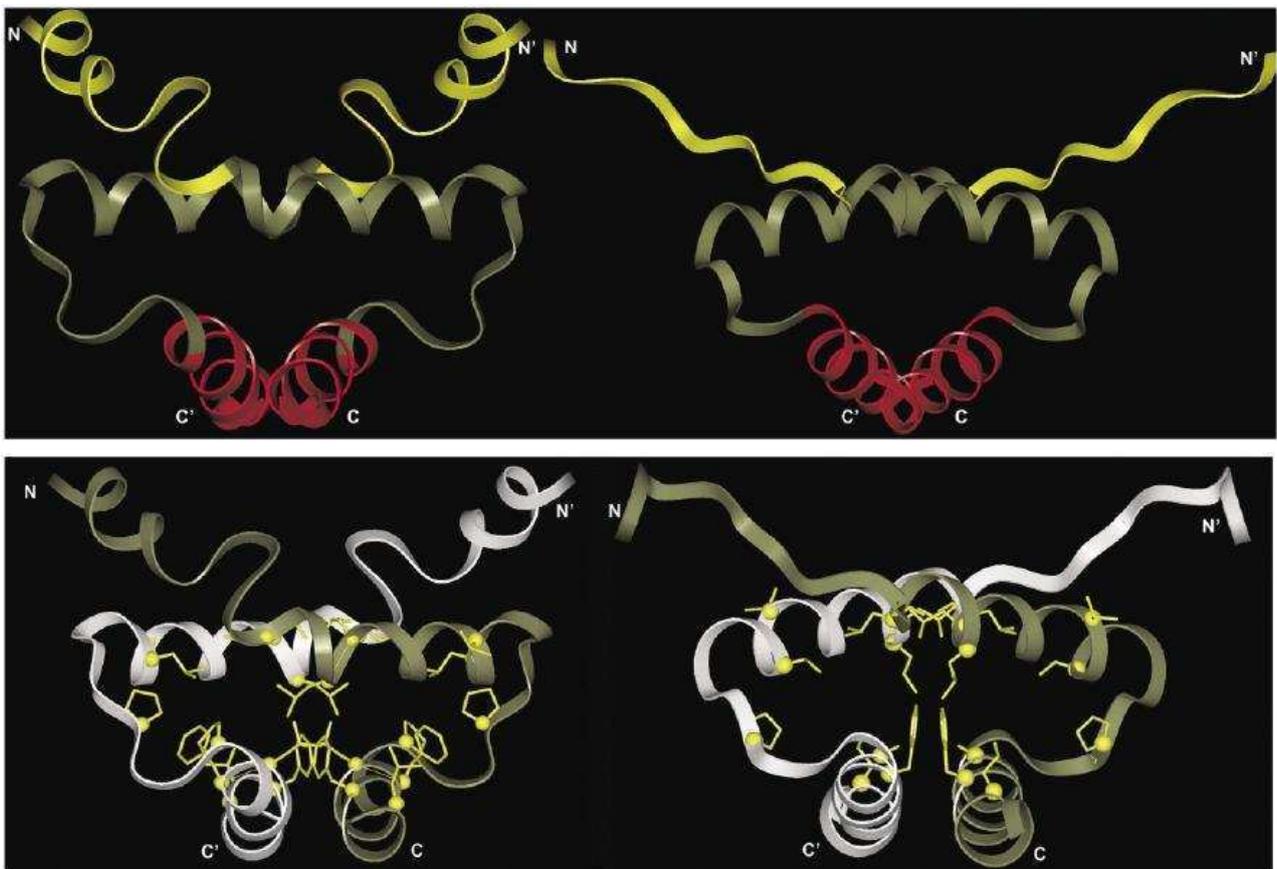


Figura 1.11: Estructura en solución de los dominios D/D de la PKA de mamíferos.

Estructuras de alta resolución obtenidas por NMR de los dominios D/D de RI α bovina (izquierda) y de RII α murina (derecha). La primera corresponde a los residuos 12-61 (pdb: 2EZW [46]) y la segunda a los residuos 1-44 (pdb: 1R2A [44]). En el panel superior se indican con código de colores los diferentes elementos: el segmento N-terminal que precede a la hélice 1 en amarillo (que en RI α forma una hélice adicional denominada hélice N-1, y en RII α se encuentra desestructurado), la hélice 1 en marrón, y la hélice 2 en rojo. Se indican asimismo los extremos N-terminal y C-terminal (que es donde continúa el *linker* en la proteína completa) con o sin primas para identificar a las dos cadenas que conforman el dímero. En el panel inferior se muestra una vista similar, ligeramente rotada para

enfaticar los aminoácidos hidrofóbicos que constituyen el núcleo de dimerización (se muestran en amarillo). Las dos cadenas se representan en diferentes colores por claridad. (Figura tomada de [46]).

La α -hélice N-terminal se denomina por convención hélice 1, mientras que la α -hélice C-terminal se denomina hélice 2. Si bien en ambas α -hélices existen residuos que resultan de importancia en mantener los contactos en el dímero, en general es reconocido que la hélice 1 constituye lo que comúnmente se conoce como subdominio 1, y que en esta región residen principalmente los determinantes de anclaje a AKAPs, mientras que en la hélice 2, o subdominio 2, se localizan principalmente los determinantes involucrados en la dimerización.

En contraste con esta región helicoidal bien conservada, existen diferencias importantes en la estructura del segmento que precede a la hélice 1 hacia su N-terminal [50]. En esta región, el dominio D/D de RI α presenta un segmento helicoidal adicional (hélice N-1) que le confiere una estructura en forma de Y, mientras que el dominio D/D de RII α contiene un segmento desestructurado que le da una apariencia más abierta (**Figura 1.12**). Estas diferencias tienen importantes implicancias en la unión a AKAPs, dado que la misma ocurre a través de esta superficie.

Existen notables diferencias en las superficies de unión a AKAPs de los dos tipos de isoformas. En los dominios D/D de las subunidades RII α se aprecia un surco hidrofóbico que sirve de sitio de anclaje para las AKAPs; en las subunidades RI α esta cavidad se encuentra parcialmente obstruida por las dos hélices N-1 a ambos lados. Otra diferencia importante entre los dos tipos de subunidades en lo que respecta al dominio D/D consiste en que las subunidades RI presentan dos puentes disulfuro intercatenarios, orientados de forma antiparalela, que se establecen entre dos residuos de cisteína muy conservados en las subunidades RI pero que no se encuentran presentes en las subunidades RII (**Figura 1.13**).

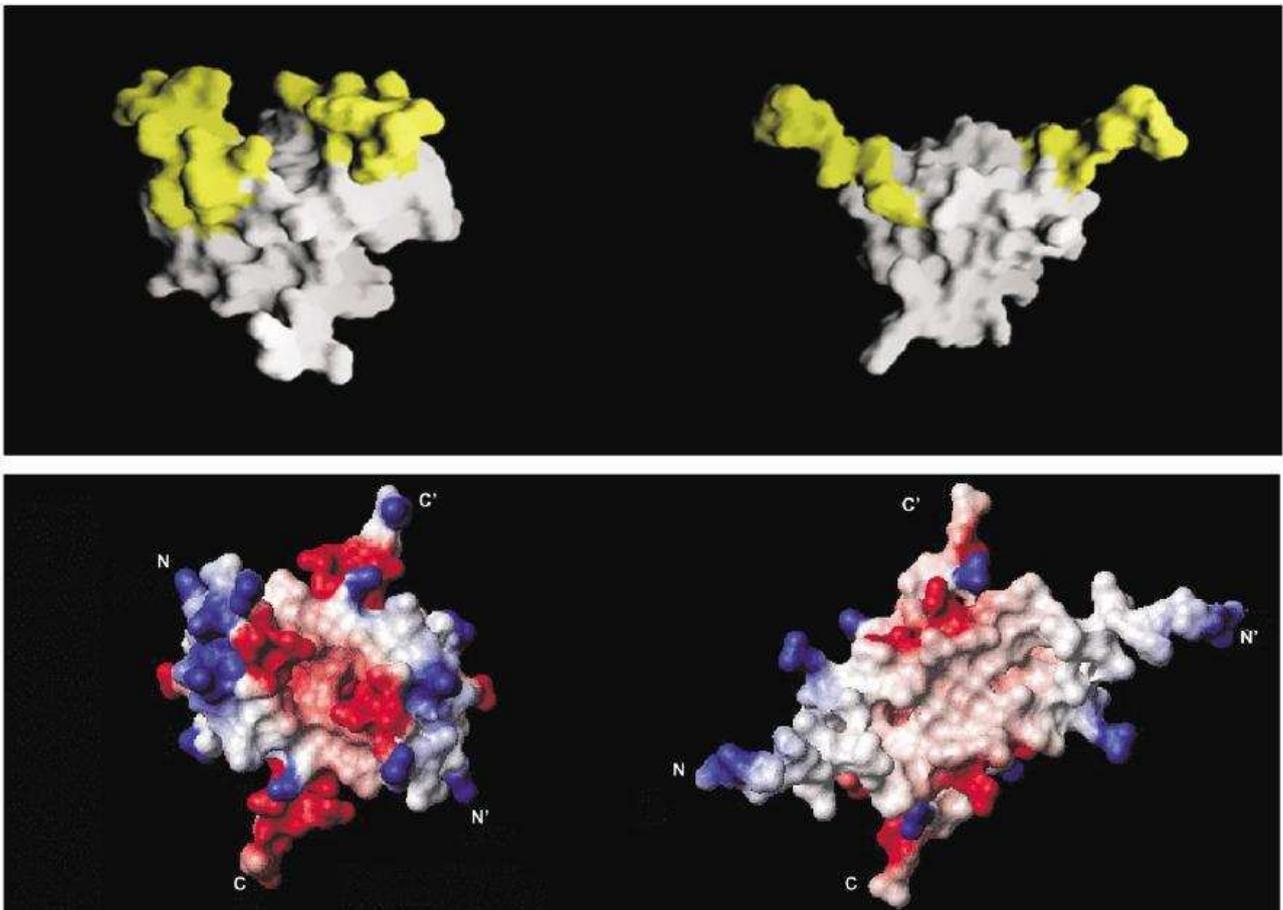


Figura 1.12: Superficie de interacción con AKAPs de los dominios D/D.

Estructuras de alta resolución obtenidas por NMR de los dominios D/D de RI α bovina (izquierda) y de RII α murina (derecha). La primera corresponde a los residuos 12-61 (pdb: 2EZW [46]) y la segunda a los residuos 1-44 (pdb: 1R2A [44]). Las imágenes del panel superior corresponden a una representación de la superficie para una vista similar a la de la **Figura 1.11**; el segmento N-terminal que precede a la hélice 1 se muestra en amarillo. En el panel inferior se muestra el potencial electrostático superficial para una vista que se encuentra rotada 90° en sentido horario respecto del eje paralelo a los bordes horizontales de la figura, exponiendo las superficies de unión a AKAPs (surcos en las figuras del panel superior). El código de colores es el comúnmente aceptado para este tipo de representaciones: residuos ácidos en rojo, residuos básicos en azul, regiones sin carga en blanco. (Figura tomada de [46]).

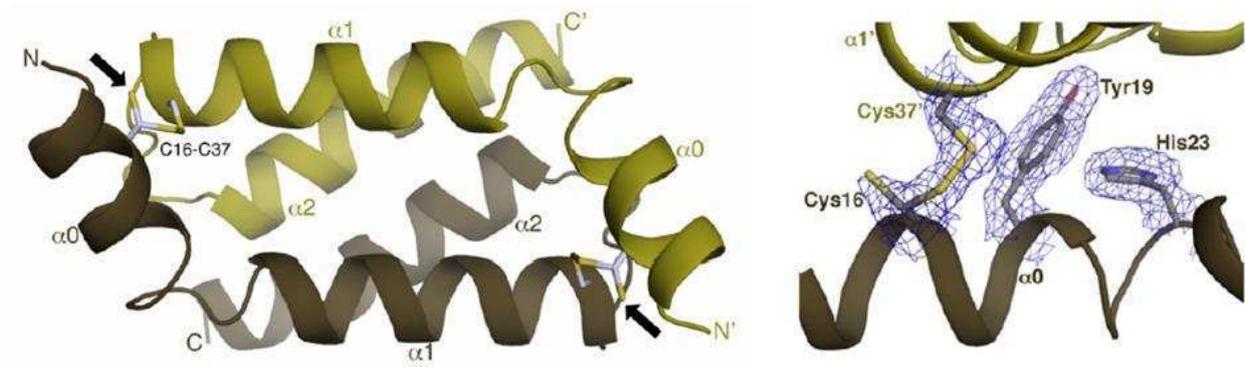


Figura 1.13: Puentes disulfuro intercatenarios de las subunidades RI.

Estructura cristalina del dominio D/D de la subunidad RI α de *Bos taurus* (pdb: 3IM3 [47]). En el panel de la izquierda se muestra una vista de la superficie de unión a AKAPs; cada cadena se representa en un color distinto, y se indican los extremos N y C terminales de cada una. α_0 denota la hélice N-1, α_1 la hélice 1, y α_2 la hélice 2. Las flechas señalan las posiciones de los dos puentes disulfuro intercatenarios que se establecen entre el residuo Cys 16 de una cadena y Cys 37 de la otra. En el panel de la derecha se muestra una ampliación de esa zona, y se indican también los contornos de densidad electrónica de los dos residuos de cisteína y de Tyr 19 e His 23, que estabilizan los puentes disulfuro por interacciones de tipo *stacking*. (Figura tomada de [47]).

Como ya se discutió, uno de los mecanismos por los cuales la PKA logra especificidad es su localización subcelular mediada por las proteínas de anclaje denominadas AKAPs. Estas proteínas se unen al dominio D/D de la subunidad R de la PKA mediante interacciones hidrofóbicas que se establecen entre las superficies que se muestran en el panel inferior de la **Figura 1.12** y en el panel de la izquierda de la **Figura 1.13**, y la cara hidrofóbica de una α -hélice anfipática que está presente en todas las AKAPs (dominio AKB). La dimerización es por lo tanto un requisito indispensable para garantizar la interacción entre la holoenzima PKA y sus proteínas de anclaje, ya que la superficie de interacción involucra a ambos monómeros. De hecho, se considera, que ésta es la principal razón por la cual la proteína es dimérica. En la siguiente sección se discute la estructura de estos complejos.

Estructura de los complejos D/D-AKB

Los dominios AKB de todas las AKAPs son segmentos de entre 14 y 18 aminoácidos que comparten una baja similitud de secuencia. Sin embargo, todos presentan dos características

distintivas: se pliegan en forma de α -hélice, y presentan un patrón repetitivo de residuos hidrofóbicos que determinan que esta α -hélice sea anfipática, dado que por su distribución en la secuencia los residuos hidrofóbicos se ubican mayoritariamente sobre una de las caras [51] (**Figura 1.14**). Esto genera una superficie de unión de alta afinidad que encaja dentro del surco hidrofóbico formado por los dominios D/D de las subunidades R de mamíferos. Cualquier mutación que perturbe la conformación de esta α -hélice (por ejemplo, la introducción de un residuo de prolina que la interrumpa), tiene efectos deletéreos sobre el anclaje de la PKA.

Las AKAPs se clasifican típicamente de acuerdo a la selectividad que presentan por cada una de las isoformas de la subunidad R; existen AKAPs específicas de RI, AKAPs específicas de RII, y AKAPs duales que pueden interactuar con los dos tipos de subunidades.

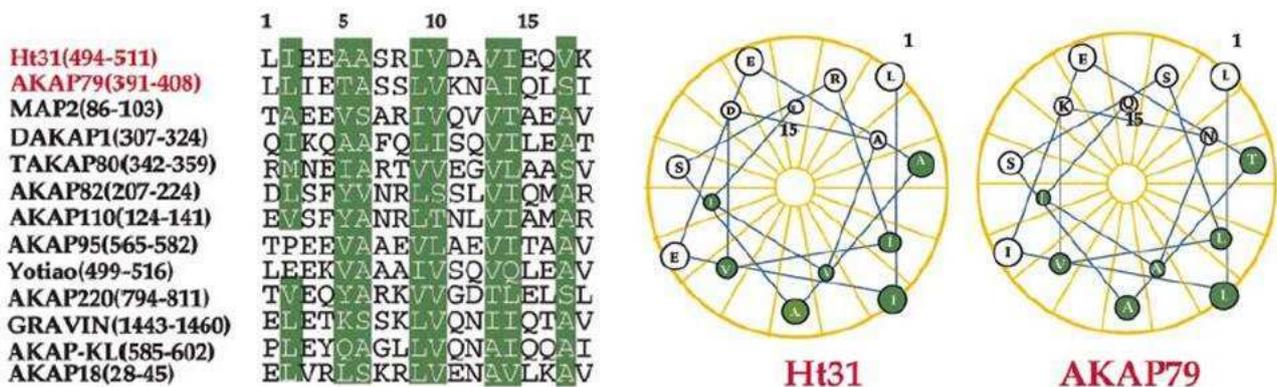


Figura 1.14: Estructura de los dominios AKB de las AKAPs.

En el panel de la izquierda se muestra un alineamiento de las secuencias de los dominios AKB de 13 AKAPs conocidas. Las posiciones correspondientes a residuos hidrofóbicos importantes para la interacción se encuentran resaltadas en verde. En el panel de la derecha se muestran los modelos correspondientes a los péptidos de unión a R de dos AKAPs muy estudiadas (indicadas en rojo en el panel de la izquierda). Se observa que los residuos hidrofóbicos resaltados en el alineamiento se agrupan en una de las caras de la α -hélice (indicados también en verde). (Figura tomada de [48]).

La mayoría de las AKAPs exhiben alta afinidad (K_a aproximadamente 10 nM) por las subunidades RII y prácticamente no interactúan con las subunidades RI o lo hacen con muy baja afinidad, por lo que se conocen como AKAPs específicas de RII [52]. Por otro lado, existen unas pocas AKAPs que apenas han empezado a ser identificadas en los últimos años, que son específicas de RI [53, 54]. Finalmente, existen AKAPs de especificidad dual (por ejemplo D-

AKAP1, D-AKAP2 y ezrina), que son capaces de interactuar tanto con RI como con RII. En general, la afinidad de los dominios AKB de estas AKAPs duales es por lo menos un orden de magnitud mayor para RII que para RI [52, 55]; sin embargo, recientemente se ha reportado una AKAP dual, Opa1, que presenta mayor afinidad por RI que por RII [56].

Las AKAPs de especificidad dual contienen una secuencia adicional fuera de la α -hélice anfipática, denominada región específica para RI (RISR), que fue originalmente identificada en la ezrina. Esta región de unos 30 aminoácidos tiene la capacidad de unirse a RI por sí misma, y contiene dos *clusters* de residuos básicos que son críticos para esa interacción con RI. Por otro lado, por experimentos de dicroísmo circular se determinó que RISR tiene estructura de α -hélice en solución, al igual que los dominios AKB. El sitio de interacción con RISR se encuentra hacia el C-terminal del dominio D/D. [57].

La elucidación del rol que desempeñan las AKAPs en distintos procesos fisiológicos se lleva a cabo frecuentemente utilizando péptidos sintéticos que emulan el dominio AKB, y que tienen la capacidad de interrumpir la interacción PKA-AKAP tanto *in vitro* como *in vivo*. El primero de estos péptidos que se desarrolló fue Ht31 (*human thyroid 31*), que corresponde al dominio AKB de la AKAP-Lbc [58] (ver **Figura 1.14**). Posteriormente, utilizando *arrays* de péptidos que permitieron explorar numerosos cambios de residuos sobre este templatado, se desarrollaron otros péptidos que presentan selectividad por una u otra isoforma de la subunidad R: el péptido AKAP-IS (AKAP-*in silico*), el cual es selectivo para RII por sobre RI [59], y el péptido disruptor del anclaje a RI (RIAD) selectivo para RI [60].

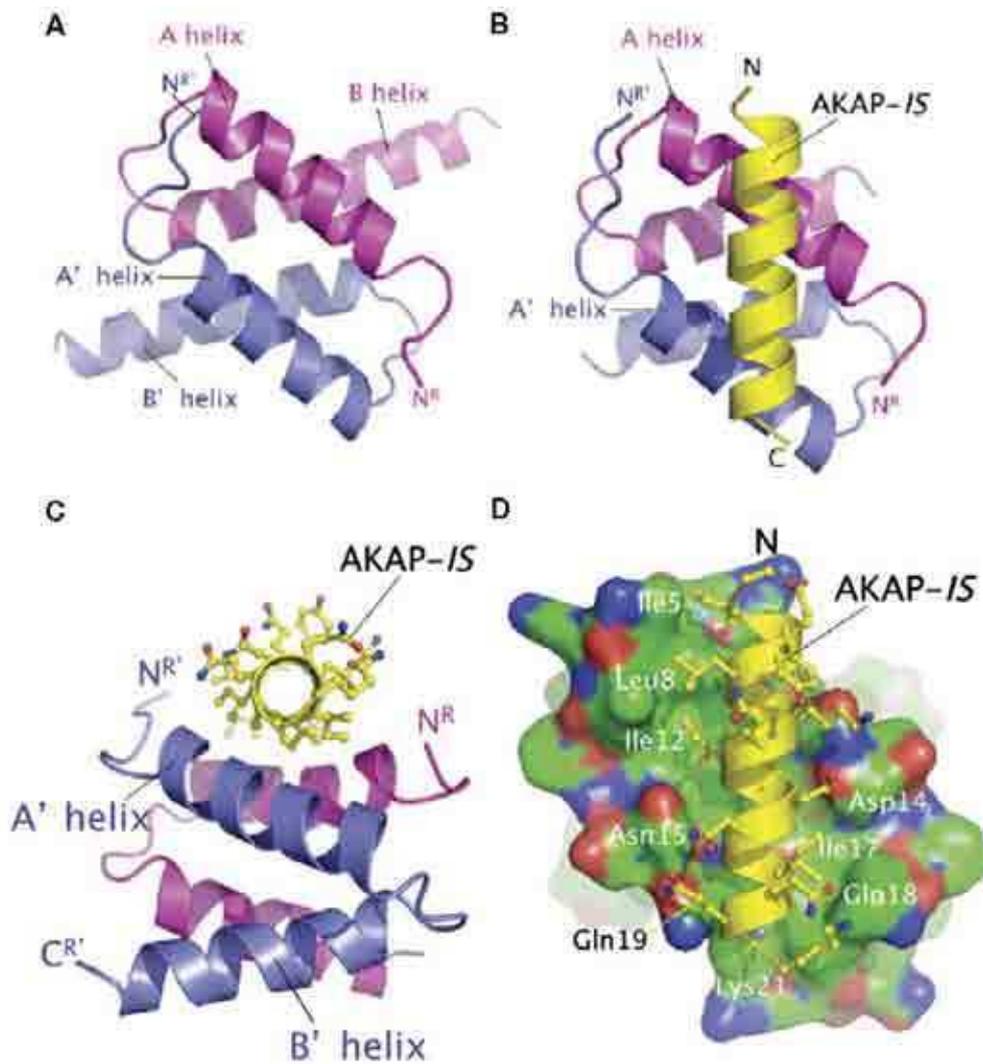


Figura 1.15: Estructura del complejo formado por el dominio D/D de RII α y el péptido AKAP-IS.

(A) Vista de la superficie de interacción con las AKAPs del dominio D/D de RII α . Las dos cadenas (R y R') se representan en azul y en magenta. Se indican también sus extremos N-terminales, y la denominación de las α -hélices: *A helix*, *A' helix* (hélices 1) y *B helix*, *B' helix* (hélice 2). El modelo fue obtenido por DRX (pdb: 2IZY [45]). (B) Misma vista que en el panel (A), pero para el complejo D/D-AKAP-IS; el péptido se muestra en amarillo con sus extremos N y C terminales identificados. (C) Vista lateral, rotada 90° en sentido antihorario respecto del eje contenido en el plano de la hoja que es a su vez perpendicular al péptido en el panel (B). Además del diagrama de cintas, se muestran las cadenas laterales de los aminoácidos, con el siguiente código de colores: carbono en amarillo, nitrógeno en azul, oxígeno en rojo. Se aprecia que todos los residuos del péptido que quedan sumergidos en la interfaz son alifáticos. (D) Misma vista que en el panel (B), pero el dominio D/D se muestra en forma de superficie con el siguiente código de colores: carbono en verde, nitrógeno en azul, oxígeno en rojo. Puede apreciarse que el surco en el que encajan los residuos de la AKAP es preponderantemente hidrofóbico. (Figura tomada de [45]).

En el año 2001 se reportaron las primeras estructuras de alta resolución correspondientes a complejos entre un dominio D/D y un dominio AKB, correspondientes al dominio D/D de RII α y los péptidos Ht31 y el AKB de la AKAP 79 [48]. Desde entonces, se han reportado numerosas estructuras tanto de RI α como de RII α , en complejo con distintos péptidos de AKAPs. Dos trabajos muy importantes son los que reportaron las estructuras de los dominios D/D de las dos isoformas en complejo con el mismo péptido derivado de una AKAP dual, la D-AKAP2; esto permitió elucidar las diferencias en la especificidad de la interacción [47, 49].

Otro trabajo muy importante consistió en la elucidación de la estructura de alta resolución del dominio D/D de la subunidad RII α en complejo con el péptido sintético AKAP-IS, selectivo para RII [45] (**Figura 1.15**). Como ya se señaló, las hélices 1 y 1' (o A y A') correspondientes a cada monómero del dominio D/D (R y R') se encuentran alineadas de forma antiparalela, generando de esta manera el surco hidrofóbico en el que encajan las cadenas laterales del péptido, que complementan la superficie expuesta por RII α ; esta superficie es muy rígida y esencialmente no sufre cambios como consecuencia de la unión a las AKAPs. La estructura permitió determinar que la interacción entre los dos componentes está dominada por residuos hidrofóbicos, dado que los aminoácidos alifáticos que constituyen la cara no polar de la α -hélice anfipática del dominio AKB se encuentran sumergidos en el surco hidrofóbico del dominio D/D. El segundo elemento importante lo constituyen los residuos Ile 3 e Ile 5 de RII α . Estos dos residuos forman parte del segmento extendido que precede a la hélice 1, que en ausencia de AKAPs se encuentra desestructurado. Trabajos previos habían demostrado que al mutar estos dos residuos por alanina la interacción se veía seriamente comprometida [61, 62]. La estructura reveló que en el complejo D/D-AKB, el segmento desestructurado se acomoda para facilitar interacciones importantes entre estas dos cadenas laterales y residuos hidrofóbicos a los lados del péptido.

Interacción entre Bcy1 e Ira2

Como ya se discutió no se han reportado hasta la fecha AKAPs en *S. cerevisiae*. Sin embargo, la proteína Ira2, que forma parte de la vía de señalización de cAMP-PKA en este organismo, presenta ciertas características que sugieren que podría ser clasificada como miembro de esa familia. La interacción entre Ira2 y Bcy1 depende de los primeros 85 aminoácidos de la segunda, lo que sugiere que el dominio D/D podría encontrarse

involucrado en la misma. Reforzando esa idea, un péptido de Ira2 de 25 aminoácidos que resulta suficiente para garantizar la interacción con Bcy1, se pliega de acuerdo a predicciones bioinformáticas en forma de una α -hélice anfipática, al igual que los dominios AKB de las AKAPs de mamíferos. [36].

Por otro lado, en el mismo trabajo se elucidaron cuáles son los residuos del péptido que resultan clave en mediar la interacción de Bcy1 mediante experimentos de *ala scanning*. Inesperadamente, los resultados mostraron que la mutación de los residuos hidrofóbicos que constituyen la cara no polar de la α -hélice anfipática no afectó la interacción; por el contrario, la mutación de tres residuos básicos, (dos lisinas y una arginina) ubicados hacia el extremo C-terminal del péptido comprometió seriamente la interacción. Se observó un comportamiento similar para los péptidos derivados de Hsp60 y Myo2, otras dos proteínas interactoras de Bcy1 reportadas en ese trabajo. Como control, se utilizó el dominio AKB de la AKAP7 de mamíferos, que no fue capaz de interactuar con Bcy1. [36].

Estos resultados se encuentran en fuerte contraste con el comportamiento del sistema D/D-AKB en mamíferos, dado que en este caso la interacción se suprime cuando se mutan los residuos de la cara hidrofóbica de la hélice anfipática de las AKAPs [51]; por lo tanto, los dominios de interacción de las proteínas interactoras de Bcy1 presentan ciertas características diferenciales con las AKAPs de mamíferos.

Finalmente, en este trabajo se propuso que los residuos de serina fosforilados localizados en el extremo N-terminal de Bcy1 podrían mediar interacciones electrostáticas con los residuos positivamente cargados de esos péptidos. En esa línea, se observó que la variante de Bcy1 que tiene mutados ambos *clusters* de serina por alanina, la asociación de Bcy1 a membranas y vesículas se ve seriamente comprometida; este comportamiento no se observa para las variantes que tiene mutado solamente uno de los dos *clusters* [36]. Esta observación sugiere que podrían existir varios puntos de contacto entre Bcy1 y sus proteínas interactoras.

Dominios de unión a cAMP

El cAMP es una molécula que se encuentra ampliamente distribuida como segundo mensajero en numerosos organismos, y por lo tanto la estructura de los dominios de unión a cAMP (y los dominios de unión a nucleótidos cíclicos en general) está muy conservada de bacterias a humanos; de hecho, esta es la región de la subunidad R más similar en cuanto a su secuencia entre las distintas isoformas y entre los distintos organismos [12].

Se han resuelto hasta la fecha estructuras de alta resolución de mutantes de delección de la subunidad R que contienen ambos dominios de unión a cAMP para R1 α (Δ 1-91) [63], R11 β (Δ 1-107) [64], y Bcy1 (Δ 1-167) (esta última forma parte del trabajo realizado en nuestro grupo de investigación) [65]; cabe recordar que el desorden intrínseco del *linker* atenta contra la cristalización y que por lo tanto hasta muy recientemente no se habían podido resolver estructuras de la subunidad R entera. La arquitectura general de cada uno de los dos dominios de unión a cAMP individuales se encuentra muy bien conservados entre las distintas proteínas, con un RMSD (raíz cuadrada de la desviación cuadrática media) para átomos de carbono α equivalentes de 0,9 Å para el CBD-A y 1,0 Å para el CBD-B [65]. Todas ellas constan de dos α -hélices en su extremo N-terminal: α N y α A, seguidas de 8 hebras β y dos α -hélices C-terminales: α B y α C. Estos elementos de estructura secundaria forman dos subdominios: un subdominio α formado por las α -hélices, y un subdominio β formado por las 8 hebras que se pliegan en forma de barril β (**Figura 1.16**).

El subdominio helicoidal es flexible, y contiene en su extremo N-terminal un motivo llamado N3A, que está formado por las hélices α N y α A, separadas por un *loop* de 8 residuos, que tiene estructura de *turn* 3_{10} y que por lo tanto se denomina 3_{10} *loop*. Por el contrario, el subdominio β es relativamente rígido, y contiene un motivo muy conservado llamado casete de unión a fosfato (PBC), al cual se une el cAMP. El PBC es un motivo de catorce aminoácidos que está presente en todos los dominios de unión a nucleótidos cíclicos. Está formado por un *loop* y una α -hélice corta, de tres giros. Tiene tres componentes clave de interacción con la molécula de cAMP, muy conservados (**Figura 1.17**): una arginina que se une al fosfato, un ácido glutámico que se une al oxidrilo de la ribosa, y un residuo tapa o *capping* que interactúa con el anillo de la adenina. [66].

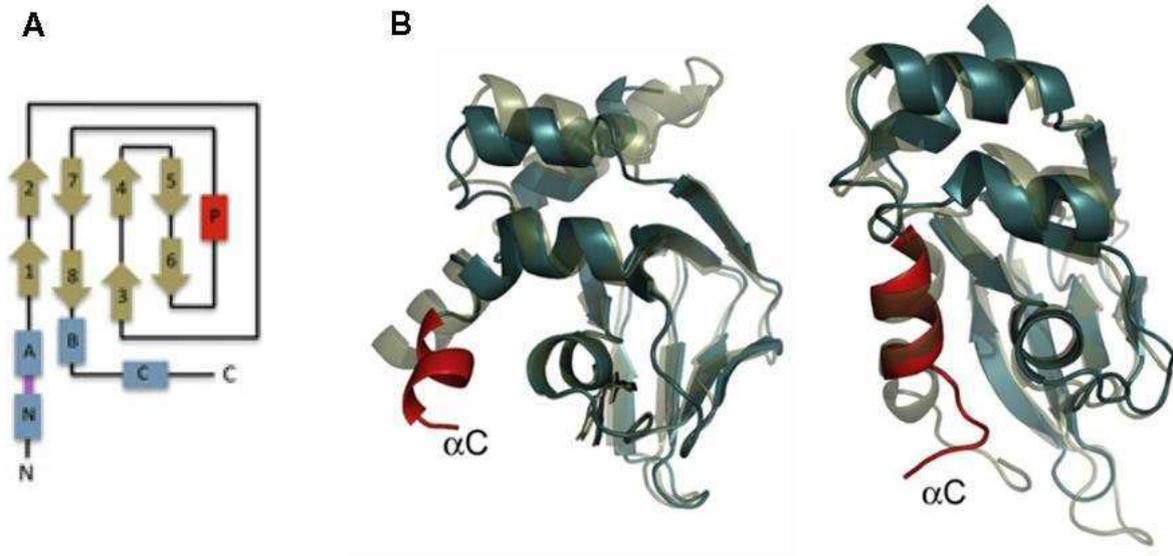


Figura 1.16: Estructura de los dominios de unión a cAMP de la subunidad reguladora de la PKA.

(A) Distribución de elementos de estructura secundaria. Las α -hélices se muestran en azul y las hebras β en amarillo. El segmento rojo representa el casete de unión a fosfato (PBC). (B) Superposición de los CBD de Bcy1 (verde azulado) con los CBD de R1 α (marrón grisáceo, con transparencia). El panel de la izquierda muestra el CBD-A, y el de la derecha, el CBD-B. Se resaltan en rojo las hélices α C de Bcy1. (Figura tomada de [65]).

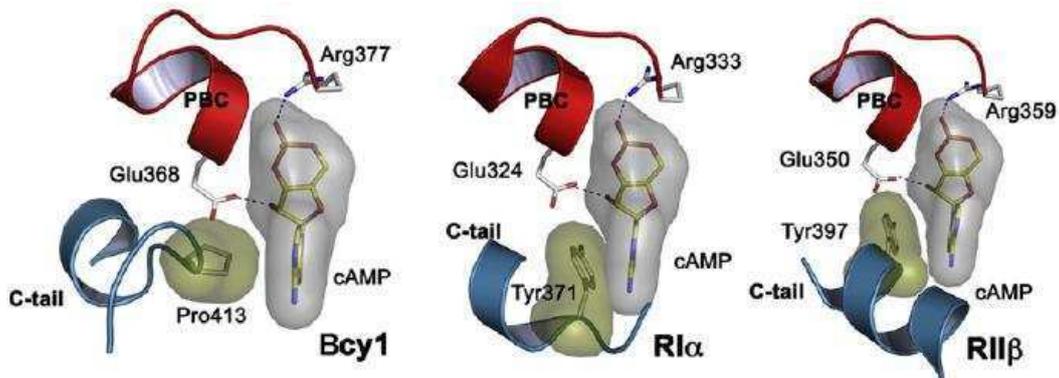


Figura 1.17: Estructura de los casete de unión a fosfato del CBD-B de la subunidad R de la PKA.

Se muestra el PBC en rojo, y los residuos conservados: la arginina que coordina el fosfato, y el ácido glutámico que coordina el oxidrilo de la ribosa. A diferencia de los mamíferos que tienen una tirosina, el residuo *capping* de Bcy1 es una prolina. (Figura tomada de [65]).

A pesar de las numerosas similitudes, existen dos diferencias esenciales entre la estructura de los CBD de Bcy1 y la de las proteínas de mamíferos. La primera es justamente el residuo *capping*; en el CBD-B, en las dos subunidades R de mamíferos es una tirosina, mientras que en Bcy1 es una prolina (**Figura 1.17**). La segunda diferencia notable entre la estructura de los CBD de Bcy1 y las subunidades R de mamíferos es la orientación relativa de los dos dominios (**Figura 1.18**).

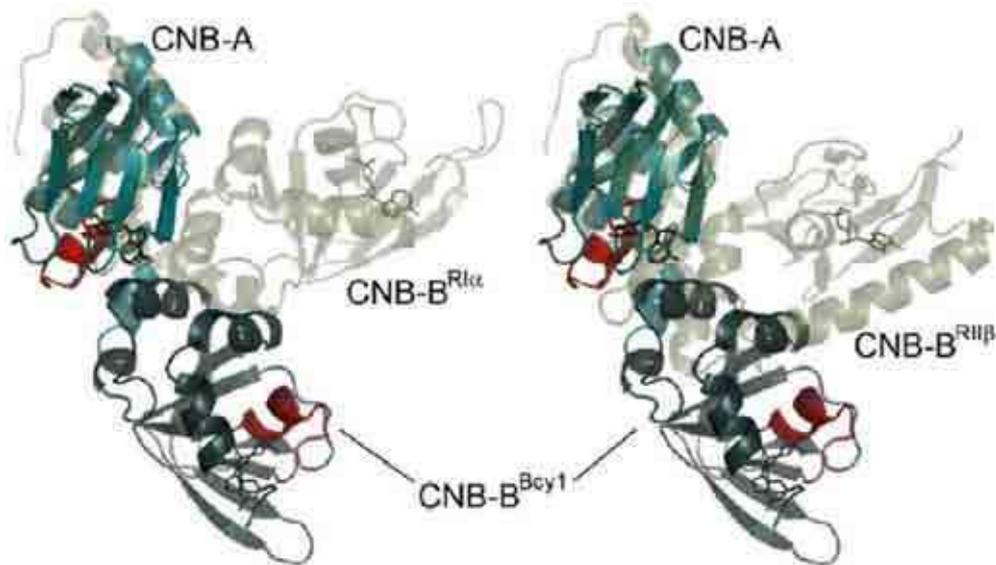


Figura 1.18: Orientación relativa de los CBD de Bcy1 y de las subunidades R de mamíferos.

Se muestran en verde azulado los dos dominios de unión a cAMP de Bcy1, y en marrón grisáceo con transparencia los de las subunidades R de mamíferos: RI α a la derecha y RII β a la izquierda. Se muestran en rojo los PBC de Bcy1. (Figura tomada de [65]).

Zona bisagra (*Linker*)

Como ya se señaló, entre el dominio D/D en el extremo N-terminal y los dos CBDs ubicados hacia el extremo C-terminal, existe una región intrínsecamente desordenada denominada *linker* o bisagra, que incluye el sitio inhibitorio (IS) de la subunidad C. Este motivo de tan solo cuatro aminoácidos (RRXS/A) emula el sitio consenso de fosforilación de los sustratos de PKA, y en la holoenzima inactiva ocupa el sitio catalítico de la misma, contribuyendo a su inhibición. Las subunidades RI tienen una alanina en la cuarta posición de ese sitio, mientras que las subunidades RII tienen una serina, que corresponde al sitio de

autofosforilación [67]; por lo tanto, las subunidades RII son fosforiladas por la subunidad C, mientras que las subunidades RI no, lo que constituye una de las principales diferencias entre las dos isoformas de R. En Bcy1 la cuarta posición la ocupa un residuo de serina fosforilable, lo que la asemeja más a las subunidades RII.

Como ya se señaló, el *linker* es una región intrínsecamente desordenada, que presenta gran flexibilidad. Su secuencia es altamente repetitiva y muy poco conservada entre las distintas isoformas, y se han predicho sitios de fosforilación y de interacción proteína-proteína en esta región. En los últimos años, se ha tomado dimensión acerca de su relevancia en cuanto a la estructura del dímero de subunidades R y a la estructura global de la holoenzima de la PKA, como se discutirá en las siguientes secciones. [40].

Estructura global del dímero de subunidades R

Como ya se señaló, durante mucho tiempo no fue posible obtener una estructura de alta resolución de la subunidad R completa, por lo que los primeros modelos se basaron en las estructuras de los dominios individuales (**Figura 1.19**). Posteriormente, utilizando la técnica de dispersión de rayos X a bajo ángulo (SAXS), que permite obtener información estructural de baja resolución de macromoléculas en solución (tamaño y forma global), se demostró que las diferentes isoformas de la subunidad R presentan notables diferencias conformacionales en cuanto a la disposición relativa de sus dominios [40]. Concretamente, mientras que el dímero de RI α es compacto y en forma de “Y”, los homodímeros de RII α y RII β son extendidos y de forma alargada. Estas diferencias se atribuyeron a la región del *linker*, dado que la estructura de los CBDs y del dominio D/D se encuentran bastante conservadas entre las distintas isoformas. Apoyando esta idea, utilizando anisotropía de fluorescencia seguida en el tiempo de fluoróforos unidos a residuos de cisteína del *linker*, o bien residuos de triptofano de esa región, se demostró que los *linker* de las subunidades RII son mucho más flexibles en solución que los de las subunidades RI, que parecen presentar un nivel superior de ordenamiento [40]; este concepto también fue demostrado por espectrometría de masa de intercambio H-D, que permite identificar péptidos de la molécula que se encuentran en regiones flexibles y/o accesibles al solvente [40]. Más aun, el análisis de las secuencias de los *linkers* de las distintas isoformas de la subunidad R de mamíferos reveló que, pese a que ambos se predicen desordenados, en la subunidad RI α no se predicen regiones de baja complejidad, mientras que en las subunidades RII se predicen dos de estas regiones. La

diferencia es que las regiones de alta complejidad tienden a tener más estructura residual y más probablemente se encuentran involucradas en interacciones proteína-proteína.

Se ha propuesto que la mayor flexibilidad de los dímeros de RII en contraste con la rigidez de RI puede proporcionar un nivel adicional de regulación de la especificidad en el ensamblado de grandes complejos macromoleculares dirigidos por las AKAPs, dado que las subunidades RII tendrían mayor posibilidad de acomodarse y buscar potenciales sustratos debido a su flexibilidad [12].

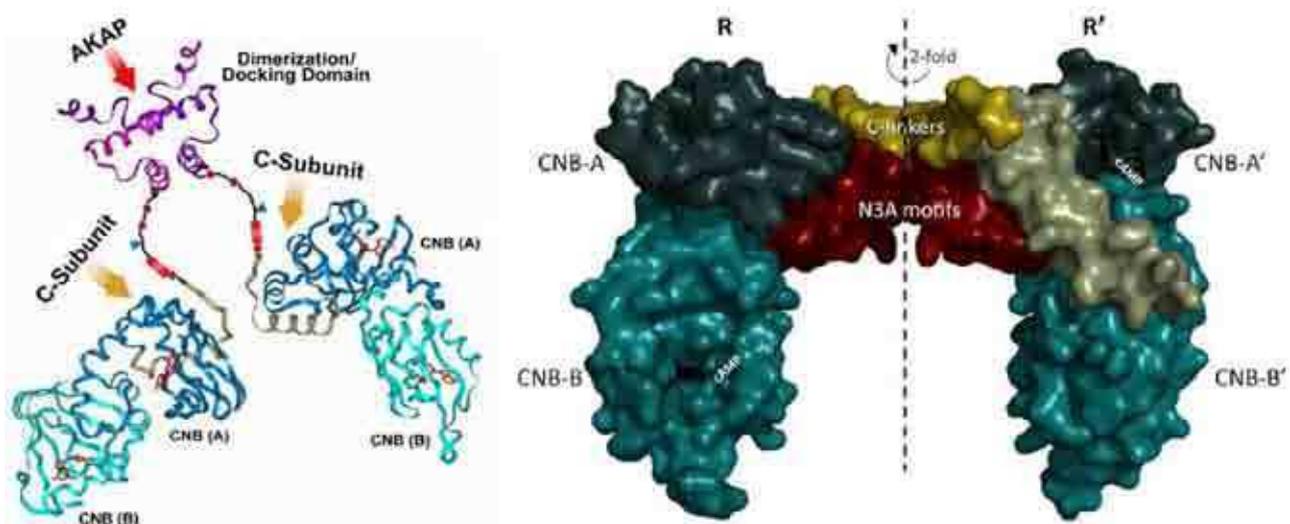


Figura 1.19: Estructura del dímero de subunidades R1α.

En el panel de la izquierda se muestra un modelo construido en base a las estructuras de alta resolución de los dominios D/D [46] y CBD [63]. El primero se muestra en magenta (*Dimerization/Docking Domain*) y los segundos en azul y turquesa (CNB (A) y CNB (B)). Se indican también las regiones de interacción con las AKAPs y con las subunidades C. En el panel de la derecha se muestra la estructura cristalina del dímero. La región mostrada corresponde a los residuos 105-376 para la cadena R y 108-376 para la cadena R', lo que incluye 8 y 5 residuos del *linker*, respectivamente. Si bien se cristalizó la proteína entera, el resto de la molécula no mostró densidad electrónica visible en el mapa debido al desorden intrínseco del *linker*. Se muestran en verde oscuro y claro los CBDs B A y B (CNB-A y CNB.B, respectivamente), en rojo los motivos N3A, y en amarillo la región del *linker* que se pudo apreciar en los mapas de densidad electrónica. En marrón se muestra la hélice $\alpha_{B/C}$. También se muestra el eje de simetría 180°. (La figura de la izquierda fue tomada de [4], y la de la derecha de [41]).

Recientemente, se ha logrado resolver la estructura cristalina de un dímero de RI α [41]. Si bien en esta estructura el dominio D/D y la mayor parte del *linker* se encuentran ausentes, se observó que los últimos residuos del extremo C-terminal del *linker* participan en el establecimiento de contactos entre las dos subunidades R que conforman el dímero (**Figura 1.19**). Adicionalmente, el motivo N3A presente en el CBD-A de las dos subunidades también contribuye a la generación de una superficie de interacción entre los dos monómeros. En este trabajo se propone que la utilización de estos dos elementos en el establecimiento de contactos entre monómeros es exclusivo de las subunidades RI, dado que no se observa en las subunidades RII, y que por lo tanto permite explicar las diferencias conformacionales entre los dos dímeros discutidas anteriormente [41].

Hasta este punto se han descrito en detalle las estructuras de las subunidades R y C por separado. Sin embargo, para comprender cómo ocurre a nivel molecular la activación de la PKA y la regulación de su actividad, fue necesario elucidar las estructuras de las holoenzimas. Este punto se describe en la siguiente sección.

Estructura de la holoenzima y aspectos estructurales de la activación de la PKA

Como ya se discutió, el mecanismo clásico de activación de la PKA postula que, como consecuencia de la unión del cAMP a la holoenzima tetramérica, la misma se disocia liberando las subunidades C catalíticamente activas. Actualmente se acepta que el mecanismo de activación es sustancialmente más complejo, en base a numerosos trabajos que han demostrado que el cAMP provoca solamente la relajación de ciertas interacciones R-C, pero no la disociación total de la holoenzima. Por ejemplo, resultados de nuestro grupo de investigación han demostrado que la PKA del hongo dimórfico *Mucor rouxii* forma un complejo ternario R-C-cAMP, en el que solamente está ocupado el CBD-B; sólo por adición de NaCl 0,5 M o de un sustrato se promueve la ocupación del CBD-A con la consecuente disociación de la holoenzima [68-70]. Por otro lado, la activación por cAMP de la PKA de *S. cerevisiae* resulta diferente dependiendo del sustrato, dado que éstos sensibilizan a la holoenzima de manera diferencial al efecto del segundo mensajero [21]. Este fenómeno se ha observado también en las holoenzimas de mamíferos [71]: en este sistema, el cAMP no sólo no logra disociar completamente a la holoenzima requiriendo de la presencia del sustrato, sino que también este último tiene un efecto diferencial sobre la activación de las distintas

holoenzimas; por ejemplo, en la holoenzima conformada por las subunidades RII β , un porcentaje significativo permanece sin disociar aun en presencia de cAMP y de sustrato [72].

En base a estas observaciones, se ha propuesto un modelo según el cual el sustrato desempeña un papel importante en la activación, siendo requerida su presencia para competir con el IS de la subunidad R por el sitio activo de la subunidad C, y generar de esta manera la disociación total de la holoenzima [72]. El sistema presenta en realidad un nivel adicional de complejidad, dado que se ha reportado que el cAMP puede inducir la actividad de la PKA sin requerir la disociación de la holoenzima [73].

Las holoenzimas de PKA se clasifican de acuerdo a la isoforma de la subunidad R que las conforman. La holoenzima contiene dos subunidades C unidas a homo o heterodímeros de RI (RI α /RI α , RI α /RI β , RI β /RI β) o a homodímeros de RII (RII α /RII α , RII β /RII β) [74]. Las cuatro isoformas de la subunidad R son funcionalmente no redundantes y, como ya se señaló, pese a presentar la misma organización de dominios existen numerosas diferencias entre ellas, que se manifiestan como diferencias entre las correspondientes holoenzimas. Estas diferencias se enumeran a continuación:

1) Como ya se indicó, las distintas isoformas de la subunidad R presentan un patrón de expresión diferencial en los distintos tipos celulares: mientras que RI α y RII α se expresan constitutivamente en todos los tejidos, RII β es la isoforma predominante en tejido adiposo, tejido endócrino, órganos reproductores y en el cerebro, y RI β se expresa principalmente en neuronas [1, 12]. Además, las cuatro subunidades son funcionalmente no redundantes, dado que sus respectivas delecciones presentan fenotipos característicos asociados; esto es válido incluso entre las subunidades α y β de un mismo tipo, pese a que presentan una elevada identidad de secuencia [12].

2) En lo que respecta a la localización subcelular, RI α es predominantemente citoplasmática mientras que RI β se encuentra principalmente asociada a mitocondrias. Por otro lado, alrededor del 75% de la PKA de tipo II se encuentra asociada a organelas y a estructuras celulares específicas mediante AKAPs; este hecho posiblemente está relacionado con la elevada cantidad de AKAPs que son específicas de RII, en contraste con la escasas AKAPs duales y específicas de RI. [1, 12]. La localización subcelular de Bcy1 ya fue discutida en secciones anteriores.

3) La PKA de tipo I es más sensible al cAMP, con una constante de activación aparente *in vitro* (concentración requerida para disociar el 50% de la holoenzima) de 50-100 nM de cAMP, mientras que para la PKA de tipo II esta constante es de 200-400 nM [1]. La sensibilidad de la PKA de *S. cerevisiae* al cAMP *in vitro* es similar a la de la PKA de mamíferos de tipo I [22].

4) Desde un punto de vista estructural, existen diferencias en cuanto a la estructura del dominio D/D, como ya se discutió, y en cuanto a la secuencia del IS: mientras que las subunidades RI presentan un residuo de glicina o alanina no fosforilable y son por lo tanto pseudosustratos de la subunidad C, las subunidades RII presentan una serina y por lo tanto son fosforiladas por la subunidad C. Esto resulta relevante dado que la fosforilación del IS disminuye la afinidad de C por R, y resulta relevante dado que los ciclos de disociación y reasociación de la holoenzima de tipo II requiere de la defosforilación del IS [12].

5) Finalmente, los dos tipos de subunidad R tienen distintos requerimientos de cofactores para la formación de las holoenzimas. Las dos subunidades RI requieren de la unión de ATP y de los dos iones Mg^{2+} que contribuyen a la catálisis a las subunidades C para formar la holoenzima [75]; sin embargo, la formación de las holoenzimas de tipo II no requiere de ATP. Consecuentemente, en las distintas holoenzimas las subunidades C se encuentran en diferentes conformaciones: cerrada en RI, ya que están presentes todos los elementos para la catálisis (ATP, Mg^{2+} , y el pseudosustrato representado por el sitio IS de la subunidad R) y abierta en RII α , con su bolsillo de unión a ATP vacío y su sitio catalítico ocupado por el IS fosforilado de la subunidad R [12].

Con el objetivo de entender a nivel molecular cómo la subunidad C es inhibida por la subunidad R, cómo el cAMP activa a la holoenzima, y cuáles son las diferencias estructurales que explican la regulación diferencial de las holoenzimas de tipo I y tipo II (es decir, comprender los mecanismos moleculares que median la especificidad), se estudió la estructura de complejos R-C formados por distintas isoformas de la subunidad R [12].

Las primeras estructuras consistieron en holoenzimas heterodiméricas RC que contienen variantes truncadas de la subunidad R sin el dominio D/D, que por lo tanto son monoméricas, y sin el CBD-B. La estructura del complejo de la mutante RI α (91-244) con la subunidad C α demostró que existe una extensa superficie de contacto entre ambas subunidades [76]. En particular, se observó por primera vez cómo el IS de la subunidad R ocupa el sitio activo de la

subunidad C en la holoenzima, bloqueando el acceso a los sustratos. Por otro lado, la región del *linker* que conecta al IS con el CBD-A (*linker II* o *C-linker*) se ordena como consecuencia de la interacción con el lóbulo mayor de la subunidad C. Mediante el estudio de las distintas isoformas de la subunidad R se observó que el IS y el *C-linker* se ordenan de la misma manera en todas ellas, mientras que el *N-linker* (que es la región que precede al IS hacia el extremo N-terminal) se posiciona de manera diferencial, desempeñando un papel fundamental en establecer las diferencias de estructura cuaternaria observadas para las diferentes holoenzimas, como se discutirá más adelante. [12].

La activación de la holoenzima de PKA por cAMP es un proceso alostérico que ocurre secuencialmente, dado que en la holoenzima inactiva solo el CBD-B (el dominio de unión a cAMP más ubicado hacia el C-terminal) se encuentra expuesto y disponible para la unión del segundo mensajero, mientras que el CBD-A permanece inaccesible. Cuando el CBD-B es ocupado, promueve la unión de cAMP al CBD-A a través de un mecanismo alostérico intramolecular [12]. Por lo tanto, las holoenzimas diméricas que contienen subunidades R que presentan solamente el CBD-A no resultan apropiadas para estudiar el mecanismo de activación alostérico de la PKA.

Para obtener información acerca de los cambios conformacionales que sufre la holoenzima como consecuencia de la activación, se resolvió la estructura cristalina de un complejo heterodimérico RC en el que la subunidad R contiene los dos CBDs, formado por la mutante de RI α (91-379), y la subunidad C α [77]. Esta estructura reveló que la subunidad RI α se caracteriza por una notable plasticidad conformacional, sufriendo cambios muy importantes como consecuencia de la unión al cAMP y la consecuente liberación de la subunidad C. En base a estos resultados, se definieron dos estados conformacionales muy característicos para las subunidades R: uno con elevada afinidad por cAMP y otro con elevada afinidad por la subunidad C. En el primero existe amplia superficie alostérica de contacto entre los dos CBDs, mientras que en los heterodímeros RC la mayor interfaz se da entre las subunidades R y C (**Figura 1.20**).

La transición entre estos dos estados se explica por el comportamiento dinámico de hélices α B y α C localizadas hacia el extremo N-terminal del CBD-A (ver **Figura 1.20**): en la holoenzima RC, estas dos α -hélices se encuentran fusionadas formando una única α -hélice extendida, la hélice α B/C, que mantiene separados los dos CBDs. Como consecuencia de la

unión al cAMP, esta hélice se quiebra, acercando los dos CBDs para generar la superficie de interacción entre ambos, y liberando la subunidad C. [12].

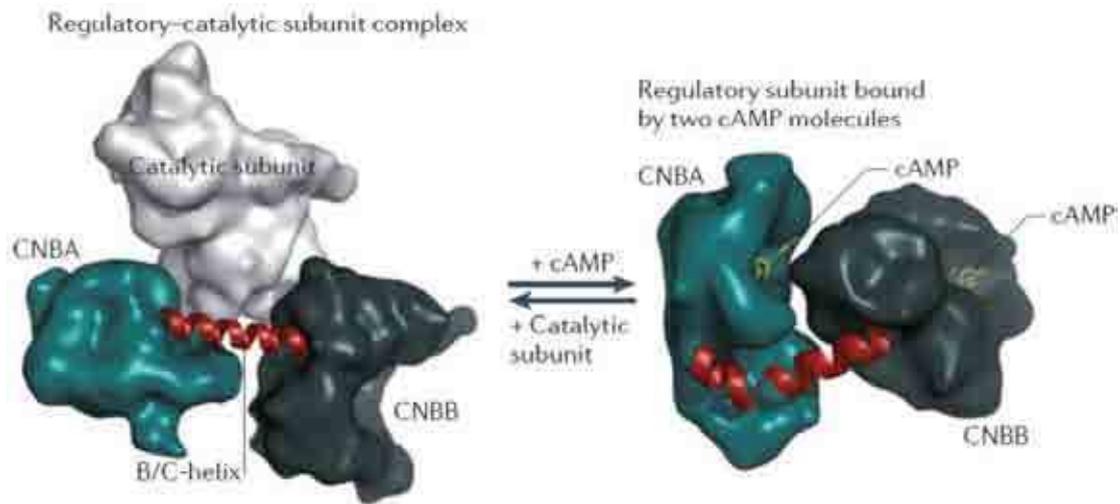


Figura 1.20: Mecanismo de activación de la holoenzima dimérica RC de la PKA.

Se muestran representaciones de las superficies de los dos CNB (aminoácidos 91-379) en tonos diferentes de verde, separados por la hélice $\alpha B/C$ (en rojo). El panel de la izquierda representa la holoenzima dimérica RC inactiva, con la subunidad C en blanco. Como consecuencia de la unión de cAMP, ocurren cambios conformacionales mayores que se traducen en el quiebre de la hélice $\alpha B/C$ en las hélices αB y αC , lo que genera una superficie de interacción entre los dos CBD, que como consecuencia rompe la superficie de interacción RC y libera la subunidad C. (Figura tomada de [12]).

Como se mencionó anteriormente, en Bcy1 la orientación relativa de los dos CBDs en el estado unido a cAMP (que representa la única estructura conocida hasta el momento) es muy diferente a la de todas las subunidades R de mamíferos [65]. En el marco de los mecanismos discutidos en los párrafos anteriores, esta observación sugiere que los cambios conformacionales que median la activación alostérica de la PKA en *S. cerevisiae* podrían ser muy distintos a los conocidos para mamíferos.

Los complejos heterodiméricos RC de las distintas isoformas de R son muy similares. Por lo tanto, si bien estos estudios permitieron comprender el mecanismo de activación de la PKA, no proporcionaron información acerca de la activación diferencial de las distintas holoenzimas, ni permitieron explicar por qué las diferentes isoformas de la subunidad R son funcionalmente no redundantes. Como consecuencia, fue necesario abordar el estudio de las

estructuras de las holoenzimas tetraméricas enteras. Estos estudios revelaron que el *N-linker*, que en contraste con el *C-linker* que se ordena como consecuencia de la interacción con la subunidad C, se encuentra siempre desordenado, desempeña un papel fundamental en el establecimiento de la estructura de la holoenzima; en concordancia con esta observación, el *N-linker* es la región menos conservada en secuencia entre las distintas isoformas [12].

Mediante la combinación de estructuras cristalinas de holoenzimas parcialmente completas [41, 77-80], e información sobre las estructuras en solución obtenida mediante experimentos de dispersión de rayos X (SAXS) y de neutrones (SANS) a bajo ángulo [40, 41, 81], se generaron modelos de las holoenzimas tetraméricas formadas por las distintas isoformas de la subunidad R. Estos estudios demostraron que las holoenzimas son estructuralmente muy diferentes (**Figura 1.21**).

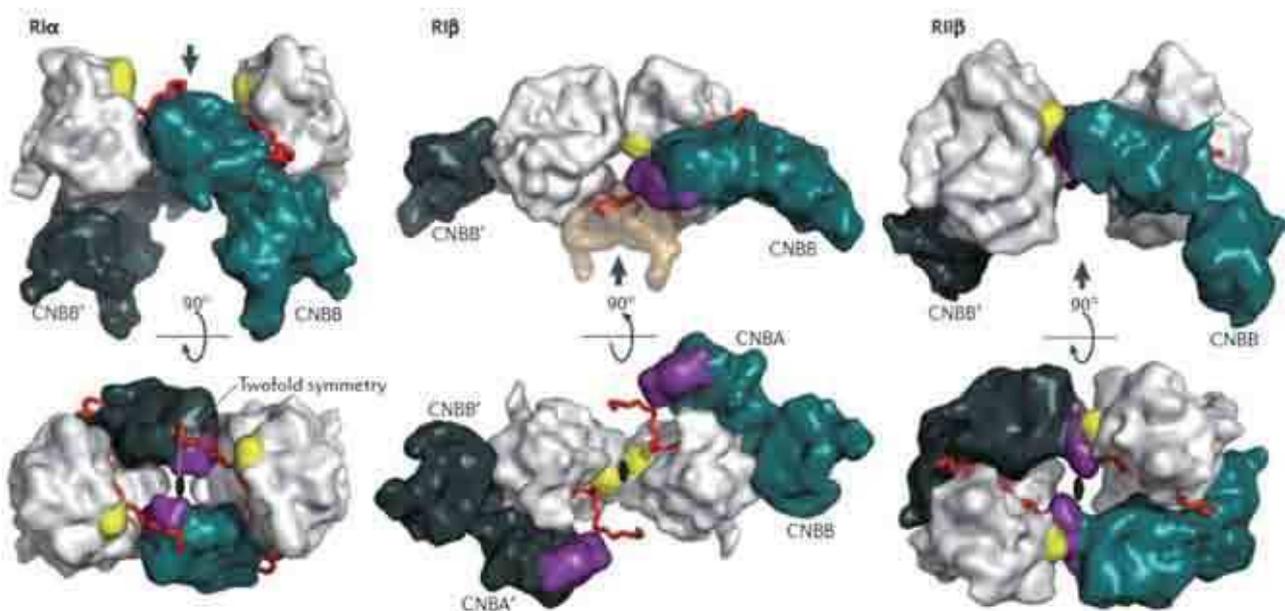


Figura 1.21: Diferencias estructurales entre las distintas holoenzimas de la PKA.

Pese a que los heterodímeros RC son muy similares para todas las isoformas de R, estos se ensamblan en holoenzimas tetraméricas que presentan diferencias muy significativas. Se muestran representaciones del contorno de las estructuras cuaternarias de las holoenzimas RI α , RI β y RII β , enfatizando las diferencias en las arquitecturas globales de los tetrámeros y también en el posicionamiento del dominio D/D, que se indica con una flecha. El *N-linker* se muestra como un segmento rojo. En púrpura se muestra el loop β 4- β 5 del CNB-A. En amarillo se muestra el motivo FDDY (Phe-Asp-Asp-Tyr) ubicado en el extremo C-terminal de la subunidad C, que es una parte integral del sitio de unión a ATP. Estos tres motivos determinan la arquitectura global de la

holoenzima. En los paneles inferiores se muestra como un punto el eje de simetría C_2 , presente en todas las estructuras. (Figura tomada de[12]).

Pese a presentar un 85% de similitud de secuencia y dominios estructurales estrictamente conservados, las holoenzimas de PKA conformadas por las subunidades $RI\alpha$ y $RI\beta$ difieren sustancialmente en sus estructuras cuaternarias, siendo la primera en forma de "Y" (al igual que el dímero de subunidades $RI\alpha$) con una disposición tal que las dos subunidades C se encuentran bien separadas. En contraposición, la segunda es mucho más elongada, y existe una importante superficie de interacción entre los dos lóbulos menores de las subunidades C (ver **Figura 1.21**). Por otro lado, como consecuencia de las diferentes disposiciones del *N-linker* (en contraste con el sitio IS y el *C-linker*, que en todas las R se encuentra ordenado), el dominio D/D asume disposiciones relativas respecto a los otros dominios completamente diferentes: la posición del dominio D/D en la holoenzima $RI\beta$ sugiere que podría formar parte integral del complejo, interactuando con las subunidades C y con los CNB, por lo que la interacción con las AKAPs podría ser detectada por el complejo entero y tener importantes consecuencias sobre su estructura global. [12].

En lo que respecta a las holoenzimas RII, ambas son también muy distintas a pesar de que los dímeros de las subunidades R de las dos son muy similares, adoptando una conformación extendida, como ya se discutió: mientras que la holoenzima $RII\alpha$ es extendida al igual que el dímero de $RII\alpha$, la holoenzima $RII\beta$ es mucho más compacta (ver **Figura 1.21**).

OBJETIVOS GENERALES

Como se señaló en el Prólogo, el objetivo general del trabajo consiste en contribuir al conocimiento de los mecanismos que median la especificidad en la vía de transducción de señales del cAMP-PKA en *S. cerevisiae* desde una perspectiva estructural. En particular, el principal aporte radica en completar el panorama de la estructura de Bcy1, la subunidad R de la PKA de este organismo. El enfoque adoptado es similar al que se utilizó para elucidar la estructura de las subunidades R de mamíferos. Como ya se señaló, este consistió en estudiar los distintos dominios por separado dado que, en una primera aproximación, la proteína no puede caracterizarse completa (al menos, no con alta resolución) debido a la desestructuración intrínseca del *linker*.

El antecedente más directo es la determinación de la estructura cristalina de la región correspondiente a los dominios de unión a cAMP de Bcy1, que fue llevada a cabo por nuestro grupo de investigación [65]. Por lo tanto, para completar la caracterización estructural de la proteína, en este trabajo se abordó el estudio de la estructura del dominio D/D.

La importancia de este dominio en la especificidad radica esencialmente en dos aspectos, tanto en mamíferos como en *S. cerevisiae*: la determinación de la localización subcelular y la determinación del estado oligomérico de la subunidad R y, consecuentemente, de la holoenzima. A su vez, el estado oligomérico tiene implicancias sobre las características alostéricas del mecanismo de activación de la PKA.

El trabajo está organizado en 4 capítulos de resultados. En cada uno de ellos se plantean en primer lugar los antecedentes más directos en lo que respecta a los temas abordados en ese capítulo. Posteriormente, se enumeran los objetivos específicos y los abordajes teóricos y experimentales utilizados para alcanzarlos. A continuación, aquellos capítulos que involucran técnicas que no son de rutina y que por lo tanto no necesariamente resultan familiares para todos los científicos, incluyen una Introducción metodológica que explica los fundamentos teóricos y prácticos mínimos que se requieren para interpretar los resultados. El cuerpo principal de cada capítulo está formado por una sección integrada de resultados y discusión, mientras que en las conclusiones solamente se enumeran los hallazgos más relevantes y se plantean algunas perspectivas futuras. A continuación se encuentra la sección de Materiales y métodos, en la que se detallan las cuestiones experimentales que no son centrales para la comprensión de los resultados. Finalmente, cada capítulo cuenta con su propia lista de referencias.

En el Capítulo 2 se discute la identificación, producción y caracterización preliminar de la región de Bcy1 correspondiente al dominio D/D. En el Capítulo 3 se analiza su estructura en solución mediante distintos abordajes experimentales, mientras que en el Capítulo 4 se discute su estructura cristalina. Por último, el Capítulo 5 presenta algunos resultados preliminares del análisis de la interacción entre Bcy1 y el péptido de Ira2 que interactúa con ella.

Finalmente, el trabajo concluye con una sección de Consideraciones finales, en la que se discuten de manera general los hallazgos más relevantes, se proponen algunas posibles implicancias de esos resultados, y se plantean perspectivas de futuros experimentos y temas de interés que surgen a partir de puntos inconclusos de este trabajo.

REFERENCIAS

1. Pidoux G, Tasken K: **Specificity and spatial dynamics of protein kinase A signaling organized by A-kinase-anchoring proteins.** *Journal of molecular endocrinology* 2010, **44**(5):271-284.
2. Johnson DA, Akamine P, Radzio-Andzelm E, Madhusudan M, Taylor SS: **Dynamics of cAMP-dependent protein kinase.** *Chemical reviews* 2001, **101**(8):2243-2270.
3. Taylor SS, Keshwani MM, Steichen JM, Kornev AP: **Evolution of the eukaryotic protein kinases as dynamic molecular switches.** *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2012, **367**(1602):2517-2528.
4. Taylor SS, Kim C, Vigil D, Haste NM, Yang J, Wu J, Anand GS: **Dynamics of signaling by PKA.** *Biochimica et biophysica acta* 2005, **1754**(1-2):25-37.
5. Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E, Anand G: **PKA: a portrait of protein kinase dynamics.** *Biochimica et biophysica acta* 2004, **1697**(1-2):259-269.
6. Wong W, Scott JD: **AKAP signalling complexes: focal points in space and time.** *Nat Rev Mol Cell Biol* 2004, **5**(12):959-970.
7. Walsh DA, Perkins JP, Krebs EG: **An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle.** *J Biol Chem* 1968, **243**(13):3763-3765.
8. Zubay G, Schwartz D, Beckwith J: **Mechanism of activation of catabolite-sensitive genes: a positive control system.** *Proceedings of the National Academy of Sciences of the United States of America* 1970, **66**(1):104-110.
9. Klein PS, Sun TJ, Saxe CL, 3rd, Kimmel AR, Johnson RL, Devreotes PN: **A chemoattractant receptor controls development in Dictyostelium discoideum.** *Science (New York, NY)* 1988, **241**(4872):1467-1472.
10. Goulding EH, Ngai J, Kramer RH, Colicos S, Axel R, Siegelbaum SA, Chess A: **Molecular cloning and single-channel properties of the cyclic nucleotide-gated channel from catfish olfactory neurons.** *Neuron* 1992, **8**(1):45-58.
11. de Rooij J, Zwartkruis FJ, Verheijen MH, Cool RH, Nijman SM, Wittinghofer A, Bos JL: **Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP.** *Nature* 1998, **396**(6710):474-477.
12. Taylor SS, Ilouz R, Zhang P, Kornev AP: **Assembly of allosteric macromolecular switches: lessons from PKA.** *Nat Rev Mol Cell Biol* 2012, **13**(10):646-658.
13. Tasken K, Aandahl EM: **Localized effects of cAMP mediated by distinct routes of protein kinase A.** *Physiological reviews* 2004, **84**(1):137-167.
14. Santangelo GM: **Glucose signaling in Saccharomyces cerevisiae.** *Microbiology and molecular biology reviews : MMBR* 2006, **70**(1):253-282.
15. Dechant R, Peter M: **Nutrient signals driving cell growth.** *Current opinion in cell biology* 2008, **20**(6):678-687.
16. Broach JR: **Nutritional control of growth and development in yeast.** *Genetics* 2012, **192**(1):73-105.
17. Conrad M, Schothorst J, Kankipati HN, Van Zeebroeck G, Rubio-Teixeira M, Thevelein JM: **Nutrient sensing and signaling in the yeast Saccharomyces cerevisiae.** *FEMS microbiology reviews* 2014, **38**(2):254-299.
18. Sobko A: **Systems biology of AGC kinases in fungi.** *Science's STKE : signal transduction knowledge environment* 2006, **2006**(352):re9.
19. Toda T, Cameron S, Sass P, Zoller M, Wigler M: **Three different genes in S. cerevisiae encode the catalytic subunits of the cAMP-dependent protein kinase.** *Cell* 1987, **50**(2):277-287.
20. Thevelein JM, de Winde JH: **Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast Saccharomyces cerevisiae.** *Molecular microbiology* 1999, **33**(5):904-918.
21. Galello F, Portela P, Moreno S, Rossi S: **Characterization of substrates that have a differential effect on Saccharomyces cerevisiae protein kinase A holoenzyme activation.** *J Biol Chem* 2010, **285**(39):29770-29779.
22. Hixson CS, Krebs EG: **Characterization of a cyclic AMP-binding protein from bakers' yeast. Identification as a regulatory subunit of cyclic AMP-dependent protein kinase.** *J Biol Chem* 1980, **255**(5):2137-2145.

23. Uno I, Matsumoto K, Ishikawa T: **Characterization of cyclic AMP-requiring yeast mutants altered in the regulatory subunit of protein kinase.** *J Biol Chem* 1982, **257**(23):14110-14115.
24. Thompson-Jaeger S, Francois J, Gaughran JP, Tatchell K: **Deletion of SNF1 affects the nutrient response of yeast and resembles mutations which activate the adenylate cyclase pathway.** *Genetics* 1991, **129**(3):697-706.
25. Beene DL, Scott JD: **A-kinase anchoring proteins take shape.** *Current opinion in cell biology* 2007, **19**(2):192-198.
26. Hadad M, Bresler-Musikant T, Neuman-Silberberg FS: **Drosophila spoonbill encodes a dual-specificity A-kinase anchor protein essential for oogenesis.** *Mechanisms of development* 2011, **128**(7-10):471-482.
27. Angelo R, Rubin CS: **Molecular characterization of an anchor protein (AKAPCE) that binds the RI subunit (RCE) of type I protein kinase A from Caenorhabditis elegans.** *J Biol Chem* 1998, **273**(23):14633-14643.
28. Tudisca V, Recouvreux V, Moreno S, Boy-Marcotte E, Jacquet M, Portela P: **Differential localization to cytoplasm, nucleus or P-bodies of yeast PKA subunits under different growth conditions.** *European journal of cell biology* 2010, **89**(4):339-348.
29. Griffioen G, Anghileri P, Imre E, Baroni MD, Ruis H: **Nutritional control of nucleocytoplasmic localization of cAMP-dependent protein kinase catalytic and regulatory subunits in Saccharomyces cerevisiae.** *J Biol Chem* 2000, **275**(2):1449-1456.
30. Harootunian AT, Adams SR, Wen W, Meinkoth JL, Taylor SS, Tsien RY: **Movement of the free catalytic subunit of cAMP-dependent protein kinase into and out of the nucleus can be explained by diffusion.** *Molecular biology of the cell* 1993, **4**(10):993-1002.
31. Griffioen G, Thevelein JM: **Molecular mechanisms controlling the localisation of protein kinase A.** *Current genetics* 2002, **41**(4):199-207.
32. Griffioen G, Branduardi P, Ballarini A, Anghileri P, Norbeck J, Baroni MD, Ruis H: **Nucleocytoplasmic distribution of budding yeast protein kinase A regulatory subunit Bcy1 requires Zds1 and is regulated by Yak1-dependent phosphorylation of its targeting domain.** *Molecular and cellular biology* 2001, **21**(2):511-523.
33. Garrett S, Broach J: **Loss of Ras activity in Saccharomyces cerevisiae is suppressed by disruptions of a new kinase gene, YAKI, whose product may act downstream of the cAMP-dependent protein kinase.** *Genes & development* 1989, **3**(9):1336-1348.
34. Smith A, Ward MP, Garrett S: **Yeast PKA represses Msn2p/Msn4p-dependent gene expression to regulate growth, stress response and glycogen accumulation.** *The EMBO journal* 1998, **17**(13):3556-3564.
35. Rossio V, Yoshida S: **Spatial regulation of Cdc55-PP2A by Zds1/Zds2 controls mitotic entry and mitotic exit in budding yeast.** *The Journal of cell biology* 2011, **193**(3):445-454.
36. Galello F, Moreno S, Rossi S: **Interacting proteins of protein kinase A regulatory subunit in Saccharomyces cerevisiae.** *J Proteomics* 2014, **109C**:261-275.
37. Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM: **Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase.** *Science (New York, NY)* 1991, **253**(5018):407-414.
38. Knighton DR, Zheng JH, Ten Eyck LF, Xuong NH, Taylor SS, Sowadski JM: **Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase.** *Science (New York, NY)* 1991, **253**(5018):414-420.
39. Mashhoon N, Carmel G, Pflugrath JW, Kuret J: **Structure of the unliganded cAMP-dependent protein kinase catalytic subunit from Saccharomyces cerevisiae.** *Archives of biochemistry and biophysics* 2001, **387**(1):11-19.
40. Vigil D, Blumenthal DK, Heller WT, Brown S, Canaves JM, Taylor SS, Trewthella J: **Conformational differences among solution structures of the type Ialpha, IIalpha and IIbeta protein kinase A regulatory subunit homodimers: role of the linker regions.** *J Mol Biol* 2004, **337**(5):1183-1194.
41. Bruystens JG, Wu J, Fortezzo A, Kornev AP, Blumenthal DK, Taylor SS: **PKA RIalpha homodimer structure reveals an intermolecular interface with implications for cooperative cAMP binding and Carney complex disease.** *Structure (London, England : 1993)* 2014, **22**(1):59-69.

42. Gunzburg J, Veron M: **A cAMP-dependent protein kinase is present in differentiating Dictyostelium discoideum cells.** *The EMBO journal* 1982, **1**(9):1063-1068.
43. Carlson GL, Nelson DL: **The 44-kDa regulatory subunit of the Paramecium cAMP-dependent protein kinase lacks a dimerization domain and may have a unique autophosphorylation site sequence.** *The Journal of eukaryotic microbiology* 1996, **43**(4):347-356.
44. Newlon MG, Roy M, Morikis D, Hausken ZE, Coghlan V, Scott JD, Jennings PA: **The molecular basis for protein kinase A anchoring revealed by solution NMR.** *Nat Struct Biol* 1999, **6**(3):222-227.
45. Gold MG, Lygren B, Dokurno P, Hoshi N, McConnachie G, Tasken K, Carlson CR, Scott JD, Barford D: **Molecular basis of AKAP specificity for PKA regulatory subunits.** *Mol Cell* 2006, **24**(3):383-395.
46. Banky P, Roy M, Newlon MG, Morikis D, Haste NM, Taylor SS, Jennings PA: **Related protein-protein interaction modules present drastically different surface topographies despite a conserved helical platform.** *J Mol Biol* 2003, **330**(5):1117-1129.
47. Sarma GN, Kinderman FS, Kim C, von Daake S, Chen L, Wang BC, Taylor SS: **Structure of D-AKAP2:PKA RI complex: insights into AKAP specificity and selectivity.** *Structure (London, England : 1993)* 2010, **18**(2):155-166.
48. Newlon MG, Roy M, Morikis D, Carr DW, Westphal R, Scott JD, Jennings PA: **A novel mechanism of PKA anchoring revealed by solution structures of anchoring complexes.** *The EMBO journal* 2001, **20**(7):1651-1662.
49. Kinderman FS, Kim C, von Daake S, Ma Y, Pham BQ, Spraggon G, Xuong NH, Jennings PA, Taylor SS: **A dynamic mechanism for AKAP binding to RII isoforms of cAMP-dependent protein kinase.** *Mol Cell* 2006, **24**(3):397-408.
50. Banky P, Newlon MG, Roy M, Garrod S, Taylor SS, Jennings PA: **Isoform-specific differences between the type Ialpha and IIalpha cyclic AMP-dependent protein kinase anchoring domains revealed by solution NMR.** *J Biol Chem* 2000, **275**(45):35146-35152.
51. Carr DW, Stofko-Hahn RE, Fraser ID, Bishop SM, Acott TS, Brennan RG, Scott JD: **Interaction of the regulatory subunit (RII) of cAMP-dependent protein kinase with RII-anchoring proteins occurs through an amphipathic helix binding motif.** *J Biol Chem* 1991, **266**(22):14188-14192.
52. Herberg FW, Maleszka A, Eide T, Vossebein L, Tasken K: **Analysis of A-kinase anchoring protein (AKAP) interaction with protein kinase A (PKA) regulatory subunits: PKA isoform specificity in AKAP binding.** *J Mol Biol* 2000, **298**(2):329-339.
53. Kovanich D, van der Heyden MA, Aye TT, van Veen TA, Heck AJ, Scholten A: **Sphingosine kinase interacting protein is an A-kinase anchoring protein specific for type I cAMP-dependent protein kinase.** *Chembiochem : a European journal of chemical biology* 2010, **11**(7):963-971.
54. Burgers PP, Ma Y, Margarucci L, Mackey M, van der Heyden MA, Ellisman M, Scholten A, Taylor SS, Heck AJ: **A small novel A-kinase anchoring protein (AKAP) that localizes specifically protein kinase A-regulatory subunit I (PKA-RI) to the plasma membrane.** *J Biol Chem* 2012, **287**(52):43789-43797.
55. Burns LL, Canaves JM, Pennypacker JK, Blumenthal DK, Taylor SS: **Isoform specific differences in binding of a dual-specificity A-kinase anchoring protein to type I and type II regulatory subunits of PKA.** *Biochemistry* 2003, **42**(19):5754-5763.
56. Pidoux G, Witczak O, Jarnaess E, Myrvold L, Urlaub H, Stokka AJ, Kuntziger T, Tasken K: **Optic atrophy 1 is an A-kinase anchoring protein on lipid droplets that mediates adrenergic control of lipolysis.** *The EMBO journal* 2011, **30**(21):4371-4386.
57. Jarnaess E, Ruppelt A, Stokka AJ, Lygren B, Scott JD, Tasken K: **Dual specificity A-kinase anchoring proteins (AKAPs) contain an additional binding region that enhances targeting of protein kinase A type I.** *J Biol Chem* 2008, **283**(48):33708-33718.
58. Carr DW, Hausken ZE, Fraser ID, Stofko-Hahn RE, Scott JD: **Association of the type II cAMP-dependent protein kinase with a human thyroid RII-anchoring protein. Cloning and characterization of the RII-binding domain.** *J Biol Chem* 1992, **267**(19):13376-13382.
59. Alto NM, Soderling SH, Hoshi N, Langeberg LK, Fayos R, Jennings PA, Scott JD: **Bioinformatic design of A-kinase anchoring protein-in silico: a potent and selective peptide antagonist of type II protein kinase A anchoring.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(8):4445-4450.

60. Carlson CR, Lygren B, Berge T, Hoshi N, Wong W, Tasken K, Scott JD: **Delineation of type I protein kinase A-selective signaling events using an RI anchoring disruptor.** *J Biol Chem* 2006, **281**(30):21535-21545.
61. Hausken ZE, Coghlan VM, Hastings CA, Reimann EM, Scott JD: **Type II regulatory subunit (RII) of the cAMP-dependent protein kinase interaction with A-kinase anchor proteins requires isoleucines 3 and 5.** *J Biol Chem* 1994, **269**(39):24245-24251.
62. Li Y, Rubin CS: **Mutagenesis of the regulatory subunit (RII beta) of cAMP-dependent protein kinase II beta reveals hydrophobic amino acids that are essential for RII beta dimerization and/or anchoring RII beta to the cytoskeleton.** *J Biol Chem* 1995, **270**(4):1935-1944.
63. Su Y, Dostmann WR, Herberg FW, Durick K, Xuong NH, Ten Eyck L, Taylor SS, Varughese KI: **Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains.** *Science (New York, NY)* 1995, **269**(5225):807-813.
64. Diller TC, Madhusudan, Xuong NH, Taylor SS: **Molecular basis for regulatory subunit diversity in cAMP-dependent protein kinase: crystal structure of the type II beta regulatory subunit.** *Structure (London, England : 1993)* 2001, **9**(1):73-82.
65. Rinaldi J, Wu J, Yang J, Ralston CY, Sankaran B, Moreno S, Taylor SS: **Structure of yeast regulatory subunit: a glimpse into the evolution of PKA signaling.** *Structure (London, England : 1993)* 2010, **18**(11):1471-1482.
66. Berman HM, Ten Eyck LF, Goodsell DS, Haste NM, Kornev A, Taylor SS: **The cAMP binding domain: an ancient signaling module.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(1):45-50.
67. Canaves JM, Taylor SS: **Classification and phylogenetic analysis of the cAMP-dependent protein kinase regulatory subunit family.** *J Mol Evol* 2002, **54**(1):17-29.
68. Pastori RL, Kerner N, Moreno S, Passeron S: **cAMP-dependent protein kinase from *Mucor rouxii*: physical evidence of a ternary complex holoenzyme-cAMP.** *Biochemical and biophysical research communications* 1981, **101**(2):663-671.
69. Moreno S, Pastori R, Passeron S: **Protein kinase from *Mucor rouxii*. Unshielding of new cyclic AMP binding sites upon dissociation of the ternary complex holoenzyme-cyclic AMP.** *Molecular and cellular biochemistry* 1983, **52**(1):13-16.
70. Guthmann M, Pastori R, Moreno S: **Polyamines and basic proteins stimulate activation by cAMP and catalytic activity of *Mucor rouxii* cAMP-dependent protein kinase.** *Cellular signalling* 1990, **2**(4):395-402.
71. Kopperud R, Christensen AE, Kjarland E, Viste K, Kleivdal H, Doskeland SO: **Formation of inactive cAMP-saturated holoenzyme of cAMP-dependent protein kinase under physiological conditions.** *J Biol Chem* 2002, **277**(16):13443-13448.
72. Vigil D, Blumenthal DK, Brown S, Taylor SS, Trewhella J: **Differential effects of substrate on type I and type II PKA holoenzyme dissociation.** *Biochemistry* 2004, **43**(19):5629-5636.
73. Yang S, Fletcher WH, Johnson DA: **Regulation of cAMP-dependent protein kinase: enzyme activation without dissociation.** *Biochemistry* 1995, **34**(19):6267-6271.
74. Carlson CR, Ruppelt A, Tasken K: **A kinase anchoring protein (AKAP) interaction and dimerization of the RIalpha and RIIbeta regulatory subunits of protein kinase a in vivo by the yeast two hybrid system.** *J Mol Biol* 2003, **327**(3):609-618.
75. Herberg FW, Taylor SS: **Physiological inhibitors of the catalytic subunit of cAMP-dependent protein kinase: effect of MgATP on protein-protein interactions.** *Biochemistry* 1993, **32**(50):14015-14022.
76. Kim C, Xuong NH, Taylor SS: **Crystal structure of a complex between the catalytic and regulatory (RIalpha) subunits of PKA.** *Science (New York, NY)* 2005, **307**(5710):690-696.
77. Kim C, Cheng CY, Saldanha SA, Taylor SS: **PKA-I holoenzyme structure reveals a mechanism for cAMP-dependent activation.** *Cell* 2007, **130**(6):1032-1043.
78. Wu J, Brown SH, von Daake S, Taylor SS: **PKA type IIalpha holoenzyme reveals a combinatorial strategy for isoform diversity.** *Science (New York, NY)* 2007, **318**(5848):274-279.
79. Brown SH, Wu J, Kim C, Alberto K, Taylor SS: **Novel isoform-specific interfaces revealed by PKA RIIbeta holoenzyme structures.** *J Mol Biol* 2009, **393**(5):1070-1082.

80. Boettcher AJ, Wu J, Kim C, Yang J, Bruystens J, Cheung N, Pennypacker JK, Blumenthal DA, Kornev AP, Taylor SS: **Realizing the allosteric potential of the tetrameric protein kinase A R1alpha holoenzyme.** *Structure (London, England : 1993)* 2011, **19**(2):265-276.
81. Vigil D, Blumenthal DK, Taylor SS, Trewhella J: **Solution scattering reveals large differences in the global structures of type II protein kinase A isoforms.** *J Mol Biol* 2006, **357**(3):880-889.

CAPÍTULO 2

MAPEO, ANÁLISIS BIOINFORMÁTICO Y PRODUCCIÓN DEL DOMINIO D/D DE BCY1

“ I need a new voice, a new law, a new way
Take the time, reevaluate
It's time to pick up the pieces
Go back to square one
I think it's time for a change ”

Dream Theater, “Take the time”

ANTECEDENTES

1) En las subunidades R de mamíferos, los dominios D/D se localizan en el extremo N-terminal de la molécula. En esta región se encuentran los determinantes moleculares necesarios para la dimerización y para la interacción con las proteínas de anclaje AKAPs. Los residuos responsables del mantenimiento de la estructura dimérica son principalmente aminoácidos alifáticos de cadena larga.

2) Estructuralmente, estas regiones presentan un motivo hélice-giro-hélice, que en el dímero se encuentran organizados en una estructura altamente simétrica de tipo manojo de cuatro α -hélices dispuestas en forma antiparalela.

3) Se ha reportado que, al igual que las proteínas de mamíferos, Bcy1 es dimérica *in vitro*, tanto obtenida de levaduras [1] como sobreexpresada en bacterias [2]. Resultados previos de nuestro laboratorio obtenidos mediante experimentos de ultracentrifugación en gradiente de sacarosa sugieren que una mutante de delección de Bcy1 a la que le faltan los primeros 85 aminoácidos de su extremo N-terminal (Bcy1 $\Delta(1-85)$) es monomérica (Tesis Doctoral de Jimena Rinaldi, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi).

4) Los primeros 48 aminoácidos del extremo N-terminal de Bcy1 participan en la localización subcelular de la proteína, en concordancia con la funcionalidad conocida para los dominios D/D de mamíferos [3].

OBJETIVOS Y ABORDAJES UTILIZADOS

Los antecedentes consignados en la sección anterior sugieren fuertemente que el dominio D/D de Bcy1 está contenido en los 85 aminoácidos del extremo N-terminal de la molécula. En base a esta hipótesis, se plantearon los siguientes objetivos para este capítulo:

1) Identificar la región de Bcy1 responsable de su estado oligomérico. Para ello, se utilizaron herramientas bioinformáticas como predicciones de estructura secundaria y alineamientos múltiples de secuencias en comparación con los dominios D/D de las subunidades R de mamíferos. Por otro lado, se realizó un análisis del estado oligomérico de

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

Bcy1 WT y de la mutante de delección Bcy1 $\Delta(1-85)$ mediante experimentos de entrecruzamiento químico.

2) Obtener el fragmento correspondiente al dominio D/D de Bcy1 en cantidad y con elevado grado de pureza para su posterior caracterización estructural. Para ello se clonó el fragmento seleccionado en un vector de expresión bacteriano, y se desarrolló un protocolo de sobreexpresión y purificación de la proteína recombinante correspondiente. Finalmente, se analizó su capacidad de establecer dímeros mediante experimentos de entrecruzamiento químico.

RESULTADOS Y DISCUSIÓN

Análisis bioinformático de la región N-terminal de Bcy1

En base a los antecedentes, se espera que el dominio D/D se encuentre localizado hacia el extremo N-terminal de la molécula. Como ya se señaló, estos dominios se definen como las regiones que son necesarias y suficientes para garantizar la dimerización y la interacción con AKAPs; para RI α comprende los aminoácidos 12-61 [4], y para RII α , 1-44 [5]. Para evaluar esta hipótesis, como primer abordaje se realizó un análisis de la secuencia de Bcy1 empleando herramientas bioinformáticas. En primer lugar, la base de datos de familias de proteínas Pfam (<http://pfam.xfam.org/>) predice que en esta región de la molécula existe un dominio perteneciente a la familia RII α . Para ahondar en este punto, se construyó un alineamiento múltiple que contiene secuencias de dominios D/D de mamíferos de estructura conocida. Por otro lado, se realizaron predicciones de estructura secundaria de Bcy1 para obtener información acerca de la posible existencia en esta región de un motivo hélice-giro-hélice, característico de estos dominios. Estos resultados se muestran en la **Figura 2.1**.

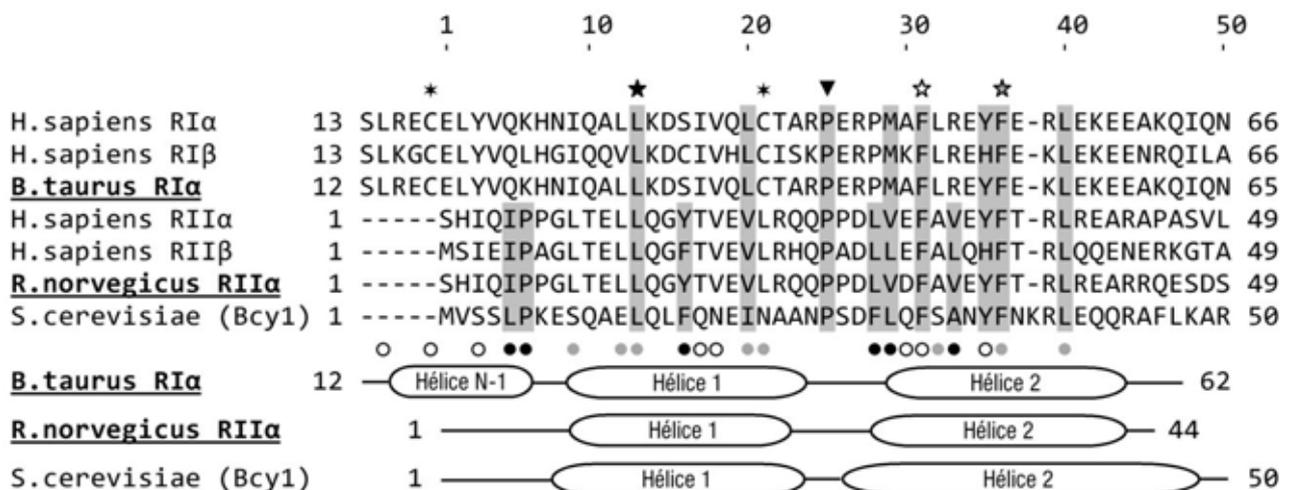


Figura 2.1: Alineamiento múltiple de secuencias y estructura secundaria de dominios D/D.

Se muestra un alineamiento de secuencias de la región N-terminal de las subunidades R, tomando como miembros representativos del grupo las proteínas humanas, bovinas y murinas; la numeración corresponde a la secuencia de Bcy1. Las proteínas para las cuales existen estructuras de alta resolución se muestran subrayadas y en negrita. Algunas posiciones importantes se encuentran señaladas en la parte superior del alineamiento: residuos de cisteína involucrados en la formación de puentes disulfuro intercatenarios en las subunidades RI (asterisco), residuo de prolina central que

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

separa las dos α -hélices (flecha), y residuos cuya mutación por alanina impide la dimerización de las subunidades R de mamíferos (estrella blanca si afecta a RI pero no a RII, negra si afecta a RII pero no a RI, y gris si afecta a ambas). Los círculos en la parte inferior del alineamiento representan residuos hidrofóbicos que establecen contactos clave para la dimerización en base a las estructuras de alta resolución en solución de RI α bovina [4] (pdb: 2EZW) y de RII α murina [5] (pdb: 1R2A); el código de grises para estos residuos es el mismo que para las mutaciones por alanina. Los residuos sombreados en gris son aquellos de entre los anteriores que se encuentran conservados en Bcy1. En la parte inferior se representa esquemáticamente la estructura secundaria experimental de los dominios D/D de mamíferos, o bien la predicción para Bcy1; las cajas indican regiones de α -hélice.

Las predicciones sugieren la existencia de un motivo hélice-giro-hélice en el extremo N-terminal de Bcy1; en particular, la hélice 1 comprende los aminoácidos 8-22, mientras que la hélice 2 consta de los aminoácidos 26-48. Este motivo se encuentra flanqueado por regiones desestructuradas correspondientes al *linker I* hacia su C-terminal, y a un pequeño segmento extendido hacia el N-terminal. De acuerdo con estas predicciones, el dominio D/D de Bcy1 es más similar al de RII α que al de RI α , debido a que el segundo presenta una α -hélice adicional (hélice N-1) que precede a la hélice 1, mientras que el primero presenta en su lugar un segmento desestructurado, que es lo que se predice para Bcy1. Por otro lado, una característica distintiva del dominio D/D de Bcy1 es que su hélice 2 se extiende en base a las predicciones unos cinco aminoácidos más (un giro y medio) que las de mamíferos; en concordancia con lo anterior, se observa que fue necesario introducir un *gap* en esa zona para optimizar el alineamiento. De acuerdo a estas observaciones, el segmento de Bcy1 8-48 contiene los elementos de estructura secundaria característicos de los dominios D/D.

El alineamiento fue analizado en el contexto de la bibliografía sobre la estructura de los dominios D/D de mamíferos. Numerosos trabajos han establecido los principales determinantes responsables de la dimerización de las distintas isoformas de las subunidades R. Los primeros trabajos consisten en estudios de mutagénesis que han permitido la identificación de residuos críticos cuyo reemplazo por alanina impide la dimerización. Por otro lado, posteriormente se han reportado las estructuras de alta resolución de estos dominios, estableciendo cuáles son los residuos que participan en contactos hidrofóbicos que son esenciales para el mantenimiento de la estructura dimérica.

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

En primer lugar, se analizaron los residuos característicos identificados por mutagénesis, que se indican con símbolos en la parte superior del alineamiento de la **Figura 2.1**. La prolina central que se encuentra en el *loop* que separa las dos α -hélices en todos los dominios D/D de mamíferos está presente en Bcy1 (Pro25). Por otro lado, existen tres residuos cuya mutación por alanina impide la dimerización de las subunidades R de mamíferos: las mutaciones Phe47Ala y Phe52Ala en RI α [6], y las mutaciones Leu13Ala y Phe36Ala en RII β [7] generan proteínas monoméricas (ver estrellas en la **Figura 2.1**); estos tres residuos están estrictamente conservados en Bcy1 (Leu13, Phe31, Phe36). La conservación estricta de estos tres residuos clave, además de la prolina central, y la predicción de estructura secundaria, sugieren fuertemente que en esta región de Bcy1 se encuentra presente un dominio D/D con características similares a los de mamíferos.

En segundo lugar, se analizaron los residuos que constituyen el núcleo hidrofóbico responsable del mantenimiento de la estructura dimérica en base a las estructuras de alta resolución de los dominios D/D de mamíferos; estos residuos se indican en círculos en la parte central de la **Figura 2.1**. La **Tabla 2.1** resume algunas observaciones importantes al respecto.

Tabla 2.1: Conservación de residuos de RI y RII en Bcy1.

En la columna 2 se indica la cantidad y el porcentaje del total de los residuos que constituyen el núcleo hidrofóbico de dimerización del dominio D/D de cada isoforma de R que se encuentra conservado en Bcy1. La columna 2 muestra el porcentaje de similitud entre las secuencias completas de los dominios D/D de cada una de las isoformas y Bcy1.

	Núcleo hidrofóbico	Todos
RI	37,5% (6/16)	32% (16/50)
RII	71,4% (10/14)	38% (18/50)

Se observan diferencias evidentes respecto de la conservación en Bcy1 de los residuos del núcleo hidrofóbico de cada una de las dos isoformas: sólo el 37,5% de los residuos importantes para la dimerización de RI α está presente en Bcy1, mientras que el 71,4% de los residuos importantes para la dimerización de RII α está presente en Bcy1 (**Tabla 2.1**, segunda columna). Es interesante comparar estos números con el grado de similitud entre la secuencia

completa de la región de Bcy1 (1-50) y los dominios D/D de mamíferos: para RI α los números son similares (32% de similitud de secuencia y 37,5% de similitud del núcleo hidrofóbico), mientras que para RII α el núcleo hidrofóbico está muy conservado en Bcy1, contrariamente a lo que ocurre con la secuencia completa (38% de similitud de secuencia y 71,4% de similitud del núcleo hidrofóbico). Estas observaciones sugieren que, a pesar de ser una región no demasiado conservada globalmente, el dominio D/D de Bcy1 contiene un núcleo de aminoácidos hidrofóbicos bastante similar al de los dominios D/D de las subunidades RII α , reforzando la hipótesis de que en esta región residen los determinantes responsables de la dimerización. En concordancia con esa idea, se ha observado con respecto a los núcleos de dimerización de los dominios D/D de RI α y RII α , que la región correspondiente a las hélices 1 y 2 alinea estructuralmente muy bien pese a que la similitud de secuencias no es demasiado alta (35%) [8].

Cabe destacar que la mayoría de estos residuos no se encuentran conservados estrictamente en Bcy1, sino sustituidos conservativamente. En los trabajos de mutagénesis de los dominios D/D de las subunidades R de mamíferos se observó que la mayor parte de los residuos que constituyen el núcleo de dimerización pueden ser mutados por otros residuos alifáticos sin interrumpir la formación del dímero [6, 7]. Esto indica que, pese a participar en interacciones importantes, su conservación estricta no es requerida para garantizar su función, y por lo tanto el hecho de que no se encuentren estrictamente conservados en Bcy1 no se contradice con su potencial rol en la dimerización.

Para concluir esta parte del análisis, es interesante remarcar que el dominio D/D de Bcy1 es más similar al de RII α que al de RI α respecto de numerosas características. En primer lugar, como ya se señaló, Bcy1 presenta una topología de estructura secundaria similar a la de RII α . Por otro lado, en Bcy1 no están presentes los dos residuos de cisteína que establecen puentes disulfuro intercatenarios en las subunidades RI α pero no en las RII α (**Figura 2.1**). Finalmente, ya se comentó también que la similitud de secuencia de Bcy1 para todo el dominio D/D (considerando la región de Bcy1 (1-50)) es ligeramente más alta con RII α que con RI α (38% y 32%); mucho más significativa resulta la comparación de los residuos del núcleo hidrofóbico (71,4% y 37,5%).

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

Para completar el análisis bioinformático, se procedió a analizar la posible participación de las α -hélices predichas para Bcy1, junto con los residuos hidrofóbicos conservados, en *clusters* que se pudieran formar en la interfaz entre ambas constituyendo un núcleo de dimerización. Para ello se utilizó el modelo de “*Helical Wheel*”, que consiste en representar las α -hélices considerando que cada giro consta de 3,6 aminoácidos, y analizar la distribución resultante de aminoácidos en sus distintas caras. Para ello se utilizó el programa “HeliQuest” [9]. El programa calcula además el vector momento hidrofóbico de la α -hélice, que es la suma vectorial de las hidrofobicidades de las cadenas laterales de los residuos que la constituyen; su dirección y su módulo son una medida de la anfipaticidad de la α -hélice. En la **Figura 2.2** se muestran los resultados obtenidos para los modelos de las α -hélices del dominio D/D de Bcy1 en comparación con las de RII α .

De acuerdo con las estructuras de alta resolución de los dominios D/D de RII α , las caras constituidas principalmente por residuos no polares que se localizan aproximadamente en la dirección del momento hidrofóbico de las hélices se encuentran sumergidas en el interior del núcleo de dimerización del dominio (residuos Leu9, Leu12, Leu13, Tyr16 y Val20 para la hélice 1, y Leu28, Val29, Ala32, Val33, Phe36 y Val39 para la hélice 2). Se observa que las hélices de Bcy1 presentan momentos hidrofóbicos aproximadamente en la misma dirección que las correspondientes hélices de RII α , presentando caras no polares similares que podrían participar en el mismo tipo de contactos. En particular, los residuos Leu13, Phe16 e Ile20 de la hélice 1 podrían constituir una superficie análoga, aunque menos extendida, que la que corresponde a la hélice 1 de RII α . Por otro lado, en lo que respecta a la hélice 2, la situación es algo distinta: si bien el momento hidrofóbico de la hélice es similar, los residuos no polares parecen estar agrupados en dos *clusters*, a uno y otro lado de tres residuos polares. Existen tres residuos, Leu40, Phe46 y Leu47, que no forman parte del núcleo hidrofóbico del D/D de las subunidades RII α , pero que sí parecen hacerlo en Bcy1; los dos últimos se encuentran localizados hacia el final de la hélice 2. Estos residuos podrían aportar características distintivas al núcleo de dimerización de Bcy1, en concordancia con lo ya observado previamente en las predicciones de estructura secundaria, según las cuales la hélice 2 de Bcy1 presenta casi dos giros adicionales a las de los dominios D/D de mamíferos.

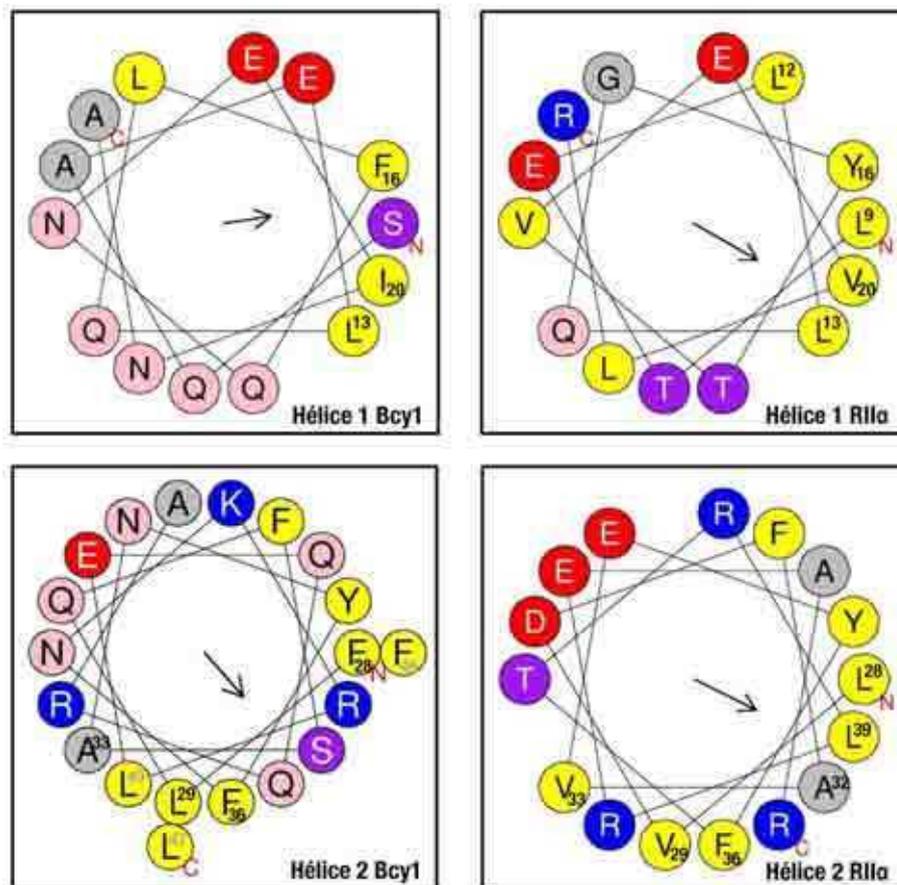


Figura 2.2: Análisis de las hélices 1 y 2 de Bcy1 mediante el modelo de “*Helical Wheel*”.

Izquierda: modelos generados para las dos α -hélices predichas de Bcy1 (hélice 1: 9-22, hélice 2: 28-47). Derecha: representación de las α -hélices de RII α . Se indican los extremos N y C terminal de cada una; en todos los casos se ubicó el extremo N-terminal hacia la derecha para facilitar la comparación. Los residuos alifáticos y los residuos aromáticos se muestran en amarillo, las amidas en rosa, los alcoholes en lila, los residuos cargados positivamente en rojo, los cargados negativamente en azul, y la alanina en gris. Las flechas indican la dirección del momento hidrofóbico de la α -hélice. Los residuos involucrados en interacciones importantes para la dimerización en las dos hélices de RII α que se indicaron en la **Figura 2.1** están numerados, junto con los residuos correspondientes conservados en Bcy1. Se indican numerados en gris residuos hidrofóbicos distintivos de Bcy1, no presentes en RII α , pero que parecen aportar a la posible interfaz de dimerización.

Mapeo experimental del dominio D/D de Bcy1

Para confirmar experimentalmente las predicciones que se desprenden del análisis de la sección anterior y determinar si la región N-terminal de Bcy1 es suficiente para garantizar su correcto estado oligomérico en solución, se realizaron experimentos de entrecruzamiento

químico utilizando el reactivo etilenglicol-bis (succinimidil succinato) (EGS). El EGS es un reactivo homobifuncional que reacciona selectivamente con los extremos N-terminal o con los grupos ϵ -amino de los residuos de lisina de las proteínas. Presenta en cada extremo de la molécula un grupo N-hidroxisuccinimida, y ambos están unidos a la cadena carbonada que los separa por un enlace éster. Su estructura y reactividad se muestran en la **Figura 2.3**.

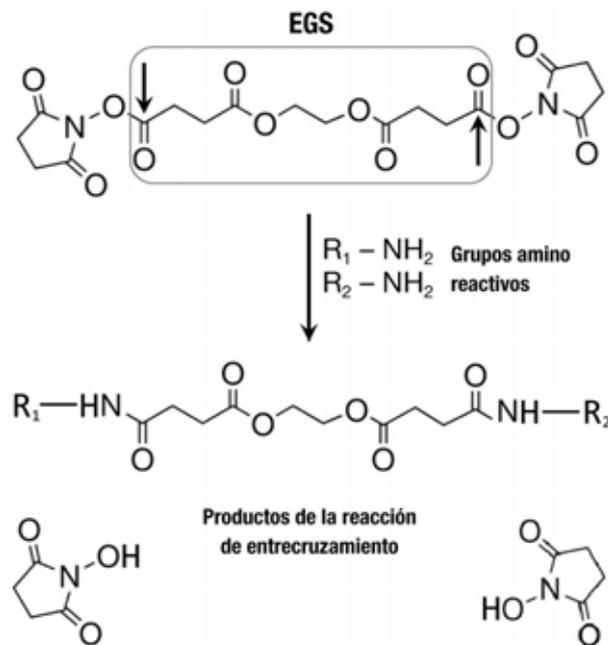


Figura 2.3: Estructura y reactividad del EGS.

Las flechas indican las posiciones en las cuales ocurre el ataque nucleofílico por parte de los grupos ϵ -amino de los residuos de lisina, o bien de los grupos N-terminal de las cadenas polipeptídicas. El recuadro gris muestra el brazo separador cuya longitud es de 1,6 nm.

Los dos grupos carbonilo a ambos lados de la molécula se comportan como electrófilos frente a un ataque nucleofílico de los grupos amino de las proteínas. Cuando dos de estos grupos se encuentran lo suficientemente cerca en el espacio (la longitud de la cadena carbonada en el EGS es de 1,6 nm), cada uno de ellos puede reaccionar con uno de los extremos del reactivo, quedando de esta manera conectados. La función del reactivo es entonces establecer uniones covalentes entre regiones de la molécula que se encuentran relativamente cercanas en el espacio; de esta manera pueden capturarse covalentemente subunidades que interactúan entre sí por fuerzas intermoleculares.

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

Para analizar la participación de la región N-terminal de Bcy1 en su dimerización, la proteína WT y una mutante de delección a la que le faltan los primeros 85 aminoácidos del extremo N-terminal (Bcy1 $\Delta(1-85)$) fueron tratadas con EGS, y los productos de la reacción se analizaron por SDS-PAGE y tinción con Coomassie Blue coloidal (este método de tinción es compatible con la espectrometría de masa, que se discute más adelante). Los resultados se presentan en la **Figura 2.4**.

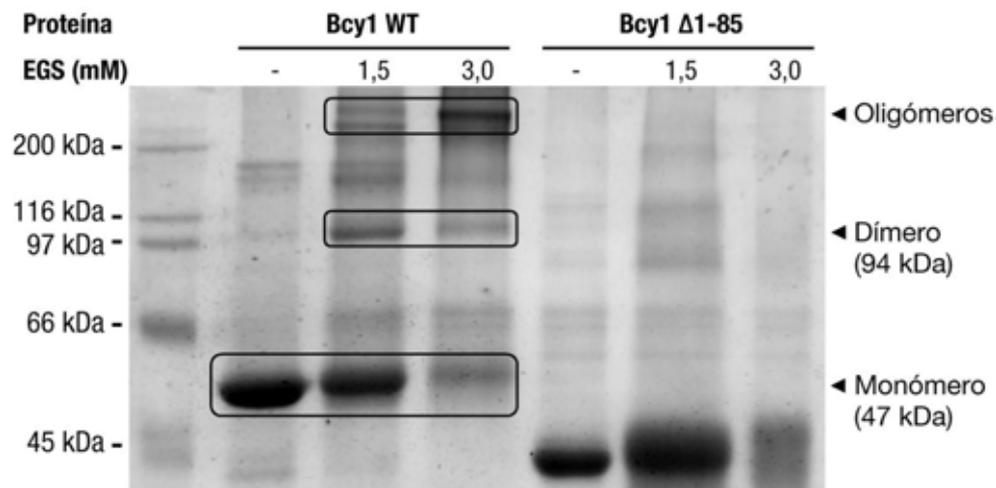


Figura 2.4: Experimentos de entrecruzamiento químico de Bcy1.

Muestras purificadas de Bcy1 WT y Bcy1 $\Delta(1-85)$ fueron tratadas con EGS en las concentraciones indicadas en la parte superior de la figura. Los productos de la reacción fueron separados por SDS-PAGE 10% y visualizados por tinción con Coomassie Blue coloidal. Se indican con un recuadro las bandas correspondientes a monómero, dímero, y oligómeros de mayor masa molecular de Bcy1 WT. La banda de la calle 2 correspondiente al monómero y la banda de la calle 3 correspondiente al dímero fueron escindidas del gel para su análisis posterior por espectrometría de masa.

Como consecuencia del tratamiento con EGS, se observa la aparición de bandas de alta masa molecular (MM) para la variante WT, pero no para la mutante. En particular, como era esperable, se aprecia una banda muy definida correspondiente a un dímero. Las bandas de mayor MM podrían en principio deberse a agregados inespecíficos; sin embargo, en las concentraciones utilizadas para los ensayos (0,02 mg/mL) no es posible que la muestra se encuentre agregada, dado que, de acuerdo a resultados no publicados de dispersión de rayos X a bajo ángulo (SAXS) de nuestro laboratorio, Bcy1 puede manejarse en concentraciones de

hasta 4 mg/mL sin agregación apreciable (Tesis Doctoral de Lucas Fernández Núñez, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi, en preparación). Por lo tanto, es improbable que estas bandas se deban a agregados inespecíficos.

En conclusión, estos resultados indican que los primeros 85 residuos del extremo N-terminal de Bcy1 están involucrados en el establecimiento no solamente de dímeros, como se esperaba en base al análisis bioinformático, sino también de oligómeros con mayor cantidad de subunidades, lo que sugiere que podrían existir asociaciones de mayor orden previamente no reportadas. Por otro lado, los resultados demuestran que el dominio D/D se encuentra restringido a esta región de la molécula, en concordancia con las predicciones bioinformáticas.

En nuestro laboratorio se ha observado que la región N-terminal de Bcy1 es necesaria para la interacción con proteínas responsables de la localización en distintos compartimientos celulares. La mutante Bcy1 $\Delta(1-85)$, que carece del dominio D/D, es incapaz de unirse a proteínas interactoras como Ira2 y Hsp60 [10]. El hecho de que tampoco pueda establecer oligómeros sugiere que, al igual que en las proteínas de mamíferos, en Bcy1 los determinantes de oligomerización y de interacción se encuentran localizados en la misma región de la molécula.

La espectrometría de masa (MS) combinada con los experimentos de entrecruzamiento químico descritos en los párrafos anteriores resulta a menudo de mucha utilidad para obtener información estructural acerca de complejos macromoleculares [11]. En particular, puede permitir identificar qué regiones de las moléculas participan en las interacciones que mantienen unido el complejo. El abordaje más común consiste en el tratamiento de las bandas cortadas del gel con tripsina y posterior análisis de los fragmentos resultantes por MS. De esta manera se obtiene el patrón de los fragmentos trípticos que las caracterizan. El análisis comparativo de los espectros de las bandas correspondientes a los complejos y a las unidades que los conforman por separado permite identificar, a partir de las diferencias, péptidos que podrían encontrarse involucrado en el establecimiento de interacciones.

A efectos de obtener más información respecto del rol estructural que podrían desempeñar las distintas regiones del extremo N-terminal de Bcy1 en la oligomerización, las bandas correspondientes a monómero y dímero de Bcy1 WT que se indican al pie de la **Figura 2.4** fueron cortadas del gel para su análisis por MS. Las bandas correspondientes a

asociaciones de mayor MM también se procesaron, pero por ser muy tenues los espectros resultaron insatisfactorios por presentar muy baja relación señal/ruido. En la **Figura 2.5** se muestran los resultados obtenidos.

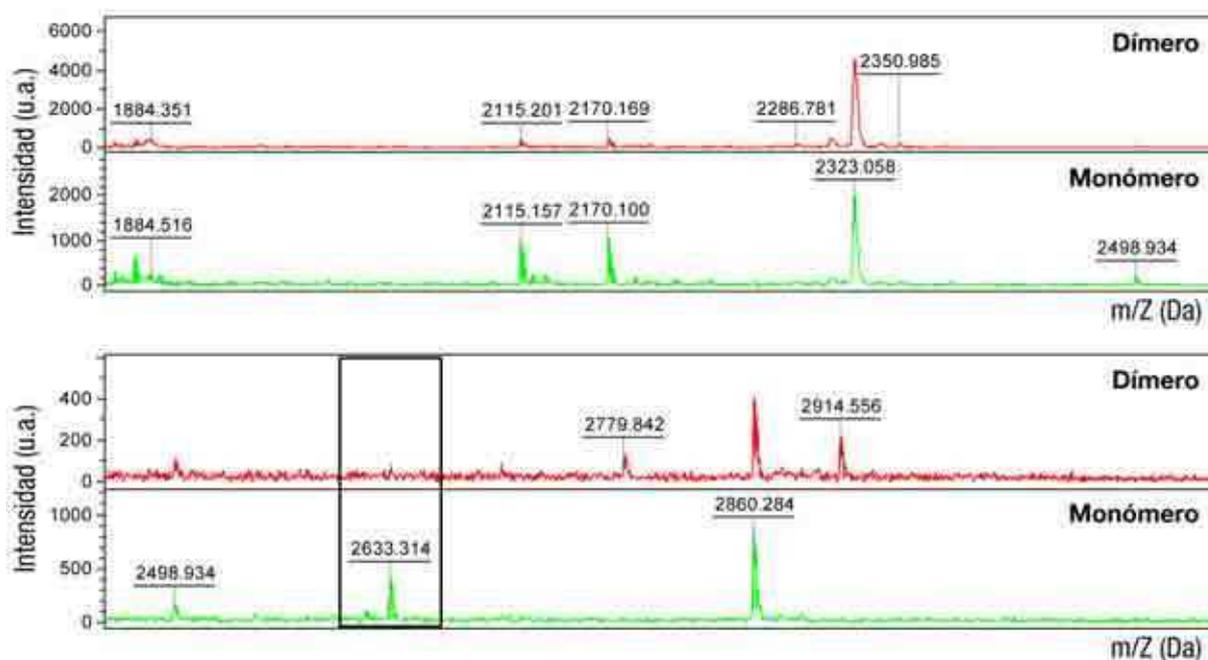


Figura 2.5: Espectros de masa de monómero y dímero de Bcy1.

Se muestra el análisis de los fragmentos trípticos de las bandas cortadas del gel de la **Figura 2.4**. En verde se muestra el espectro del monómero, y en rojo el espectro del dímero capturado covalentemente. Para cada pico se indica el valor de su m/z. El recuadro señala la posición del pico de 2633,3 Da.

Los espectros de monómero y dímero resultaron casi idénticos, a excepción de un pico de $m/Z = 2633,3$ Da (indicado en la **Figura 2.5** con un recuadro), que está ausente en el espectro correspondiente al dímero capturado por el tratamiento con EGS, pese a ser bastante abundante en la banda correspondiente al monómero. Este péptido se sometió a espectrometría de masa en tándem (MS-MS) a efectos de verificar su identidad; se corroboró que corresponde a la secuencia 58-80 de Bcy1 (57-K.NIVLFPEPEESFSRPQSAQSQR.-80), que se encuentra a continuación del dominio D/D predicho. Dado que el aminoácido que está ubicado inmediatamente antes de este fragmento tríptico es una lisina, una explicación posible para su desaparición en la banda correspondiente al dímero es que ese residuo haya

reaccionado con el EGS durante la reacción de entrecruzamiento, contribuyendo a establecer el dímero unido covalentemente que se logró capturar en estos experimentos. Esto impediría que la tripsina (que hidroliza péptidos hacia el C-terminal de residuos de lisina o arginina) corte en esa posición durante el procesamiento de las bandas, y como consecuencia no se generaría el fragmento tríptico correspondiente.

Estos resultados sugieren fuertemente que los péptidos que comprenden los aminoácidos 58-80 de cada monómero del dímero se encuentran muy cercanos en el espacio. Este péptido se encuentra ubicado en el *linker I*, que como ya se discutió en la Introducción, es una región poco conservada, intrínsecamente desordenada en mamíferos [12], y que el análisis bioinformático presentado efectivamente predijo como una región de Bcy1 desestructurada. Es interesante remarcar también que este péptido incluye el *cluster II* de residuos de serina fosforilables que se mencionó en la Introducción. El hecho de que se encuentren cercanos en el espacio permite hipotetizar que esta región podría tener algún tipo de rol regulatorio.

Es importante resaltar que no es posible determinar si ocurrió entrecruzamiento en la región 1-50 (el probable dominio D/D de Bcy1) mediante este abordaje, dado que si bien hay residuos de lisina en esta zona, no se observaron péptidos correspondientes a esta región en los espectros de masa. Esto podría deberse a que los péptidos que se generan son de muy pocos aminoácidos como consecuencia de la distribución de los residuos de lisina y arginina en esta zona de la molécula, y por lo tanto aparecen en la región de bajo m/z de los espectros, que es una zona difícil de analizar porque allí se observan numerosos picos correspondientes a la matriz utilizada para la ionización de los péptidos por MALDI.

Producción del dominio D/D de Bcy1

Si bien las observaciones de la sección anterior podrían sugerir que en la región 58-80 existen elementos adicionales al dominio D/D que contribuyen a la dimerización, se decidió seleccionar el fragmento 1-50 de Bcy1, correspondiente dominio D/D predicho, para proceder a su clonado y sobreexpresión. Esta decisión se basa en que uno de los principales objetivos de este trabajo fue determinar la estructura cristalina del dominio D/D de Bcy1 y en que, como ya se discutió en la Introducción, se ha observado que en las proteínas de mamíferos la desestructuración del *linker I* atenta contra la cristalización. De esta manera, se eligió el mínimo fragmento posible con capacidad de dimerizar para maximizar las posibilidades de

cristalización. Otra observación que apoya la elección del fragmento 1-50 es que el residuo 52 de Bcy1 es una prolina, lo que sugiere que cualquier elemento de estructura secundaria previo (es decir, la hélice 2) debería interrumpirse en esa posición.

Se procedió por lo tanto al clonado de este fragmento; para ello se utilizó el vector de expresión en bacterias pET32a, que permite producir una proteína recombinante de interés bajo un promotor inducible por IPTG. Se generó una construcción que codifica para una proteína de fusión que se representa esquemáticamente en la **Figura 2.6**. Una vez obtenida la construcción y verificada su identidad por secuenciación, se procedió a la optimización de su sobreexpresión y purificación.

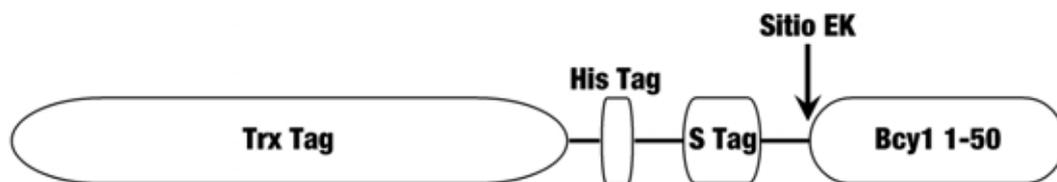


Figura 2.6: Representación esquemática de la proteína de fusión generada en este trabajo.

De N a C terminal consta de los siguientes segmentos: *tag* de tiorredoxina (Trx; 109 aminoácidos) que permite mejorar la solubilidad de las proteínas recombinantes expresadas en bacterias, minimizando su localización en cuerpos de inclusión; *tag* de polihistidina (His; 6 aminoácidos) para su purificación por afinidad; *tag* S (15 aminoácidos), que consta de un péptido derivado de la ribonucleasa, que permite el reconocimiento por anticuerpos específicos; sitio de corte para enteroquinasa (Sitio EK; 5 aminoácidos), que permite remover los *tags* anteriores; y Bcy1 (1-50) (50 aminoácidos). La MM de la proteína de fusión completa es 22,9 kDa, mientras que la del fragmento resultante del corte con enteroquinasa, que deja un residuo adicional de alanina adicional en el extremo N-terminal de Bcy1 (1-50), es de 5,9 kDa.

Se desarrolló un protocolo de purificación que consta de los siguientes pasos: obtención de un extracto libre de células; purificación de la proteína de fusión mediante un paso de incubación con una resina de afinidad de Ni²⁺; remoción del *tag* de la proteína de fusión inmovilizada en la resina por corte con enteroquinasa; purificación del sobrenadante de la incubación anterior por cromatografía de exclusión molecular (SEC); purificación del eluido del paso anterior mediante un paso de intercambio iónico.

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

La identidad del fragmento generado de acuerdo a este protocolo fue verificada por MS (ver Materiales y métodos). En la **Figura 2.7** se muestra el resultado de una purificación característica, seguida por SDS-PAGE de alícuotas de cada uno de los pasos.

Se observa que la proteína de fusión se expresa en gran cantidad como resultado de la inducción del cultivo con IPTG (calles 1 y 2). Por otro lado, resulta evidente la utilidad de cada uno de los pasos de la purificación (calles 7, 8 y 9). Se logró obtener Bcy1 (1-50) con un alto grado de pureza; su identidad fue verificada por MS de (ver Materiales y métodos).

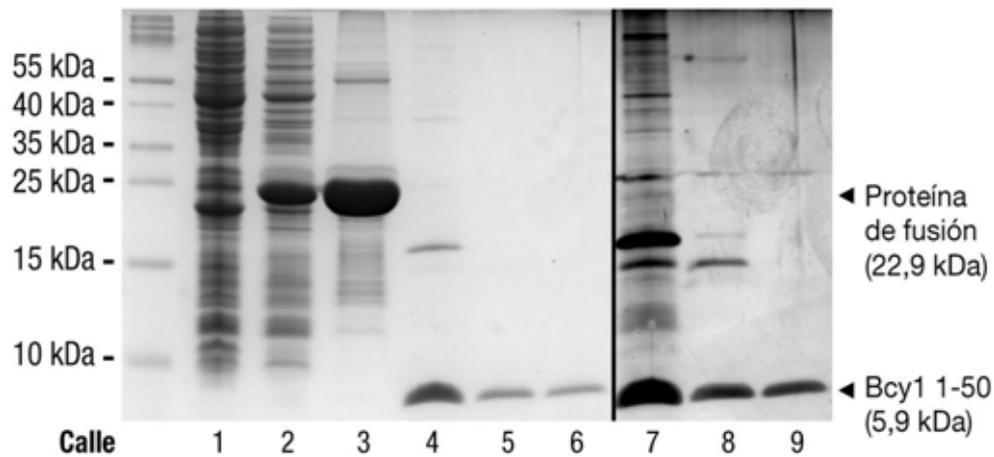


Figura 2.7: Purificación de Bcy1 (1-50).

SDS-PAGE 15% en buffer Tricina de alícuotas de las distintas fracciones de una purificación representativa de Bcy1 (1-50). El panel de la izquierda muestra una tinción con Coomassie Blue mientras que el de la derecha fue teñido con plata para resaltar la utilidad de los dos últimos pasos de la purificación. Las calles son: 1) Extracto libre de células obtenido a partir de un cultivo no inducido de una cepa de *E. coli* transformada con la construcción pET32a-Bcy1(1-50). 2) Extracto libre de células de un cultivo de la misma cepa luego de inducción por 4 horas con IPTG. 3) Proteínas retenidas en la resina de Ni²⁺ luego de la incubación con (2) y lavados. 4) Sobrenadante de la digestión con enteroquinasa de (3). 5) Pico correspondiente a Bcy1 (1-50) obtenido por SEC utilizando una columna Superdex-75. 6) Sobrenadante de la incubación en *batch* de (5) con DEAE-celulosa. 7, 8 y 9 son las mismas calles que 4, 5 y 6, pero visualizadas por tinción con plata.

Finalmente, para evaluar la capacidad del fragmento producido de dimerizar y verificar así que corresponde efectivamente al dominio D/D de Bcy1, se realizaron experimentos de entrecruzamiento químico con EGS. En la **Figura 2.8** se muestran los resultados obtenidos.

Capítulo 2 Caracterización preliminar del dominio D/D de Bcy1

Como consecuencia del tratamiento con EGS, se observa la aparición de bandas de mayor MM que el monómero. En particular, resulta evidente la presencia de una banda correspondiente a un dímero en varias de las calles. Inesperadamente, se observan también bandas correspondientes a formas de mayor grado de asociación, cuya MM coincide con un trímero y un tetrámero. No se observaron formas de mayor grado de oligomerización que tetrámero.

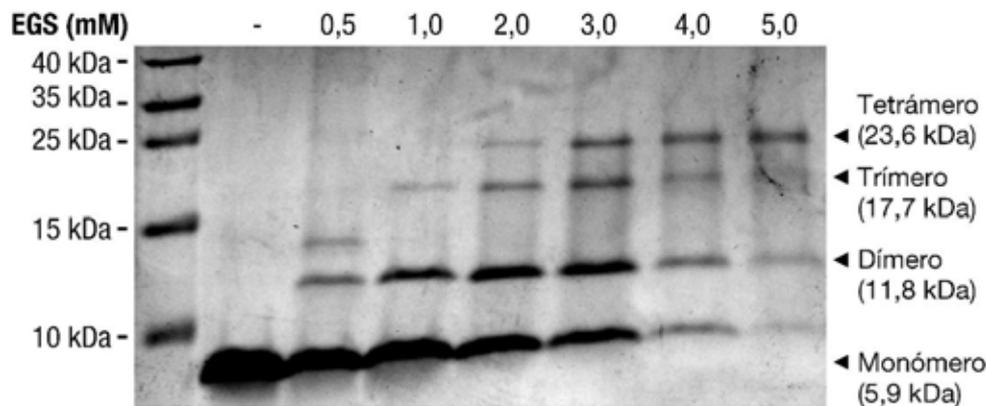


Figura 2.8: Experimentos de entrecruzamiento químico de Bcy1 (1-50).

Muestras purificadas de Bcy1 (1-50) fueron tratadas con EGS en las concentraciones indicadas en la parte superior de la figura. Los productos de la reacción fueron separados por SDS-PAGE 15% utilizando el sistema de buffers Tricina, y visualizados por tinción con Coomassie Blue. Se indican con flechas las bandas correspondientes a monómero, dímero, trímero y tetrámero.

Las bandas correspondientes a formas intermedias que se evidencian en los experimentos de entrecruzamiento químico no son evidencia de que esas especies existan efectivamente en solución, dado que en estos experimentos es común observar formas intermedias debido a que el entrecruzamiento es ineficiente (por eso es que se trabaja con concentraciones crecientes del reactivo). Por lo tanto, estos resultados permiten afirmar que el fragmento producido tiene la capacidad de formar oligómeros de hasta 4 subunidades, pero no es posible obtener en base a estos experimentos información acerca de la población de estados oligoméricos que existen en solución; concretamente, no proporcionan evidencia acerca de la existencia del trímero, dímero o monómero.

De todas maneras, los resultados confirman que el fragmento Bcy1 (1-50) tiene la capacidad de establecer oligómeros en solución por sí mismo, y a pesar de haber encontrado estados de oligomerización inesperados, se definió a esta región como el dominio D/D de Bcy1. Estas observaciones podrían coincidir con los resultados de entrecruzamiento químico de Bcy1 WT completa, en los cuales se observaron bandas de alta MM indicativas de asociaciones superiores a dímero (ver **Figura 2.4**). Esta posibilidad resulta muy novedosa, dado que todas las subunidades R que contienen dominio D/D se consideran siempre diméricas. Cabe señalar, por otro lado, que en algunos trabajos en los que se reporta la optimización de las condiciones para la resolución de las estructuras de los dominios D/D de mamíferos por RMN, se han observado evidencias de asociaciones mayores que dímero. Para RI α [13] se observó por SEC que, a pHs cercanos a la neutralidad, existe una especie con una MM aparente de 18 kDa, cuando la MM esperada para el dímero es de 10 kDa. Asimismo, experimentos de dispersión dinámica de la luz (DLS) revelaron que esa muestra es polidispersa, indicando una heterogeneidad de poblaciones moleculares en esas condiciones. A pH = 4, por otro lado, la muestra se comportó como era esperado para un dímero. Para RI β [14] se realizaron observaciones similares. En ambos casos, las especies de alta MM fueron atribuidas a agregados inespecíficos no representativos de la estructura nativa, por lo que no se ha vuelto a indagar en ese aspecto, y los dominios D/D de todos los organismos se asumen de naturaleza dimérica.

CONCLUSIONES Y PERSPECTIVAS

En este capítulo se logró identificar la región de Bcy1 responsable de la oligomerización de la molécula, y producir en cantidad y calidad suficientes para su posterior caracterización estructural ese fragmento recombinante.

Por otro lado, se identificaron nuevos estados de oligomerización previamente no reportados tanto del dominio D/D de Bcy1 como de la proteína entera. En el siguiente capítulo se ahondará en este punto, a efectos de averiguar cuáles de las formas reportadas son representativas de verdaderas especies que existen en solución, y cuáles son las características estructurales que determinan su estado oligomérico.

Un aspecto interesante que no será abordado en este trabajo, pero que resulta como punto de partida para trabajos futuros, es el rol que puede desempeñar en la oligomerización el segmento que se encuentra hacia el C-terminal del D/D de Bcy1, en el *linker I*, que resulta de interés porque contiene residuos de serina fosforilables, y que por lo tanto podría ser funcionalmente relevante. Este punto está siendo abordado actualmente como parte de la Tesis Doctoral de Enzo Tofolón, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi.

MATERIALES Y MÉTODOS

Análisis de secuencias

Las secuencias de las proteínas RI α (P10644), RI β (P31321), RII α (P13861), RII β (P31323) de *Homo sapiens*, RI α de *Bos taurus* (P00514), RII α de *Rattus norvegicus* (P12368), y Bcy1 de *Saccharomyces cerevisiae* (P07278), fueron colectadas de la base de datos UniProtKB (www.uniprot.org). El alineamiento múltiple de secuencias fue generado utilizando el comando “align2d” implementado en el paquete de modelado por homología MODELLER [15]. A diferencia de la mayoría de los algoritmos utilizados para generar alineamientos que sólo consideran la identidad o similitud de aminoácidos, este método tiene en cuenta información estructural de las proteínas cuya estructura de alta resolución es conocida. En particular, considera el contexto en el que se localizan las inserciones y las deleciones, tendiendo a colocar estas últimas en regiones expuestas al solvente (dado que típicamente los núcleos de las estructuras se encuentran más conservados estructuralmente que las superficies), fuera de los elementos de estructura secundaria, y entre posiciones que se encuentran cercanas en el espacio.

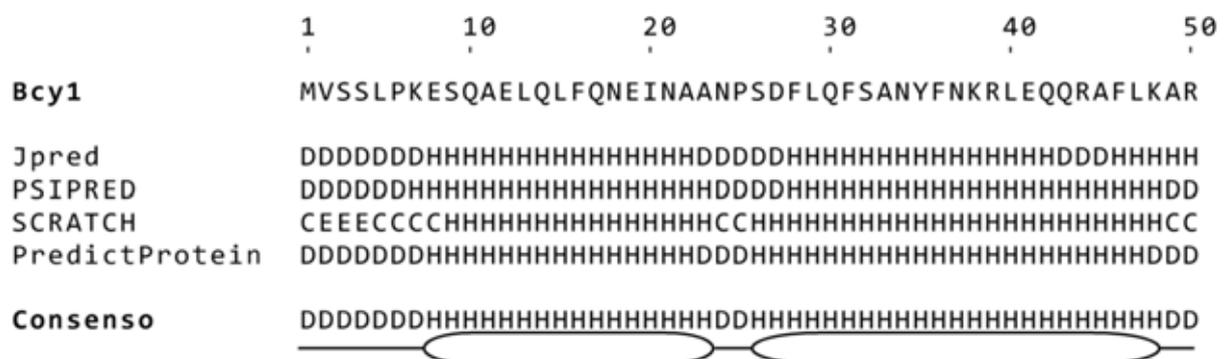


Figura 2.S.1: Predicción de estructura secundaria de la región N-terminal de Bcy1.

Se indican con códigos de letras los resultados proporcionados por los diferentes métodos. D: desordenado, H: α -hélice, E: hoja β , C: giro. Las cajas del esquema de la parte inferior representan las dos α -hélices predichas. Si bien no se muestra en la figura, a partir del aminoácido 51 se predice un largo segmento desestructurado correspondiente a la región del *linker*.

Las predicciones de estructura secundaria de Bcy1 fueron generadas como consenso de cuatro métodos independientes que utilizan distintos algoritmos, todos ellos ampliamente difundidos y muy citados en la bibliografía: PSIPRED [16] (<http://bioinf.cs.ucl.ac.uk/psipred/>), SCRATCH [17] (<http://scratch.proteomics.ics.uci.edu/>), Jpred [18] (www.compbio.dundee.ac.uk/www-jpred/) y PredictProtein [19] (<http://ppopen.rostlab.org/>). La predicción consenso que se reporta en la **Figura 2.1** se generó considerando para cada aminoácido de Bcy1 el elemento de estructura secundaria más frecuente de entre los cuatro métodos (**Figura 2.S.1**).

Los modelos de “*Helical Wheel*” fueron construidos utilizando el servidor HeliQuest [9] (heliquest.ipmc.cnrs.fr). Además de modelar las hélices, el servidor calcula su vector momento hidrofóbico medio, cuyo módulo se obtiene según la ecuación [20]:

$$\langle \mu H \rangle = \frac{1}{N} \left\langle \left[\sum_{n=1}^N H_n \sin \delta \right]^2 + \left[\sum_{n=1}^N H_n \cos \delta \right]^2 \right\rangle^{\frac{1}{2}}$$

donde μH es el momento hidrofóbico medio, N es la cantidad de aminoácidos de la hélice, H_n es la hidrofobicidad del n -ésimo aminoácido en la secuencia de acuerdo a su coeficiente de partición octanol/agua, y δ es el ángulo formado por cadenas laterales consecutivas en la secuencia a lo largo del esqueleto de la α -hélice ($\delta=100^\circ$).

Plásmidos y cepas

Tabla 2.S.1: Plásmidos utilizados en este trabajo.

Plásmido	Organismo	Descripción	Fuente
YEp51-BCY1	<i>S. cerevisiae</i>	Contiene el gen de Bcy1 WT bajo el promotor GAL10.	[21]
YEp51-Δ85BCY1	<i>S. cerevisiae</i>	Contiene el gen de Bcy1 con los primeros 85 aminoácidos deletados bajo el promotor GAL10.	[10]
pET32a	<i>E. coli</i>	-	Comercial
pET32a-Bcy1(1-50)	<i>E. coli</i>	Contiene un fragmento que codifica para los primeros 50 aminoácidos del extremo N-terminal de Bcy1 entre los sitios NcoI y SacI.	Este trabajo

Medios de cultivo y condiciones de crecimiento

Los medios de cultivo de *S. cerevisiae* fueron preparados según se ha descrito [22]. Los cultivos de *E. coli* fueron crecidos en medio Luria-Bertani (LB). La composición de los mismos se muestra en la **Tabla 2.S.3**. Los medios sólidos utilizados para crecer las cepas tanto de bacterias como de levaduras a partir de sus stocks almacenados a -80° C fueron preparados con la misma composición, con un agregado de 2% de ágar.

Para sobreexpresar en levaduras de las proteínas Bcy1 WT y Bcy1 Δ(1-85) se utilizaron las cepas de *S. cerevisiae* 1115-BCY1 y 1115-Δ85BCY1, que contienen los plásmidos YEp51-BCY1 y YEp51-Δ85BCY1 respectivamente. Típicamente, para generar masa celular se crecieron 100 mL de precultivo en medio SD sin leucina para mantener la presión de selección sobre los plásmidos. Luego de 16 horas de cultivo, las células fueron cosechadas, lavadas una vez con agua para remover el exceso de glucosa (que reprime el promotor), y usadas para inocular 1L de medio YPGal. Este cultivo se creció 16 horas para lograr la sobreexpresión de la proteína de interés y las células fueron cosechadas, lavadas una vez con agua para remover el remanente de medio de cultivo, congeladas y almacenadas hasta el momento de ser usadas para la purificación de proteínas.

Tabla 2.S.3: Medios de cultivo utilizados en este trabajo.

Medio	Organismo	Composición	Utilidad
Medio mínimo sintético (SD)	<i>S. cerevisiae</i>	0,67% base nitrogenada sin aminoácidos para levaduras, 2% glucosa, aminoácidos necesarios para completar los requerimientos auxotróficos y mantener la presión de selección sobre los plásmidos.	Mantenimiento de los plásmidos en los precultivos.
Medio rico (YPGal)	<i>S. cerevisiae</i>	2% bactopectona, 2% extracto de levadura, 2% galactosa.	Inducción del promotor GAL10.
Luria-Bertani (LB)	<i>E. coli</i>	1% bactopectona, 0,5% extracto de levadura, 0,5% cloruro de sodio.	Mantenimiento de los plásmidos en los precultivos (suplementado con antibióticos) o bien inducción del sistema de expresión pET (suplementado con IPTG).

Para sobreexpresar la proteína recombinante Bcy1 (1-50) se utilizó la cepa de *E. coli* BL21-CodonPlus (DE3)-RIPL conteniendo el plásmido pET32a-Bcy1(1-50). Para producir aproximadamente 1 mg de Bcy1 (1-50) pura, se crecieron ON 50 mL de un precultivo en

medio LB suplementado con 100 µg/mL de ampicilina y 34 µg/mL de cloranfenicol para el mantenimiento de los plásmidos. Las células fueron cosechadas y usadas para inocular 1 L de medio LB fresco, suplementado con los mismos antibióticos. Este nuevo cultivo fue crecido por 2 horas, y cuando se alcanzó una densidad óptica de 0,6/0,8 (fase exponencial), se indujo la sobreexpresión de la proteína recombinante por adición de IPTG 0,1 mM (Sigma). Luego de 4 horas de sobreexpresión, las células fueron cosechadas, lavadas una vez con agua para remover el remanente de medio de cultivo, congeladas y almacenadas hasta el momento de ser usadas para la purificación de proteínas.

Clonado de Bcy1 (1-50)

El plásmido pET32a-Bcy1(1-50) se construyó por inserción de un fragmento de DNA que codifica para los aminoácidos 1-50 de Bcy1 entre los sitios NcoI y SacI del vector comercial pET32a (Novagen). En la Figura 2.S.3 se muestra la región de clonado.

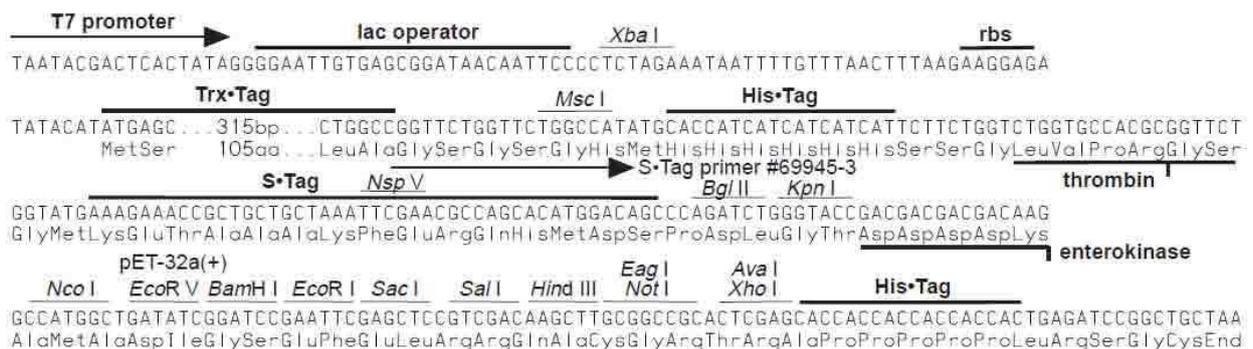


Figura 2.S.3: Sitio de múltiple clonado del plásmido pET32a.

Los oligonucleótidos:

forward (5'-GACGACAAGGCCATGGTATCTTCTTTGCCCAAG-3')

reverse (5'-CAAGCTTGTGACG-GTTACCTGGCCTTGAGGAACG-3')

fueron utilizados para amplificar por PCR la región deseada de BCY1, usando como templado el plásmido YEp51-BCY1. El oligonucleótido reverse introduce un codón stop artificial inmediatamente después del aminoácido 50, generando de esta manera una proteína de fusión que contiene, de N a C terminal, un tag de tiorredoxina, un tag de polihistidina, un tagS,

y un sitio de corte para enteroquinasa que permite remover el *tag*. El clonado del fragmento amplificado se realizó por incubación del vector previamente linealizado por digestión en simultáneo con las enzimas de restricción NcoI y SacI (Promega) con el fragmento utilizando el sistema de clonado InFusion (Clontech). La secuencia de la construcción fue verificada por secuenciación; la secuencia de la proteína de fusión generada se muestra en la **Figura 2.S.4**. El clonado se realizó en *E. coli* DH5 α utilizando el método de obtención de células competentes y transformación de Inoue [23]. Posteriormente, el vector recombinante fue transformado en células BL21-CodonPlus (DE3)-RIPL para posterior sobreexpresión de la proteína de fusión utilizando el método clásico de Hanahan [24].

```
MSDKIIHLTDDSFDTDLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQ  
NPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGMHMHHHHHSSGL  
VPRGSGMKETAALKFERQHMDSPDLGTDKAMVSSLPKESQAELQLFQNEINAANPSDFLQ  
FSANYFNKRLEQQRAFLKAR*
```

Figura 2.S.4: Secuencia de la proteína de fusión.

Código de colores: *tag* de tiorredoxina (109 aminoácidos); *tag* de polihistidina (6 aminoácidos); *tagS* (15 aminoácidos); sitio de corte para enteroquinasa (5 aminoácidos); Bcy1 (1-50) (50 aminoácidos).

Purificación de proteínas

Bcy1 WT y Bcy1 Δ (1-85)

Para producir estas proteínas se utilizó una versión modificada de un procedimiento reportado previamente [10]. Las cepas de *S.cerevisiae* 1115-BCY1 y 1115- Δ 85BCY1 crecidas en YPGal, que sobreexpresan cada una de las proteínas de interés, fueron utilizadas como punto de partida. Las células fueron resuspendidas en buffer de lisis (20 mM fosfato, 100 mM NaCl, 10 mM DTT, 5 mM EDTA, pH = 7,4) suplementado con cóctel de inhibidores de proteasas (Roche), y se obtuvieron extractos libres de células por ruptura de esta suspensión con perlas de vidrio en tubos de microcentrífuga de 2 mL, utilizando un vórtex. Los restos celulares fueron eliminados por centrifugación y los extractos clarificados fueron incubados en *batch* con una resina de afinidad N6-cAMP-agarosa (BioLog) previamente equilibrada con

buffer de lisis, ON a 4° C (para minimizar la proteólisis) con agitación leve. Esta resina retiene específicamente las proteínas con capacidad de unir cAMP. Luego de la incubación, la resina fue lavada exhaustivamente con buffer de lisis, y las variantes de Bcy1 fueron eluídas por incubación con buffer de elución (20 mM MES, 40 mM cAMP, pH = 6); el cAMP compite con los ligandos presentes en la resina por los sitios de unión a cAMP de Bcy1, despegándola. La pureza de las proteínas fue verificada por SDS-PAGE 10% teñido con Coomassie Blue; en la **Figura 2.S.5** se muestran los resultados de una purificación representativa.

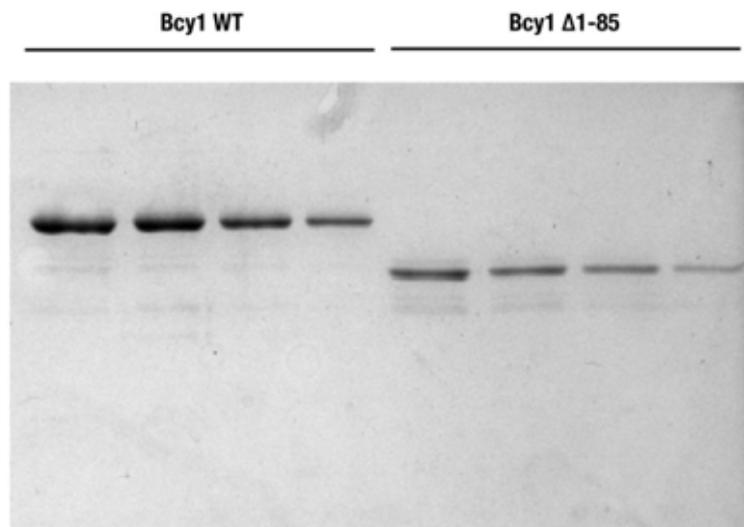


Figura 2.S.5: Purificación de Bcy1 WT y Bcy1 $\Delta(1-85)$.

SDS-PAGE 10% visualizado por tinción con Coomassie Blue de alícuotas de una purificación representativa de Bcy1 WT y Bcy1 $\Delta(1-85)$.

Las alícuotas de mejor calidad se juntaron y su concentración se estimó por comparación de las bandas del gel con la banda de un patrón de BSA de cantidad conocida. Cabe destacar que no es posible cuantificar por absorbancia a 280 nm con exactitud debido a que la proteína obtenida mediante este protocolo de purificación contiene cAMP unido y en la solución que la contiene, que absorbe a esa longitud de onda. Finalmente, se agregó glicerol hasta una concentración de 20% final a las muestras, que fueron alícuotadas, congeladas y almacenadas hasta su uso.

Bcy1 (1-50)

Como punto de partida se utilizó la cepa de *E. coli* BL21-CodonPlus (DE3)-RIPL-pET32a-Bcy1(1-50), que sobreexpresa la proteína de interés. Las células fueron resuspendidas en buffer de lisis (50 mM HEPES, 100 mM NaCl, 50 mM imidazol, pH = 7) suplementado con cóctel de inhibidores de proteasas (Roche); el imidazol cumple la función de minimizar el pegado inespecífico de proteínas indeseadas a la resina de Ni²⁺ (ver más adelante). Se obtuvieron extractos libres de células por ruptura de esta suspensión con perlas de vidrio en tubos de microcentrífuga de 2 mL, utilizando un vórtex, o bien por sonicación. Los restos celulares fueron eliminados por centrifugación para obtener los extractos clarificados.

El primer paso de la purificación consistió en una incubación en *batch* del extracto clarificado con una resina de afinidad de Ni²⁺ HisLink (Promega) previamente equilibrada con buffer de lisis, durante 1 hora a 20° C con agitación leve. Esta resina retiene específicamente las proteínas con capacidad de complejar Ni²⁺, en particular, la proteína de fusión de interés, a través de su *tag* de polihistidina. Luego de la incubación, la resina fue lavada exhaustivamente con buffer de lisis, y se le realizó un último lavado con buffer de enteroquinasa (20 mM Tris-HCl, 50 mM NaCl, 2 mM CaCl₂, pH = 7,4) para equilibrarla para el próximo paso.

El segundo paso de la purificación consistió en la remoción del *tag*, lo cual se logró por tratamiento de la proteína de fusión inmovilizada con enteroquinasa (Novagen). Típicamente, aproximadamente 5 mg de proteína de fusión contenidos en 250 µL de resina decantada con agregado de 1 volumen de buffer de enteroquinasa, se incubaron con 5 U de la enzima durante 16 horas a 24° C con agitación suave. De esta manera se logró la liberación de Bcy1 (1-50) parcialmente purificada de la resina, con un residuo de alanina adicional en su extremo N-terminal como consecuencia de la construcción utilizada. Por otro lado, el *tag* permanece en su mayor parte unido a la resina; sin embargo, algunos fragmentos de proteólisis se liberan al sobrenadante, y por lo tanto fueron requeridos pasos adicionales de purificación.

El tercer paso consistió en una SEC del sobrenadante del paso anterior, utilizando una columna Superdex-75 GL 10/300 (GE Healthcare), de 25 mL de capacidad, acoplada a un HPLC Knauer Smartline Pump 1050 con un detector UV-Vis 2520. La columna fue equilibrada con buffer PBS (20 mM fosfato, 150 mM NaCl, pH = 7,4) y alícuotas de 500 µL de proteína parcialmente purificada fueron inyectadas en el equipo utilizando un *loop* apropiado. La cromatografía fue desarrollada a 20° C con un flujo de 1,0 mL/min en condiciones isocráticas de solvente, y la absorbancia se monitoreó a 280 nm. Las fracciones correspondientes a Bcy1

(1-50) fueron colectadas, juntadas y concentradas usando concentradores Vivaspín 500 (GE Healthcare).

Una pequeña cantidad de proteína correspondiente a restos del *tag* eluye en el mismo pico que Bcy1 (1-50). Por lo tanto, un cuarto y último paso de purificación fue requerido para eliminar estos restos. Aprovechando las diferencias de punto isoeléctrico entre Bcy1 (1-50) y estos fragmentos contaminantes (la identidad de éstos fue estudiada por MS), se realizó una incubación en *batch* con una resina de intercambio aniónico DEAE-celulosa (Sigma) previamente equilibrada con HEPES 50 mM, pH = 7. Luego de 30 minutos de incubación a 20°C con agitación suave, el sobrenadante conteniendo Bcy1 (1-50) purificada fue recuperado; los contaminantes quedan retenidos en la resina. La pureza de la proteína fue evaluada por SDS-PAGE 15% en buffer Tricina con tinción por plata (ver **Figura 2.7**). Las alícuotas fueron juntadas y cuantificadas por absorbancia a 280 nm en un espectrofotómetro NanoDrop utilizando el coeficiente de absorción molar calculado de 1490 M⁻¹cm⁻¹, y finalmente alicuotadas y congeladas hasta su uso. La identidad de Bcy1 (1-50) fue verificada por MS-MS, como se muestra en la **Figura 2.S.6**.

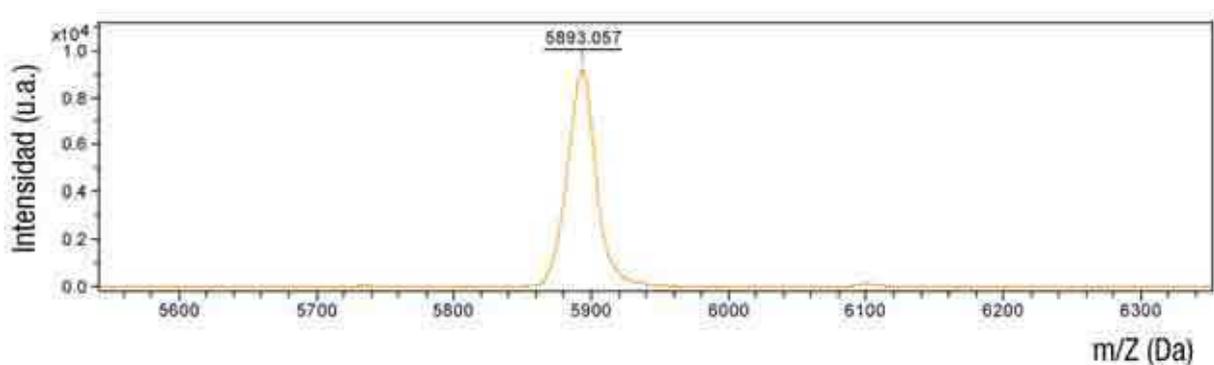


Figura 2.S.6: Espectro de masa de Bcy1 (1-50) purificado.

Se indica el valor de m/z del pico correspondiente al ion molecular [M+H]⁺. La MM teórica esperada para el fragmento es de 5890,64 Da.

Electroforesis

Las electroforesis en geles de poliacrilamida en condiciones desnaturizantes se desarrollaron según el método clásico de Laemmli [25] para el caso de las proteínas Bcy1 WT y Bcy1 $\Delta(1-85)$, y utilizando el sistema de buffers a base de Tricina [26] para Bcy1 (1-50), dado que este permite obtener mejor resolución para proteínas de muy baja MM. Previo a la

siembra, las muestras se diluyeron 1:4 en buffer de siembra 4X (Tris-HCl 0,2 M, 8% SDS, 40% glicerol, 0,4 M β -mercaptoetanol, azul de bromofenol, pH = 6,8) y se hirvieron durante 2 minutos para desnaturalizar las proteínas.

Para las tinciones con Coomassie Blue, los geles se incubaron durante 30 minutos con solución de fijación/tinción (0,25% Coomassie brilliant blue R-250, 45% metanol, 10% ácido acético) y posteriormente se eliminaron los restos de colorante mediante sucesivos lavados durante varias horas con solución desteñidora (30% metanol, 7% ácido acético).

Para las tinciones con Coomassie Blue coloidal para MS, se trabajó con soluciones cuidadosamente preparadas para no contaminar las muestras con queratina humana, lo que representa un inconveniente típico en este tipo de análisis. En primer lugar, los geles se incubaron 3 veces durante 30 minutos con solución de fijación (30% etanol, 2% ácido fosfórico). Posteriormente, se realizaron 3 lavados de 20 minutos con solución de ácido fosfórico 2% para eliminar el exceso de etanol, seguido de una incubación de 30 minutos con solución para preparar el gel para la tinción (18% etanol, 2% ácido fosfórico, 15% sulfato de amonio). Finalmente, se agregó sobre la última solución 1mL de Coomassie brilliant blue G-250 2% y se incubó durante 24 horas. Los restos de colorante se eliminaron mediante sucesivos lavados con agua.

Para las tinciones con plata, se incubaron los geles con solución de fijación (40% etanol, 10% ácido acético) durante 1 hora. Posteriormente, se realizó una incubación de 30 minutos con solución de sensibilización (30% etanol, 6,8% acetato de sodio, 0,2% tiosulfato de sodio). Luego de tres lavados de 5 minutos con agua para remover el exceso de la última solución, se realizó una incubación de 20 minutos con solución de tinción (0,25% nitrato de plata). Se realizaron dos lavados de 1 minuto con agua para remover el exceso de plata, y se incubó entre 3 y 5 minutos con solución de revelado (2,5% carbonato de sodio, 0,04% formaldehído), teniendo cuidado de no exceder el tiempo de incubación para evitar exceso de tinción. Finalmente, se incubó durante 10 minutos con solución de stopping (0,04 M EDTA) y se realizaron varios lavados breves con agua.

Experimentos de entrecruzamiento químico con EGS

Unos 20 μ g de Bcy1 WT, Bcy1 Δ (1-85), o Bcy1 (1-50), fueron diluidas en buffer de reacción (HEPES 20 mM, 0,2 M KCl, 10 mM DTT, y 1mM EDTA, pH = 7,5). Una solución concentrada de EGS (Sigma) en DMSO (el DMSO cumple el rol de cosolvente, dado que el EGS

es poco soluble en solución acuosa), se agregó a la muestra hasta una concentración final de EGS de 0,5-5,0 mM, y una concentración de DMSO no superior al 10%. La reacción procedió durante 30 minutos a 20° C en un volumen final de 1 mL, y fue detenida por agregado de Tris-HCl pH = 9 hasta una concentración final de 50 mM (el Tris al ser una amina reacciona con el EGS, inactivándolo). Los productos de reacción fueron precipitados con ácido tricloroacético, los pellets lavados exhaustivamente con acetona fría para remover el exceso de ácido, secados, y analizados por SDS-PAGE.

Espectrometría de masa

El análisis por espectrometría de masa fue realizado por el servicio de proteómica CEQUIBIEM (FCEN-UBA), dirigido por la Dra. Silvia Moreno.

Para el análisis de fragmentos trípticos de las variantes monómero y dímero entrecruzado por EGS de Bcy1 WT, las bandas fueron cortadas del gel y lavadas para eliminar el colorante. Posteriormente, las proteínas contenidas en la banda se redujeron con DTT y se alquilaron con iodoacetamida para bloquear los residuos de cisteína. Finalmente, las proteínas se digirieron con tripsina *in gel*. El líquido sobrenadante de la incubación, que contiene los fragmentos trípticos, fue analizado en un espectrómetro de masa MALDI-TOF-TOF Ultraflex II (Bruker), utilizando como matriz ácido α -ciano-4-hidroxicinámico.

La lista de picos correspondiente a los fragmentos trípticos de cada una de las muestras fue utilizada como *input* para el programa MASCOT, que permite contrastar el patrón de fragmentos obtenidos con la base de datos de todos los fragmentos trípticos posibles para las proteínas de cualquier organismo de interés, y de esta manera verificar la identidad de la proteína (y de los fragmentos). El análisis informático de los fragmentos trípticos confirmó que todas las muestras corresponden a Bcy1, como era esperable. El programa permite además asignar a cada pico el fragmento que le dio origen en base a su valor de m/z. La identidad del pico que resultó ser de interés para este trabajo (m/z = 2633 Da) se verificó por MS/MS y secuenciación *de novo*, confirmando que corresponde al fragmento 58-80 de Bcy1 (NIVLFPEPEESFSRPQSAQSQR).

El análisis de la muestra líquida purificada de Bcy1 (1-50) para verificar su MM se llevó a cabo sembrando directamente la muestra en el equipo utilizando como matriz ácido sinapínico, que es la matriz apropiada para observar péptidos de alta MM.

REFERENCIAS

1. Hixson CS, Krebs EG: **Characterization of a cyclic AMP-binding protein from bakers' yeast. Identification as a regulatory subunit of cyclic AMP-dependent protein kinase.** *J Biol Chem* 1980, **255**(5):2137-2145.
2. Johnson KE, Cameron S, Toda T, Wigler M, Zoller MJ: **Expression in Escherichia coli of BCY1, the regulatory subunit of cyclic AMP-dependent protein kinase from Saccharomyces cerevisiae. Purification and characterization.** *J Biol Chem* 1987, **262**(18):8636-8642.
3. Griffioen G, Anghileri P, Imre E, Baroni MD, Ruis H: **Nutritional control of nucleocytoplasmic localization of cAMP-dependent protein kinase catalytic and regulatory subunits in Saccharomyces cerevisiae.** *J Biol Chem* 2000, **275**(2):1449-1456.
4. Banky P, Roy M, Newlon MG, Morikis D, Haste NM, Taylor SS, Jennings PA: **Related protein-protein interaction modules present drastically different surface topographies despite a conserved helical platform.** *J Mol Biol* 2003, **330**(5):1117-1129.
5. Newlon MG, Roy M, Morikis D, Hausken ZE, Coghlan V, Scott JD, Jennings PA: **The molecular basis for protein kinase A anchoring revealed by solution NMR.** *Nat Struct Biol* 1999, **6**(3):222-227.
6. Banky P, Huang LJ, Taylor SS: **Dimerization/docking domain of the type Ialpha regulatory subunit of cAMP-dependent protein kinase. Requirements for dimerization and docking are distinct but overlapping.** *J Biol Chem* 1998, **273**(52):35048-35055.
7. Li Y, Rubin CS: **Mutagenesis of the regulatory subunit (RII beta) of cAMP-dependent protein kinase II beta reveals hydrophobic amino acids that are essential for RII beta dimerization and/or anchoring RII beta to the cytoskeleton.** *J Biol Chem* 1995, **270**(4):1935-1944.
8. Gold MG, Lygren B, Dokurno P, Hoshi N, McConnachie G, Tasken K, Carlson CR, Scott JD, Barford D: **Molecular basis of AKAP specificity for PKA regulatory subunits.** *Mol Cell* 2006, **24**(3):383-395.
9. Gautier R, Douguet D, Antony B, Drin G: **HELIQUEST: a web server to screen sequences with specific alpha-helical properties.** *Bioinformatics* 2008, **24**(18):2101-2102.
10. Galello F, Moreno S, Rossi S: **Interacting proteins of protein kinase A regulatory subunit in Saccharomyces cerevisiae.** *J Proteomics* 2014, **109C**:261-275.
11. Stengel F, Aebersold R, Robinson CV: **Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes.** *Mol Cell Proteomics* 2012, **11**(3):R111 014027.
12. Taylor SS, Ilouz R, Zhang P, Kornev AP: **Assembly of allosteric macromolecular switches: lessons from PKA.** *Nat Rev Mol Cell Biol* 2012, **13**(10):646-658.
13. Banky P, Newlon MG, Roy M, Garrod S, Taylor SS, Jennings PA: **Isoform-specific differences between the type Ialpha and IIalpha cyclic AMP-dependent protein kinase anchoring domains revealed by solution NMR.** *J Biol Chem* 2000, **275**(45):35146-35152.
14. Newlon MG, Roy M, Hausken ZE, Scott JD, Jennings PA: **The A-kinase anchoring domain of type IIalpha cAMP-dependent protein kinase is highly helical.** *J Biol Chem* 1997, **272**(38):23637-23644.
15. Madhusudhan MS, Marti-Renom MA, Sanchez R, Sali A: **Variable gap penalty for protein sequence-structure alignment.** *Protein Eng Des Sel* 2006, **19**(3):129-133.
16. McGuffin LJ, Bryson K, Jones DT: **The PSIPRED protein structure prediction server.** *Bioinformatics* 2000, **16**(4):404-405.
17. Cheng J, Randall AZ, Sweredoski MJ, Baldi P: **SCRATCH: a protein structure and structural feature prediction server.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W72-76.
18. Cole C, Barber JD, Barton GJ: **The Jpred 3 secondary structure prediction server.** *Nucleic Acids Res* 2008, **36**(Web Server issue):W197-201.
19. Yachdav G, Kloppmann E, Kajan L, Hecht M, Goldberg T, Hamp T, Honigschmid P, Schafferhans A, Roos M, Bernhofer M *et al*: **PredictProtein--an open resource for online prediction of protein structural and functional features.** *Nucleic Acids Res* 2014, **42**(Web Server issue):W337-343.
20. Eisenberg D, Weiss RM, Terwilliger TC: **The helical hydrophobic moment: a measure of the amphiphilicity of a helix.** *Nature* 1982, **299**(5881):371-374.
21. Portela P, Zarembeg V, Moreno S: **Evaluation of in vivo activation of protein kinase A under non-dissociable conditions through the overexpression of wild-type and mutant regulatory subunits in Saccharomyces cerevisiae.** *Microbiology* 2001, **147**(Pt 5):1149-1159.

22. Sherman F, Fink GR, Hicks JB: **Laboratory Course Manual for Methods in Yeast Genetics.** *Cold Spring Harbor Laboratory* 1986.
23. Inoue H, Nojima H, Okayama H: **High efficiency transformation of Escherichia coli with plasmids.** *Gene* 1990, **96**(1):23-28.
24. Hanahan D: **Studies on transformation of Escherichia coli with plasmids.** *J Mol Biol* 1983, **166**(4):557-580.
25. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, **227**(5259):680-685.
26. Schagger H: **Tricine-SDS-PAGE.** *Nat Protoc* 2006, **1**(1):16-22.

CAPÍTULO 3

ESTRUCTURA EN SOLUCIÓN DEL DOMINIO D/D DE BCY1

“ Medrano es quizá el único que lo sabe: somos triviales y lo pagamos con felicidad o con desgracia, la felicidad de la marmota envuelta en grasa, la sigilosa desgracia de Raúl Costa que aprieta contra su pijama negro un cisne de ceniza, y hasta cuando nacemos para preguntar y otear las respuestas, algo infinitamente desconcertante que hay en la levadura del pan argentino (...)”

Julio Cortázar, “Los Premios”

ANTECEDENTES

1) La región N-terminal de Bcy1 contiene características similares a los dominios D/D de mamíferos en cuanto a secuencia y estructura, lo que permite inferir que se encuentra involucrada en el establecimiento de dímeros en solución (resultados del Capítulo 2).

2) Inesperadamente, cuando se evaluó la capacidad para establecer oligómeros del fragmento recombinante Bcy1 (1-50) (que representa el dominio D/D putativo de Bcy1), se detectaron especies de dos, tres, y cuatro subunidades, sugiriendo que podrían existir notables diferencias entre las estructuras cuaternarias de las proteínas de mamíferos y levaduras. Esta variedad de estados oligoméricos podría atribuirse a sutiles diferencias de secuencia y estructura localizadas hacia el extremo C-terminal de los dominios D/D de mamíferos y levaduras. (resultados del Capítulo 2).

OBJETIVOS Y ABORDAJES UTILIZADOS

En base a los antecedentes, se procedió a estudiar la estructura en solución del fragmento recombinante Bcy1 (1-50) utilizando una combinación de métodos biofísicos, a efectos de elucidar de manera inequívoca su estructura cuaternaria en solución, y obtener información acerca de los determinantes que la sostienen. Para ello, se plantearon los siguientes objetivos para este capítulo:

1) Determinar la cantidad de subunidades que conforman la estructura del fragmento Bcy1 (1-50) en su estado nativo en solución, y comparar con el estado oligomérico de Bcy1 WT en condiciones similares, para descartar artificios experimentales. Para ello se utilizaron técnicas como cromatografía de exclusión molecular y dispersión estática de la luz.

2) Obtener información acerca de la estructura secundaria de Bcy1 (1-50) en solución, utilizando espectropolarimetría de dicroísmo circular.

3) Obtener parámetros hidrodinámicos que proporcionan información acerca del tamaño y la forma global de la molécula en solución, es decir, acerca de su estructura terciaria con

baja resolución. Para ello se utilizaron técnicas como cromatografía de exclusión molecular, ultracentrifugación analítica en gradiente de densidad, y dispersión de rayos X a bajo ángulo.

4) Construir un modelo consistente con la información experimental recopilada en los objetivos anteriores utilizando herramientas de modelado molecular, para completar la caracterización estructural de baja resolución, y obtener de esta manera algún indicio acerca de cuáles podrían ser las características distintivas de Bcy1 (1-50) que expliquen eventuales diferencias con las proteínas de mamíferos en lo que respecta al estado oligomérico.

INTRODUCCIÓN METODOLÓGICA

En esta sección se describe el marco teórico y ciertos aspectos prácticos de algunas técnicas relevantes para el desarrollo de los resultados de este capítulo.

Dispersión estática de la luz (SLS)

La dispersión estática de la luz es actualmente la técnica más confiable para determinar la MM y, consecuentemente, el estado oligomérico de macromoléculas en solución [1]. Típicamente, los detectores que se utilizan para registrar las señales experimentales que se requieren para este análisis se encuentran acoplados a la salida de una columna de exclusión molecular, por lo que comúnmente se la denomina SEC-SLS.

La técnica se basa en que la intensidad de la luz dispersada a 90° del flujo de la muestra (LS) es directamente proporcional a la MM y a la concentración de la proteína en solución, y completamente independiente de su forma y de sus dimensiones moleculares:

$$LS = K_{LS} \cdot C \cdot MM \cdot \left(\frac{dn}{dC} \right)^2 \quad (\text{Ecuación 3.1})$$

donde K_{LS} es una constante de calibración del instrumento, C es la concentración de la muestra, y dn/dC representa la variación del índice de refracción de la muestra con la concentración. Además del detector de LS, se utiliza un detector de índice de refracción (RI) conectado en tándem al primero, cuya señal puede escribirse como:

$$RI = K_{RI} \cdot C \cdot \left(\frac{dn}{dC} \right) \quad (\text{Ecuación 3.2})$$

donde K_{RI} es otra constante de calibración del instrumento. Para proteínas que no contienen carbohidratos, la variación del índice de refracción del soluto con la concentración es una constante, $dn/dC = 0,186 \text{ mL/g}$. De esta manera, reemplazando este valor y combinando la **Ecuación 3.1** con la **Ecuación 3.2**, puede determinarse directamente la MM a partir del cociente de la señal de los dos detectores:

$$MM = K' \cdot \left(\frac{LS}{RI} \right) \quad (\text{Ecuación 3.3})$$

donde K' es un cociente de constantes. Este es el llamado “método de dos detectores”, que es el más comúnmente utilizado [1]. Para determinar K' , se calibra el equipo utilizando una proteína de MM conocida, registrando sus señales de LS y RI.

Espectropolarimetría de dicroísmo circular (CD)

Esta técnica permite evaluar el contenido de estructura secundaria de proteínas en solución, tanto cualitativamente (es decir, en lo que respecta a la identificación de los tipos de estructura secundaria que la conforman) como cuantitativamente (o sea, en cuanto a la determinación del porcentaje de aminoácidos de la molécula que se encuentran involucrados en cada uno de esos elementos) [2].

La técnica se basa en la absorción diferencial de la radiación UV circularmente polarizada a izquierda y a derecha por parte de los grupos cromóforos presentes en las proteínas, como consecuencia de su ubicación en entornos moleculares asimétricos debido a la estructura tridimensional de la molécula. El experimento más básico de CD consiste en hacer incidir sobre la muestra un haz de radiación linealmente polarizada, que puede interpretarse como la suma de dos componentes circularmente polarizadas, una a la izquierda y la otra a la derecha, de igual frecuencia y que se encuentran en fase. Como consecuencia de la absorción diferencial de estos dos componentes por parte de la muestra, la luz transmitida resulta ser elípticamente polarizada. La magnitud que representa este fenómeno es la elipticidad (θ), que es una medida de la diferencia entre los coeficientes de absorción molar de los dos componentes de la luz linealmente polarizada (ϵ_L , ϵ_R). La elipticidad expresada en grados es la inversa de la tangente de b/a , donde b y a son los ejes menor y mayor de la elipse resultante. [2].

En proteínas, el principal cromóforo que contribuye a la absorción en el UV lejano (que es la región del espectro electromagnético que contiene información acerca de la estructura secundaria) es el enlace peptídico (que absorbe por debajo de 240 nm). En esta región existen dos transiciones características: una transición $n-\pi^*$ centrada en 220 nm, y una más intensa $\pi-\pi^*$ centrada en 190 nm. Cuando una cantidad de cromóforos del mismo tipo se encuentran próximos en el espacio y en entornos moleculares similares (es decir, formando parte del mismo elemento de estructura secundaria), dan lugar a bandas de CD características que son indicativas del tipo de estructura secundaria del cual forman parte. De esta manera, los

diferentes tipos de estructura secundaria son identificables de forma cualitativa por esta técnica, por presentar espectros de CD muy característicos. Por otro lado, es posible determinar el porcentaje del total de la molécula que se encuentra formando parte de cada uno de los distintos elementos mediante un análisis de deconvolución, utilizando como referencia bases de datos de proteínas con contenidos de estructura secundaria bien caracterizados. [2].

Determinación de parámetros hidrodinámicos

En este capítulo se presenta la determinación experimental de cinco propiedades que proporcionan información acerca de las dimensiones moleculares y de la forma global de una macromolécula en solución: el radio de Stokes (R_s), el coeficiente de sedimentación en agua a 20° C ($S_{20,W}$), el coeficiente friccional (f), el radio de giro (R_g) y la máxima distancia lineal de la partícula (D_{max}). Los métodos que permiten la determinación de las tres primeras se presentan en esta sección, mientras que los que corresponden a las dos últimas se introducen más adelante, dado que la técnica involucrada requiere de un trasfondo teórico más profundo para su comprensión.

El R_s se define como el radio de una esfera que tiene las mismas propiedades hidrodinámicas (en particular, el mismo coeficiente de difusión traslacional, D , y el mismo coeficiente friccional, f) que la partícula [3]. El R_s y el f están relacionados por la ecuación de Stokes-Einstein:

$$f = 6. \pi. \eta. R_s \quad (\text{Ecuación 3.4})$$

donde f es el coeficiente friccional de la partícula, que es una medida de su resistencia al movimiento traslacional, y η es la viscosidad del solvente. Para una determinada masa de proteína, o para un volumen determinado, el f aumenta en la medida en que la molécula se hace más elongada o asimétrica; se relaciona con el coeficiente de difusión traslacional (D) mediante:

$$D = k. T / f \quad (\text{Ecuación 3.5})$$

donde k es la constante de Boltzmann y T es la temperatura absoluta. Tanto el R_s como el f dependen del tamaño y de la forma de la molécula: las proteínas de mayor MM, más grandes y más elongadas o asimétricas tienen R_s y f mayores, mientras que las proteínas de menor MM, más pequeñas, compactas y esféricas tienen R_s y f menores.

Una magnitud que a menudo resulta de utilidad es el radio de la mínima esfera no hidratada que contiene una determinada masa de proteína (R_{min}) [3]. Este valor puede calcularse a partir de la fórmula para el volumen de una esfera, la MM de la proteína, y su volumen parcial específico (v). Considerando para este último un valor de $0,73 \text{ cm}^3/\text{g}$ (que es un valor promedio utilizado comúnmente para proteínas), el R_{min} depende solamente de la MM, y puede calcularse como:

$$R_{min} = 0,066 \cdot (MM)^{1/3} \text{ nm} / D\alpha^{1/3} \quad (\text{Ecuación 3.6})$$

Se define análogamente el f_{min} , que es el mínimo coeficiente friccional que puede tener una proteína de determinada MM, y que corresponde a una molécula perfectamente esférica y no hidratada, cuyo radio es entonces R_{min} . De esta manera, R_{min} y f_{min} se encuentran relacionadas mediante la **Ecuación 3.4**.

Por otro lado, el coeficiente de sedimentación a 20°C en agua ($S_{20,w}$) es una medida de la velocidad con que la partícula migra por centrifugación a través de un gradiente de densidad [3]. Al igual que el R_s , depende principalmente de tres factores: MM, tamaño y forma. Las proteínas de menor MM y mayor tamaño y/o asimetría sedimentan con menor velocidad y por lo tanto tienen $S_{20,w}$ menores, mientras que las proteínas de mayor MM y más compactas (menor tamaño y mayor globularidad) sedimentan con mayor velocidad, y en consecuencia tienen $S_{20,w}$ mayores. De esta manera, contrariamente a lo que ocurre con el R_s , el $S_{20,w}$ guarda relación inversa con el f , como lo indica la fórmula de Svedberg:

$$S_{20,w} = \frac{MM \cdot (1 - v \cdot \rho)}{N_A \cdot f} \quad (\text{Ecuación 3.7})$$

donde ρ es la densidad del solvente (1,0 g/cm³ para el agua), y N_A es el número de Avogadro. Además de los factores mencionados inherentes a la estructura de la proteína, el $S_{20,W}$ depende de su concentración.

Por analogía con el R_{min} , puede definirse el máximo $S_{20,W}$ que puede tener una proteína de determinada MM (S_{max}), que corresponde a una molécula perfectamente esférica y no hidratada (o sea, a una partícula de $R_s = R_{min}$ y $f = f_{min}$) [3]. Este valor puede calcularse reemplazando la **Ecuación 3.4** para R_{min} y f_{min} , y la **Ecuación 3.6**, en la **Ecuación 3.7**, y suponiendo los valores para v y ρ ya mencionados. De esta manera, se obtiene:

$$S_{max} = 0,00361.(MM)^{2/3} S/Da^{2/3} \quad (\text{Ecuación 3.8})$$

donde 1 S = 10⁻¹³ segundos (Svedberg) que es la unidad típicamente empleada para los coeficientes de sedimentación.

Los valores de estas tres propiedades (R_s , $S_{20,W}$, y f) no resultan particularmente informativos en sí mismos. En consecuencia, se define el cociente friccional, f/f_{min} , que resulta de utilidad dado que su valor es indicativo de si la molécula es globular, ligeramente elongada, o muy elongada [3]. Utilizando la **Ecuación 3.4** para R_s , f y R_{min} , f_{min} , y la **Ecuación 3.7** para $S_{20,W}$, f y S_{max} , f_{min} , puede demostrarse que:

$$f/f_{min} = R_s/R_{min} = S_{max}/S_{20,W} \quad (\text{Ecuación 3.9})$$

Resulta evidente que siempre será $f > f_{min}$, dado que la forma de cualquier proteína jamás resulta perfectamente esférica, sino más bien ligeramente elipsoidal o alargada. Por otro lado la superficie jamás es suave sino rugosa, lo que confiere resistencia adicional al movimiento traslacional. Finalmente, las moléculas siempre se encuentran hidratadas por una capa de una o dos moléculas de agua que se encuentran virtualmente inmovilizadas en la superficie; este agua de hidratación aumenta el radio efectivo y por lo tanto aumenta f [3].

Se han tabulado valores para el cociente friccional de muchas proteínas de estructura conocida. Como regla general, se ha observado que para proteínas razonablemente

globulares, $f/f_{min} \sim 1,2/1,3$. Para proteínas moderadamente elongadas se encuentra en el rango de 1,5/1,9, mientras que para proteínas muy elongadas (por ejemplo, la tropomiosina) el valor puede encontrarse entre 2,0 y 3,0 [3]. De esta manera, la determinación experimental tanto del R_s como del $S_{20,w}$ permite obtener información acerca de la forma de la molécula en solución. Más aun, como se describirá en la siguiente sección, la combinación de los dos parámetros permite obtener una estimación de la MM de la proteína en solución.

La estimación del R_s de una partícula en solución puede realizarse mediante dos técnicas: dispersión dinámica de la luz (DLS) o SEC. Por otro lado, el $S_{20,w}$ se estima típicamente mediante ultracentrifugación analítica en un gradiente de densidad, construido con soluciones de sacarosa o de glicerol de distintas concentraciones.

Determinación de la MM en solución utilizando el R_s y el $S_{20,w}$

En el año 1966, Siegel y Monty desarrollaron un método para determinar la MM de una proteína en solución que aun en la actualidad es muy utilizado por su simplicidad [4]. Este método se basa en la idea de que, a diferencia de lo que ocurre con el R_s y el $S_{20,w}$ individualmente (que son dependientes del tamaño y de la forma de la molécula, además de su MM), su combinación posibilita independizarse de los dos primeros factores. Combinando la **Ecuación 3.4** con la **Ecuación 3.7**, y suponiendo los valores típicos para la densidad del agua y el volumen parcial específico de la proteína, se obtiene:

$$MM = 4,205 \cdot S \cdot R_s \cdot \frac{Da}{S \text{ nm}} \quad (\text{Ecuación 3.10})$$

Dispersión de Rayos X a bajo ángulo (SAXS)

Esta técnica resulta de gran utilidad para la caracterización estructural con baja resolución de macromoléculas en solución. En primer lugar, permite obtener parámetros relacionados con las dimensiones moleculares (el radio de giro $-R_g-$ y la máxima distancia lineal de la partícula $-D_{max}-$); en segundo lugar, proporciona información acerca de la forma global, ya que permite obtener un histograma de distribución de distancias intramoleculares de la partícula que es representativo de su forma, y, en ocasiones, modelos de su contorno superficial. Adicionalmente, puede contribuir a la determinación del estado oligomérico de proteínas en solución. Finalmente, resulta muy útil para comparar estructuras de alta resolución obtenidas por cristalografía de rayos X con las estructuras nativas, por lo que las

dos funcionan como técnicas complementarias. En contraste con la cristalografía, como la técnica se aplica sobre muestras en solución, no requiere de cristalización y la adquisición y el análisis de los datos es relativamente sencillo, directo, y rápido. [5, 6].

El fenómeno se basa en la dispersión de fotones correspondiente a la región de los Rayos X del espectro electromagnético por parte de los electrones de la muestra: cuando un haz monocromático incide sobre ésta, los electrones de sus átomos se vuelven a su vez fuentes de ondas secundarias, que son colectadas en un detector 2D de la misma manera que ocurre en cristalografía de Rayos X (**Figura 3.1**). El diseño experimental de las dos técnicas es, a grandes rasgos, similar; la diferencia radica en que, como consecuencia del estado de agregación de la muestra en las dos técnicas, la información que se obtiene es diferente. En cristalografía, las moléculas se encuentran posicionadas regularmente en el cristal, capturadas en una misma (o en pocas) conformaciones, y orientadas todas de la misma manera. De este modo, las ondas secundarias emitidas interfieren constructivamente en el detector para dar un patrón de puntos en 2D; este patrón se relaciona, mediante una transformación de Fourier, con el mapa de densidad electrónica 3D que permite construir el modelo molecular. En SAXS, por el contrario, las moléculas pueden moverse libremente dado que la muestra se encuentra en solución (existe, sin embargo, SAXS en estado sólido, pero no forma parte del análisis de rutina de biomoléculas); por lo tanto, se encuentran orientadas al azar y en un determinado instante cada una de ellas adopta una conformación distinta. Como consecuencia, no hay interferencia en el detector; el perfil de dispersión es isotrópico y la información de alta resolución se pierde, obteniéndose además valores promedio de las magnitudes de interés entre todas las posibles conformaciones y orientaciones que adoptan las distintas moléculas en el tiempo que dura el experimento. El análogo de SAXS del mapa de densidad electrónica es la función de distribución de distancias, $P(r)$, que se relaciona con el perfil de dispersión por transformación de Fourier (ver más adelante). [5, 6]

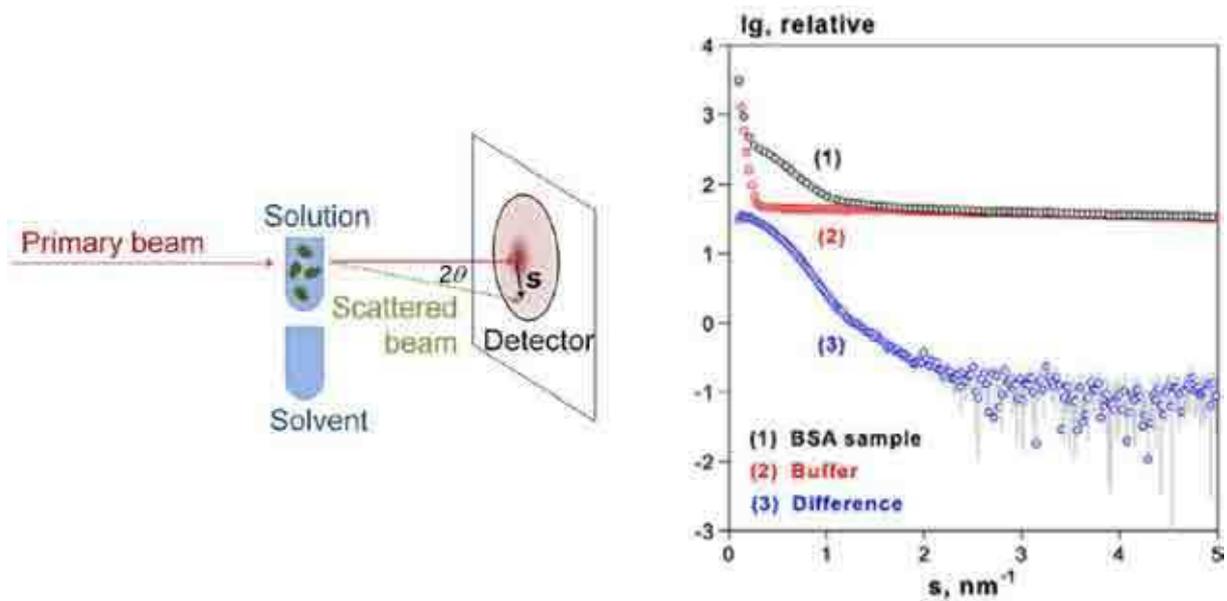


Figura 3.1: Diseño experimental de SAXS.

En el panel de la izquierda se muestra un diseño experimental típico de SAXS, en el que se hace incidir un haz monocromático de Rayos X (Primary beam) sobre la muestra o sobre el solvente, y se registra en un detector 2D el haz dispersado (Scattered beam) en función del vector de onda Q (“s” en la figura). Los datos se reducen a 1D en la forma de un gráfico de intensidad dispersada [I(Q)] en función del vector de onda Q (panel de la derecha). En el panel de la derecha se muestra también cómo se obtiene el perfil de dispersión de la macromolécula (azul) por sustracción de la señal del solvente (rojo) a la señal de la solución (negro). En este caso, la muestra corresponde a una muestra de albúmina sérica bovina (BSA). (La figura del panel de la izquierda fue tomada de [7] y la del panel de la derecha de [5]).

En un experimento típico de SAXS, se registran el perfil de dispersión de la solución que contiene la muestra y el del solvente que la contiene por separado. El patrón de dispersión de la macromolécula se obtiene por sustracción del segundo al primero (**Figura 3.1**). Dado que, como ya se señaló, el perfil de dispersión es isotrópico, solo una dirección del detector es informativa (la dirección radial). Como la segunda dirección contiene información redundante, el patrón de dispersión 2D se reduce a 1D integrando y promediando la señal a lo largo de toda la circunferencia del detector. Luego de la sustracción del blanco y la reducción, el resultado final de un experimento de SAXS es un gráfico 2D de intensidad dispersada [I(Q)] en función del vector de onda Q (**Figura 3.1**):

$$Q = 4. \pi. \text{sen } \theta / \lambda \quad (\text{Ecuación 3.11})$$

donde 2θ es el ángulo que forma el haz dispersado respecto de la normal, y λ es la longitud de onda del haz incidente.

Cuando se utiliza una fuente de Rayos X de laboratorio y no un sincrotrón (que es el caso de este trabajo) es necesario introducir una corrección adicional previo al análisis de los datos, que está relacionada con que la sección de la muestra iluminada no es puntual sino que tiene un determinado ancho (*slit-smearing effects*) [8]. Existen básicamente dos maneras de llevar a cabo esta corrección (ambas son procedimiento matemáticos): o bien se procesa el perfil de dispersión $[I(Q)]$ vs Q para obtener un perfil corregido, o bien la corrección puede formar parte del procedimiento para obtener la función de distribución de distancias, $P(r)$, a partir del perfil de dispersión (ver más adelante).

La señal resultante del tratamiento de datos expuesto anteriormente es proporcional al número de partículas en el volumen iluminado (es decir, a la concentración del soluto), a su MM, y al contraste entre la muestra y el solvente (que es la diferencia entre las densidades electrónicas de ambas). Dado que las macromoléculas biológicas están constituidas por átomos livianos, el contraste en SAXS de biomoléculas es muy bajo, y por lo tanto generalmente se requieren concentraciones elevadas para obtener resultados confiables (sobre todo en el caso de proteínas de baja MM, como es el caso de este trabajo), por lo que los problemas de agregación de la muestra pueden ser limitantes. [5, 6].

La extracción de información estructural a partir de los datos de SAXS requiere que tanto la muestra como el conjunto de datos cumplan con ciertos requisitos: la muestra debe presentar alto grado de pureza (sobre todo en lo que respecta a impurezas de elevada MM -lo que incluye agregados- que resultan particularmente problemáticas dado que, como ya se señaló, la señal de SAXS es proporcional a la MM). Por otro lado, la solución debe ser monodispersa e ideal (en el sentido de que las partículas del soluto no deben interactuar entre sí). Por lo tanto, es necesario llevar a cabo una cuidadosa evaluación de la calidad de los datos antes de extraer ninguna conclusión. [5, 6]

En cualquier estudio de SAXS, el primer análisis consiste en la aplicación de la aproximación de Guinier, que establece que para una solución ideal y monodispersa de

macromoléculas globulares idénticas, para valores suficientemente bajos de Q es válida la siguiente relación [5, 6]:

$$I(Q) = I_0 \cdot \exp\left(-\frac{1}{3} R_g^2 Q^2\right) \quad (\text{Ecuación 3.12})$$

donde I_0 es la intensidad dispersada a ángulo 0, y R_g es el radio de giro de la partícula. Esta aproximación es válida en general en el rango $Q \cdot R_g < 1,3$, y requiere que el perfil de dispersión haya sido corregido por *smearing*, como ya se señaló. En la práctica, se realiza un gráfico de $\ln [I(Q)]$ vs Q^2 , lo que permite la determinación directa de los parámetros I_0 y R_g a partir de la pendiente y la ordenada al origen de la recta. Este análisis constituye una herramienta muy sencilla para obtener rápidamente información acerca del tamaño, el estado oligomérico, y la calidad de los datos, como se discutirá a continuación.

En primer lugar, para muestras monodispersas el gráfico de Guinier debe ser lineal en ese rango de Q ; esto constituye un buen indicador de la calidad de los datos, ya que un gráfico no lineal puede ser indicativo de numerosos defectos como por ejemplo sustracción inadecuada del blanco, polidispersión de la muestra, o presencia de interacciones atractivas o repulsivas entre las partículas [5, 6]. Las interacciones atractivas o la agregación se manifiestan como un aumento pronunciado de la señal a bajos ángulos y consecuentemente una sobreestimación del R_g y de I_0 , mientras que la repulsión entre partículas se manifiesta como una caída abrupta a bajos ángulos y por lo tanto resulta en una subestimación de los parámetros. (Figura 3.2).

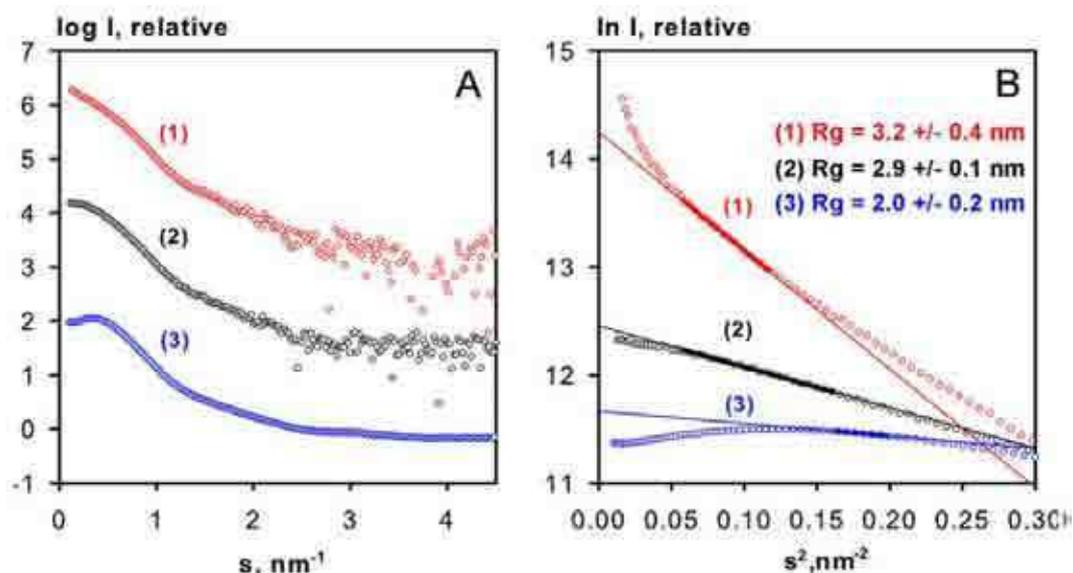


Figura 3.2: Verificación de la calidad de los datos de SAXS y análisis de Guinier.

Se muestran los perfiles de dispersión (panel de la izquierda) y los gráficos de Guinier (panel de la derecha) para muestras de BSA registradas en distintos buffers y condiciones que ilustran (1) agregación (rojo), (2) datos de buena calidad (negro), y (3) repulsión entre partículas (azul). Notar que aun en los casos en que existen efectos indeseados apreciables pueden obtenerse los parámetros a partir de las regiones lineales del gráfico. Se muestran los R_g obtenidos en cada caso, mientras que la I_0 resulta de la intersección de la recta de Guinier con el eje y en el panel de la derecha. (La figura fue tomada de [5]).

En segundo lugar, la aproximación de Guinier resulta de mucha utilidad para la determinación de la I_0 , que no puede ser registrada experimentalmente porque a un ángulo de 0° incide sobre el detector el remanente del haz incidente. Este parámetro, normalizado por la concentración del soluto, es directamente proporcional a su MM. En la práctica, la constante de proporcionalidad se determina colectando un conjunto de datos para una muestra de MM y concentración conocidas (por ejemplo, BSA o lisozima). De esta manera, SAXS permite estimar la MM del soluto en solución [5, 6]; sin embargo, en numerosas oportunidades esta posibilidad se encuentra limitada por errores en la determinación de la concentración de la muestra. Si la proteína tiene una MM lo suficientemente grande, en general este análisis es suficiente para distinguir entre las distintas formas oligoméricas; para proteínas de baja MM (como es el caso de este trabajo) el análisis puede resultar insatisfactorio.

Finalmente, el gráfico de Guinier permite estimar el R_g de la partícula a partir de la pendiente de la recta. El R_g se define como la distancia cuadrática media al centro de masa de la molécula de todos los elementos de volumen que contribuyen a la dispersión de Rayos X, pesadas por la densidad electrónica de cada uno de ellos. Por lo tanto proporciona información acerca de la distribución de masa en la molécula, y es un indicador de sus dimensiones globales [9].

El segundo análisis en cualquier estudio de SAXS consiste en la obtención de la función de distribución de distancias de la partícula, $P(r)$. Esta función es la representación de los datos de SAXS en el espacio real (el espacio de las distancias), así como el perfil de dispersión es la representación de los datos en el espacio recíproco (el espacio del vector de onda Q). Ambas se encuentran relacionadas por una transformación de Fourier [5]:

$$P(r) = r^2 / 2\pi^2 \int_0^\infty [Q^2 I(Q) \text{sen}(Qr) / Qr] dQ \quad (\text{Ecuación 3.13})$$

Esta función puede interpretarse como un histograma de distancias intramoleculares, lo que resulta informativo de la forma global de la molécula (**Figura 3.3**); además, es muy sensible a cambios conformacionales como por ejemplo perturbaciones en la orientación relativa de los dominios o aparición de ligeras asimetrías. Por otro lado, permite obtener otro parámetro indicativo de las dimensiones moleculares, la máxima distancia lineal entre dos puntos de la partícula (D_{\max}). Esta distancia corresponde al punto en el cual la función corta al eje x (además del origen de coordenadas). [5, 6]. Una vez obtenida la función $P(r)$, además de la información cualitativa acerca de la forma de la molécula, pueden calcularse por otro método independiente al análisis de Guinier el R_g y la I_0 por definición, a partir de la curva [9]:

$$R_g^2 = \frac{\int_0^{D_{\max}} [P(r)r^2] dr}{2 \int_0^{D_{\max}} [P(r)] dr} \quad (\text{Ecuación 3.14})$$

$$I_0 = 4\pi \int_0^{D_{\max}} [P(r)] dr \quad (\text{Ecuación 3.15})$$

Estos valores son más confiables que los que se obtienen por el análisis de Guinier, dado que este método utiliza la totalidad del conjunto de datos, en contraposición con el primero, que solo toma los datos correspondientes a valores de Q bajos (donde es válida la aproximación de Guinier) [6].

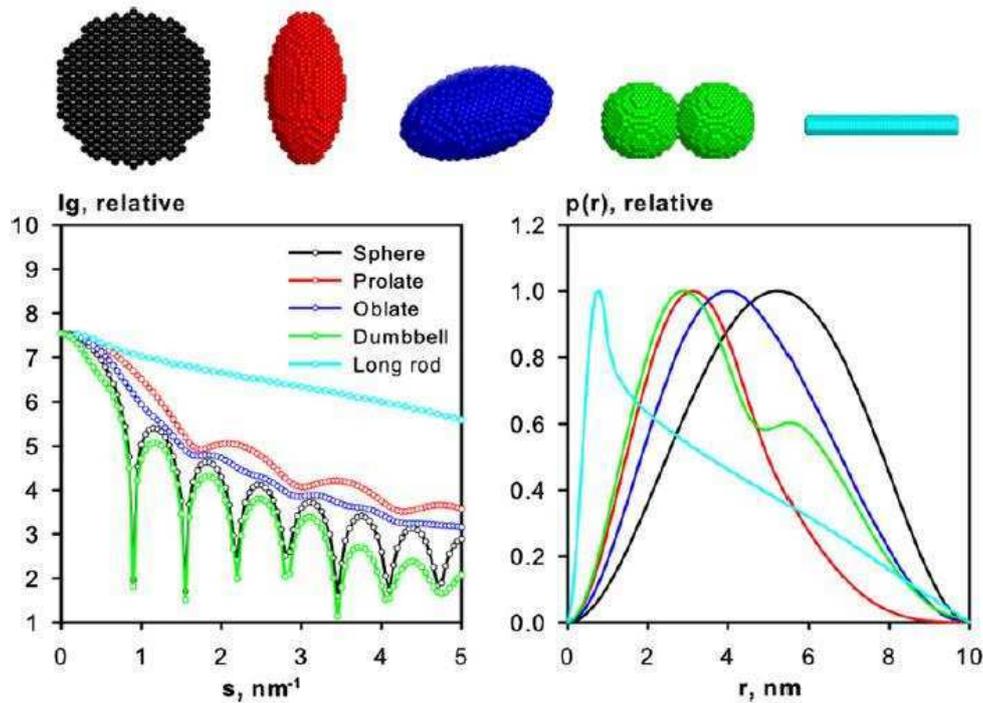


Figura 3.3: Distintas formas de la función $P(r)$.

Se muestran para diferentes formas geométricas los perfiles de dispersión (panel de la izquierda) y sus funciones de distribución de distancias (panel de la derecha): esfera (negro), elipsoide prolato (rojo), elipsoide oblató (azul), pesa (verde) y cilindro alargado (turquesa). Todas las formas tienen la misma D_{\max} , lo que se evidencia en que sus funciones $P(r)$ cortan al eje x en el mismo punto (10 nm). Notar que la esfera se caracteriza por una función simétrica en forma de campana, cuyo máximo se encuentra aproximadamente en $r = D_{\max}/2$. Por otro lado, la forma de la función del cilindro es fuertemente asimétrica; el pico a bajas distancias corresponde a su diámetro. En la pesa, el pico mayor corresponde al radio de cada una de las esferas mientras que el pico menor corresponde a la distancia entre los centros de las esferas. Este tipo de interpretaciones es muy común en proteínas con múltiples dominios, lo que permite obtener información acerca de las dimensiones de los dominios individuales así como también de las distancias que los separan. (La figura fue tomada de [5]).

Dado que el conjunto de datos experimentales siempre es acotado, el cómputo de la función no puede hacerse de manera directa a partir de la **Ecuación 3.13**, y en general se resuelve por métodos numéricos indirectos en los que $P(r)$ se representa como una combinación lineal de funciones ortogonales definidas en el rango $[0, D_{\max}]$. Por lo tanto, la D_{\max} es un parámetro que debe ser proporcionado por el usuario. El método consiste en hallar los coeficientes de la combinación lineal que generan una función $P(r)$ tal que la concordancia entre los datos experimentales $[I(Q)]_{\text{exp}}$ y los datos teóricos $[I(Q)]_{\text{teo}}$ calculados por

transformada de Fourier de la función $P(r)$ “candidata”, sea lo mejor posible. A su vez, otro factor tenido en consideración en la resolución del sistema es que la solución debe ser “suave”. [6].

El programa más popular en el que se implementa este tipo de métodos es GNOM [10], del paquete de análisis de datos de SAXS ATSAS [11]. En este programa, el usuario proporciona como *input* el perfil de dispersión de SAXS no corregido por *smearing*. El haz de Rayos X se modela como un perfil de forma trapezoidal caracterizado por dos parámetros: su altura (AH) y su base (LH). Además del perfil de SAXS, el usuario proporciona entonces tres parámetros: AH, LH, y la D_{\max} para evaluar la función $P(r)$. En la práctica, la mejor solución se encuentra por prueba y error de distintos valores de D_{\max} , y la calidad de cada solución se evalúa en función de una serie de parámetros estadísticos que son proporcionados por el programa, y por inspección visual de la función $P(r)$ resultante considerando ciertos criterios (por ejemplo, la función nunca puede adoptar valores negativos dado que el eje y representa frecuencias). Adicionalmente, el método corrige los *slit-smearing effects*.

Un tercer nivel de análisis en SAXS consiste en la reconstrucción *ab initio* de modelos del contorno molecular a partir de los perfiles de dispersión experimentales. Estos métodos se basan en expresar el patrón de SAXS como una suma de armónicos esféricos, lo que permite computar rápidamente los perfiles de dispersión teóricos a partir de modelos moleculares. A partir de esta idea, se desarrollaron algoritmos que permiten generar estos modelos mediante una minimización de la discrepancia entre el perfil experimental y el perfil teórico calculado a partir del modelo. [5, 6].

El programa más popular de generación de modelos *ab initio* es DAMMIN (dummy atom model minimization) [12] del paquete ATSAS [11]. El programa representa la partícula como una colección de esferas (*dummy atoms*) densamente empacadas dentro de un volumen de búsqueda esférico, que tiene un diámetro igual a la D_{\max} experimental de la partícula; cada una de estas esferas se asigna arbitrariamente a la molécula o al solvente. El algoritmo consiste en un protocolo de *simulated annealing*, en el cual en cada paso una esfera elegida al azar es reasignada (ya sea del solvente a la molécula, o viceversa) y se evalúa una función que describe la discrepancia entre el perfil de SAXS experimental y el del modelo obtenido en ese paso. Esa función contiene un término de penalidad adicional que contempla que las esferas que forman parte de la molécula deben estar conectadas entre sí. El protocolo termina cuando

se alcanza un mínimo local en la función de discrepancia, es decir, cuando luego de un cierto número de pasos su valor se mantiene constante. [5, 6].

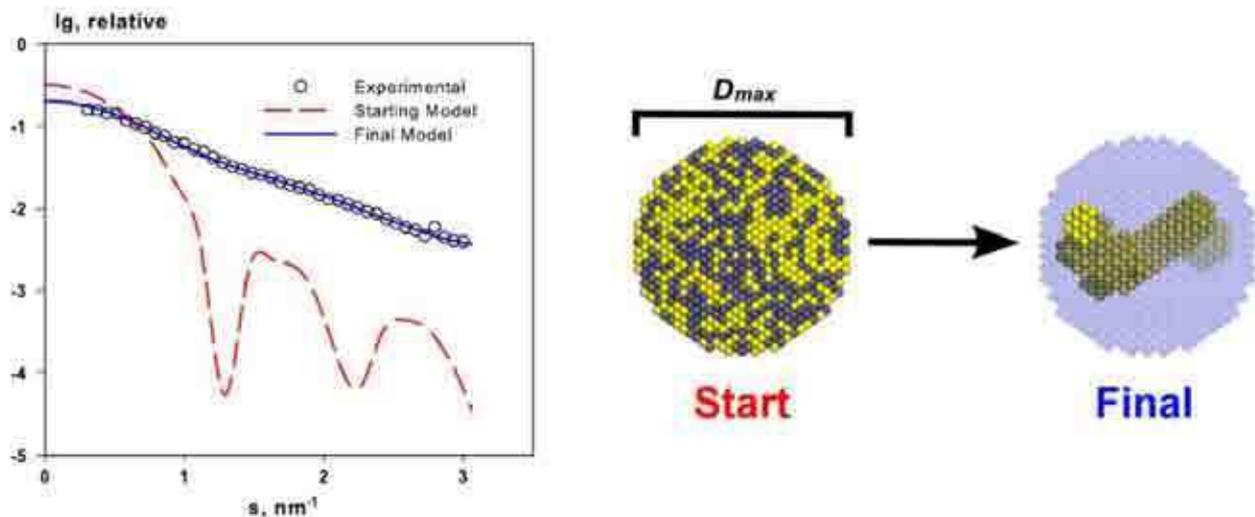


Figura 3.4: Modelado *ab initio* utilizando el programa DAMMIN.

En el panel de la izquierda se muestran los perfiles de dispersión experimental (puntos blancos), del modelo de partida (línea punteada roja) y del modelo final (línea azul). En el panel de la derecha se esquematiza el procedimiento de generación del modelo: partiendo de un volumen esférico de diámetro igual a D_{max} que contiene esferas de solvente (violeta) y de la molécula (amarillo) aleatoriamente distribuidas, se genera mediante un protocolo de *simulated annealing* un modelo molecular consistente con los datos experimentales. (La figura fue tomada de [5]).

Otro programa muy utilizado, desarrollado posteriormente a partir de DAMMIN, es DAMMIF [13]. Este programa funciona esencialmente de la misma manera que el anterior, pero introduce una serie de modificaciones en el protocolo de búsqueda que lo hacen mucho más rápido; de esta manera, pueden calcularse numerosos modelos en tiempos accesibles, para su posterior comparación. Adicionalmente, en este programa el volumen de búsqueda no se encuentra restringido, sino que puede expandirse durante el transcurso del cálculo; esta mejora permite evitar efectos de borde artificiosos que pueden ocurrir cuando se utiliza un volumen de búsqueda restringido por una D_{max} ligeramente subestimada. [5, 6]. Este fue el programa utilizado en este trabajo.

Es importante mencionar que los métodos *ab initio* no proporcionan una solución única, dado que no hay información suficiente en un conjunto de datos experimentales 1D para definir unívocamente una estructura 3D. Por lo tanto, si se corre el programa numerosas veces se obtienen modelos ligeramente distintos, por lo que existe cierta ambigüedad en el resultado final. Para sortear esta dificultad, pueden generarse múltiples modelos y observar que se conserven las características generales más allá de diferencias sutiles en la forma. Por otro lado, para minimizar *a priori* el abanico de posibles modelos, algunos programas permiten proporcionar información proveniente de otros métodos experimentales; por ejemplo, en DAMMIN el usuario puede proporcionar como *input* información acerca de la simetría de la molécula, la cual puede ser conocida a partir de las estructuras cristalinas. [5, 6].

Finalmente, como ya se mencionó, SAXS resulta una herramienta muy poderosa como método complementario de otras técnicas. Por ejemplo, los modelos atómicos obtenidos por difracción de Rayos X o por modelado por homología pueden ser superpuestos con los modelos de contorno molecular generados *ab initio* por SAXS. Por otro lado, como ya se señaló, existen métodos que permiten obtener los perfiles teóricos de SAXS de modelos atómicos de distintas fuentes; estos perfiles teóricos pueden ser comparados con los datos experimentales de SAXS de esas moléculas. [5, 6].

Este tipo de abordajes permiten validar modelos generados por homología, dado que un modelo que proporciona un buen ajuste con los datos experimentales de SAXS es considerado una buena descripción de la estructura real en solución en esas condiciones experimentales [5]. Por otro lado, permiten evaluar la compatibilidad entre las estructuras cristalinas y en solución. También resulta de utilidad para discernir cuáles de los posibles estados oligoméricos observados en estructuras cristalinas son relevantes en solución, lo cual generalmente es imposible de determinar en base a las estructuras cristalinas solamente, dado que en los cristales siempre se generan interacciones artificiales, no nativas, como consecuencia del proceso de cristalización. [5]. Este tipo de abordajes fueron utilizados en este trabajo, tanto en este capítulo como en el siguiente.

Existen numerosos métodos que permiten el cómputo de las curvas teóricas de SAXS a partir de modelos atómicos. Estos métodos difieren en numerosos aspectos, principalmente en lo que respecta al tratamiento del volumen de solvente excluido por la partícula, y al

modelado de la capa de agua de hidratación de las proteínas en solución, que al ser un 10% más densa que el solvente *bulk* contribuye a la dispersión de Rayos X. Como consecuencia de estas diferencias, a menudo los diferentes métodos arrojan resultados ligeramente distintos, por lo que esto debe ser tenido en consideración al momento de analizar los resultados. [5].

RESULTADOS Y DISCUSIÓN

Estado oligomérico de Bcy1 (1-50) en solución

Para elucidar el estado de oligomerización de Bcy1 (1-50) en solución, en primer lugar se analizó su comportamiento en una columna de SEC, aprovechando que esta técnica forma parte del protocolo de purificación. En este método separativo la difusión diferencial de las macromoléculas a través de la matriz depende de su MM, de su tamaño, y de su forma: las proteínas más grandes y no globulares recorren trayectorias más cortas dado que no pueden acceder a todos los poros de la matriz, y por lo tanto eluyen antes de la columna, mientras que las más pequeñas y esféricas recorren trayectorias más largas y por lo tanto eluyen más tarde. La magnitud que mejor representa el comportamiento de las moléculas en esta técnica es el radio de Stokes (R_s), que depende de ambas características moleculares (tamaño y forma) [3].

Dado que la forma de la macromolécula influye marcadamente en la velocidad de migración a través de la matriz, en general no es posible determinar la MM (y consecuentemente el estado oligomérico) de una proteína en solución solamente mediante la comparación directa de su volumen de elución (V_e) con los de los estándares de MM. Sin embargo, cuando tanto los estándares como la muestra pueden considerarse aproximadamente globulares, existe una correlación razonable entre el V_e y la MM [3].

Se registraron cromatogramas de una muestra de Bcy1 (1-50) purificada de concentración 100 μ M, utilizando proteínas estándar de MM conocida. El fragmento purificado eluyó como un único pico con un volumen de elución (V_e) ligeramente inferior al del marcador de 25 kDa (α -quimiotripsina) (**Figura 3.5**). Considerando que en este caso se utilizaron estándares que son proteínas globulares bien conocidas (ver **Tabla 3.S.1**), y teniendo como posibilidades las distintas variantes encontradas en los experimentos de entrecruzamiento químico del Capítulo 2, si Bcy1 (1-50) fuera también globular en solución, su MM aparente resultaría demasiado elevada en comparación con lo esperado para una proteína monomérica (5,9 kDa), dimérica (11,8 kDa), o trimérica (17,7 kDa), siendo más cercana a la de una forma tetramérica (23,6 kDa). Por lo tanto, en un primer análisis, estos resultados sugieren que, en estas condiciones experimentales, existe en solución una única especie predominante conformada por cuatro subunidades.

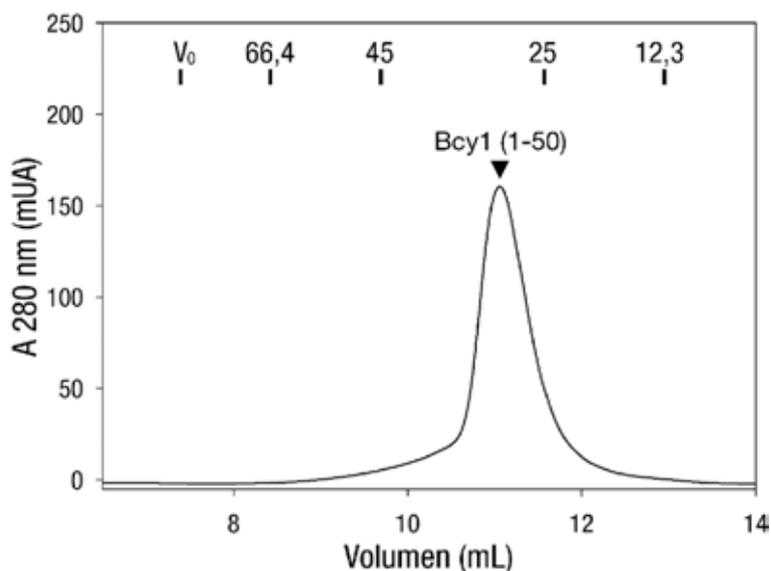


Figura 3.5: Caracterización de Bcy1 (1-50) por SEC.

Cromatograma obtenido para una muestra purificada 100 μM de Bcy1 (1-50) a 25 $^{\circ}$ C en buffer PBS utilizando una columna Superdex-75. El eje y se representa en unidades de absorbancia ($\times 10^3$). Se indican el volumen excluido de la columna (V_0) y la posición de los estándares: BSA (MM = 66,4 kDa), ovoalbúmina (MM = 45 kDa), α -quimiotripsina (MM = 25 kDa) y citocromo C (MM = 12,3 kDa).

La observación de un único pico en SEC no garantiza que exista una sola forma oligomérica en solución; podrían coexistir en equilibrio rápido dos o más especies que, en ese caso, no podrían separarse por esta técnica, dado que durante el desarrollo de la cromatografía se encontrarían en constante interconversión, y por lo tanto migrarían juntas. Para analizar esta posibilidad, se registraron cromatogramas de diluciones sucesivas de la muestra original 100 μM de Bcy1 (1-50), volviendo a cargar en la columna el eluido de cada corrida, considerando que en la columna la muestra se diluye aproximadamente en un factor de 1:2. En caso de haber equilibrios de oligomerización involucrados, si las constantes de equilibrio asociadas son similares a las concentraciones de trabajo, la dilución debería alterar la composición de la mezcla original favoreciendo las formas con menor cantidad de subunidades. En la **Figura 3.6** se muestran los resultados obtenidos.

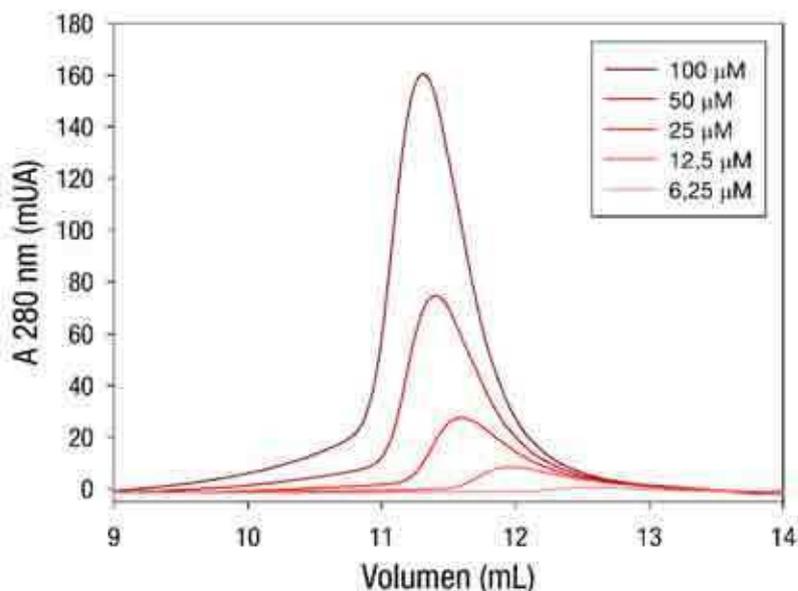


Figura 3.6: Equilibrio de los distintos estados de oligomerización de Bcy1 (1-50) en solución.

Cromatogramas de diluciones sucesivas de una muestra purificada de Bcy1 (1-50), obtenidos en las mismas condiciones de la Figura 3.1, volviendo a inyectar luego de cada corrida el eluido de la anterior.

Se observa un corrimiento constante del pico de Bcy1 (1-50) hacia volúmenes mayores (es decir, menor MM) como consecuencia de las diluciones sucesivas, lo que sugiere que existe una mezcla de formas con distintos estados de oligomerización, y que las muestras más diluidas se encuentran enriquecidas en las especies de menor cantidad de subunidades respecto de la muestra original. Estos resultados sugieren también que la estructura cuaternaria de este fragmento es dinámica en solución, y que la concentración de proteína podría ser un mecanismo regulatorio del estado oligomérico.

Para decidir si esta posibilidad es factible en un contexto fisiológico, se procedió a obtener una estimación de la concentración de Bcy1 WT *in vivo*, lo que representa una tarea compleja por numerosos motivos. En primer lugar, no es sencillo contar con un dato confiable de la cantidad de moléculas por célula. En segundo lugar, el volumen celular es muy variable y depende de numerosos factores, como por ejemplo de las condiciones de cultivo, del estadio del ciclo celular, del genotipo, de si se analizan las células madre o las hijas, etc. [14]. Por otro lado, como ya se discutió en la Introducción, la localización subcelular de Bcy1 se encuentra sometida a regulación por diferentes factores, encontrándose en distintos compartimientos

subcelulares dependiendo de las condiciones de crecimiento [15, 16]. Finalmente, no todas las estructuras celulares son accesibles a todas las proteínas.

Habiendo hecho estas salvedades, considerando que una célula madre haploide de una cepa WT creciendo en fase exponencial en medio rico tiene un volumen celular de 42 fL [14], y que se ha reportado un valor promedio de 4280 moléculas/célula para Bcy1 [17], la concentración estimada es del orden de 0,1 μ M. Este valor se encuentra seguramente bastante subestimado, dado que para calcularlo se supuso que todo el volumen celular es accesible a Bcy1. Más aun, se ha reportado que en fase estacionaria la cantidad de Bcy1 aumenta entre 8 y 9 veces respecto a la que existe en fase exponencial [18]. En este contexto, resulta razonable suponer que en ciertas situaciones (por ejemplo, cuando Bcy1 se localiza en el núcleo, que representa aproximadamente un 10% del volumen celular) las concentraciones intracelulares de Bcy1 podrían llegar a ser del orden de las utilizadas en los experimentos en los que se observan indicios de coexistencia de los distintos estados de oligomerización.

Cabe destacar que la MM aparente observada podría explicarse también por la existencia de alguna especie elongada, no globular; por ejemplo, el pico observado en la **Figura 3.5** podría corresponder a una mezcla de un tetrámero aproximadamente globular en equilibrio rápido con un trímero o un dímero que por su estructura tuvieran un R_s similar al del tetrámero. Para determinar de manera inequívoca el estado de oligomerización de Bcy1 (1-50) en solución, se realizaron mediciones de dispersión estática de la luz (SLS) acoplada a SEC, en condiciones similares de concentración de proteína, temperatura y buffer, a las utilizadas para registrar el cromatograma de la **Figura 3.5**. Esta técnica es actualmente la más confiable para determinar con alta exactitud la MM de una proteína en solución [1], y se basa en que la intensidad de la luz dispersada a 90° del flujo de la muestra (LS) es directamente proporcional a la MM de la molécula y a su concentración, y completamente independiente de su forma (ver Introducción metodológica para más detalles). En la **Figura 3.7** muestran los resultados obtenidos.

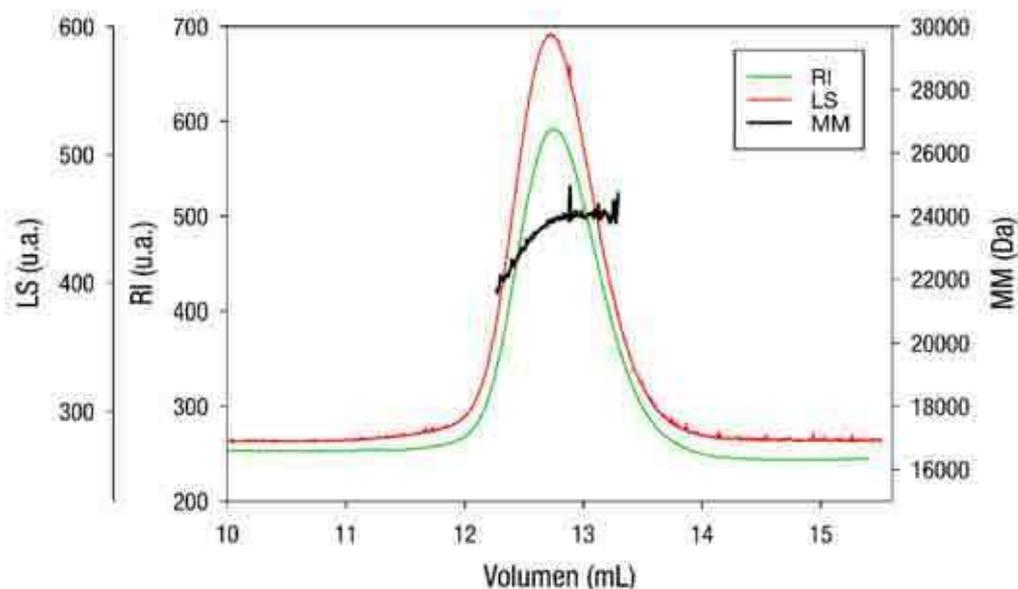


Figura 3.7: Caracterización de Bcy1 (1-50) por SEC-SLS.

Las mediciones fueron realizadas utilizando una muestra purificada 100 μM de Bcy1 (1-50) a 25 $^{\circ}$ C en buffer PBS utilizando una columna Superdex-75. Los ejes de la izquierda representan las señales de dispersión a 90 $^{\circ}$ (LS; línea roja) y de índice de refracción (RI; línea verde) en unidades arbitrarias. La MM se muestra en el eje de la derecha (línea negra).

La MM promedio del único pico del cromatograma fue de $(23,5 \pm 0,7)$ kDa (esperado para el tetrámero: 23,6 kDa), lo que corresponde a un número de subunidades de $(4,0 \pm 0,1)$. Estos resultados confirman claramente que, en estas condiciones experimentales, la proteína existe en solución como una única especie que consta de cuatro subunidades. Por otro lado, refuerzan la hipótesis de que el corrimiento del pico observado en los cromatogramas de la **Figura 3.6** como consecuencia de la dilución se debe al establecimiento de equilibrios que dan lugar a especies de menor cantidad de subunidades.

Estado oligomérico de Bcy1 WT en solución

Como ya se mencionó en la Introducción y en el Capítulo 2, los primeros trabajos sobre Bcy1 han demostrado que, al igual que las subunidades R de mamíferos, la proteína es dimérica *in vitro* [19, 20]; estos resultados se encuentran en discrepancia con los de la sección anterior. Esto podría deberse a diferencias entre esos estudios y este trabajo en cuanto a las condiciones experimentales utilizadas, o bien a que los resultados reportados en la sección

anterior sean artificiosos como consecuencia del sistema de expresión utilizado (es decir, por haber trabajado con la proteína expresada en bacterias y no en levaduras), o bien del fragmento producido (por ejemplo, podría suceder que en Bcy1 WT existieran secuencias adicionales, no contenidas en Bcy1 (1-50), que tengan efecto sobre el estado de oligomerización de la proteína completa).

En el primero de estos trabajos [19] se determinó la MM de Bcy1 WT endógena utilizando el método de Siegel y Monty [4], que se basa en la relación existente entre el R_s y el coeficiente de sedimentación a 20° C en agua ($S_{20,w}$) con la MM (se describe el método más adelante en la sección “parámetros hidrodinámicos”). En el segundo trabajo [20] se determinó la MM de Bcy1 WT expresada en bacterias en base a su $S_{20,w}$ únicamente. En cuanto a las condiciones experimentales, el primer trabajo se realizó en buffer MES 10 mM, 100 mM NaCl, pH = 6,9, mientras que el segundo en buffer fosfato 50 mM, 100 mM NaCl, pH = 6,5; en ambos casos las condiciones son similares a las utilizadas en este estudio (buffer PBS: 20 mM fosfato, 150 mM NaCl, pH = 7,4). En ninguno de los dos trabajos se aclara la concentración de la proteína utilizada.

Resulta difícil evaluar las interpretaciones de los resultados de esos estudios para compararlas con los que se presentan en este trabajo, dado que en algunos casos la información no se presenta; por ejemplo, en el trabajo de levaduras no se muestran los resultados de la calibración de la columna de SEC utilizada para obtener el R_s ni la del gradiente de sacarosa utilizada para obtener el $S_{20,w}$. Lo primero resulta llamativo, dado que en el texto se comenta que la proteína eluye en el volumen excluido de la columna, por lo que no sería posible determinar el R_s en esas condiciones. Por otro lado, en el trabajo de bacterias no se determinó el R_s , y la MM se infirió solamente por el $S_{20,w}$, que resultó similar al del trabajo anterior, de lo que se dedujo que debía tratarse en ambos casos del mismo estado oligomérico (dímero). Como se discutió en la Introducción metodológica, la estimación de la MM realizada sólo a partir del $S_{20,w}$ puede conducir a interpretaciones erróneas.

En cuanto a la posibilidad de que los resultados de este trabajo puedan resultar artificiosos debido al sistema de expresión, es importante tener en consideración que los trabajos estructurales previos de los dominios D/D de mamíferos fueron desarrollados utilizando siempre construcciones recombinantes sobreexpresadas en bacterias análogas a la utilizada para la sobreexpresión de Bcy1 (1-50) [21-27] (ya se mencionó en el capítulo anterior que en algunos casos se obtuvieron indicios de estructuras de mayor estado de oligomerización que no fueron investigados posteriormente).

Visto lo anterior, dado que no se han reportado hasta la fecha estructuras tetraméricas ni de las subunidades R enteras ni de sus dominios D/D en ningún organismo, y que los trabajos que demostraron el estado dimérico de Bcy1 WT fueron realizados en las mismas condiciones experimentales que este estudio y presentan algunos aspectos que son poco claros, se decidió revisar este concepto mediante SEC-SLS que, como ya se señaló, es la técnica más difundida en la actualidad para determinar el estado oligomérico de una proteína en solución. Esta técnica no se ha utilizado previamente para estudiar la MM de las subunidades R de ningún organismo. Cabe mencionar también que existen resultados previos de nuestro laboratorio de ultracentrifugación analítica en gradiente de sacarosa que sugieren que en solución podrían coexistir numerosas formas oligoméricas de Bcy1 (Tesis Doctoral de Jimena Rinaldi, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi).

En consecuencia, se realizaron experimentos de SEC-SLS en condiciones similares de temperatura y buffer a las utilizadas para estudiar el fragmento Bcy1 (1-50), utilizando preparaciones de Bcy1 WT purificadas a partir de la cepa de levaduras 1115-BCY1, que sobreexpresa esta proteína bajo el promotor inducible por galactosa GAL10 (ver Capítulo 2, sección Materiales y Métodos). Por otro lado, para obtener una aproximación a lo que podría ser el estado oligomérico en una situación más similar al contexto celular, se realizaron experimentos de entrecruzamiento químico con EGS similares a los del Capítulo 2, pero utilizando como fuente de proteína extractos libres de células preparados a partir de esa cepa en vez de la proteína purificada. En este caso, los resultados se analizaron por *western blot* utilizando anticuerpos específicos contra Bcy1. Los resultados se muestran en la **Figura 3.8**.

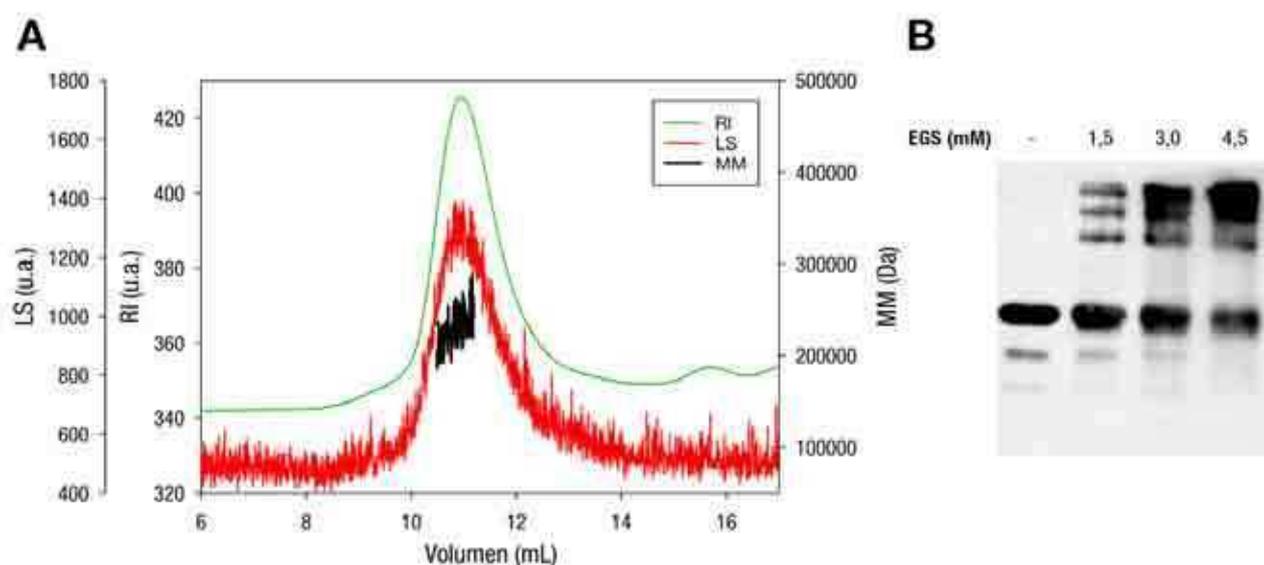


Figura 3.8: Estado oligomérico de Bcy1 WT en solución y en su contexto celular.

(A) Las mediciones fueron realizadas utilizando una muestra purificada 20 μ M de Bcy1 WT a 25° C en buffer PBS utilizando una columna Superdex-200. Los ejes de la izquierda representan las señales de dispersión a 90° (LS; línea roja) y de índice de refracción (RI; línea verde) en unidades arbitrarias. La MM se muestra en el eje de la derecha (línea negra). (B) Extractos crudos de la cepa 1115-BCY1 fueron sometidos a tratamiento con el reactivo entrecruzante de lisinas EGS en las concentraciones indicadas en la parte superior de la figura. Los productos de la reacción fueron capturados utilizando una resina de afinidad cAMP-agarosa, separados por SDS-PAGE 10% y visualizados por *western blot* con anticuerpos específicos contra Bcy1.

La proteína eluyó de la columna como un solo pico bien definido con un V_e de aproximadamente 11 mL; se esboza un segundo pico a un V_e ligeramente inferior a 16 mL, pero por ser muy débil la señal de LS no se pudo caracterizar. La MM promedio del pico mayoritario fue de (230 ± 20) kDa; dado que la MM de Bcy1 es 47,2 kDa, el número de subunidades determinado por esta técnica es de $(4,9 \pm 0,4)$. Por lo tanto, en estas condiciones experimentales Bcy1 es o bien tetramérica o bien pentamérica en solución. La existencia de una especie pentamérica sería inconsistente con los resultados obtenidos para Bcy1 (1-50), que indicaron claramente que existe como una única especie tetramérica en las mismas condiciones experimentales. Es probable que la elevada MM observada para Bcy1 esté asociada con un mayor error experimental en este caso, debido principalmente a dos factores. El primero es la utilización de una concentración de proteína 5 veces más baja para Bcy1 WT que para Bcy1 (1-50). El segundo es que la columna Superdex-200 utilizada para la primera se encontraba bastante más deteriorada que la columna Superdex-75 empleada para la segunda. Estos dos factores se evidencian en una señal mucho más ruidosa de LS para Bcy1 WT que para Bcy1 (1-50) (comparar las líneas rojas de las **Figuras 3.7 y 3.8 (A)**).

Por otro lado, en los experimentos de entrecruzamiento químico realizados sobre el extracto libre de células se observa claramente la aparición de tres bandas como consecuencia del tratamiento con el reactivo (**Figura 3.8 (B)**). Estas bandas podrían atribuirse a especies dimérica, trimérica y tetramérica, en consonancia con el resultado del párrafo anterior, sugiriendo que en condiciones similares a su contexto celular Bcy1 WT tiene la capacidad de formar tetrámeros en solución. En este punto cabe recordar que en los experimentos de

entrecruzamiento químico de Bcy1 WT purificada del Capítulo 2 (**Figura 2.4**) se habían observado bandas de alrededor de 200 kDa como consecuencia del tratamiento con EGS.

En conjunto, estos resultados y los de las secciones anteriores apoyan fuertemente la idea de que existe una forma tetramérica de Bcy1 previamente no reportada, y que su oligomerización ocurre a través de su dominio D/D, que se encuentra restringido a los 50 aminoácidos N-terminales de la molécula; esta forma es fisiológicamente relevante dado que existe en un entorno similar al contexto celular.

Posible rol fisiológico de la oligomerización de Bcy1

Para evaluar si el estado de oligomerización de Bcy1 tiene alguna incidencia sobre el estado fisiológico de las levaduras, se decidió sobreexpresar el fragmento Bcy1 (1-50) en una cepa de *S. cerevisiae* WT, y analizar si esto tiene algún efecto sobre los fenotipos característicos asociados a la actividad de PKA. Para ello, el fragmento recombinante se subclonó en el vector de expresión de levaduras pYES2 (ver Materiales y métodos para los detalles experimentales) y el vector recombinante se transformó en la cepa de *S. cerevisiae* WT KT-1115 (ver Materiales y métodos del Capítulo 2). Se procedió a analizar los siguientes fenotipos que, como se discutió en la Introducción, correlacionan con la actividad de PKA: curva de crecimiento, acumulación de glucógeno, y resistencia al *heat shock*. Los resultados se muestran en la **Figura 3.9**.

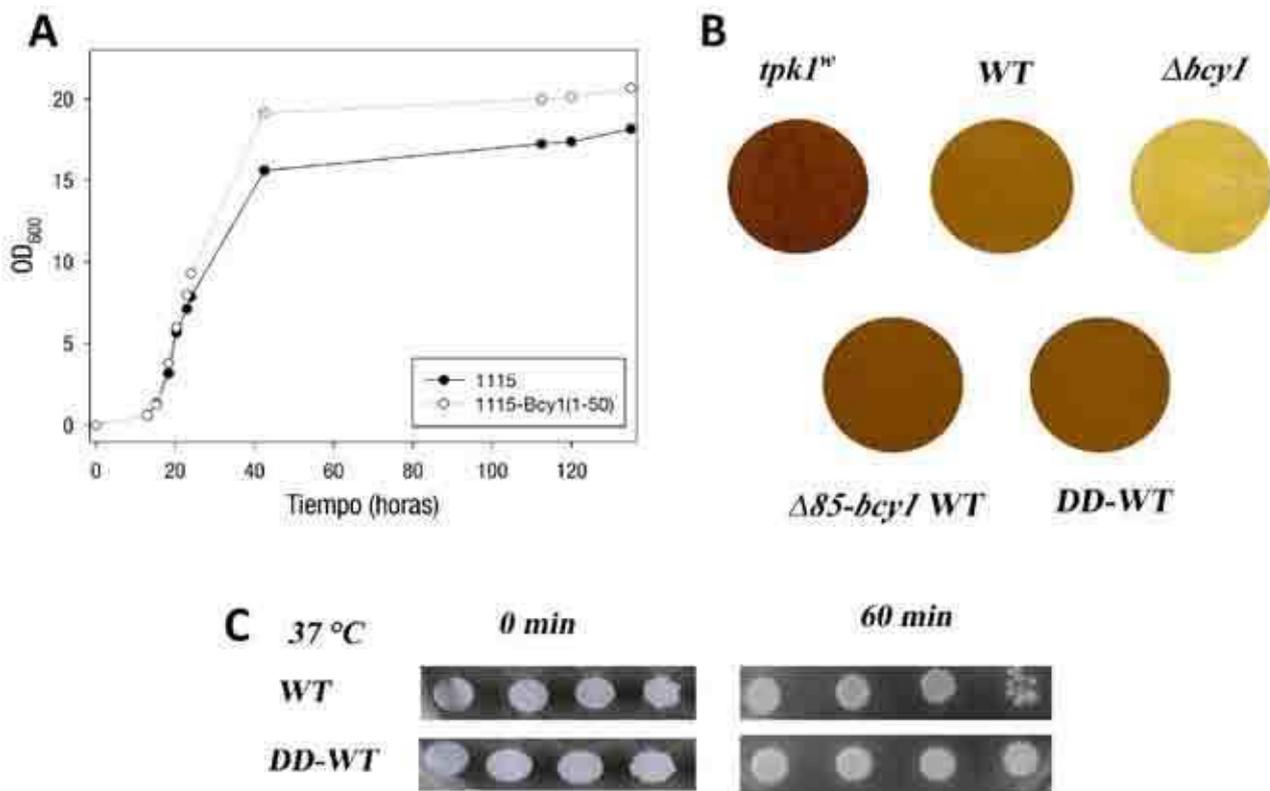


Figura 3.9: Fenotipos asociados a la sobreexpresión de Bcy1 (1-50).

(A) Curvas de crecimiento en medio YPGal de las cepas 1115 y 1115-Bcy1 (1-50). Las curvas se monitorearon por densidad óptica a 600 nm. (B) Ensayos de acumulación de glucógeno. Se utilizaron como controles las cepas *tpk1^w* (baja actividad PKA), *Δbcy1* (alta actividad PKA) y *Δ85bcy1 WT*, que sobreexpresa la variante Bcy1 $\Delta(1-85)$ sobre un fondo Bcy1 WT, y que se caracteriza por una acumulación de glucógeno más elevada que la cepa WT. Además, se utilizaron las cepas de interés: 1115 (WT) y 1115-Bcy1 (1-50) (DD-WT). Los parches más oscuros indican mayor acumulación de glucógeno. (C) Ensayos de *heat shock*. Las células creciendo en fase exponencial se sometieron (o no) a tratamiento a 37° C durante 60 minutos y se plaquearon en medio YPGal para evaluar su viabilidad. En todos los casos se muestra un experimento representativo de varios en los que se observaron resultados similares.

Para evaluar el primero de estos fenotipos, se creció la cepa transformada con el vector recombinante en medio YPGal para inducir la sobreexpresión del fragmento a partir del promotor GAL10, y la cepa sin transformar como control. Se monitorearon las curvas de crecimiento de las dos cepas por turbidimetría, representada por la densidad óptica del cultivo a 600 nm (OD₆₀₀), que es una medida de la cantidad de biomasa presente en el medio (Figura 3.9 (A)). Por otro lado, se tomaron alícuotas de los dos cultivos creciendo en fase

exponencial, luego del *shift diaúxico* (ver más adelante) y se evaluó su resistencia al *heat shock* sometiéndolas a un tratamiento a 37° C durante 1 hora; posteriormente, se plaquearon las células en medio rico para evaluar su viabilidad (**Figura 3.9 (C)**). Finalmente, se midió la acumulación de glucógeno mediante la exposición a vapores de yodo de placas con medio rico conteniendo parches de las células de las cepas de interés. Como control, se utilizaron las cepas *tkp1^w*, que se caracteriza por tener delecionadas las subunidades *tpk2* y *tpk3*, y expresar una subunidad *tpk1* mutante que tiene muy baja actividad de PKA, y por lo tanto presenta elevada acumulación de glucógeno; y una cepa Δ *bcy1*, que al no expresar la subunidad R de la PKA presenta actividad PKA constitutiva elevada y por lo tanto no acumula glucógeno (**Figura 3.9 (B)**)

En primer lugar, se observa que la cepa que sobreexpresa el fragmento recombinante presenta niveles de acumulación de glucógeno ligeramente superiores a los de la cepa WT (**Figura 3.9 (B)**) y es más resistente al *heat shock* (**Figura 3.9 (C)**). Como se discutió en la Introducción, ambos fenotipos son característicos de una actividad de PKA atenuada.

Por otro lado, se observa en las curvas de crecimiento (**Figura 3.9 (A)**) que la sobreexpresión del fragmento recombinante genera un aumento en el crecimiento de la cepa transformada alrededor de las 20 horas de cultivo, comparada con la WT; también se observa que el cultivo alcanza mayor densidad al llegar a fase estacionaria en el primer caso. Asimismo, a partir del momento en que se manifiestan las diferencias en las curvas de crecimiento, las células de las dos cepas presentan distintas características al ser observadas al microscopio óptico, dado que en algunas de las células de la cepa transformada se aprecian formas irregulares que no son típicas de levaduras (no se muestra). Sin embargo, se observó que las células de las dos cepas presentaban el mismo tamaño, por lo que la diferencia en la OD₆₀₀ no se debe a una diferencia en el tamaño celular sino a la cantidad de células. El hecho de que la cepa que sobreexpresa Bcy1 (1-50) pueda alcanzar la fase estacionaria resulta también relevante dado que, como se indicó en la Introducción, la deleción de Bcy1 (o sea, actividad PKA elevada) trae como consecuencia la imposibilidad de alcanzar esta fase del cultivo.

El momento a partir del cual se manifiestan las diferencias podría coincidir con el *shift diaúxico* (por ejemplo, se ha observado para algunas cepas que éste se alcanza a una OD₆₀₀ de alrededor de 7 [18]) que es un estadio del cultivo típico de *S. cerevisiae*, caracterizado por un

arresto transiente del crecimiento asociado a la adaptación de las células a crecer en base a la respiración del etanol acumulado en el medio como consecuencia del agotamiento de la fuente primaria de carbono por fermentación. Consistentemente, se ha reportado que la vía de señalización del cAMP se encuentra involucrada en la regulación de este proceso [28]. Más aun, se ha asociado el *shift diáuxico* con un cambio en el estado de fosforilación de Bcy1 y, como ya se mencionó, con un aumento en los niveles intracelulares de Bcy1 [18].

En conclusión, estos resultados demuestran que la sobreexpresión del fragmento Bcy1 (1-50) tiene consecuencias sobre el estado fisiológico de las células, y sugieren que el fenotipo asociado es el de una actividad de PKA ligeramente atenuada. Una posible explicación para esta observación es que el fragmento recombinante podría interferir en la oligomerización de Bcy1 endógena alterando su función y/o su localización subcelular, y como consecuencia resultar en una actividad de PKA desregulada (en particular, atenuada). Para apoyar esta hipótesis, se verificó que la construcción se expresa adecuadamente en las condiciones del cultivo (ver **Figura 3.S.3**). En la misma línea, se ha observado también que Bcy1 (1-50) expresada en levaduras tiene la capacidad de formar hetero oligómeros con Bcy1 WT *in vitro*, respaldando esta hipótesis (Tesis Doctoral de Enzo Tofolón, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi, en curso).

A continuación, se procedió a llevar a cabo una caracterización detallada de la estructura de esta novedosa entidad tetramérica en solución, para obtener información acerca de sus características generales así como también acerca de los determinantes que podrían mediar la tetramerización.

Caracterización estructural del tetrámero Bcy1 (1-50)

Este objetivo se desarrolló en tres etapas: determinación de la estructura secundaria, determinación de propiedades hidrodinámicas que dan información acerca del tamaño y la forma de la molécula en solución, y finalmente construcción de un modelo molecular consistente con la información experimental recolectada. En las siguientes secciones se describen estos puntos.

Estructura secundaria en solución

En base al análisis bioinformático realizado en el Capítulo 2, se predijo la existencia de un motivo hélice-giro-hélice para Bcy1 (1-50) (**Figura 2.1**). Para verificar estas predicciones y evaluar el contenido de estructura secundaria, se utilizó la técnica de espectropolarimetría de dicroísmo circular (CD). Esta técnica proporciona información cualitativa y cuantitativa acerca de la estructura secundaria de una proteína en solución; concretamente, permite determinar qué elementos de estructura secundaria se encuentran presentes en la molécula, y estimar la cantidad de residuos que se encuentran formando parte de cada uno de ellos (ver Introducción metodológica para más detalles).

Se registraron espectros de CD de muestras purificadas de Bcy1 (1-50) en condiciones experimentales similares a las que se vienen utilizando en los análisis de las secciones anteriores. Como puede observarse en la **Figura 3.10 (A)**, el espectro de CD en el UV lejano de la proteína nativa a temperatura ambiente presenta dos mínimos bien definidos a 210 y 222 nm, lo que es típico de proteínas con un elevado contenido de α -hélice. Por otro lado, este resultado es indicativo de que la proteína se encuentra bien plegada en solución. Cabe aclarar que no se realizó la deconvolución para evaluar el porcentaje de α -hélice porque la concentración de la proteína no estaba determinada con exactitud, y esto es un requisito necesario para obtener resultados confiables [2]. De todos modos, el análisis cualitativo del aspecto del espectro resulta suficiente para asegurar que el contenido de α -hélice es elevado, en concordancia con el análisis bioinformático, según el cual se predijo que aproximadamente el 70% de los residuos presentan conformación de α -hélice.

Por otra parte, se decidió aprovechar la capacidad de la técnica de funcionar como sensor de la estructura de la proteína para evaluar su estabilidad térmica. Para ello, se registraron espectros de CD a 90° C y a 25° C luego del tratamiento de alta temperatura. En primer lugar, se observa que a 90° C se pierde la estructura secundaria como consecuencia de la desnaturalización, lo que se manifiesta en un cambio cualitativo del espectro de CD, que adquiere un aspecto típico de proteínas desplegadas (línea gris punteada en la **Figura 3.10 (A)**). Llamativamente, cuando la muestra fue llevada nuevamente a temperatura ambiente se observó recuperación total de la estructura en escasos segundos, indicando que el proceso de desnaturalización térmica es reversible (línea sólida gris en la **Figura 3.10 (A)**). Esto proporciona evidencia adicional de que la proteína se encuentra bien plegada y de la alta estabilidad de su plegamiento.

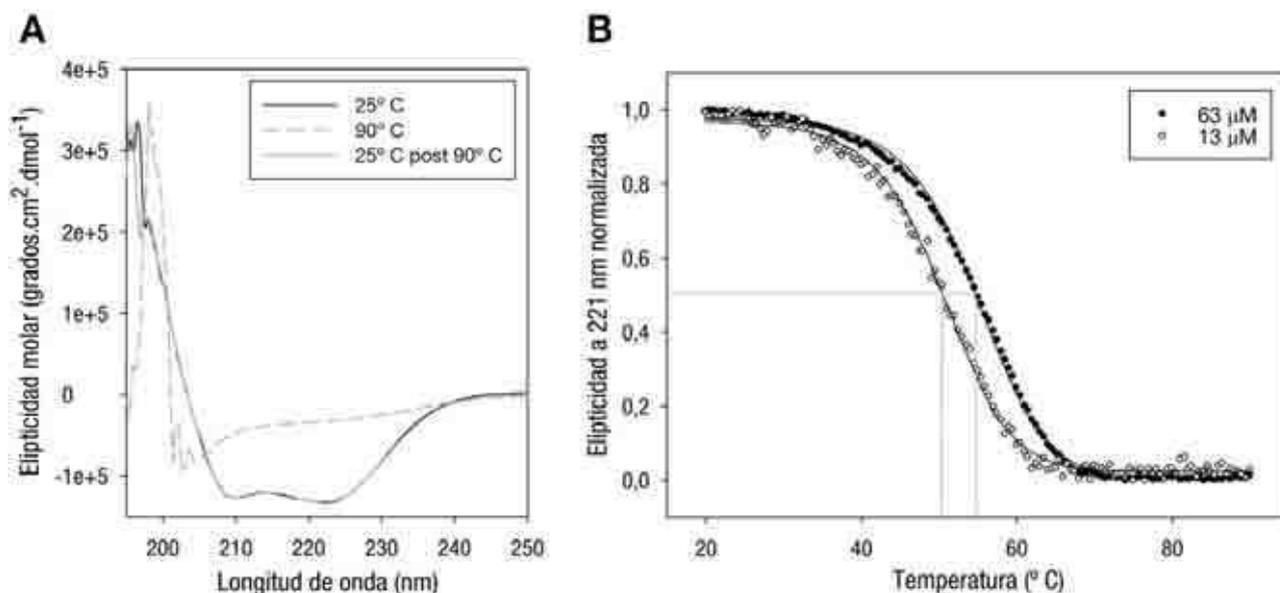


Figura 3.10: Experimentos de dicroísmo circular de Bcy1 (1-50).

(A) Espectros de CD en el UV lejano de una muestra de Bcy1 (1-50) de concentración $12,7 \mu\text{M}$, registrados en buffer PBS en tres condiciones distintas: 25°C , estado nativo (línea negra sólida); 90°C , desnaturalizada (línea gris punteada); 25°C , replegada a 25°C luego de la desnaturalización a 90°C (línea gris sólida). (B) Curvas de desnaturalización térmica de dos muestras de Bcy1 (1-50) de distintas concentraciones en buffer PBS, seguidas por elipticidad a 221 nm. Las curvas se normalizaron por el valor mínimo de elipticidad de cada una de ellas (correspondiente a la muestra nativa) para poder comparar las dos concentraciones. Los datos experimentales se muestran en puntos, y el ajuste de las curvas en líneas sólidas. Las líneas grises representan la temperatura de *melting* de cada muestra, correspondiente al valor al cual la elipticidad normalizada es la mitad del valor inicial.

Para continuar con este análisis, se registraron curvas de desnaturalización térmica para dos concentraciones de proteína (Figura 3.10 (B)). En estos experimentos, las muestras se calentaron gradualmente hasta 90°C , y se registró la elipticidad a 221 nm en función de la temperatura como indicador de la estructura (cabe aclarar en primer lugar que las dos muestras presentaron un comportamiento similar en cuanto al aspecto de sus espectros de CD a 25°C , 90°C , y 25°C post 90°C). Las curvas se normalizaron para compensar las diferencias en la señal absoluta como consecuencia de sus diferentes concentraciones. Asimismo, considerando que el proceso de desnaturalización térmica es reversible (como se evidencia en la recuperación total de la estructura luego de la desnaturalización), se ajustaron los datos experimentales a un modelo reversible de dos estados (nativo y desplegado) como

se detalla en la sección de Materiales y métodos, a efectos de calcular la temperatura de desnaturalización característica o temperatura de *melting* (T_m).

Se observa en primer lugar que la curva presenta una forma sigmoidea unimodal, lo que indica que el proceso de desnaturalización se caracteriza por un comportamiento cooperativo; esta es una característica típica de proteínas que se encuentran adecuadamente plegadas en solución. Por otro lado, se observa que la T_m depende marcadamente de la concentración total de proteína, siendo de 50,4^o C para la muestra 13 μ M, y de 54,7^o C para la muestra 63 μ M. Este comportamiento es típico de proteínas oligoméricas, dado que a mayores concentraciones el equilibrio se encuentra desplazado hacia la forma de mayor cantidad de subunidades, y por lo tanto la muestra más concentrada resulta más resistente a la desnaturalización térmica.

Finalmente, se intentó evaluar por esta técnica la existencia de equilibrios entre los distintos estados de oligomerización, para completar los resultados del análisis por SEC presentado en la **Figura 3.6**. En principio, si los equilibrios de oligomerización estuvieran involucrados en el proceso de desnaturalización, podrían haberse observado varios saltos en las curvas de desnaturalización térmica, correspondientes a transiciones graduales de complejos de mayor cantidad de subunidades a menor cantidad de subunidades. Sin embargo, este no fue el caso. Existen dos posibles explicaciones: o bien los distintos estados oligoméricos presentan esencialmente el mismo contenido de estructura secundaria, de forma tal que al desarmarse el tetrámero se generan entidades que presentan la misma señal de CD y que por lo tanto no son distinguibles por esta técnica, o bien el proceso de desnaturalización ocurre de manera cooperativa en un solo paso, de tetrámero a monómero. Para ahondar en este punto, se registraron espectros de CD de diluciones sucesivas, de manera análoga al análisis realizado por SEC. Los resultados se muestran en la **Figura 3.11**.

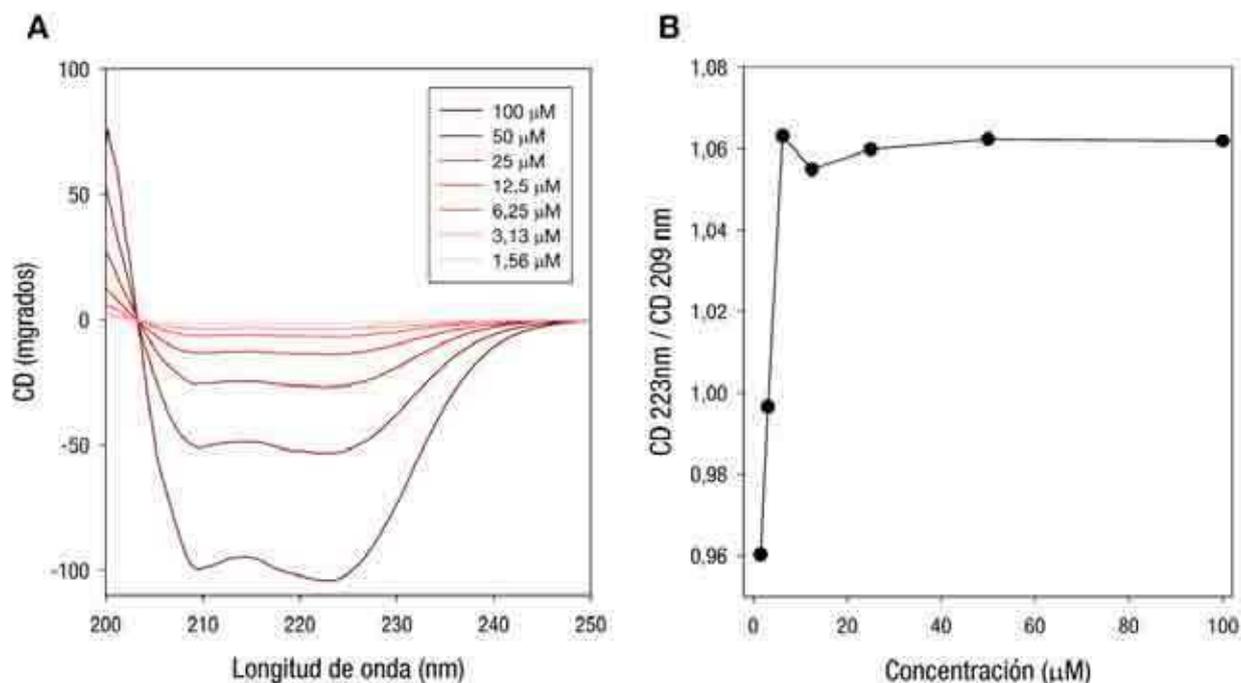


Figura 3.11: Análisis de equilibrios de oligomerización de Bcy1 (1-50) por dicroísmo circular.

(A) Espectros de CD en el UV lejano de diluciones seriadas preparadas a partir de una muestra de Bcy1 (1-50) 100 μM registrados en buffer PBS a 25 $^{\circ}$ C. (B) Intensidad de la señal de CD a 223 nm normalizada por la intensidad de la señal de CD a 209 nm, calculado para las mismas muestras del panel (A).

En primer lugar, se aprecia que en el rango de concentraciones evaluadas, el aspecto general de los espectros es muy similar para todas las muestras, correspondiente a estructuras predominantemente α -helicoidales, lo que indica que la estructura secundaria de la proteína no se ve notablemente alterada como consecuencia de la dilución (Figura 3.11 (A)). Sin embargo, existen diferencias cualitativas sutiles en los espectros que no se observan claramente a simple vista.

A efectos de poner en evidencia estas diferencias, se graficó el cociente de las señales correspondientes a los dos mínimos de los espectros (209 y 223 nm) en función de la concentración (Figura 3.11 (B)). Como puede observarse, para concentraciones de aproximadamente entre 10 y 100 μM el cociente es el mismo, lo que indica que los espectros son cualitativamente idénticos. Por el contrario, para las dos muestras más diluidas se invierte la intensidad relativa de los picos. Este cambio cualitativo en los espectros podría ser

indicativo de un cambio estructural, que presumiblemente se encuentra asociado a un cambio en el estado de oligomerización para estas diluciones.

Estas observaciones podrían encontrarse en consonancia con los resultados de las diluciones sucesivas seguidas por SEC. Como se observa en la **Figura 3.6**, los cambios más significativos en el V_e de la muestra se observan al pasar de 25 μM a 12,5 μM , y aun en mayor medida de ésta última a la 6,25 μM . Estas observaciones sugieren que a concentraciones menores a 10 μM las especies de menor cantidad de subunidades que el tetrámero podrían volverse apreciables en condiciones *in vitro*.

Propiedades hidrodinámicas de Bcy1 (1-50)

Como segunda etapa de la caracterización estructural de Bcy1 (1-50) en solución, se determinaron parámetros hidrodinámicos característicos que son informativos de la estructura global de la molécula (es decir, de su forma y de sus dimensiones moleculares): el radio de Stokes (R_s), el coeficiente de sedimentación a 20^o C en agua ($S_{20,w}$), el cociente friccional (f/f_{min}), el radio de giro (R_g), y la máxima distancia lineal de la partícula (D_{max}) (estos parámetros se resumen en la **Tabla 3.1** más adelante). Para detalles sobre la información que proporcionan estos parámetros y sobre la manera de calcularlos, consultar la Introducción metodológica.

Se define el R_s como el radio de una esfera hidratada que presenta las mismas propiedades hidrodinámicas que la molécula. Para estimar el R_s de Bcy1 (1-50) se utilizó SEC, aprovechando que esta técnica forma parte del protocolo de purificación. Existe una correlación lineal entre la posición del pico correspondiente a una determinada proteína en el cromatograma (representada por la constante de partición, K_{av}), y su R_s . Por lo tanto, se calibró la columna según el cromatograma de la **Figura 3.5**, utilizando los R_s conocidos para los estándares, junto con compuestos que permiten conocer el volumen excluido (V_0) y el volumen total (V_t) de la columna, que son necesarios para el cálculo del K_{av} . En la **Figura 3.12** se muestra la curva de calibración. Se obtuvo un valor de $R_s = 2,3$ nm.

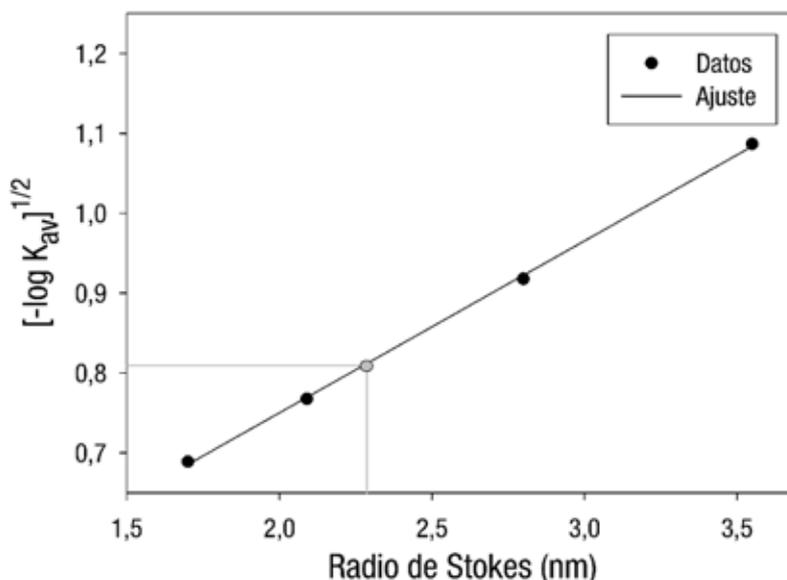


Figura 3.12: Determinación del R_s de Bcy1 (1-50).

El cromatograma de la **Figura 3.1** fue calibrado utilizando los siguientes estándares: BSA ($R_s = 3,55$ nm), ovoalbúmina ($R_s = 2,8$ nm), α -quimiotripsina ($R_s = 2,09$ nm) y citocromo C ($R_s = 1,7$ nm). Se define la constante de partición como $K_{av} = (V_e - V_0)/(V_t - V_0)$, donde V_e es el volumen de elución de cada proteína, y V_0 y V_t son los volúmenes excluido y total de la columna respectivamente. Se indican en puntos negros los datos de los estándares, y con un punto gris el valor correspondiente a la muestra. La ecuación del ajuste es: $[-\log K_{av}]^{1/2} = 0,2153 R_s + 0,3196$ ($R^2 = 0,9995$). V_e para Bcy1 (1-50) fue de 11,0 mL, lo que corresponde a $[-\log K_{av}]^{1/2} = 0,81$, y a $R_s = 2,3$ nm.

Por otro lado, se calculó el radio de la mínima esfera no hidratada capaz de contener la masa de la proteína, R_{min} , según la **Ecuación 3.6**. El valor obtenido fue de 1,89 nm, considerando que Bcy1 (1-50) es un tetrámero en solución ($MM = 23560$ Da). A partir de estos dos valores, se estimó el cociente friccional f/f_{min} como el cociente de los dos radios, de acuerdo a la **Ecuación 3.9**, obteniéndose de esta manera un valor de 1,22. Este parámetro se encuentra tabulado para una gran cantidad de proteínas de estructura conocida, y se ha observado que su valor es indicativo de la globularidad de la molécula en solución [3]. El valor obtenido se encuentra dentro de los valores comunes para proteínas globulares (1,2/1,3), lo cual es indicativo de que Bcy1 (1-50) en su forma tetramérica presenta una estructura globular en solución.

En segundo lugar, se procedió a la estimación del $S_{20,w}$ de Bcy1 (1-50) utilizando ultracentrifugación analítica en gradiente de sacarosa. Este parámetro es indicativo de la

velocidad de migración de la proteína a través del gradiente, y constituye un método alternativo para estimar el cociente friccional. Adicionalmente, en combinación con el R_s , permite estimar la MM de la proteína en solución [4] (ver más adelante).

La técnica consiste en sembrar una solución que contiene las proteínas estándar y la muestra en la parte superior de un tubo en el que se ha formado el gradiente, y posteriormente utilizar una ultracentrífuga analítica para forzar a las moléculas a moverse a través del mismo. Una vez finalizada la centrifugación, el gradiente se fracciona, y se analizan las proteínas presentes en cada fracción por SDS-PAGE. De manera análoga a lo discutido para el R_s y SEC, existe una correlación lineal entre la posición de las distintas proteínas en el gradiente y su $S_{20,w}$. Cabe aclarar que, pese a que los experimentos no se realizan en condiciones estándar (es decir, en agua a 20° C), bajo ciertas suposiciones es posible de todas maneras calcular el $S_{20,w}$ de la muestra por calibración directa, utilizando los valores de $S_{20,w}$ de las proteínas estándar, que se encuentran típicamente tabulados. Este análisis es válido siempre que todas las proteínas utilizadas (incluyendo la muestra) puedan considerarse aproximadamente globulares, y de volúmenes parciales específicos similares [29]. Como ya se discutió en los párrafos anteriores, Bcy1 (1-50) resulta ser globular a juzgar por su valor de f/f_{min} , y como todas las proteínas estándar se encuentran bien caracterizadas estructuralmente y son globulares a su vez (ver **Tabla 3.S.1**), este abordaje resulta apropiado.

En la **Figura 3.13** se muestran los resultados obtenidos para los experimentos de centrifugación, que fueron realizados en condiciones similares de buffer y concentración de proteína a las de los experimentos anteriores. Bcy1 (1-50) fue fraccionada en un solo pico, correspondiente a un valor de $S_{20,w} = 2,4$ S. Por otro lado, se estimó el máximo coeficiente de sedimentación posible para una proteína de la MM de Bcy1 (1-50) tetramérica de acuerdo a la **Ecuación 3.8**. El valor obtenido fue de $S_{max} = 2,97$ S. A partir de estos dos valores, se obtuvo una segunda estimación del f/f_{min} de acuerdo a la **Ecuación 3.9**, obteniéndose de esta manera un valor de 1,24, el cual se encuentra en absoluta concordancia con el valor obtenido por SEC (1,22).

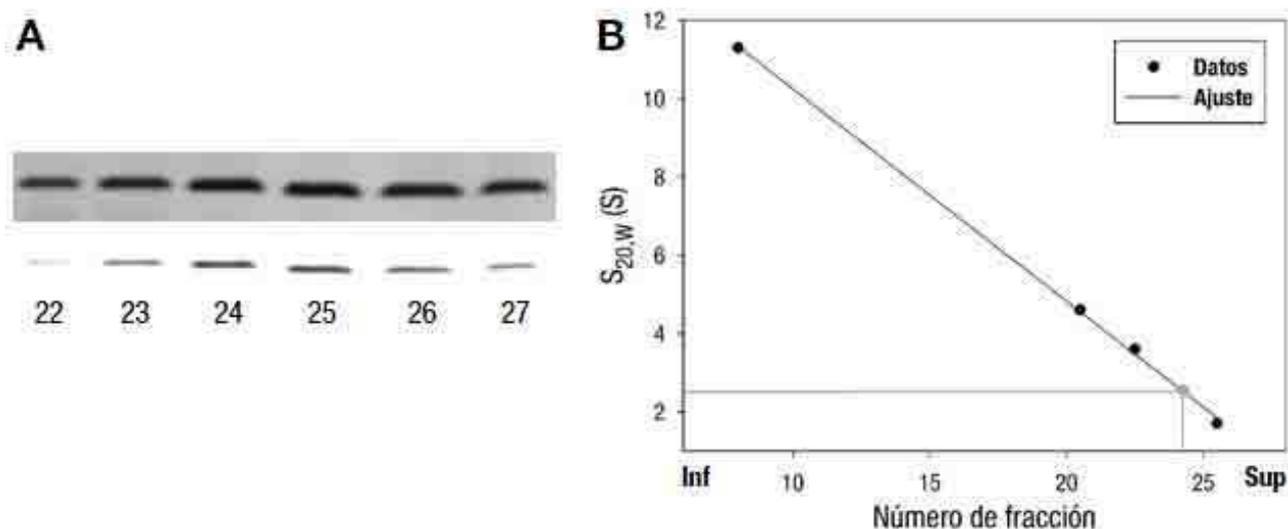


Figura 3.13: Determinación del $S_{20,w}$ de Bcy1 (1-50).

En la parte superior de un gradiente de sacarosa 5%-20% construido en buffer PBS se sembraron 200 μ L del mismo buffer conteniendo 100 μ M de Bcy1 (1-50) y cantidades adecuadas de las proteínas estándar para poder ser visualizadas. La centrifugación se desarrolló a 4^o C, y posteriormente se colectaron fracciones de aproximadamente 150 μ L desde la parte inferior del gradiente. Las fracciones fueron analizadas por SDS-PAGE 15% en buffer Tricina, y visualizadas por tinción con Coomassie Blue. **(A)** Fracciones correspondientes a Bcy1 (1-50). El panel superior corresponde a la imagen tal como fue obtenida, mientras que el panel inferior muestra una imagen con atenuación del fondo para determinar con mayor facilidad la posición de la proteína en el gradiente. **(B)** Curva de calibración. Las proteínas usadas fueron: catalasa ($S_{20,w} = 11,3$ S), BSA ($S_{20,w} = 4,6$ S), ovoalbúmina ($S_{20,w} = 3,6$ S), y citocromo C ($S_{20,w} = 1,7$ S). Se indican en puntos negros los datos de los estándares, y con un punto gris el valor correspondiente a la muestra. La ecuación del ajuste es $S_{20,w} = -0,5425$ (N^o de fracción) + 15,6758 ($R^2 = 0,9992$). El pico de Bcy1 (1-50) fue colectado entre las fracciones 24 y 25, por lo que se le asignó el valor de 24,5 a su máximo, lo que corresponde a un $S_{20,w} = 2,4$ S.

Finalmente, se calculó la MM de Bcy1 (1-50) en solución de acuerdo al método de Siegel y Monty, mencionado anteriormente [4]; el método utiliza como datos el R_s y el $S_{20,w}$, y proporciona una estimación alternativa a la obtenida por SEC-SLS. Utilizando la **Ecuación 3.10**, se obtuvo un valor de $MM = 23,2$ kDa, lo que corresponde al MM de un tetrámero con un error de aproximadamente el 1,5 %, que se encuentra dentro del rango comúnmente aceptado para este método [3], y en absoluta concordancia con los resultados de SEC-SLS.

Para completar esta parte de la caracterización estructural, se realizaron experimentos de dispersión de rayos X a bajo ángulo (SAXS). Como ya se señaló en la Introducción metodológica, esta técnica permite obtener información estructural de baja resolución de

partículas en solución, en términos de ciertos parámetros que son informativos de sus dimensiones moleculares (el radio de giro, R_g , y la máxima distancia lineal de la partícula, D_{max}), y también en lo que respecta a su forma global (representada por su función de distribución de distancias moleculares y por modelos de contorno generados a partir de la curva de dispersión). La técnica consiste en hacer incidir un haz monocromático de rayos X sobre una muestra pura de la proteína en solución, y registrar la intensidad dispersada en función del vector de onda Q ; en este perfil de dispersión se encuentra codificada la información estructural.

Dado que la intensidad de la señal de SAXS depende (entre otros factores) de la MM de la macromolécula, para una proteína relativamente pequeña como Bcy1 (1-50) es necesario utilizar concentraciones relativamente altas a efectos de obtener resultados de SAXS confiables. Por lo tanto, no fue posible realizar los experimentos en las mismas concentraciones que se vienen utilizando en el desarrollo del capítulo. Se registraron datos para muestras 1,0 mM y 0,5 mM (es decir, 10 y 5 veces más concentrada que para los experimentos de SEC-SLS). Cabe mencionar que para estas concentraciones es esperable que la formación del tetrámero se encuentre aún más favorecida que antes, por lo que en principio no existiría ambigüedad en cuanto a las formas oligoméricas presentes, dado que para las muestras 100 μ M ya existía solamente el tetrámero en base a los experimentos de SEC-SLS. Por otro lado, se trabajó en un buffer distinto al de los experimentos anteriores, dado que no pudo emplearse PBS porque el fosfato resulta una interferencia importante en los experimentos de dispersión de rayos X.

En primer lugar, se procedió a analizar la calidad de los datos; dado que ambas muestras resultaron satisfactorias (ver Materiales y métodos), se decidió trabajar con la muestra más concentrada, que presentó mejor relación señal/ruido. Para un análisis apropiado por SAXS, se requiere que las muestras sean monodispersas, sin efectos evidentes de agregación o de repulsión entre partículas; la técnica resulta muy sensible a estos dos efectos, dado que en ambos casos se observan severas desviaciones en el perfil de dispersión a ángulos bajos (ver **Figura 3.2 (A)**). Se observa en la **Figura 3.14 (A)** que los datos no presentaron evidencias de ninguno de estos efectos. Por otro lado, se realizó el análisis de Guinier, que consiste en representar $\ln [I(Q)]$ vs Q^2 (**Figura 3.14 (B)**); la linealidad de este gráfico para valores de $Q \times R_g$ menores a 1,3 es un requisito indispensable para el posterior análisis de los datos, dado que es fuerte evidencia de la monodispersión de la muestra (ver **Figura 3.2 (B)**). El gráfico de Guinier resultó ser lineal hasta $Q \times R_g = 1,6$. Como primera conclusión, se desprende de este

análisis que tanto la muestra como los datos obtenidos son de buena calidad, y que por lo tanto es válido obtener información estructural a partir de este conjunto de datos.

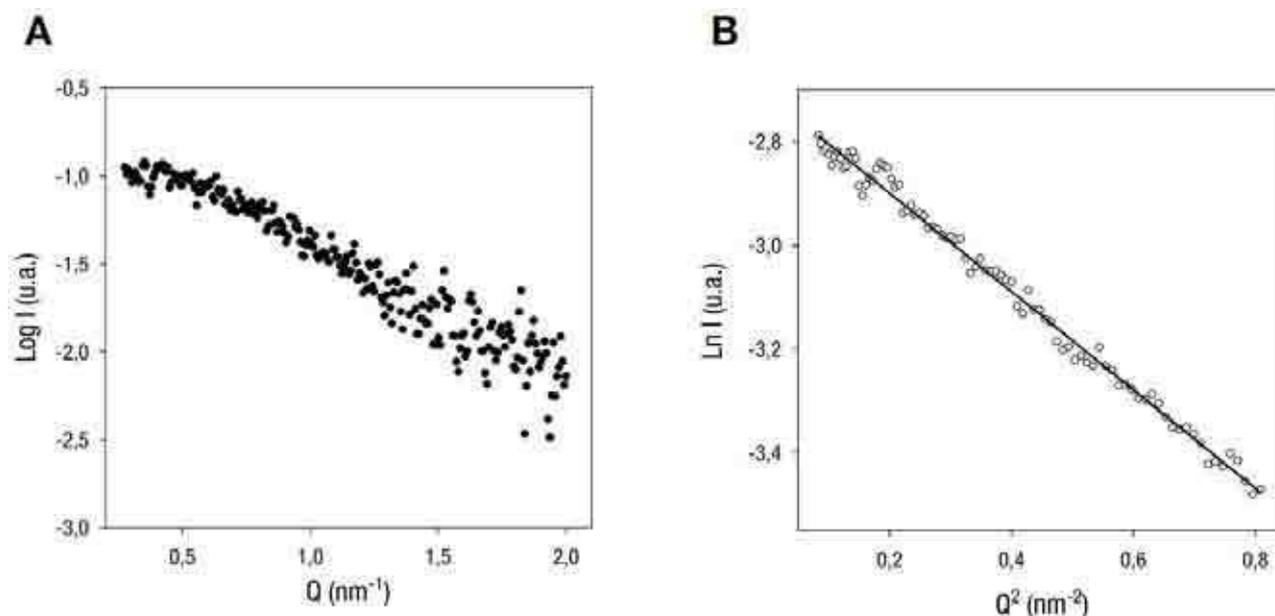


Figura 3.14: Análisis preliminar de los datos de SAXS.

Los datos fueron colectados para una muestra purificada 1 mM de Bcy1 (1-50) a 12° C en buffer HEPES 25 mM, 50 mM NaCl, pH = 7. Luego de la integración y de la sustracción de la contribución correspondiente al solvente, se obtuvo **(A)** el perfil de dispersión, y **(B)** el gráfico de Guinier. Q es el vector de onda ($Q = 4\pi\text{sen}\theta/\lambda$, donde 2θ es el ángulo que forma el haz dispersado con la normal, y λ es la longitud de onda del haz incidente) e I es la intensidad. El gráfico de Guinier corresponde a la región $Q \times R_g < 1,6$ ($Q < 0,9 \text{ nm}^{-1}$). Los dos gráficos fueron corregidos por *slit-smearing effects* como se describe en la Introducción metodológica. Los puntos en el gráfico de Guinier representan los datos experimentales, mientras que la línea sólida representa el ajuste de la recta $\ln[I(Q)] = -2,1431 - 0,9865Q^2$ ($R^2 = 0,9912$); la ordenada al origen es $\ln [I_0]$ y la pendiente es $-1/3 R_g^2$.

De la pendiente de esta recta se obtuvo un valor para el radio de giro, que es una medida de la distribución de masa de la molécula ($R_g = 1,72 \text{ nm}$) y un valor para la intensidad dispersada a ángulo 0, que es proporcional a la MM de la molécula en solución ($I_0 = 0,117$). (Para una descripción más rigurosa de estos parámetros, consultar la Introducción metodológica).

Luego de verificada la calidad de los datos, se procedió a computar la función de distribución de distancias ($P(r)$) de la molécula en solución (**Figura 3.15**). Esta función puede interpretarse como un histograma de distancias intramoleculares, que es informativo de la

forma de la molécula (ver **Figura 3.3**). La función $P(r)$ fue calculada a partir del perfil de dispersión de la **Figura 3.14 (A)** mediante el método indirecto de transformada de Fourier implementado en el programa GNOM [10].

La función $P(r)$ presenta una forma de campana muy característica de proteínas compactas y globulares en solución, sin asimetrías evidentes. La máxima distancia entre dos puntos cualesquiera de la partícula, D_{\max} , corresponde al valor de r para el cual la curva cae a cero, como se indica en la **Figura 3.15**; este valor fue de 5,5 nm. Por otro lado, el método permite obtener una estimación más exacta del R_g y de la I_0 que la que proporciona el análisis de Guinier, dado que en este caso se utiliza el conjunto completo de datos, a diferencia del segundo, que solamente considera los datos correspondientes a valores bajos de Q (que corresponden al rango en el cual es válida la aproximación de Guinier). Estos dos parámetros pueden calcularse por definición a partir del histograma como se indica en la Introducción metodológica; los valores obtenidos fueron: $R_g = 1,83$ nm (ligeramente superior al obtenido por la aproximación de Guinier) e $I_0 = 0,116$ (muy similar al anterior).

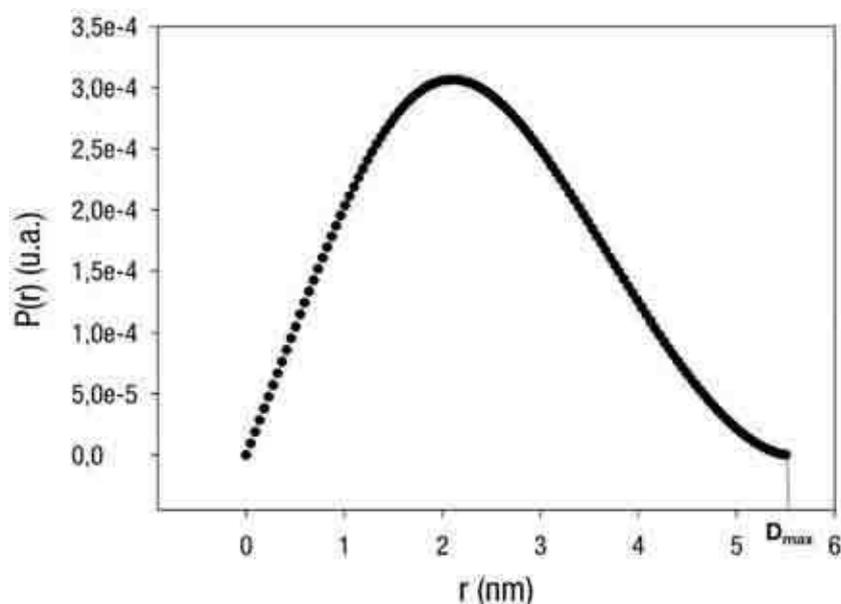


Figura 3.15: Función de distribución de distancias de Bcy1 (1-50).

La función $P(r)$ se calculó utilizando el perfil de dispersión de la **Figura 3.14 (A)** con una $D_{\max} = 5,5$ nm. El método de cálculo corrige los *slit-smearing effects*, como se explica en la Introducción metodológica. La desviación estándar de los puntos es menor que el tamaño de los símbolos, por lo que no se muestra.

Así como el cociente friccional f/f_{min} permite obtener información acerca de la forma de la molécula en solución, el valor del cociente R_g/R_s esperado para proteínas globulares es de aproximadamente 0,775 [30]. El valor obtenido para Bcy1 (1-50) es de $1,83 \text{ nm}/2,3 \text{ nm} = 0,796$, lo que resulta muy cercano al valor esperado.

Como se mencionó en la introducción metodológica, además de los parámetros relacionados con la dimensión de la macromolécula y la información acerca de su forma, SAXS permite obtener información acerca del estado oligomérico de una proteína en solución, dado que la I_0 es proporcional a la MM. Sin embargo, este abordaje requiere de una estimación precisa de la concentración de la proteína, y para proteínas difíciles de cuantificar y de baja MM, la diferencia entre la MM de las distintas formas puede no ser suficiente como para ser diferenciadas. Efectivamente, en este trabajo no se pudo determinar con precisión la MM de Bcy1 (1-50) en solución por este método dado que, además de presentar una baja MM (5890 Da), la molécula es difícil de cuantificar con exactitud, dado que al no contener residuos de triptofano su absorptividad molar es muy baja.

Modelo estructural del tetrámero Bcy1 (1-50)

Para concluir la caracterización estructural, se construyó un modelo del tetrámero de Bcy1 (1-50) utilizando la información experimental disponible. Para ello, se trabajó bajo la premisa de que la estructura clásica de los dominios D/D diméricos de mamíferos debe encontrarse subyacente en la estructura del tetrámero. Esta idea se desprende de numerosas evidencias recopiladas en este capítulo y en el anterior: los experimentos de CD revelaron que Bcy1 (1-50) presenta un alto contenido de α -hélice al igual que los dominios D/D de mamíferos (**Figura 3.10**), las predicciones de estructura secundaria sugirieron la existencia de un motivo hélice-giro-hélice, típico de los dominios D/D de mamíferos, en esta región de la molécula (**Figura 2.1**), y finalmente el análisis comparativo de las secuencias de los dominios D/D reveló que el núcleo de aminoácidos hidrofóbicos que son cruciales para determinar la estructura de los dominios D/D diméricos de mamíferos se encuentra presente en Bcy1 (1-50) (**Figura 2.1**).

En consecuencia, el tetrámero fue concebido como un “dímero de dímeros”. La estrategia consistió en generar en primer lugar un modelo por homología del dímero Bcy1 (1-50) utilizando como plantados las estructuras disponibles de los dominios D/D de mamíferos, y posteriormente utilizar este modelo dimérico como plantado para generar modelos posibles

del tetramero utilizando métodos de *docking*. Finalmente, se aprovechó la información experimental disponible (principalmente los resultados de SAXS) para seleccionar, de entre los modelos tetraméricos generados, el más representativo de la estructura real en solución.

El modelo del dímero fue generado con el programa MODELLER [31]. Este programa toma como *input* la secuencia de la proteína cuyo modelo se quiere construir (*target*), el o los archivos pdb correspondiente/s a la/s estructura/s de alta resolución de las proteínas relacionadas que servirán como molde para la construcción del modelo (templados), y un alineamiento de las secuencias de ambas. A partir de esta información, el programa básicamente superpone la secuencia del *target* a la estructura del templado, cuidando de satisfacer ciertas restricciones.

Para generar el modelo del dímero Bcy1 (1-50), se utilizaron como templados las estructuras cristalina y en solución del dominio D/D de la subunidad RII α . Se eligió este templado dado que, como se señaló en el Capítulo 2, el dominio D/D de Bcy1 comparte más características con éste que con el dominio D/D de RI α . Posteriormente, los modelos fueron sometidos a refinamiento y la calidad de los modelos resultantes fue evaluada utilizando numerosos programas (ver Materiales y métodos para más detalles).

Pese a que la similitud entre las secuencias de los dominios D/D de Bcy1 y RII α no es demasiado elevada (38%), el modelo es bastante confiable debido a que las características estructurales más distintivas de estos dominios sí se encuentran bien conservadas, como ya se discutió en el Capítulo 2; reforzando esta idea, también se comentó que los dominios D/D de RI α y RII α alinean estructuralmente muy bien pese a presentar una baja similitud de secuencia (35%) [32].

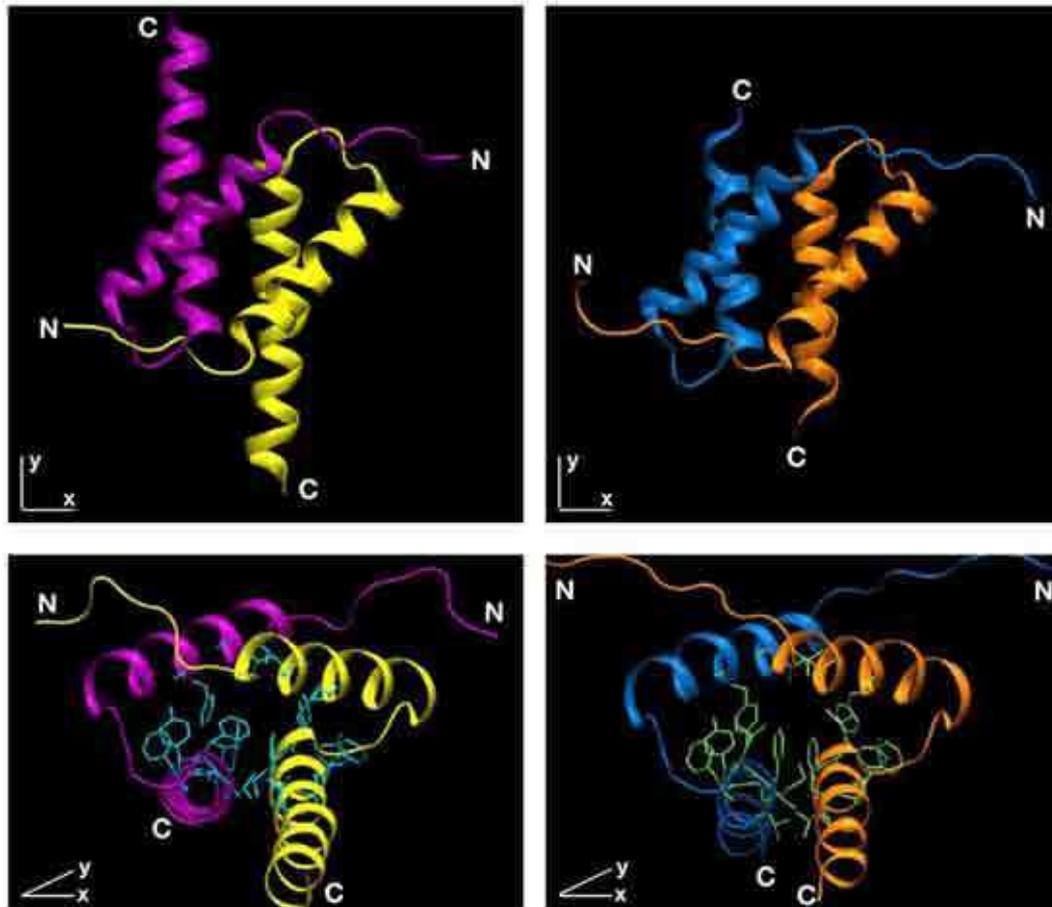


Figura 3.16: Modelo dimérico de Bcy1 (1-50).

La estructura del modelo del dímero Bcy1 (1-50) (paneles de la izquierda) y el modelo atómico de la estructura en solución del dominio D/D de RII α obtenido por NMR (pdb: 1R2A) (paneles de la derecha) se muestran en diagrama de cintas. Las cadenas se representan en diferentes colores por claridad (Bcy1: violeta y amarillo, RII α : naranja y azul), y se indican también los extremos N y C terminales de cada una de las cadenas. Los paneles superiores muestran una vista de la superficie expuesta al solvente que constituye el sitio de unión a AKAPs, mientras que los paneles inferiores se encuentran rotados 90° respecto del eje x para mostrar el núcleo de aminoácidos hidrofóbicos involucrados en la dimerización. Los residuos hidrofóbicos más importantes que constituyen este núcleo se muestran como líneas en los paneles inferiores, en turquesa para Bcy1 y en verde para RII α .

La distribución de elementos de estructura secundaria en el modelo es similar a la del dominio D/D de RII α , y coincide con los resultados de las predicciones del Capítulo 2; la principal diferencia entre ambas, como ya se discutió en ese capítulo, se localiza hacia el final de la hélice 2, donde se modelaron dos giros adicionales en Bcy1 (1-50) (Figura 3.16). El modelo muestra un núcleo bien definido de residuos hidrofóbicos en la interfaz entre los dos

monómeros, similar al del dominio D/D de RII α , que contiene principalmente residuos aromáticos.

Para generar el modelo del tetramero, se utilizó el servidor automatizado de *docking* de proteínas ClusPro [33]. Este programa genera, para un receptor y un ligando, un cierto número de modelos del complejo entre ambos (del orden de 100), que son clasificados en distintas categorías en función del tipo de interacciones intermoleculares predominantes entre las unidades que lo conforman (electrostático, hidrofóbico, Van der Waals/electrostático, equilibrado). El usuario debe posteriormente seleccionar en base a información experimental u otro tipo de criterios cuáles de esos posibles complejos resultan biológicamente relevantes. Al igual que la mayoría de los métodos de *docking*, el algoritmo consiste en dividir el espacio en una grilla tridimensional, sobre la cual se evalúan un elevado número de rotaciones y traslaciones del ligando respecto del receptor, que se mantiene fijo; de esta manera, se genera un altísimo número (del orden de 10⁷) de potenciales complejos. A continuación, el programa utiliza distintas funciones de *scoring* para evaluar la calidad de los complejos resultantes y seleccionar los mejores; estas funciones típicamente tienen en cuenta la complementariedad entre las superficies de los dos componentes del complejo, en términos de sus superficies de Van der Waals, junto con otros factores que pueden incluir las energías de desolvatación de los componentes asociadas a la formación del complejo.

El modelo del dímero de Bcy1 (1-50) fue proporcionado como *input* para el programa como receptor y como ligando; el servidor generó 93 modelos que exhibieron una amplia diversidad de estructuras como consecuencia de las diferentes orientaciones relativas de los dímeros en cada uno de ellos. Para seleccionar los modelos más representativos de la estructura real en solución, se aprovechó la información experimental de SAXS, dado que de las técnicas utilizadas resulta la más informativa en cuanto a la forma de las moléculas en solución. Por lo tanto, se llevó a cabo un protocolo de selección para seleccionar los modelos que mejor concordancia presentaran con los datos experimentales de SAXS.

Los modelos fueron enviados al servidor FoXS [34]; este programa permite calcular rápidamente perfiles teóricos de SAXS de proteínas a partir de sus archivos pdb, y compararlos con los datos experimentales. El servidor proporciona un parámetro de bondad del ajuste de los datos teóricos del modelo a los experimentales (χ^2), y además una estimación del R_g de la estructura contenida en el archivo pdb. Dado que muchos modelos presentaron

estructuras poco compactas como consecuencia de la orientación relativa de los dímeros, sus R_g resultaron anormalmente altos en comparación con el dato experimental (1,83 nm); por lo tanto, se decidió seleccionar un valor de corte de $R_g = 1,9$ nm, y se descartaron los 48 modelos que presentaron valores de R_g más altos que el corte. De los 45 modelos restantes, aquel cuyo perfil teórico de SAXS ajustó mejor a los datos experimentales de acuerdo al criterio proporcionado por FoXS, fue seleccionado como el más representativo de la estructura en solución; se obtuvo un valor de $\chi^2 = 1,83$ para el mejor modelo (**Figura 3.17 (A)**), lo que es indicativo de la buena concordancia entre ambos conjuntos de datos (típicamente, el ajuste se considera razonable para $\chi^2 < 2$).

Como ya se señaló en la Introducción metodológica, SAXS permite obtener modelos del contorno de la macromolécula en solución. Un segundo nivel de validación consistió en la comparación del modelo atómico generado por ClusPro con el modelo de la forma global proporcionado por SAXS. Para construirlo, se utilizó el método *ab initio* implementado en el programa DAMMIF [13], del paquete de análisis de datos de SAXS ATSAS [11] (para detalles sobre el funcionamiento del programa consultar la Introducción metodológica). Este programa utiliza como *input* la función $P(r)$ de la partícula generada por GNOM [10] a partir del perfil de dispersión de SAXS, y opcionalmente información adicional acerca de la simetría de la partícula; en este caso, se seleccionó la simetría P4 (tetramero).

Se generaron 20 modelos de contorno, que presentaron formas globales similares. De manera análoga a como se procedió para la selección del modelo de ClusPro más adecuado, se eligió el que mejor ajustó a los datos experimentales de SAXS; para ello se utilizó nuevamente el servidor FoXS, que también permite computar el perfil de dispersión teórico para este tipo de estructuras. El ajuste del perfil del mejor modelo a los datos experimentales resultó satisfactorio ($\chi^2 = 1,86$) (**Figura 3.17 (A)**). El modelo final presenta una forma aproximadamente cilíndrica, ligeramente alargada en la dirección de la D_{\max} (**Figura 3.17 (B)**). Cabe aclarar que los detalles que pueden observarse acerca de la rugosidad de la superficie no son informativos, dado que la resolución de la técnica de SAXS (que se estima como $d = 2\pi/Q_{\max}$, donde Q_{\max} es el máximo valor que toma el vector de onda para el conjunto de datos; en este caso, $d \sim 3$ nm) no permite obtener esa información; los detalles que se observan solamente son consecuencia de la forma elegida para representar la molécula.

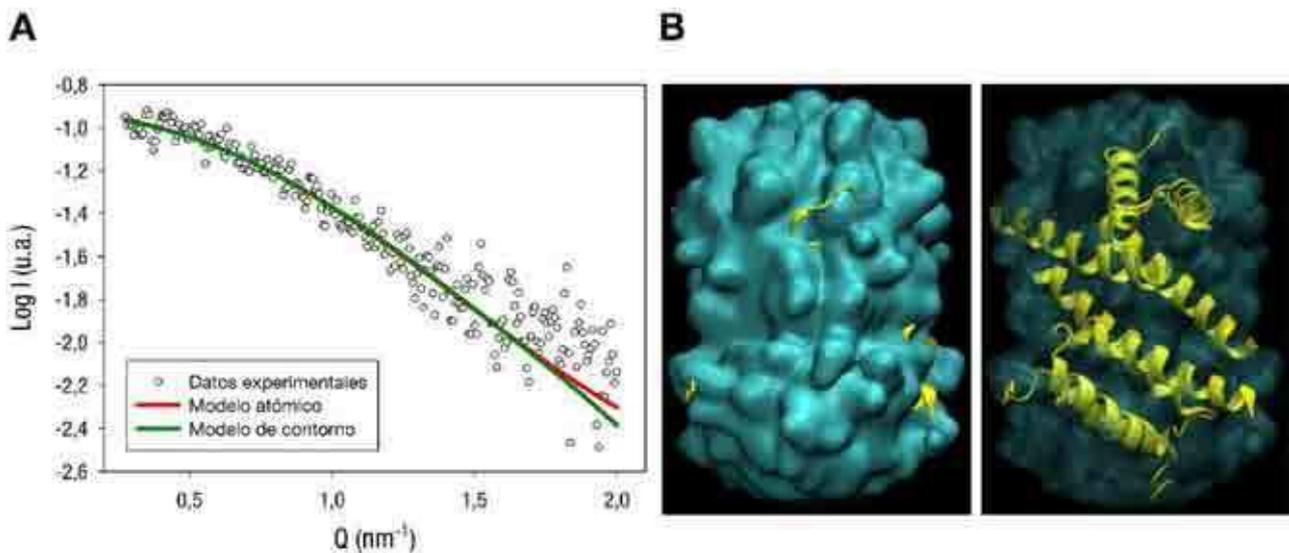


Figura 3.17: Validación de los modelos del tetrámero Bcy1 (1-50).

(A) Perfil experimental de SAXS de Bcy1 (1-50) corregido por *slit-smearing effects* (puntos blancos), superpuesto con la curva teórica de los modelos atómico (línea roja; $\chi^2 = 1,83$) y de contorno (línea verde; $\chi^2 = 1,86$). Los R_g calculados a partir de las curvas teóricas de los dos modelos fueron 1,87 nm y 1,89 nm respectivamente. (B) Superposición de los modelos de contorno (turquesa) y atómico (diagrama de cintas amarillo) generada usando el programa SUPCOMB [35]. Se muestran dos vistas para resaltar que el segundo se encuentra bien contenido en el primero; en la primera vista el modelo de superficie se muestra sólido, y en la segunda, traslúcido.

Finalmente, para comparar ambos modelos, se utilizó el programa SUPCOMB [35], que permite encontrar la mejor superposición espacial entre modelos atómicos y de contorno. Como se observa en la **Figura 3.17 (B)**, la superposición entre los dos modelos resulta bastante buena, dado que el modelo de contorno contiene casi en su totalidad al modelo atómico.

El modelo del tetrámero seleccionado de acuerdo al procedimiento descrito está dominado por interacciones de Van der Waals/electrostáticas entre las dos unidades diméricas, de acuerdo con la clasificación establecida por el servidor ClusPro. Es importante mencionar que 4 modelos de esa misma categoría y con una orientación relativa de los dímeros muy similar se ubicaron entre los 10 mejores modelos de acuerdo al criterio de χ^2 proporcionado por FoXS. Por el contrario, los modelos dominados por otros tipos de interacciones que también presentaron un buen ajuste a los datos experimentales resultaron

ser únicos en cuanto a su plegamiento. El hecho de que se hayan identificado varios modelos similares dominados por el mismo tipo de interacciones refuerza la idea de que estos modelos son los más representativos de la estructura real en solución.

Si bien estos modelos no resultan óptimos para obtener información estructural acerca de interacciones puntuales, algunas características generales sí pueden ser analizadas. Los dos dímeros que constituyen el tetrámero interactúan a través de la cara expuesta al solvente de la hélice 2 de los dos monómeros que los constituyen (**Figura 3.18; panel de la izquierda**), y se encuentran alineados de forma antiparalela mediante un arreglo simétrico (en otras palabras, el extremo N-terminal de la hélice 2 de una de las cadenas de uno de los dímeros se encuentra alineado con el extremo C-terminal de la hélice 2 de una de las cadenas del otro dímero) (**Figura 3.18; panel de la derecha**). Más aun, consistentemente con la clasificación de este modelo proporcionada por el servidor ClusPro, puede observarse que numerosos residuos polares y cargados se acumulan en el centro de este núcleo de 4 hélices; esto contrasta con los núcleos de residuos hidrofóbicos que sostienen a los dímeros en sí mismos (**ver Figura 3.16**).

En base a este modelo, es posible concebir la idea de una Bcy1 WT tetramérica en la cual el núcleo formado por los dominios D/D diméricos yace en el centro de la estructura, y los *linkers* de las cuatro subunidades irradian hacia el exterior desde el extremo C-terminal del segmento (1-50) de cada cadena (**Figura 3.18, panel de la derecha**). Finalmente, es interesante destacar que en estos modelos las posibles superficies de unión a AKAPs de cada uno de los dímeros que conforman el tetrámero permanecen expuestas al solvente, con sus caras apuntando hacia lados opuestos del complejo (ver por ejemplo las superficies generadas por las dos hélices N-terminales de cada dímero en la **Figura 3.18, panel de la izquierda**). Más aun, como consecuencia de la naturaleza tetramérica de esta proteína, existen dos de estas superficies por cada complejo (a diferencia de lo que ocurre en las subunidades R diméricas de mamíferos), y se forman nuevas superficies expuestas al solvente en los sitios en los que los dos dímeros entran en contacto, que podrían proporcionar nuevos sitios de interacción.

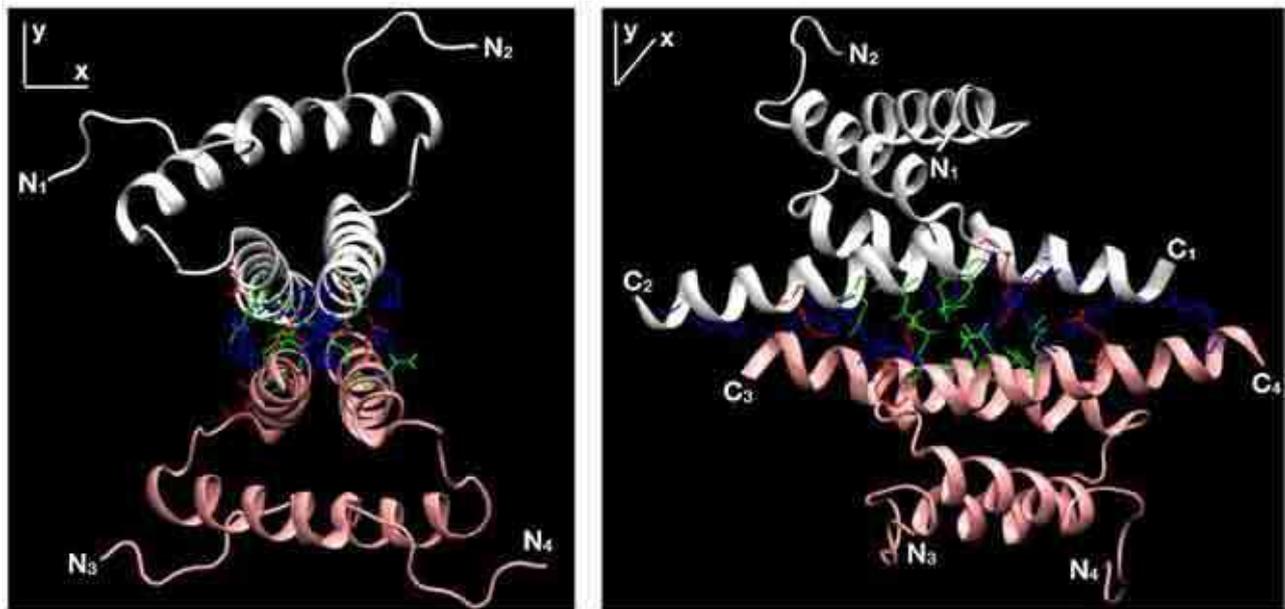


Figura 3.18: Modelo tetramérico de Bcy1 (1-50).

Se muestran en diagramas de cintas dos vistas del modelo del tetrámero Bcy1 (1-50). Los dos dímeros que constituyen el tetrámero se encuentran coloreados en rosa y blanco, y los extremos N y C terminales de cada una de las cadenas se encuentran indicados con la numeración para cada una de las cadenas en subíndice. Los residuos polares y cargados que yacen en la interfaz de tetramerización se muestran explícitamente como líneas según el código de colores comúnmente aceptado (verde: polar, no cargado; rojo: ácido; azul: básico). El panel de la izquierda muestra el arreglo de las cuatro hélices 2 que constituyen el núcleo de tetramerización. El panel de la derecha se encuentra rotado 90° alrededor del eje y respecto del anterior, para mostrar con mayor claridad las características de los aminoácidos que se encuentran en esa interfaz.

Comparación de parámetros hidrodinámicos de dominios D/D

Finalmente, para completar la caracterización estructural, se procedió a comparar los distintos parámetros obtenidos experimentalmente en las últimas secciones para Bcy1 (1-50) con los correspondientes a los modelos de dímero y tetrámero, y a los dominios D/D de mamíferos. Pese a que existen numerosas estructuras de alta resolución para estos últimos, no se han reportado hasta el momento estas propiedades. Por lo tanto, se utilizó el programa HYDROPRO [36] para computar estos parámetros a partir de las estructuras de alta resolución; también se utilizó para obtener los correspondientes a los modelos dimérico y tetramérico de Bcy1 (1-50).

Este programa permite calcular una serie de propiedades hidrodinámicas utilizando como *input* un archivo pdb correspondiente a la estructura de la molécula de interés. La estrategia de cálculo se basa en que las fuerzas friccionales que tienen lugar en la superficie de las macromoléculas son las responsables del comportamiento hidrodinámico de la partícula en solución; por lo tanto, lo que se necesita para computar las propiedades hidrodinámicas es modelar la superficie [36]. Para ello, a partir del archivo de coordenadas atómicas se genera en primer lugar un modelo hidrodinámico de la partícula reemplazando todos los átomos por esferas idénticas de un determinado radio, que se superponen entre ellas para ocupar todo el espacio interior de la molécula; este modelo es una representación fiel del tamaño y la forma de la molécula hidratada en solución. En segundo lugar, este modelo se transforma en un modelo de “cáscara”, formado por esferas más pequeñas dispuestas sobre la superficie del modelo anterior, que representan detalladamente el contorno de la molécula hidratada; a partir de este modelo se computan las propiedades hidrodinámicas.

Un aspecto crucial del comportamiento hidrodinámico de las partículas es su hidratación; este efecto es particularmente importante para las proteínas, dado que en solución se encuentran fuertemente hidratadas por una capa de una o dos moléculas de agua que, a los fines prácticos, pueden considerarse unidas constitutivamente a la molécula y, por lo tanto, en lo que respecta al comportamiento hidrodinámico forman parte íntegra de la misma. Para considerar este efecto, en este programa se elige el radio de las esferas de forma tal de expandir el volumen del modelo atómico inicial; de esta manera, el modelo de cáscara tiene un volumen mayor al de la molécula aislada, lo que contempla los efectos de hidratación.

Tabla 3.1: Propiedades hidrodinámicas de los dominios D/D de las subunidades R.

Molécula	MM (kDa)	$S_{20,W}$ (S)	R_s (nm)	f/f_{min}	R_g (nm)	D_{max} (nm)
Bcy1 (1-50)	23,2 (23,5)^a	2,4 ± 0,2^b	2,3 ± 0,1^c	1,23^d	1,83 ± 0,01^e	5,5^e
D/D RI α ^g	12,0 ^f	1,46 ± 0,03	2,15 ± 0,04	1,36 ⁱ	1,59 ± 0,03	5,7 ± 0,2
D/D RII α ^g	10,8 ^f	1,31 ± 0,02	2,16 ± 0,04	1,41 ⁱ	1,63 ± 0,04	6,0 ± 0,5
Bcy1 (1-50) modelo dímero ^h	11,8 ^f	1,46	2,09	1,34 ⁱ	1,58	5,9
Bcy1 (1-50) modelo tetrameroⁱ	23,6^f	2,37 ± 0,01	2,58 ± 0,02	1,31ⁱ	1,97 ± 0,02 (1,84 ± 0,02)	7,2 ± 0,2
Ribonucleasa A	13,7 ^f	2,0 ^k	1,9 ^l	1,2 ⁿ	1,48 ^m	-
Quimiotripsinógeno A	25,7 ^f	2,6 ^k	2,09 ^k	1,2 ^k	1,81 ^m	-

- ^a Calculada por el método de Siegel y Monty [4]; el valor entre paréntesis indica el resultado de SEC-SLS.
- ^b Determinado por ultracentrifugación analítica en gradiente de sacarosa; se muestra la media de tres experimentos independientes y su desvío estándar.
- ^c Determinado por SEC; se muestra la media de tres experimentos independientes y su desvío estándar.
- ^d Calculado como el promedio de $S_{\max}/S_{20,W}$ (1,24) y R_s/R_{\min} (1,22).
- ^e Determinados por SAXS. El error del R_g es proporcionado por el programa GNOM; la D_{\max} , al ser un parámetro proporcionado por el usuario, no tiene error asociado.
- ^f Las MM teóricas fueron determinadas a partir de la secuencia de aminoácidos de las proteínas.
- ^g $S_{20,W}$, R_s , R_g y D_{\max} para las estructuras de alta resolución en solución de los dominios D/D de las subunidades R de mamíferos fueron calculados para cada uno de los modelos contenidos en los archivos pdb (2EZW para RI α y 1R2A para RII α) utilizando el programa HYDROPRO [36] con los parámetros por defecto, y el resultado presentado como la media de todos los modelos con su desvío estándar.
- ^h $S_{20,W}$, R_s , R_g y D_{\max} para el modelo dimérico de Bcy1 (1-50) fueron calculados utilizando el programa HYDROPRO [36] con los parámetros por defecto
- ⁱ $S_{20,W}$, R_s , R_g y D_{\max} para el modelo tetramérico de Bcy1 (1-50) fueron calculados para cada uno de los mejores 4 modelos que presentaron una estructura similar utilizando el programa HYDROPRO [36] con los parámetros por defecto, y el resultado presentado como la media de todos los modelos con su desvío estándar. El valor entre paréntesis para el R_g corresponde al estimado por FoXS (cita), expresado como la media de los 4 modelos con su desvío estándar.
- ^j Calculados como el promedio de $S_{\max}/S_{20,W}$ y R_s/R_{\min} .
- ^k [3].
- ^l [37].
- ^m [36].
- ⁿ Calculado como R_s/R_{\min} .

En la **Tabla 3.1** se muestran los resultados de los parámetros hidrodinámicos experimentales obtenidos para Bcy1 (1-50), junto con los calculados para los modelos de dímero y tetramero, y para los dominios D/D de las subunidades R de mamíferos, a partir de los archivos pdb de cada una de las estructuras; también se muestran como comparación los parámetros correspondientes a la ribonucleasa A y al quimiotripsinógeno A, dos proteínas de MM similar a los dominios D/D diméricos de las subunidades R de mamíferos, y al tetramero de Bcy1 (1-50) respectivamente.

En primer lugar se observa que, en líneas generales, los parámetros correspondientes al modelo del tetramero Bcy1 (1-50) concuerdan con los valores experimentales, proporcionando nueva evidencia acerca de la validez del modelo. En particular, el valor del $S_{20,W}$ del modelo se encuentra en excelente acuerdo con el valor experimental, mientras que el R_s del modelo resulta algo superior al valor experimental. Por otro lado, el R_g presenta algunas

diferencias dependiendo de si el cálculo fue efectuado con FoXS o con HYDROPRO, dado que los dos algoritmos utilizan distintos tratamientos de la capa de hidratación de la proteína [34]: el R_g proporcionado por FoXS concuerda muy bien con el valor experimental, mientras que el R_g proporcionado por HYDROPRO es un poco más elevado. Las discrepancias entre los valores calculados para el modelo y el valor experimental para estos parámetros pueden deberse, además de a imprecisiones en los cálculos, a un empaquetamiento imperfecto de los dos dímeros que conforman el tetrámero como consecuencia de errores en el modelo dimérico, generando de esta manera un tetrámero menos compacto que la estructura real, caracterizado por lo tanto por valores de los parámetros hidrodinámicos mayores a los reales.

Finalmente, el parámetro del modelo que presenta mayor discrepancia con el valor experimental es la D_{max} . En primer lugar cabe destacar que, de todos los parámetros calculados por HYDROPRO, este fue el que presentó mayor variabilidad. Durante los cálculos se observó que para los dominios D/D de las subunidades R de mamíferos, la D_{max} dependía marcadamente de la disposición de los extremos N-terminales desestructurados que preceden a la hélice 1; estos segmentos adoptan múltiples conformaciones en los diferentes modelos obtenidos por NMR (notar el valor del desvío estándar para el D/D de RII α). Por la forma en que se construyó el modelo de Bcy1 (1-50), la conformación de esta región de la molécula no fue tomada en demasiada consideración, por lo que representa una posible fuente de error en la D_{max} . Por otro lado, es posible que el valor experimental de la D_{max} de Bcy1 (1-50) (5,5 nm) se encuentre ligeramente subestimado; una evidencia de esto es que el modelo generado *ab initio* a partir de los datos de SAXS utilizando el programa DAMMIF (ver **Figura 3.17**) presentó una $D_{max} = 6$ nm.

Como era esperable, los parámetros calculados para los dominios D/D de las subunidades R de mamíferos resultaron muy similares a los correspondientes al modelo del dímero de Bcy1 (1-50) y, con la excepción de la D_{max} , menores en todos los casos que los correspondientes al modelo tetramérico, lo que se encuentra en concordancia con la estructura oligomérica de las distintas proteínas en solución.

Finalmente, resulta interesante comparar el R_g de todas las proteínas estudiadas con los de otras proteínas globulares que se encuentran estructuralmente muy bien caracterizadas. El R_g de los dominios D/D diméricos es muy similar al de la ribonucleasa A, que tiene una MM similar a ellas; cabe destacar que, a juzgar por los valores de f/f_{min} , los dominios D/D

diméricos son algo más elongados que la ribonucleasa, lo que se encuentra en concordancia con el valor de R_g ligeramente más elevado para los primeros respecto de la segunda, pese a tener una MM ligeramente inferior. En lo que respecta a Bcy1 (1-50), su R_g resulta casi idéntico al del quimiotripsinógeno A, otra proteína globular de MM similar; ambos son a su vez mayores a los de las proteínas anteriores. Este análisis refleja de qué manera puede emplearse el R_g como criterio de tamaño molecular.

CONCLUSIONES Y PERSPECTIVAS

En este capítulo se demostró la existencia de una subunidad R de la PKA con una novedosa estructura tetramérica, en contraste con las subunidades R de mamíferos que son diméricas. El tetrámero de Bcy1 se sostiene mediante la interacción de los dominios D/D de sus monómeros; el estado oligomérico de esa región de la molécula (el fragmento Bcy1 (1-50)), fue verificado por distintas técnicas, y se llevó a cabo una caracterización exhaustiva de su estructura en solución. Se trata de un dominio altamente globular constituido exclusivamente por α -hélices, en el que posiblemente se encuentre presente la estructura clásica de los dominios D/D de mamíferos. Se reportaron las propiedades hidrodinámicas de este fragmento y se construyó un modelo de su estructura en solución, que se encuentra fuertemente avalado por la evidencia experimental. El modelo sugiere que la interfaz de tetramerización se encuentra constituida principalmente por residuos polares y cargados, en contraste con la interfaz de dimerización, que está dominada por aminoácidos hidrofóbicos.

El estado oligomérico de Bcy1 parece presentar un comportamiento dinámico en solución, y al ser perturbado *in vivo* tiene efectos apreciables sobre el crecimiento de las levaduras. Esto sugiere que la oligomerización de esta proteína podría constituir algún mecanismo regulatorio de procesos fisiológicos relevantes.

En el siguiente capítulo se describirá la estructura cristalina de este fragmento, con el objetivo de conocer en detalle los determinantes responsables de la tetramerización, para poder así analizar cuáles son las características distintivas del dominio D/D de Bcy1 que determinan que su estructura cuaternaria sea diferente de la de las subunidades R de mamíferos.

Los resultados de este capítulo sugieren numerosas preguntas que no serán abordadas en este trabajo pero que sirven de disparador para trabajos posteriores, relacionadas principalmente con el rol fisiológico que desempeña la oligomerización de la subunidad R de la PKA en *S. cerevisiae*. Un primer paso para avanzar en esta dirección es el estudio del estado oligomérico de Bcy1 *in vivo*, que será abordado en los próximos años en nuestro laboratorio.

MATERIALES Y MÉTODOS

Cromatografía de exclusión molecular

Los cromatogramas se registraron en el mismo equipo y bajo las mismas condiciones experimentales detalladas en el Capítulo 2 para las purificaciones de Bcy1 (1-50). Para estimar el R_s , la columna fue calibrada con proteínas de R_s conocido: BSA (3,55 nm), ovoalbúmina (2,80 nm), α -quimiotripsina (2,09 nm), y citocromo C (1,70 nm) (**Figura 3.S.1**); todos los estándares fueron de Sigma. Se construyó un gráfico de $(-\log K_{av})^{1/2}$ vs R_s , donde $K_{av} = (V_e - V_0)/(V_t - V_0)$, donde V_e es el volumen de elución de la proteína, y V_0 y V_t son los volúmenes excluidos y total de la columna respectivamente. El V_0 fue determinado usando Blue Dextran, mientras que el V_t fue determinado usando acetona. Para este tipo de determinaciones, se inyectaron solamente 100 μ L de Bcy1 (1-50) en la columna utilizando un loop apropiado, para obtener mejor resolución que en las purificaciones.

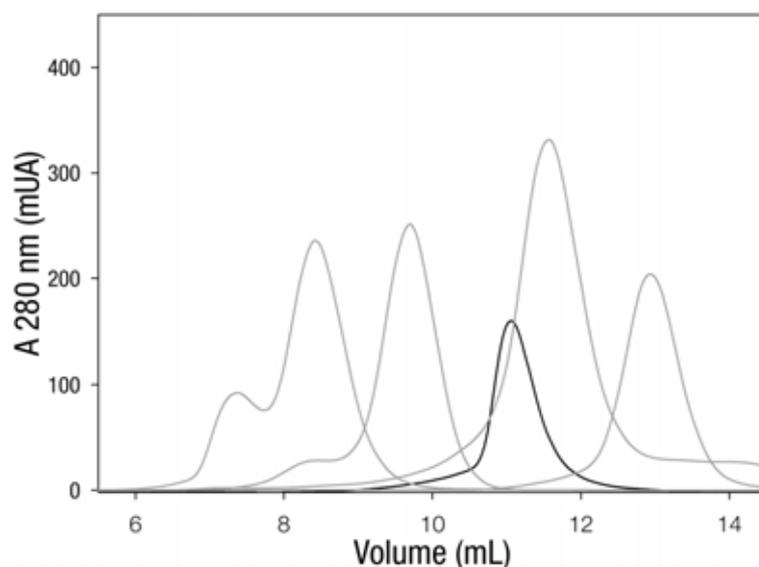


Figura 3.S.1: Cromatogramas de las proteínas utilizadas como estándares.

Se muestra en línea negra el pico correspondiente a Bcy1 (1-50), y en líneas grises los estándares; de izquierda a derecha: BSA (el primer pico, que eluye con el V_0 , corresponde a oligómeros que se observan siempre en las preparaciones de alta concentración de esta proteína, mientras que el pico siguiente corresponde al monómero), ovoalbúmina, α -quimiotripsina y citocromo C.

Tabla 3.S.1: Propiedades hidrodinámicas de las proteínas utilizadas como estándares.

Proteína	MM (kDa)	R_s (nm)	$S_{20,W}$ (S)	f/f_{min}
Catalasa	230 (tetramero)	5,20	11,3	1,20
BSA	66,4	3,55	4,6	1,29
Ovoalbúmina	45	2,80	3,6	1,27
α -	25	2,09	2,6	1,21
Quimiotripsina				
Citocromo c	12,3	1,70	1,7	1,13

Fuente: [3].

Experimentos de dispersión estática de la luz

Las MM promedio de Bcy1 (1-50) y Bcy1 WT en solución fueron determinadas utilizando un detector de dispersión de la luz a 90° (LS) Precision Detectors PD2010 y un refractómetro diferencial LKB 2142 (RI) conectados en tándem a un equipo HPLC. Las columnas utilizadas fueron: Superdex-75 GL 10/300 para Bcy1 (1-50) y Superdex-200 GL 10/300 para Bcy1 WT (GE Healthcare). 100 μ L de cada una de las proteínas purificadas fueron inyectados en la columna utilizando un loop apropiado, y los cromatogramas fueron desarrollados en buffer PBS (20 mM fosfato, 150 mM NaCl, pH = 7,4) en condiciones isocráticas a un flujo de 0,4 mL/min a 20° C. La concentración de las muestras fue de \sim 100 μ M para Bcy1 (1-50) y \sim 20 μ M para Bcy1 WT. Para cada uno de los detectores, a la señal correspondiente al pico de cada muestra se le sustrajo la correspondiente al buffer, y la MM en cada punto del cromatograma fue calculada mediante el cociente LS/RI por comparación de este valor con el que se obtuvo para la BSA (MM = 66,5 kDa), utilizando el programa Discovery32 proporcionado por el fabricante del equipo.

Experimentos de entrecruzamiento químico

Los experimentos de entrecruzamiento químico utilizados para evaluar el estado oligomérico de Bcy1 WT en su contexto celular se realizaron utilizando el mismo protocolo detallado en el Capítulo 2, con la diferencia de que como fuente de proteína se utilizaron \sim 50 μ g de proteína total contenidos en un extracto clarificado de la cepa 1115-BCY1 (para detalles

sobre la cepa utilizada y sobre el procedimiento de obtención del extracto, ver Materiales y métodos del Capítulo 2). Una vez completada la reacción de entrecruzamiento, los productos se capturaron por incubación con ~30 µg de resina de afinidad cAMP-agarosa; las proteínas se separaron de la resina por incubación con buffer de siembra de electroforesis y posterior hervor, y fueron sembradas en un gel SDS-PAGE 10%. Las electroforesis se desarrollaron como se indica en la sección Materiales y métodos del Capítulo 2. Los productos de la reacción se analizaron por *western blot* utilizando un anticuerpo primario anti-Bcy1 (Santa Cruz Biotechnology) en una dilución 1:200 del *stock*, y un anticuerpo secundario conjugado a peroxidasa anti-cabra (Sigma) en una dilución 1:2000 del *stock*.

Construcción de la cepa 1115-Bcy1(1-50)

Para generar una construcción que permita expresar Bcy1 (1-50) en levaduras, el fragmento correspondiente a la proteína de fusión se subclonó a partir del vector pET32a-Bcy1(1-50) (ver Materiales y métodos del Capítulo 2) entre los sitios EcoRI y HindIII del plásmido pYES2 (Figura 3.S.2).

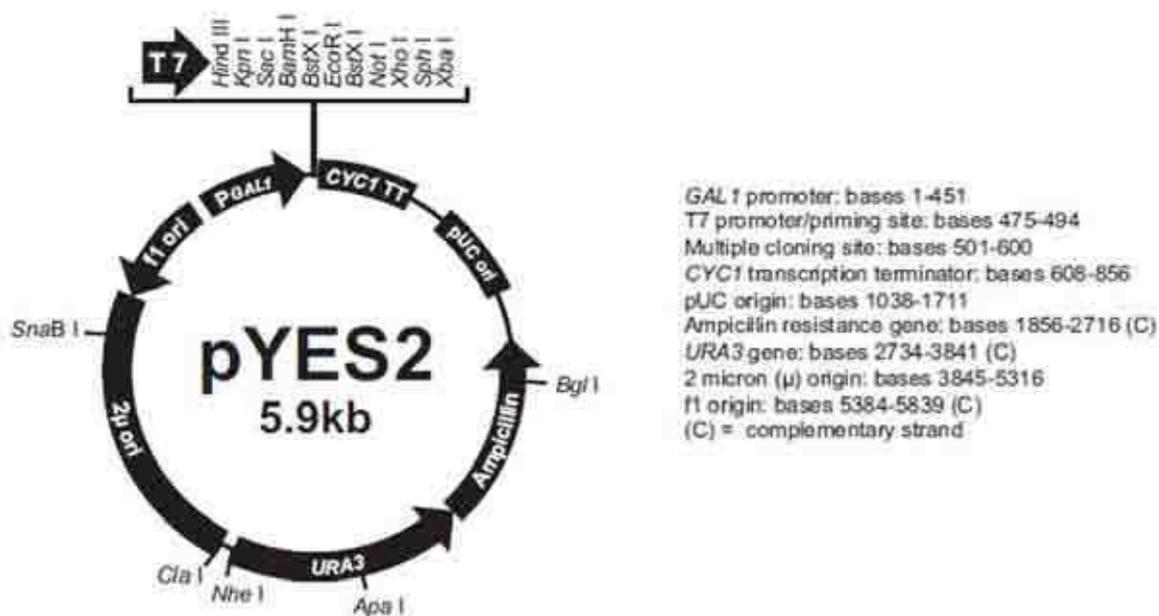


Figura 3.S.2: Mapa del plásmido pYES2.

Para ello, se utilizaron como *primers* los oligonucleótidos:

forward (5'-GTAAGCTTATGAGCGATAAAATTATTCACCTG)

reverse: (5'-GTGAATTCTTACCTGGCCTTGAGGAACG)

El clonado fue realizado por la compañía Top Gene Technologies (www.topgenetech.com). Posteriormente, el plásmido recombinante fue transformado en la cepa de *S. cerevisiae* WT KT 1115 (ver Materiales y métodos del Capítulo 2) para generar la cepa 1115-Bcy1(1-50), utilizando el método de acetato de litio [38]. La expresión del fragmento fue verificada por *western blot* de extractos clarificados obtenidos a partir de cultivos de esa cepa inducidos con galactosa (ver Materiales y métodos del Capítulo 2 para detalles sobre las condiciones de cultivo), utilizando un anticuerpo primario dirigido contra el tag de histidina (Genscript) en una dilución 1:2000 del *stock*, y un anticuerpo secundario anti-ratón conjugado a peroxidasa (Santa Cruz Biotechnology) en una dilución 1:1000 del *stock* (Figura 3.S.3).

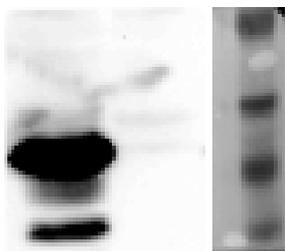


Figura 3.S.3: Verificación de la expresión de Bcy1(1-50) en levaduras.

Western blot revelado con anticuerpos anti-his de extractos clarificados de la cepa KT 1115 (segunda calle) y 1115-Bcy1(1-50) (primera calle) luego de ser crecidas en galactosa para inducir la sobreexpresión del fragmento de interés. Las MM de los marcadores de la calle 3 son (comenzando desde arriba): 50 kDa, 35 kDa, 25 kDa, 20 kDa. La MM esperada para la proteína recombinante es 22,9 kDa.

Western blot

Las proteínas fueron transferidas de los gels a membranas de nitrocelulosa durante 90 minutos a 90 V utilizando buffer de transferencia (25 mM Tris, 192 mM glicina, 20%

metanol). Las membranas fueron bloqueadas por incubación durante 1 hora con buffer de bloqueo (5% leche descremada, 0,05% Tween en buffer PBS) para evitar el pegado inespecífico del anticuerpo a la membrana, e incubadas ON a 4° C con el anticuerpo primario en buffer de bloqueo. Posteriormente, las membranas fueron lavadas con buffer de bloqueo y luego con buffer PBS, e incubadas durante 1 hora a temperatura ambiente con un anticuerpo secundario conjugado a peroxidasa, en buffer de bloqueo. Luego de la incubación, las membranas fueron lavadas con buffer de bloqueo y luego con buffer PBS. Finalmente, las membranas fueron incubadas con el reactivo Luminol, y las bandas inmunoreactivas fueron detectadas por quimioluminiscencia utilizando un analizador de imágenes Bio-Imaging Analyzer Bas-1800II FUJIFILM.

Ensayos fenotípicos

Para las curvas de crecimiento, en primer lugar, se crecieron precultivos ON de las cepas KT 1115 y 1115-Bcy1(1-50) en medio YPGal para generar masa celular; utilizando una dilución apropiada, se registró la OD₆₀₀ de los dos precultivos, y se inoculó en 100 mL del mismo medio el volumen necesario de cada uno de ellos para generar una OD₆₀₀ inicial de 0,02 (aproximadamente 100 µL de inóculo). A partir de ese momento, las curvas de crecimiento se siguieron monitoreando la OD₆₀₀ cada 2 horas, utilizando una dilución del cultivo que permitiera medir siempre un valor similar de OD₆₀₀ (aproximadamente 0,4). Dado que las curvas se realizaron en un medio sin presión de selección, fue necesario verificar el mantenimiento del plásmido en la cepa transformada para garantizar la expresión del fragmento recombinante. Para ello, se tomaron alícuotas correspondientes a cada punto de la curva, y se sembraron en placas con medio SD sin uracilo (ver Materiales y métodos del Capítulo 2 para una descripción del medio). El plásmido confiere la capacidad de sintetizar uracilo a la cepa KT 1115 que es auxotrófica para este componente, y se observó que para todos los puntos de la curva transformada la cepa fue capaz de crecer en el medio selectivo, mientras que la WT no, confirmando el mantenimiento del plásmido.

Para los ensayos de *heat shock*, alícuotas de distintos momentos de las curvas de crecimiento de las distintas cepas se sometieron a tratamiento a 37° C durante una hora, y posteriormente se plaquearon en medio YPGal diluciones seriadas de las mismas para apreciar las diferencias en la viabilidad.

Para los ensayos de acumulación de glucógeno, las cepas de interés se sembraron en parches en placas de medio YPGal. Luego de crecer aproximadamente dos días, se sometieron a vapores de yodo, que permite detectar el glucógeno acumulado, y se registraron imágenes de las placas. [39].

Espectropolarimetría de dicroísmo circular

Los espectros de CD fueron registrados en el UV lejano (190-260 nm) utilizando un espectropolarímetro Jasco J-815 equipado con un sistema de control de temperatura Peltier, en una cubeta de cuarzo de 1 mm de camino óptico. Se utilizó una velocidad de escaneo de 50 nm/min y un ancho de banda espectral de 2 nm. Para cada muestra, se registraron 5 espectros y el resultado fue presentado como el promedio de esas 5 mediciones para minimizar el ruido. Los datos fueron registrados en buffer PBS (20 mM fosfato, 150 mM NaCl, pH = 7,4). La concentración de proteína para estos experimentos fue del rango 10-100 μ M.

Las curvas de desnaturalización térmica fueron registradas calentando las muestras a una velocidad de 2^o C/min, y monitoreando la elipticidad a 221 nm en función del tiempo, con un ancho de banda espectral de 5 nm. Se utilizaron dos concentraciones de proteína: 13 μ M y 63 μ M, en buffer PBS. Las señales de cada muestra fueron normalizadas respecto de su valor de elipticidad a la temperatura inicial para poder comparar ambas curvas. Dado que el proceso de desnaturalización térmica resultó ser reversible, la temperatura de *melting* (T_m) fue calculada ajustando la siguiente ecuación a la señal normalizada, correspondiente a un modelo de desnaturalización reversible de dos estados:

$$CD(T) = (IN + bN.T) + (IU + bU.T) / (1 + \exp(-\frac{T - T_m}{v}))$$

donde CD (T) es la elipticidad a 221 nm normalizada, T es la temperatura, IN e IU son las señales normalizadas de los estados nativo y desplegado a 0^o C respectivamente, bN y bU representan la dependencia de esas señales con la temperatura, y v es un factor geométrico que considera la cooperatividad de las curvas.

Ultracentrifugación analítica en gradiente de sacarosa

Los experimentos se realizaron utilizando una ultracentrífuga Beckman Optima XL-90, equipada con un rotor SW55 Ti de tipo *swinging-bucket*. Para cada experimento, 200 μ L de buffer PBS (20 mM fosfato, 150 mM NaCl, pH = 7,4) conteniendo catalasa, BSA, ovoalbúmina y citocromo C como marcadores, y Bcy1 (1-50) 100 μ M, fueron cargados en la parte superior de un gradiente de 4,2 mL 5%-20% de sacarosa, preparado en el mismo buffer. El rotor fue operado durante 18 horas a 4^o C y 100.000 g. Luego de la centrifugación, se colectaron fracciones de 150 μ L desde la parte inferior del gradiente y las proteínas en cada fracción fueron analizadas por SDS-PAGE 10% con buffer glicina y SDS-PAGE 15% con buffer Tricina, con tinción por Coomassie Blue. En la **Figura 3.S.4** se muestran las fracciones que contienen a los marcadores.

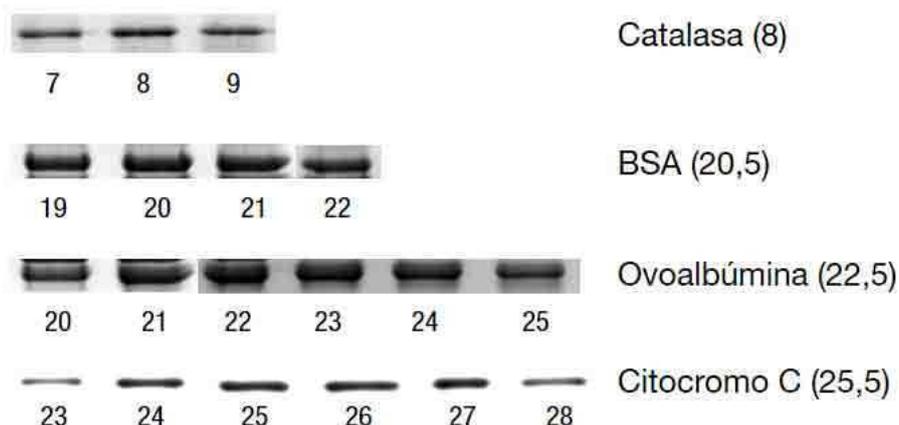


Figura 3.S.4: Determinación de la posición de los marcadores de $S_{20,w}$ en el gradiente.

Se muestran las imágenes de los gels correspondientes a las fracciones que contienen a cada uno de los marcadores. En la derecha se indica la fracción correspondiente al máximo del pico de cada uno de ellos, determinada por la intensidad de las bandas del gel.

Dispersión de Rayos X a bajo ángulo

Para registrar los datos de SAXS, se utilizó un instrumento de haz colimado (10 mm) SAXSess (Anton Paar), equipado con una fuente de radiación $\text{CuK}\alpha$ ($\lambda = 1,54 \text{ \AA}$) y un detector 2D de tipo *image plate*. Se utilizaron muestras purificadas de Bcy1 (1-50) de distintas concentraciones en buffer HEPES 25 mM, 50 mM NaCl, pH = 7,0, que fueron colocadas en una cubeta capilar de 50 μ L a 12^o C. Se evaluó la variación de los difractogramas en el tiempo para

asegurar que no existieran efectos de daño a la muestra; se observó que en los lapsos de tiempo que duraron los experimentos los patrones no variaban significativamente. Los datos finales fueron colectados durante 60 minutos.

El perfil de dispersión de rayos X de la proteína fue obtenido por sustracción del perfil del buffer al de la muestra. Se estudiaron muestras de 3 mg/mL (50 mM) y 6 mg/mL (100 mM) para descartar efectos de la concentración de proteína. Ambos gráficos de Guinier fueron lineales en el rango $Q_{\min} - Q \times R_g = 1,3$, y la forma del perfil de dispersión de ambas muestras no mostró señales ni de agregación ni efectos de repulsión entre partículas. La intensidad a ángulo 0 (I_0) resultó proporcional a la concentración de la muestra, y el valor del R_g no resultó dependiente de la misma. Estas observaciones demuestran que ambos conjuntos de datos resultaron de buena calidad; por lo tanto, se decidió trabajar con la muestra más concentrada que presentó mejor relación señal/ruido.

El gráfico de Guinier presentado en los resultados fue construido manualmente como $\ln [I(Q)]$ vs Q^2 para un rango apropiado de Q , utilizando el conjunto de datos corregido por *slit-smearing effects*. La función de distribución de distancias de la partícula, $P(r)$, se obtuvo utilizando el método numérico de transformada inversa de Fourier implementado en el programa GNOM [10] del paquete de programas de análisis de datos de SAXS ATSAS [11]. Este programa corrige los datos por *slit-smearing effects* utilizando los parámetros de la geometría del haz incidente, que fueron: $AH = 0,283 \text{ \AA}^{-1}$ y $LH = 0,181 \text{ \AA}^{-1}$. El programa proporciona, además de la función de distribución de distancias de la partícula, los parámetros I_0 , R_g y D_{\max} . Este último es elegido por el usuario como *input* para el cálculo; el método consiste en ensayar distintos valores de D_{\max} , y elegir el que proporciona el mejor ajuste entre el perfil experimental de SAXS y el perfil teórico calculado por transformada de Fourier de la función $P(r)$ "candidata". Para evaluar la calidad del ajuste, se utilizan ciertos parámetros estadísticos que son proporcionados por el programa, junto con la inspección visual de la curva de $P(r)$.

El modelo *ab initio* del contorno de la proteína en solución fue construido utilizando el programa DAMMIF [13], del mismo paquete de análisis. Este programa toma como *input* la función $P(r)$ computada utilizando GNOM, e información sobre la simetría de la molécula (en este caso, se utilizó una simetría de tipo P4, correspondiente a un tetrámero). Mediante el algoritmo explicado en la Introducción metodológica, el programa generó 20 modelos. Se seleccionó el que proporcionó un mejor ajuste entre su perfil de dispersión teórico, computado con el programa FoXS [34], y el perfil experimental de SAXS.

Construcción del modelo atómico del tetrámero

El modelo por homología del dímero de Bcy1 (1-50) fue generado con el paquete MODELLER [31], utilizando un alineamiento estructural de los aminoácidos (1-50) de Bcy1 con las formas apo de los dominios D/D de RII α (códigos pdb: 1R2A (NMR), 1L6E (NMR) y 2IZY (cristal)), generado con el comando align2D implementado en el programa. Este comando contempla información estructural de los templados para construir el alineamiento (ver Materiales y métodos del Capítulo 2). Previamente, se ensayaron distintas combinaciones de los templados y distintos alineamientos, y se compararon utilizando la función de *scoring* GA341 proporcionada por el programa, que permite comparar modelos obtenidos en diferentes tandas de modelado [31]. Se observó que el alineamiento que proporciona los mejores modelos fue, por lejos, el generado con el comando align2D, y que en todos los casos los modelos resultaron mejores cuando se utilizaron los tres templados simultáneamente. La similitud de secuencia entre los templado y el *target* es del 38%.

Para modelar los dos giros de α -hélice adicionales predichos para Bcy1 (1-50) en base al análisis bioinformático, se aprovechó que la estructura correspondiente al pdb 2IZY presenta casualmente dos giros helicoidales adicionales en esa misma región, que no se observan en las estructuras en solución obtenidas por NMR, y que por lo tanto son un artificio de la cristalización que se debe probablemente a la presencia de un *tag* de polihistidinas en el extremo C-terminal de la construcción utilizada [32].

Se utilizó un *script* fabricado a partir de los ejemplos que ofrece el programa, que sirve para construir modelos de proteínas oligoméricas a partir de múltiples templados, con imposición de restricciones de simetría sobre las dos cadenas del dímero (considerando que en las estructuras de los dominios D/D de mamíferos las dos cadenas son superponibles). Se generaron 20 modelos, que fueron examinados por inspección visual y ordenados de acuerdo a la función de *scoring* DOPE proporcionada por el programa, que resulta la más adecuada para seleccionar los mejores modelos de una misma tanda de modelado [31]. Se analizó la dependencia de la calidad de los modelos en base a esta función con la conformación de los segmentos desestructurados N-terminales, sin observarse ninguna correlación. Los mejores cuatro modelos fueron refinados utilizando el programa KoBaMIN [40], y se evaluó la calidad de los mismos antes y después del refinamiento como se indica en el siguiente párrafo, para cotejar el éxito del mismo.

Para evaluar la calidad de los modelos se utilizaron los servidores QMEAN [41] y PDBsum [42]. Todos los modelos mostraron mejoras sustanciales en sus parámetros de calidad después del protocolo de refinamiento. El mejor modelo presentó 92,4% de sus aminoácidos en las regiones más favorables del gráfico de Ramachandran, y ninguno en las regiones no permitidas. La evaluación del programa PROCHECK, que permite cotejar la calidad estereoquímica del modelo [43], reveló que no hay ningún factor G menor a -0,5, y que el factor G global fue de 0,42 (este factor indica la calidad promedio de las longitudes de enlaces y ángulos; para que un modelo sea confiable, el factor G debería ser superior a -0,5). Por otro lado el *score* QMEAN proporciona una estimación de la confiabilidad global de un modelo; toma valores entre 0 y 1, y se espera que sea cercano a la unidad para buenos modelos. El score del mejor modelo fue de 0,836, que es muy similar al de los templados. En conclusión, el modelo resultó ser satisfactorio tanto desde el punto de vista de los estimadores globales como del de los estimadores locales.

El modelo atómico de la estructura tetramérica se construyó como se indica en la sección de Resultados y discusión. Brevemente, el modelo del dímero se utilizó como receptor y como ligando en el programa automatizado de *docking* de proteínas ClusPro [33], implementado *online*. El servidor generó 93 modelos en total, que fueron posteriormente analizados para seleccionar el mejor en base a su concordancia con la información experimental disponible. Es importante mencionar que se ensayaron los cuatro mejores modelos del dímero generados por homología, tanto antes como después de su refinamiento, como *input* para el programa de *docking*; en casi todos los casos se obtuvieron resultados similares para el modelo del tetrámero luego del protocolo de selección.

REFERENCIAS

1. Wen J, Arakawa T, Philo JS: **Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions.** *Analytical biochemistry* 1996, **240**(2):155-166.
2. Kelly SM, Jess TJ, Price NC: **How to study proteins by circular dichroism.** *Biochimica et biophysica acta* 2005, **1751**(2):119-139.
3. Erickson HP: **Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy.** *Biological procedures online* 2009, **11**:32-51.
4. Siegel LM, Monty KJ: **Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases.** *Biochimica et biophysica acta* 1966, **112**(2):346-362.
5. Mertens HD, Svergun DI: **Structural characterization of proteins and complexes using small-angle X-ray solution scattering.** *Journal of structural biology* 2010, **172**(1):128-141.
6. Blanchet CE, Svergun DI: **Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution.** *Annual review of physical chemistry* 2013, **64**:37-54.
7. Petoukhov MV, Svergun DI: **Applications of small-angle X-ray scattering to biomacromolecular solutions.** *The international journal of biochemistry & cell biology* 2013, **45**(2):429-437.
8. Moore PB: **Small-angle scattering. Information content and error analysis.** *J Appl Cryst* 1980, **13**:168-175.
9. Jacques DA, Trehwella J: **Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls.** *Protein science : a publication of the Protein Society* 2010, **19**(4):642-657.
10. Svergun DI: **Determination of the regularization parameter in indirect-transform methods using perceptual criteria.** *J Appl Cryst* 1992, **25**:495-503.
11. Petoukhov MV, Konarev PV, Kikhney AG, Svergun DI: **ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis.** *J Appl Cryst* 2007, **40**:223-228.
12. Svergun DI: **Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.** *Biophysical journal* 1999, **76**(6):2879-2886.
13. Franke D, Svergun DI: **DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.** *J Appl Cryst* 2009, **42**:342-346.
14. Jorgensen P, Nishikawa JL, Breikreutz BJ, Tyers M: **Systematic identification of pathways that couple cell growth and division in yeast.** *Science (New York, NY)* 2002, **297**(5580):395-400.
15. Griffioen G, Anghileri P, Imre E, Baroni MD, Ruis H: **Nutritional control of nucleocytoplasmic localization of cAMP-dependent protein kinase catalytic and regulatory subunits in *Saccharomyces cerevisiae*.** *J Biol Chem* 2000, **275**(2):1449-1456.
16. Tudisca V, Recouvreux V, Moreno S, Boy-Marcotte E, Jacquet M, Portela P: **Differential localization to cytoplasm, nucleus or P-bodies of yeast PKA subunits under different growth conditions.** *European journal of cell biology* 2010, **89**(4):339-348.
17. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS: **Global analysis of protein expression in yeast.** *Nature* 2003, **425**(6959):737-741.
18. Werner-Washburne M, Brown D, Braun E: **Bcy1, the regulatory subunit of cAMP-dependent protein kinase in yeast, is differentially modified in response to the physiological status of the cell.** *J Biol Chem* 1991, **266**(29):19704-19709.
19. Hixson CS, Krebs EG: **Characterization of a cyclic AMP-binding protein from bakers' yeast. Identification as a regulatory subunit of cyclic AMP-dependent protein kinase.** *J Biol Chem* 1980, **255**(5):2137-2145.
20. Johnson KE, Cameron S, Toda T, Wigler M, Zoller MJ: **Expression in *Escherichia coli* of BCY1, the regulatory subunit of cyclic AMP-dependent protein kinase from *Saccharomyces cerevisiae*. Purification and characterization.** *J Biol Chem* 1987, **262**(18):8636-8642.

21. Hausken ZE, Coghlan VM, Hastings CA, Reimann EM, Scott JD: **Type II regulatory subunit (RII) of the cAMP-dependent protein kinase interaction with A-kinase anchor proteins requires isoleucines 3 and 5.** *J Biol Chem* 1994, **269**(39):24245-24251.
22. Li Y, Rubin CS: **Mutagenesis of the regulatory subunit (RII beta) of cAMP-dependent protein kinase II beta reveals hydrophobic amino acids that are essential for RII beta dimerization and/or anchoring RII beta to the cytoskeleton.** *J Biol Chem* 1995, **270**(4):1935-1944.
23. Newlon MG, Roy M, Hausken ZE, Scott JD, Jennings PA: **The A-kinase anchoring domain of type IIalpha cAMP-dependent protein kinase is highly helical.** *J Biol Chem* 1997, **272**(38):23637-23644.
24. Banky P, Huang LJ, Taylor SS: **Dimerization/docking domain of the type Ialpha regulatory subunit of cAMP-dependent protein kinase. Requirements for dimerization and docking are distinct but overlapping.** *J Biol Chem* 1998, **273**(52):35048-35055.
25. Newlon MG, Roy M, Morikis D, Hausken ZE, Coghlan V, Scott JD, Jennings PA: **The molecular basis for protein kinase A anchoring revealed by solution NMR.** *Nat Struct Biol* 1999, **6**(3):222-227.
26. Banky P, Newlon MG, Roy M, Garrod S, Taylor SS, Jennings PA: **Isoform-specific differences between the type Ialpha and IIalpha cyclic AMP-dependent protein kinase anchoring domains revealed by solution NMR.** *J Biol Chem* 2000, **275**(45):35146-35152.
27. Banky P, Roy M, Newlon MG, Morikis D, Haste NM, Taylor SS, Jennings PA: **Related protein-protein interaction modules present drastically different surface topographies despite a conserved helical platform.** *J Mol Biol* 2003, **330**(5):1117-1129.
28. Boy-Marcotte E, Perrot M, Bussereau F, Boucherie H, Jacquet M: **Msn2p and Msn4p control a large number of genes induced at the diauxic transition which are repressed by cyclic AMP in *Saccharomyces cerevisiae*.** *Journal of bacteriology* 1998, **180**(5):1044-1052.
29. Marks MS: **Determination of molecular size by zonal sedimentation analysis on sucrose density gradients.** *Current protocols in cell biology / editorial board, Juan S Bonifacino [et al]* 2001, **Chapter 5:Unit 5.3.**
30. Receveur-Brechot V, Durand D: **How random are intrinsically disordered proteins? A small angle scattering perspective.** *Current protein & peptide science* 2012, **13**(1):55-75.
31. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller.** *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]* 2006, **Chapter 5:Unit 5.6.**
32. Gold MG, Lygren B, Dokurno P, Hoshi N, McConnachie G, Tasken K, Carlson CR, Scott JD, Barford D: **Molecular basis of AKAP specificity for PKA regulatory subunits.** *Mol Cell* 2006, **24**(3):383-395.
33. Comeau SR, Gatchell DW, Vajda S, Camacho CJ: **ClusPro: a fully automated algorithm for protein-protein docking.** *Nucleic Acids Res* 2004, **32**(Web Server issue):W96-99.
34. Schneidman-Duhovny D, Hammel M, Sali A: **FoXS: a web server for rapid computation and fitting of SAXS profiles.** *Nucleic Acids Res* 2010, **38**(Web Server issue):W540-544.
35. Kozin MB, Svergun DI: **Automated matching of high- and low-resolution structural models.** *J Appl Cryst* 2001, **34**:33-41.
36. Garcia De La Torre J, Huertas ML, Carrasco B: **Calculation of hydrodynamic properties of globular proteins from their atomic-level structure.** *Biophysical journal* 2000, **78**(2):719-730.
37. Gast K, Zirwer D, Damaschun H, Hahn U, Muller-Frohne M, Wirth M, Damaschun G: **Ribonuclease T1 has different dimensions in the thermally and chemically denatured states: a dynamic light scattering study.** *FEBS letters* 1997, **403**(3):245-248.
38. Ito H, Fukuda Y, Murata K, Kimura A: **Transformation of intact yeast cells treated with alkali cations.** *Journal of bacteriology* 1983, **153**(1):163-168.
39. Zaremberg V, Moreno S: **Analysis of the mechanism of activation of cAMP-dependent protein kinase through the study of mutants of the yeast regulatory subunit.** *European journal of biochemistry / FEBS* 1996, **237**(1):136-142.
40. Rodrigues JP, Levitt M, Chopra G: **KoBaMIN: a knowledge-based minimization web server for protein structure refinement.** *Nucleic Acids Res* 2012, **40**(Web Server issue):W323-328.
41. Benkert P, Kunzli M, Schwede T: **QMEAN server for protein model quality estimation.** *Nucleic Acids Res* 2009, **37**(Web Server issue):W510-514.
42. de Beer TA, Berka K, Thornton JM, Laskowski RA: **PDBsum additions.** *Nucleic Acids Res* 2014, **42**(Database issue):D292-296.

43. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **PROCHECK: a program to check the stereochemical quality of protein structures.** *J Appl Cryst* 1993, **26**:283-291.

CAPÍTULO 4

ESTRUCTURA CRISTALINA DEL DOMINIO D/D DE BCY1

“ So we lay in the black embrace
And the seed is sown in a Holy place
And I watched, and I waited for the dawn ”

Bruce Dickinson, “Revelations”

ANTECEDENTES

En los capítulos previos se reportó la existencia de una novedosa estructura tetramérica para un dominio D/D de la subunidad R de la PKA, y mediante una combinación de distintas técnicas se logró obtener un panorama bastante completo de su estructura en solución.

OBJETIVOS Y ABORDAJES UTILIZADOS

En este capítulo se describe el análisis de la estructura cristalina del tetrámero Bcy1 (1-50), utilizando como técnica principal la cristalografía de Rayos X, y una combinación de distintas herramientas bioinformáticas complementarias. Para llevar a cabo este análisis, se plantearon los siguientes objetivos específicos.

- 1) Obtener una estructura de alta resolución para el tetrámero Bcy1 (1-50) mediante cristalografía de Rayos X.
- 2) Comparar esta estructura con el modelo de la estructura en solución y con la información experimental obtenida mediante distintas técnicas en el Capítulo 3.
- 3) Analizar exhaustivamente la estructura del tetrámero, e identificar los determinantes estructurales que son responsables de la tetramerización. Comparar con los dominios D/D diméricos de las subunidades R de mamíferos e identificar los elementos que determinan las diferencias entre ambas en cuanto a sus estados oligoméricos.
- 4) Realizar un análisis filogenético de dominios D/D de distintos organismos, a efectos de identificar los determinantes mencionados, y especular con la posible existencia de estructuras tetraméricas en otros organismos.

INTRODUCCIÓN METODOLÓGICA

En esta sección se describen brevemente los aspectos básicos de la técnica experimental utilizada en este capítulo, la cristalografía de Rayos X, en base a [1, 2]. Para la sección de validación de las estructuras, se utilizó principalmente [3].

Generalidades sobre la técnica

La resonancia magnética nuclear (NMR) y la cristalografía de rayos X de monocristal son las dos técnicas experimentales que se utilizan actualmente para obtener modelos atómicos tridimensionales de macromoléculas o complejos macromoleculares. La principal diferencia entre ambas radica en que la primera técnica proporciona modelos de la estructura en solución, mientras que en la segunda las moléculas se encuentran en estado sólido. La cristalografía de rayos X de monocristal es por lejos la técnica más difundida en la actualidad: al momento de la escritura de este trabajo, de las aproximadamente 100 000 estructuras depositadas en el *Protein Data Bank* (PDB), más del 90% fueron obtenidas por esta técnica.

El método consiste en la interpretación del patrón de difracción de rayos X generado en una placa bidimensional a partir de un monocristal de la macromolécula de interés. La información acerca de la estructura molecular se encuentra codificada en ese patrón de manera similar a cómo la información acerca de las distancias intramoleculares y la forma del contorno molecular se encuentran codificadas en la curva de dispersión de SAXS. El resultado final de un experimento de cristalografía de rayos X es un mapa de densidad electrónica de la molécula, que puede ser posteriormente interpretado en términos de un modelo molecular a nivel atómico. Para continuar con la analogía entre ambas técnicas, el mapa de densidad electrónica y la información contenida en el patrón de difracción están relacionados por transformada de Fourier, al igual que la función de distribución de distancias intramoleculares ($P(r)$) y el perfil de dispersión de SAXS. También al igual que en SAXS, la información experimental permite la reconstrucción de la estructura atómica subyacente en el cristal mediante métodos computacionales.

Propiedades de los cristales

Un monocristal es un arreglo tridimensional, ordenado, de moléculas y/o iones que se encuentran unidos por interacciones no covalentes, en el cual las especies están organizadas respetando ciertos principios de simetría y periodicidad.

La mínima unidad que mediante la aplicación de operaciones de simetría permite reproducir la totalidad del cristal es la unidad asimétrica. Por otro lado, la mínima unidad que reproduce la totalidad del cristal mediante operaciones de traslación únicamente es la celda unitaria. Según la definición del PDB, “la unidad asimétrica es la menor porción de una estructura cristalina a la cual pueden aplicarse operaciones de simetría para generar la celda unitaria completa (a su vez, la celda unitaria es la unidad repetitiva del cristal). La aplicación de operaciones de simetría cristalográficas a la unidad asimétrica genera la celda unitaria, la cual por traslación en las tres direcciones reproduce la totalidad del cristal”. Todos los archivos depositados en el PDB corresponden a la unidad asimétrica del cristal, que puede contener una unidad biológica completa, una porción de una unidad biológica, o varias unidades biológicas; esto debe ser determinado en base a otras evidencias experimentales, y en general es difícil hacerlo solamente en base a la estructura cristalina.

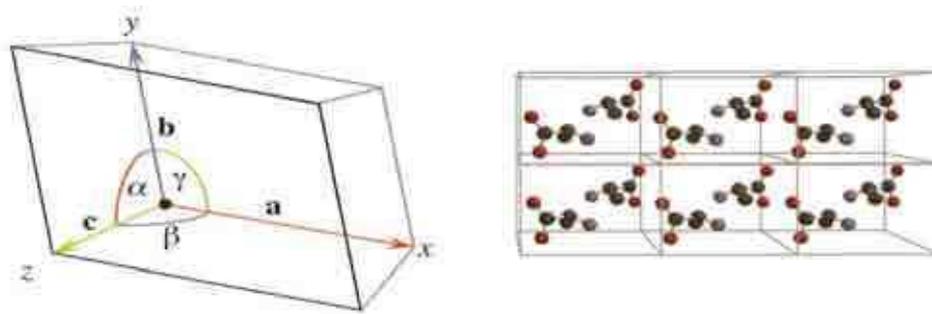


Figura 4.1: Propiedades de la celda unitaria.

En el panel de la izquierda se muestra una representación de una celda unitaria con sus parámetros de red, los vectores (**a**, **b**, **c**) y los ángulos (**α**, **β**, **γ**). En el panel de la derecha se representa esquemáticamente cómo por traslación en las tres direcciones de una celda unitaria se reproduce la totalidad del cristal. En este ejemplo, cada celda unitaria contiene dos unidades asimétricas, relacionadas por operaciones de simetría, y en cada una de ellas hay una molécula de alanina. (Figura tomada de [1]).

La celda unitaria está caracterizada por tres vectores de red (**a**, **b**, **c**) y tres ángulos (**α**, **β**, **γ**) que definen unívocamente su geometría y sus dimensiones (**Figura 4.1**). A estos parámetros se los conoce como “parámetros de red” o “dimensiones de la celda unitaria”. Los valores que pueden adoptar estos parámetros están restringidos por la estructura molecular y la simetría

según la cual las moléculas se disponen en el cristal, lo que da lugar a los siete sistemas cristalinos: cúbico, hexagonal, monoclinico, ortorrómbico, tetragonal, triclínico y trigonal.

Fundamento del fenómeno de difracción

De manera general, el fenómeno de difracción de rayos X (XRD) ocurre cuando los rayos atraviesan una rendija cuyas dimensiones son similares a la longitud de onda de la radiación incidente. En un experimento típico de XRD, se hace incidir un haz monocromático sobre un monocristal de la proteína de interés; dado que las distancias interatómicas en el cristal son del mismo orden de magnitud que la longitud de onda de la radiación incidente, los cristales cumplen el rol de las rendijas y por lo tanto los rayos X son difractados por la red cristalina. Como consecuencia del ordenamiento periódico de un número elevado de moléculas idénticas en un cristal, los haces difractados interfieren sobre un detector bidimensional que se utiliza para registrar estos haces (denominados “reflexiones”). Como resultado, se obtiene un patrón de puntos resultantes de la interferencia constructiva de una gran cantidad de haces en posiciones específicas del detector (**Figura 4.2**). La intensidad y la dirección correspondientes a cada una de esas reflexiones constituyen los datos experimentales que se utilizan posteriormente para el cómputo del mapa de densidad electrónica y para la generación del modelo molecular.

Las direcciones de las reflexiones se interpretan en términos de planos paralelos formados por átomos del cristal que las originan (por este motivo se conocen como “reflexiones”). Los planos más sencillos de visualizar son los que corresponden a las caras de la celda unitaria, pero existen virtualmente infinitos planos equiespaciados que son potenciales fuentes de reflexiones. Cada familia de planos se identifica por tres números, que se denotan h , k , y l , conocidos como “índices de Miller”. (**Figura 4.3**) Cada índice corresponde a la cantidad de planos por celda unitaria en las direcciones x , y , z , respectivamente.

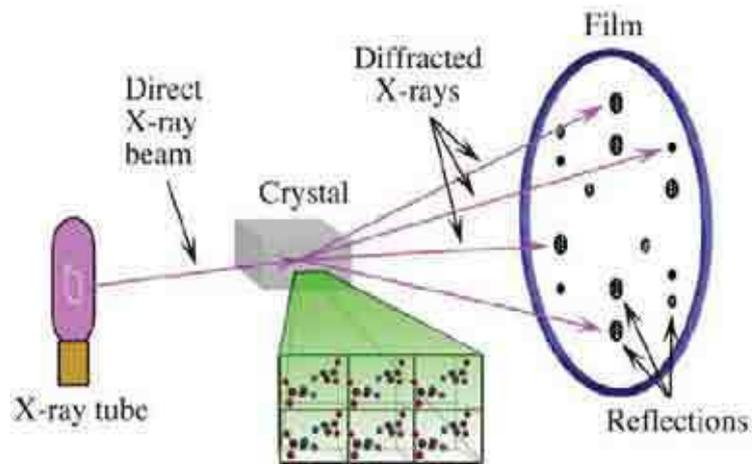


Figura 4.2: Esquema general de un experimento de difracción de rayos X.

Los principales componentes del dispositivo experimental son: la fuente de rayos X que genera el haz incidente, el cristal que genera los haces difractados, y el detector bidimensional en el que se registran las reflexiones. (Figura tomada de [1]).

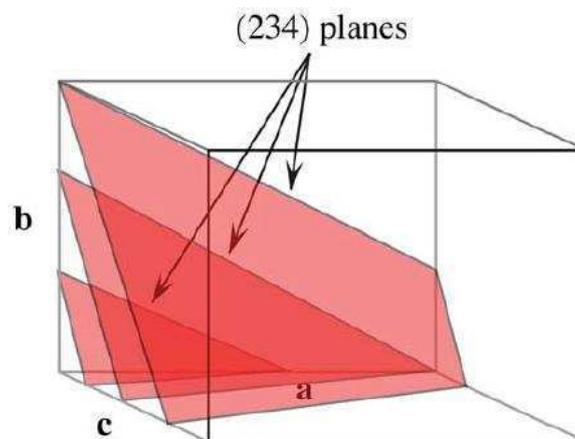


Figura 4.3: Representación de los índices de Miller.

Se muestran los índices de Miller de la familia de planos paralelos (2 3 4). Notar que los planos cortan al eje x (a) en 2 partes, al eje y (b) en 3 partes, y al eje z (c) en 4 partes. (Figura tomada de [1]).

En la interpretación de Bragg del fenómeno de difracción, como ya se indicó, cada conjunto de planos paralelos equiespaciados (separados por una distancia que se denota d_{hkl}) se considera como una única fuente de difracción que da origen a una única reflexión, que

presenta una dirección respecto del haz incidente tal que se satisface la condición de Bragg: $2d_{hkl} \sin\theta = n\lambda$, donde θ es el ángulo del haz difractado, λ es la longitud de onda de la radiación incidente, y n es un número entero (**Figura 4.4**).

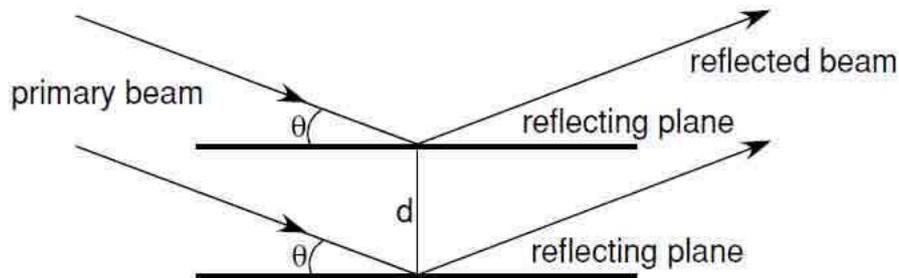


Figura 4.4: Interpretación de Bragg del fenómeno de difracción.

Se muestran las direcciones de los haces incidentes y reflejado, y la distancia d entre los planos paralelos que contribuyen a la interferencia constructiva, generando un único punto en el detector. (Figura tomada de [2]).

Esta interpretación del fenómeno resulta apropiada para determinar la geometría de la colección de los datos. Durante la etapa de la resolución de la estructura, resulta más conveniente considerar otro modelo, según el cual cada pequeño elemento de volumen de densidad electrónica en la celda unitaria se trata como un elemento difractor independiente, representado por un término en una suma de Fourier que describe esa reflexión en particular. Esta visión del fenómeno es la que está relacionada con el posicionamiento de los átomos dentro de la celda unitaria.

Esquema general de trabajo

La determinación de la estructura de una proteína por cristalografía de rayos X consiste de los siguientes pasos: (1) obtención de la proteína purificada, (2) obtención de cristales, (3) montaje del cristal en el equipo, (4) colección de los datos, (5) procesamiento de los datos, (6) resolución de la estructura, (7) refinamiento de la estructura, (8) validación de la estructura.

Obtención de la proteína purificada

Para un experimento de rayos X bien realizado, se requiere de cantidades muy altas de proteína (del orden de 10 mg) con muy elevado grado de pureza (95%). Típicamente, se trabaja en concentraciones altas, cercanas a la saturación (si bien el número es muy dependiente de cada proteína, las concentraciones de aproximadamente 10 mg/mL son comunes).

Cristalización

El proceso de cristalización consiste en la obtención de cristales de la proteína a partir de una solución concentrada de la misma, utilizando un agente precipitante. La cristalización de proteínas es un proceso empírico, dado que no existen elementos para predecir qué condiciones serán favorables para la obtención de cristales de una proteína en particular; tampoco es posible extrapolar condiciones de cristalización de proteínas relacionadas ya sea en secuencia o en estructura. De todos los pasos involucrados en la obtención de una estructura cristalina, este es el menos comprendido, y hasta el día de hoy constituye la etapa limitante en la cristalografía de proteínas.

Los procedimientos llevados a cabo para cristalizar una proteína son diferentes a los que se aplican para compuestos inorgánicos u orgánicos de baja MM, dado que en estos casos el agregado de solventes volátiles, o bien la evaporación lenta de las soluciones, son suficientes muchas veces para obtener cristales. Sin embargo, en el caso de las proteínas esas operaciones no son aplicables, dado que en la mayoría de los casos provocan la desnaturalización de la muestra; por lo tanto, la obtención de cristales de proteínas requiere de métodos particulares y de un trabajo mucho más cuidadoso.

La formación de cristales ocurre en dos etapas: nucleación y crecimiento. La nucleación es la formación de los primeros agregados de moléculas a partir de los cuales posteriormente el cristal crece; este es el paso crucial, y requiere de concentraciones de proteína y de agente precipitante más altas que la que se requiere para la etapa de crecimiento (**Figura 4.5**). Una vez formados los primeros núcleos, si las condiciones de nucleación persisten se formarán muchos núcleos y consecuentemente muchos cristales pequeños en lugar de pocos cristales grandes (que es la situación ideal para obtener datos de buena calidad). En cambio, si una vez que formados los primeros núcleos la concentración de proteína decrece lo suficiente como

para alcanzar las condiciones de crecimiento, éstos podrán crecer y generar cristales más grandes.

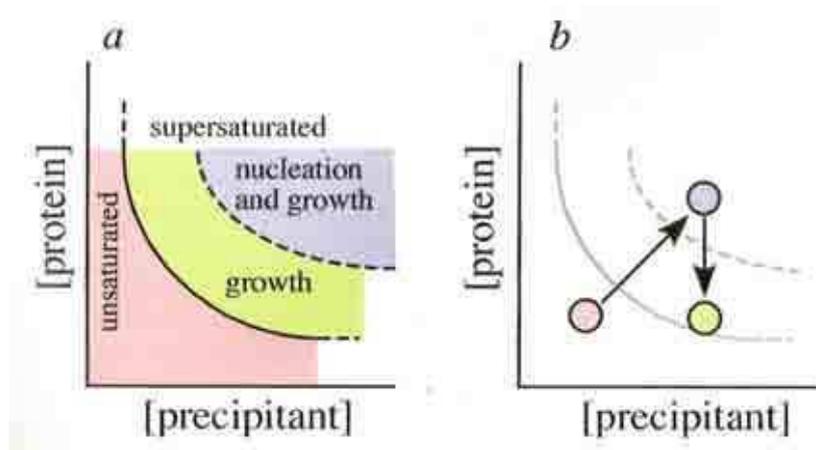


Figura 4.5: Método de cristalización por difusión de vapor.

(a) Diagrama de fases. (b) Estrategia de trabajo. Punto rosa: solución insaturada. Punto azul: formación de los primeros núcleos. Punto verde: crecimiento de los cristales. (Figura tomada de [1]).

El método de cristalización más utilizado para obtener cristales de proteínas es el de difusión por vapor (ver **Figura 4.5**). Este método consiste en alcanzar la condición de sobresaturación de la solución lentamente para favorecer la formación de una pequeña cantidad de núcleos de cristalización. Para ello, la proteína purificada se mezcla con una solución de cristalización que contiene un agente precipitante en una concentración tal que el sistema se encuentre dentro de la zona estable del diagrama de solubilidad (círculo rosa en la **Figura 4.5 (b)**); esta concentración es en principio desconocida. Posteriormente, la muestra se coloca en una cámara cerrada en forma de una gota colgante, ubicada sobre un reservorio que contiene la misma solución de cristalización que le fue agregada. La ventaja de disponer la muestra en forma de gota colgante es que existe menor probabilidad de que los cristales formados se adhieran a la superficie donde reside la gota.

En la gota colgante que contiene la proteína, los solutos de la solución de cristalización están más diluidos que en el reservorio ubicado debajo. Por lo tanto, el agua difunde lentamente desde la gota hacia el reservorio, estimulada por esta diferencia de concentraciones. En consecuencia, las concentraciones de agente precipitante y de proteína

en la gota aumentan lentamente (flecha diagonal en el esquema de la **Figura 4.5 (b)**). Cuando se alcanza la zona de nucleación se dan las condiciones para que se formen los primeros núcleos; a medida que la proteína cristaliza, su concentración en la gota disminuye (flecha descendente en la **Figura 4.5 (b)**) y consecuentemente los cristales crecen, hasta que se alcanza nuevamente la región de estabilidad del diagrama.

Dado que, como ya se señaló, el proceso de cristalización es empírico, para poder explorar un número importante de condiciones se venden *kits* comerciales que constan de decenas a centenas de diferentes formulaciones que resultaron exitosas en el pasado para cristalizar otras proteínas. La gran mayoría de las soluciones contiene sulfato de amonio o polietilenglicol (PEG) como agentes precipitantes, en distintos buffers a distintos pHs y con el agregado de distintos aditivos. Muchas veces, la necesidad de explorar una cantidad tan alta de condiciones requiere del trabajo con robots, que arman las placas de cristalización.

En general, los cristales obtenidos en primera instancia mediante el *screening* de distintas condiciones de cristalización suelen ser de calidad o tamaño insuficientes para obtener datos de difracción adecuados para resolver la estructura de la molécula. Por lo tanto, en una segunda etapa estas condiciones iniciales deben ser optimizadas ajustando más finamente las características de la solución de cristalización, partiendo de la base de las condiciones que resultaron exitosas en la primera etapa.

Montado de los cristales y colección de los datos

Una vez obtenidos cristales de calidad adecuada, se procede a coleccionar los datos de difracción; para ello, un único monocristal debe ser ubicado de forma adecuada en el equipo de difracción para ser irradiado por los rayos X. Actualmente, en XRD de proteínas se trabaja casi exclusivamente en condiciones de criocristalografía; esto significa que durante la colección de los datos el cristal se encuentra a muy bajas temperaturas (aproximadamente 100 K). Esto tiene dos ventajas fundamentales: en primer lugar, el daño por radiación es notablemente menor, ya que los radicales libres que se originan al irradiar los cristales con rayos X presentan una capacidad de difusión muy limitada; esto permite en la mayoría de los casos coleccionar un conjunto de datos completo con un único cristal. Por otro lado, las bajas temperaturas reducen la movilidad de las moléculas, lo que se ve reflejado en mapas de densidad electrónica más definidos, y consecuentemente mejores modelos moleculares.

La colección de datos a bajas temperaturas requiere que los cristales sean capturados en pequeños *loops*, en donde se mantienen sostenidos por tensión superficial del líquido en que se encuentran inmersos. El *loop* que contiene el cristal de interés es montado posteriormente en el equipo de XRD, y se le hace incidir un flujo de nitrógeno a 100 K durante toda la duración del experimento. Previo al montado del cristal en el equipo, es necesario sumergirlo en una solución crioprotectora, que generalmente está compuesta por la misma solución de cristalización adicionada con glicerol; el objetivo de este paso es evitar la formación de hielo en el *loop* al enfriar, ya que el mismo dificulta el análisis posterior pues genera una señal de difracción de fondo muy fuerte, observándose anillos característicos superpuestos a las reflexiones originadas en la muestra.

Los equipos utilizados para la colección de datos de XRD constan de una fuente de rayos X, un colimador para estrechar el haz incidente, un sistema monocromador del haz, un goniómetro para controlar la orientación y oscilación del cristal durante la colección, y un sistema de detección. El sistema se encuentra acoplado a una computadora en la que además se almacenan las imágenes de los patrones de difracción obtenidos durante el tiempo que dura la colección de los datos (*frames*) para su posterior análisis.

Las fuentes de rayos X más utilizadas en cristalografía de proteínas son los generadores de ánodo rotatorio y las estaciones de sincrotrón. Los primeros están formados por un filamento (cátodo) que se encuentra a un potencial marcadamente negativo respecto de un bloque giratorio de cobre, cromo o molibdeno (ánodo rotatorio). Se hace circular una corriente eléctrica por el cátodo, y como consecuencia de la diferencia de potencial entre ambos electrodos se genera una corriente de electrones que impacta en el ánodo generando así rayos X. Una de las fuentes de radiación más utilizadas es la línea K_{α} del cobre, que tiene una longitud de onda característica de 1,542 Å.

Los equipos de ánodo rotatorio resultan muy prácticos porque puede disponerse de ellos en un laboratorio; sin embargo, muchas veces la intensidad de la radiación generada resulta insuficiente para coleccionar datos que permitan la resolución de la estructura de una proteína dado que, como estas están formadas por átomos livianos, interactúan débilmente con los rayos X. En esos casos, se requiere del uso de generadores más potentes como las estaciones sincrotrón. En este trabajo, sin embargo, fue posible resolver la estructura utilizando un conjunto de datos coleccionado en un equipo de laboratorio.

Los detectores de rayos X que se utilizan en cristalografía de proteínas son capaces de detectar la radiación dispersada en una superficie 2D, y por este motivo se los conoce con el nombre general de detectores de área. El más común de estos detectores, que fue el que se utilizó en este trabajo, es el detector de tipo *image plate*.

Finalmente, es importante mencionar que cada *frame* bidimensional que contiene una determinada cantidad de reflexiones corresponde a una orientación relativa del cristal respecto del haz incidente. A efectos de obtener un gran número de reflexiones que sean informativas de la totalidad de la estructura tridimensional, el cristal debe ser rotado durante la colección de los datos para capturar información de la mayor cantidad de planos posible.

Procesamiento de los datos

Los datos obtenidos como se describió hasta el momento consisten en una colección de *frames*, cada uno de los cuales es un arreglo bidimensional de puntos correspondientes a distintas reflexiones originadas en el cristal. El procesamiento de los datos consiste en generar a partir de esos *frames* un archivo que contenga una lista de direcciones e intensidades de las distintas reflexiones para cada orientación relativa del cristal respecto del haz incidente. El proceso consta de tres etapas: indexación (*index*), integración (*integration*) y escalado (*scaling*).

Para comprender estos tres pasos, es necesario entender el fenómeno de difracción de acuerdo a la interpretación de Bragg que, como ya se discutió, estipula que cada reflexión observada en el patrón de difracción se origina en un grupo de planos de átomos que generan interferencia constructiva. Las coordenadas de estos planos en el llamado “espacio recíproco” (que es el espacio de coordenadas del detector, y se define en contraposición al “espacio real”, que es el de la red cristalina) se identifican de acuerdo a sus índices de Miller. Al origen de coordenadas (donde impacta el haz incidente) se le asigna arbitrariamente el valor $(h\ k\ l) = (0\ 0\ 0)$.

La indexación es la asignación para cada reflexión de sus índices de Miller, así como también la determinación de los parámetros de la celda unitaria y el grupo de espacio al que pertenece el cristal. Esto es posible ya que la posición de las reflexiones está determinada por el sistema cristalino y por las dimensiones de la celda unitaria. Concretamente, existe una relación inversa entre el espaciamiento de objetos idénticos (por ejemplo, celdas unitarias) en la red cristalina, y el espaciamiento de las reflexiones en el espacio recíproco; por lo tanto, la

distancia entre las reflexiones en la red recíproca contiene la información sobre las dimensiones de la celda unitaria. Por otro lado, otro dato muy importante es la reflexión más alejada del origen de coordenadas. La inversa de esta distancia proporciona la resolución potencial que puede tener el modelo extraído de ese conjunto de datos. En otras palabras, cuanto más lejos del centro se extiendan las reflexiones, mayor será la resolución (dado que, por las propiedades del espacio recíproco, una máxima distancia representa la mínima distancia medible en el espacio real).

La integración consiste en la digitalización de las intensidades medidas para cada reflexión, condensando todos los archivos provenientes de cada una de las imágenes en un único archivo, de acuerdo a la orientación de cada imagen. Debido a la variabilidad del poder difractor de los cristales, a la diferencia en el recorrido del haz de rayos X en las diferentes orientaciones, y a las fluctuaciones propias de la intensidad del haz, las intensidades de las reflexiones no son consistentes entre las diferentes imágenes. El proceso por el cual se ponen en una misma escala se denomina escalado.

Resolución de la estructura

El objetivo último de la resolución de la estructura de una proteína es la obtención de un mapa de densidad electrónica que permita generar un modelo atómico. Para ello, se requiere del conocimiento de los llamados “factores de estructura”, que se relacionan con el mapa de densidad electrónica dentro de la celda unitaria mediante una operación de transformada de Fourier.

A causa de su ordenamiento, la densidad electrónica de un cristal puede describirse mediante una función periódica tridimensional, $\rho(x,y,z)$. Es una función de \mathbb{R}^3 en \mathbb{R} , por lo que no puede visualizarse en el espacio; lo que se representan típicamente son superficies de nivel de densidad electrónica constante.

Como ya se señaló, para la resolución de la estructura es conveniente interpretar a cada reflexión, caracterizada por sus índices de Miller, no en términos de planos sino como contribuciones de dispersión por parte de todos los átomos contenidos en la celda unitaria. Dado que los elementos que dispersan los rayos X son en realidad los electrones (por esto es que, justamente, el resultado final del experimento es un mapa de densidad electrónica, que representa el contorno molecular), los factores de estructura se escriben como una suma de Fourier que presenta contribuciones de todos los elementos de volumen de la celda unitaria,

cada uno caracterizado por la densidad electrónica promedio en su interior. Cada reflexión tiene un factor de estructura asociado, que puede escribirse:

$$F_{hkl} = f(\rho_1) + f(\rho_2) + \dots + f(\rho_n)$$

donde F_{hkl} es el factor de estructura asociado a la reflexión con índices de Miller (h k l), y $f(\rho_i)$ representa la contribución a la difracción por parte del i-ésimo elemento de volumen, para los n elementos presentes en la celda unitaria.

De esta manera, el problema de obtener un mapa de densidad se reduce al de resolver un sistema de ecuaciones, formado por la expresión de todos los factores de estructura asociados a las reflexiones colectadas en función de $\rho(x,y,z)$; este sistema se resuelve por medio de transformada de Fourier. Los factores de estructura son magnitudes vectoriales, caracterizados por tres variables: su amplitud, su frecuencia, y su fase; la resolución del sistema requiere del conocimiento de las tres. La intensidad de las reflexiones es proporcional a la amplitud de los correspondientes factores de estructura; la posición de cada reflexión se relaciona con la frecuencia. Las fases, por otro lado, no pueden determinarse experimentalmente a partir de los datos de difracción obtenidos como se describió hasta este punto. Este es el llamado “problema de las fases”, que en cristalografía representa la etapa central de la resolución de la estructura.

La resolución de la estructura consiste por lo tanto en el cálculo de las fases de los factores de estructura, que permite el cómputo de la densidad electrónica y la construcción final del modelo. En macromoléculas existen básicamente tres métodos para alcanzar este objetivo: reemplazo isomorfo múltiple, dispersión anómala a varias longitudes de onda, y reemplazo molecular, que es por lejos el más utilizado, y el que se empleó para resolver la estructura presentada en este trabajo. En este método, la estructura se resuelve utilizando una combinación de los datos de difracción de la proteína de interés, y un modelo molecular de una proteína relacionada, que debe ser relativamente parecido al de la proteína de interés: se requiere en general al menos un 30-40% de identidad de secuencia.

El fundamento del método de reemplazo molecular se basa en que si la estructura de interés es similar al modelo de partida, las fases de los factores de estructura de las reflexiones de ambas también serán similares. Los factores de estructura del modelo (F_{calc}) se pueden computar fácilmente (lo que incluye sus fases). Dichos valores dependen fuertemente

de la orientación y de la ubicación del modelo en la celda unitaria. El objetivo del método de reemplazo molecular es entonces orientar y ubicar al modelo en la unidad asimétrica de la proteína de interés de forma tal que los módulos de F_{calc} se asemejen lo máximo posible a los módulos de los factores de estructura experimentales de la molécula de interés (F_{obs}). El último paso del método consiste en el refinamiento de cuerpo rígido o *fitting* del modelo ya orientado. Se obtiene como resultado final del reemplazo un conjunto de coordenadas ubicadas convenientemente en la celda unitaria a partir de las cuales se calculan fases iniciales para comenzar a trabajar en la obtención de la estructura final.

Este método tiene la desventaja de generar un sesgo hacia el modelo de partida debido a que las fases calculadas provenientes del modelo conocido poseen mucha influencia en el resultado final de los módulos de F_{obs} . Cabe mencionar que existen casos más complicados en los que, por ejemplo, la proteína de interés no es lo suficientemente similar al modelo de referencia, o bien aparecen múltiples copias del modelo de búsqueda en la unidad asimétrica. Estas dos situaciones se presentaron en la resolución de la estructura que se reporta en esta tesis, por lo que se requirió de la combinación del método de reemplazo molecular con otro tipo de estrategias para lograr la resolución de la estructura.

Construcción del modelo y refinamiento

La primera etapa de la construcción del modelo consiste en un ajuste manual de los F_{calc} y los F_{obs} . Para ello, se trabaja con dos mapas de densidades electrónicas: el mapa correspondiente a los datos experimentales (F_{obs}), y el mapa de simple diferencia de Fourier ($F_{\text{obs}}-F_{\text{calc}}$). El primer mapa es el mapa “real” que corresponde a la estructura experimental asociada a la etapa del refinamiento en que se está trabajando. El segundo mapa permite apreciar las diferencias entre los mapas experimental y el del modelo generado hasta esa etapa del refinamiento, y muestra aquellas regiones del espacio que presentan densidad electrónica pero que aun no están representadas en el modelo (es decir, donde el modelo necesita ser complementado) o bien aquellas en que se han modelado regiones que no están presentes en la estructura, y que por lo presentan modelo en exceso. Ambos mapas son utilizados conjuntamente para completar y corregir el modelo de acuerdo a la secuencia aminoacídica de la proteína; este proceso se conoce como *model building*.

Una vez finalizado este primer ciclo de modelado visual, se da lugar al refinamiento matemático de la estructura. En términos generales se puede definir a este proceso como un

ajuste del modelo para que exista un máximo acercamiento entre los valores calculados y observados de los módulos de los factores de estructura. A su vez, el desplazamiento de los átomos en el refinamiento se halla fuertemente condicionado por penalidades de geometría como planaridad de anillos aromáticos, quiralidad, distancias y ángulos de enlace, etc. El refinamiento es un proceso iterativo, y la salida de cada etapa contiene coordenadas refinadas a partir de las cuales se pueden calcular nuevos conjuntos de las fases, módulos de F_{calc} , y mapas de densidad electrónica. Estos son sometidos nuevamente a modelado visual, y el proceso se continúa hasta alcanzar convergencia y la mayor completitud posible para el modelo final.

El refinamiento se monitorea mediante el “factor R”, que es un parámetro estadístico que da cuenta de las discrepancias entre las amplitudes de los factores de estructura F_{obs} y F_{calc} para la etapa del refinamiento en que se está trabajando. El factor R se define como:

$$R = \frac{\sum |F_{obs} - F_{calc}|}{\sum F_{obs}}$$

Este debe presentar un descenso progresivo a medida que avanzan los ciclos de refinamiento. Una estructura correctamente resuelta y refinada debería presentar un valor no muy superior a $R = 0,20-0,25$, aunque el valor final depende de la resolución y de la calidad de los datos, entre otros factores.

Adicionalmente, se calcula otro parámetro estadístico denominado factor R libre (R_{free}). Éste se calcula de la misma manera que el factor R, pero utilizando los datos de un pequeño porcentaje de reflexiones seleccionadas al azar, que no fueron incluidas en el proceso de resolución y refinamiento de la estructura. Este procedimiento se conoce en cristalografía con el nombre de validación cruzada, y permite detectar falencias por modelado excesivo en las estructuras. Si el modelo se vuelve demasiado complejo en el sentido de incorporar información no avalada por datos experimentales, el R_{free} no mejorará (mientras que el R sí continuará disminuyendo). Este parámetro resulta de suma importancia en la validación final, y no debe exceder en más de 6-7 unidades porcentuales al valor de R.

Es importante aclarar que existen distintas variantes de refinamiento y que cada una de ellas se aplica en diferentes instancias. Por lo general se comienza con un refinamiento de cuerpo rígido, permitiendo que las coordenadas atómicas se muevan únicamente en bloques. Luego se suelen llevar a cabo uno o varios ciclos del proceso denominado templado simulado

(*simulated annealing*), que consiste en evitar que la estructura quede atrapada en mínimos locales de la superficie de energía potencial mediante dinámica molecular que involucra ciclos de calentamiento y enfriamiento. El templado simulado es muy recomendado en los primeros ciclos de refinamiento ya que permite grandes desplazamientos de las cadenas principales para ajustar el modelo a las reflexiones experimentales, siendo poco útil o inclusive perjudicial en las etapas finales. En las últimas etapas se refina para cada átomo no sólo su posición (coordenadas x, y, z), sino también un cuarto parámetro denominado factor térmico isotrópico o factor B. Este último refleja la movilidad del átomo en la estructura cristalina, observándose densidad electrónica débil o difusa para valores elevados del mismo.

Validación de la estructura

Finalmente, el último paso antes del análisis de la estructura es su validación. Este proceso tiene como objetivo detectar errores durante la resolución y el refinamiento, y determinar la calidad del modelo final en base a un conjunto de parámetros estadísticos. Para ello se comparan los resultados obtenidos para el modelo contra los valores comúnmente obtenidos, tomados de bases de datos de estructuras de alta resolución. A grandes rasgos, los estimadores de calidad pueden dividirse en dos grandes grupos: los que reflejan la calidad de los datos en sí mismos, y los que reflejan la calidad del modelo.

Estimadores de la calidad de los datos

Un parámetro fundamental es la resolución de los datos de difracción (**Figura 4.6**), que depende de la cantidad de reflexiones utilizadas para la resolución de la estructura (a mayor cantidad de reflexiones, mayor resolución). Se puede definir como la mínima distancia entre planos de la red cristalina que proporcionan reflexiones distintas y medibles. Este parámetro define el nivel de detalle que puede ser distinguido en los mapas; es decir, la mínima distancia entre características estructurales que puede ser apreciada. A partir de 2,7 Å se establece el límite de resolución media, porque hay probabilidades de detectar moléculas de agua individuales unidas por puente hidrógeno. A partir de 1,2 Å, que es la distancia de un enlace C-C, se considera resolución a nivel atómico; cabe aclarar que este no es un valor de resolución que se logre con facilidad en cristalografía de proteínas.

Otro parámetro muy importante es la relación señal ruido promedio para las intensidades de las reflexiones (I/σ). Lógicamente, cuanto más alto sea este valor, mejor es la calidad de los

datos. En la capa de máxima resolución, este valor no debe ser menor a 2; valores menores indican que la resolución real de los datos es menor que la estipulada. Si en cambio resulta ser bastante mayor, puede significar que el cristal tiene la capacidad de difractar mejor, pero que la calidad de los datos colectados estuvo limitada por el dispositivo experimental utilizado (esto es así porque el límite de resolución de los datos se define para $I/\sigma = 2,0$).

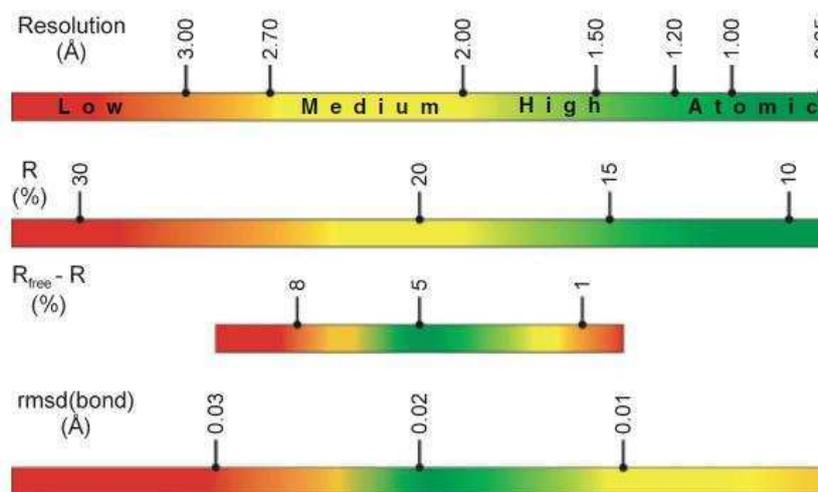


Figura 4.6: Parámetros de calidad más relevantes en cristalografía de proteínas.

Para cada estimador (resolución, R , R_{free} , y desviación media de las distancias de enlace) se muestran en código de colores los distintos valores posibles: en verde los mejores valores, en rojo los peores, y en amarillo y naranja los intermedios. (Figura tomada de [3]).

Un tercer parámetro de importancia es el denominado R_{merge} . Si bien su definición matemática es complicada, su interpretación es bastante simple. Debido a las diferentes orientaciones que adoptan los cristales durante todo el proceso de colección de los datos, la mayor parte de las reflexiones se computan más de una vez; esto se conoce como “multiplicidad” o “redundancia” del conjunto de datos, y en general es un efecto deseable dado que permite computar con mayor precisión las intensidades de una determinada reflexión. Dado que las reflexiones presentan cierta dispersión, el R_{merge} es un indicador de la dispersión de las distintas intensidades medidas para una reflexión en particular, respecto de la media. En general, su valor no debería superar 0,1.

Finalmente, un último parámetro a considerar es la completitud del conjunto de datos. Este número representa el porcentaje del total de las reflexiones posibles para ese cristal en particular, calculadas de manera teórica, que fueron registradas experimentalmente. Lógicamente, cuanto más elevado sea este número, mejor es la calidad de los datos.

Estimadores de la calidad del modelo

En primer lugar, como ya se ha mencionado los parámetros R y R_{free} son una medida de la bondad del ajuste del modelo a los datos experimentales (ver **Figura 4.6**).

Por otro lado, el factor de temperatura o factor B (que también ya fue descrito), refleja la movilidad de cada átomo en la estructura, que se origina como consecuencia de los movimientos térmicos vibracionales de los átomos alrededor de su posición de reposo. Este parámetro se obtiene como parte del protocolo de refinamiento, y se encuentra listado en la cuarta columna de todos los archivos pdb; sus unidades son Å^2 , y los valores suelen encontrarse entre 2 y 100; valores alejados de estos números son indicativos de errores. En ocasiones, puede proporcionar un primer indicio de la movilidad conformacional asociada a distintas regiones de la molécula.

Otro factor a analizar es la cantidad de moléculas de agua modeladas. Para resolución media (que es el caso de la estructura resuelta en este trabajo) no debería haber más de 0,3-0,5 moléculas de agua modeladas por cada residuo de proteína. Valores superiores son indicativos de sobreinterpretación de los datos experimentales.

Finalmente, todos los modelos deben ser evaluados en base a su estereoquímica para garantizar que no existan desviaciones serias de los valores típicamente encontrados para proteínas. En primer lugar, deben evaluarse los ángulos diedros φ y ψ mediante un gráfico de Ramachandran. Por otro lado las desviaciones medias de la idealidad para distancias de enlace y ángulos deben encontrarse dentro de los valores aceptados (ver **Figura 4.6**). En el caso particular de los enlaces, el valor debe ser cercano a $0,02 \text{ Å}$; valores superiores son indicativos de violaciones de las restricciones estereoquímicas, y valores inferiores sugieren lo contrario (imposición de demasiadas restricciones durante el refinamiento).

RESULTADOS Y DISCUSIÓN

Obtención de cristales, colección de los datos y resolución de la estructura

Se siguió el esquema de trabajo que ya fue descrito en la Introducción metodológica de este capítulo. Se produjeron grandes cantidades de Bcy1 (1-50) mediante el protocolo detallado en el Capítulo 2. En primer lugar, se realizaron los primeros ensayos de solubilidad, que permitieron determinar que una concentración de 3 mg/mL era adecuada para intentar las pruebas de cristalización. Por lo tanto, se procedió a ensayar distintas combinaciones de agentes precipitantes y buffers utilizando el método de difusión de vapor. En la **Figura 4.7** se muestra una imagen de una placa de 24 pocillos que contiene gotas de Bcy1 (1-50) en distintas condiciones de prueba.



Figura 4.7: Placa de cristalización de Bcy1 (1-50).

Se utilizó el método de difusión de vapor en gota colgante para ensayar distintas condiciones de cristalización; los ensayos se hicieron a 20° C. Cada pocillo se encuentra herméticamente cerrado; en el fondo se ubica el reservorio de 1 mL de solución precipitante, mientras que en la tapa se observa la gota colgante obtenida por mezclado de 2 μ L de Bcy1 (1-50) 3 mg/mL en buffer HEPES 25 mM pH = 6,8 50 mM NaCl, con el mismo volumen de cada solución precipitante.

Luego de estos ensayos se obtuvieron cristales de proteína en varias condiciones, utilizando como precipitantes sulfato de amonio y PEG 4000. Estos primeros cristales difractaron a muy baja resolución (4 Å), por lo que se procedió a la optimización de las condiciones de cristalización para mejorar la calidad de los mismos. En general, los cristales obtenidos con sulfato de amonio resultaron muy frágiles, por lo que se focalizó el trabajo en la utilización de PEG 4000.

Finalmente, se logró obtener cristales de proteína con buen poder de difracción mediante el método de *microseeding*, que consiste en inocular pequeños núcleos de cristalización generados en una determinada condición, en gotas con una concentración de precipitante ligeramente menor en las cuales de por sí no se generan cristales. De esta manera, se logra que los núcleos crezcan más lentamente que en la solución que les dio origen, y por lo tanto se obtienen cristales más grandes y ordenados. En la **Figura 4.8** se muestra una imagen de los cristales que se utilizaron para coleccionar los datos.

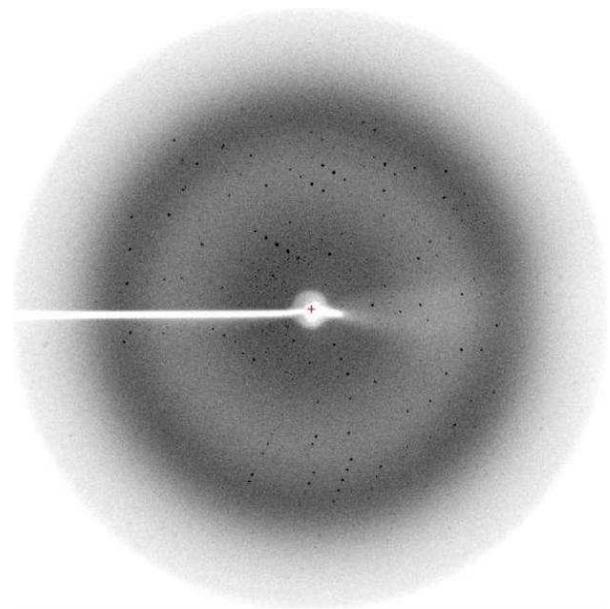


Figura 4.8: Cristales de Bcy1 (1-50).

En el panel de la izquierda se muestra una imagen de los cristales que se utilizaron para la colección de los datos, mientras que en el panel de la derecha se muestra un frame representativo del patrón de difracción. Si bien no se aprecia en la figura, los puntos se extienden hasta el borde del detector.

El cristal se colocó en una solución crioprotectora adecuada y se montó en el equipo de XRD, difractando a una resolución máxima de 2.56 Å. Se logró coleccionar un conjunto de datos completo, y luego de la indexación, integración y escalado, se pudieron obtener los parámetros de red (Tabla 4.1). El cristal pertenece al grupo de espacio P2₁ (celda monoclinica primitiva con un eje polar helicoidal). Se estimó en este punto la presencia de 8 cadenas de Bcy1 (1–50) en la unidad asimétrica del cristal.

Tabla 4.1: Estadísticas de colección de los datos y resolución de la estructura.

Colección de los datos	
Longitud de onda (Å)	1,5418
Grupo de espacio	P 1 2 ₁ 1
Parámetros de red	
(a, b, c) (Å)	56,824 52,134 68,832
(α, β, γ) (°)	90 109,7 90
Resolución (Å)	64,8 - 2,56
Reflexiones totales	50257
Reflexiones únicas	1242 (1209)
Multiplicidad	4
R _{merge}	0,125
Complejitud	99,59 (98,69)
I/σ promedio	12,03 (3,52)
Construcción del modelo y refinamiento	
Cantidad de átomos	2902
Proteína	2865
Iones inorgánicos	1
Moléculas de agua	36
Cantidad de aminoácidos (cadenas)	351 (8)
R	0,225
R _{free}	0,259
Factor B promedio (Å²)	
Proteína	32,90
Ligandos	29,00
Solvente	13,30
RMSD de la idealidad (enlaces) (Å²)	
RMSD de la idealidad (ángulos) (°)	1,89
Gráfico de Ramachandran	
Regiones más favorecidas (%)	96,5
Regiones prohibidas	0,0

Los valores entre paréntesis en la primera parte de la tabla corresponden a la capa de mayor resolución. Los datos de cantidad de átomos, cadenas y aminoácidos corresponden a la unidad asimétrica del cristal. Para obtener los números por celda unitaria hay que multiplicar esos valores por dos.

Los primeros intentos de resolver la estructura por reemplazo molecular utilizando como sondas de búsqueda las estructuras de los dominios D/D de mamíferos resultaron infructuosos, debido a la baja identidad de secuencia entre Bcy1 (1-50) y el dominio D/D de RII α (aproximadamente 40%), y sobre todo al tamaño relativamente grande de la celda unitaria, lo que siempre representa una dificultad dado que los problemas que involucran numerosas cadenas por unidad asimétrica suelen ser complejos de resolver. Para intentar sortear esta dificultad, se utilizó una combinación de métodos de modelado *ab initio* para generar sondas de búsqueda para el método de reemplazo molecular. Mediante la generación de una gran cantidad de modelos, se logró finalmente encontrar uno que resultó ser una sonda de búsqueda adecuada, y que por lo tanto permitió resolver la estructura.

Validación de la estructura

En la **Tabla 4.1** se muestran los parámetros obtenidos para la colección de datos, refinamiento del modelo y validación final del mismo. Se obtuvo una resolución final de 2,56 Å, lo que representa un valor de resolución media (ver **Figura 4.6**). El valor del R_{merge} (0,125) es ligeramente elevado, pero aceptable, dado que el valor reportado se refiere a la capa de máxima resolución. La completitud de los datos es elevada (99,59% total y 98,69% para la capa de mayor resolución). La relación señal ruido (I/σ) promedio fue alta (12,03); el valor en la capa de mayor resolución también fue elevado (3,52) lo que indica que los cristales tenían capacidad de difractar a mayor resolución.

En lo que respecta a las estadísticas del refinamiento, el factor R (22,5%) y el R_{free} (25,9%) presentaron valores más que aceptables para la resolución de los datos, como se discutió en la Introducción metodológica de este capítulo (ver **Figura 4.6**); la diferencia entre ambos fue de 3,4%, lo que también resulta más que aceptable. El factor de temperatura promedio presentó un valor que se encuentra dentro de lo esperado (32,66 Å²). La cantidad de moléculas de agua por residuo de proteína fue de 0,1, lo que es razonable para estructuras de resolución media.

En cuanto a la validación final de la estructura, el gráfico de Ramachandran presenta el 96.5% de los residuos en zonas de conformación muy favorable, y ningún residuo en las zonas no permitidas. Las desviaciones medias de la idealidad tanto para ángulos como para distancias de enlace se encuentran dentro de los valores aceptables.

Topología general de la estructura cristalina

En la estructura final, la unidad asimétrica del cristal contiene 2865 átomos de proteína correspondientes a 8 cadenas polipeptídicas, 36 átomos de oxígeno correspondientes a moléculas de agua, y un ion cloruro ubicado en el centro (**Figura 4.9**) (ver **Tabla 4.1**). Como ya se señaló, el cristal pertenece al grupo de espacio $P 1 2_1 1$, que corresponde a un sistema cristalino monoclinico. Este sistema se caracteriza por contener dos unidades asimétricas por celda unitaria; por lo tanto, la celda unitaria de esta estructura cristalina contiene 16 cadenas polipeptídicas.

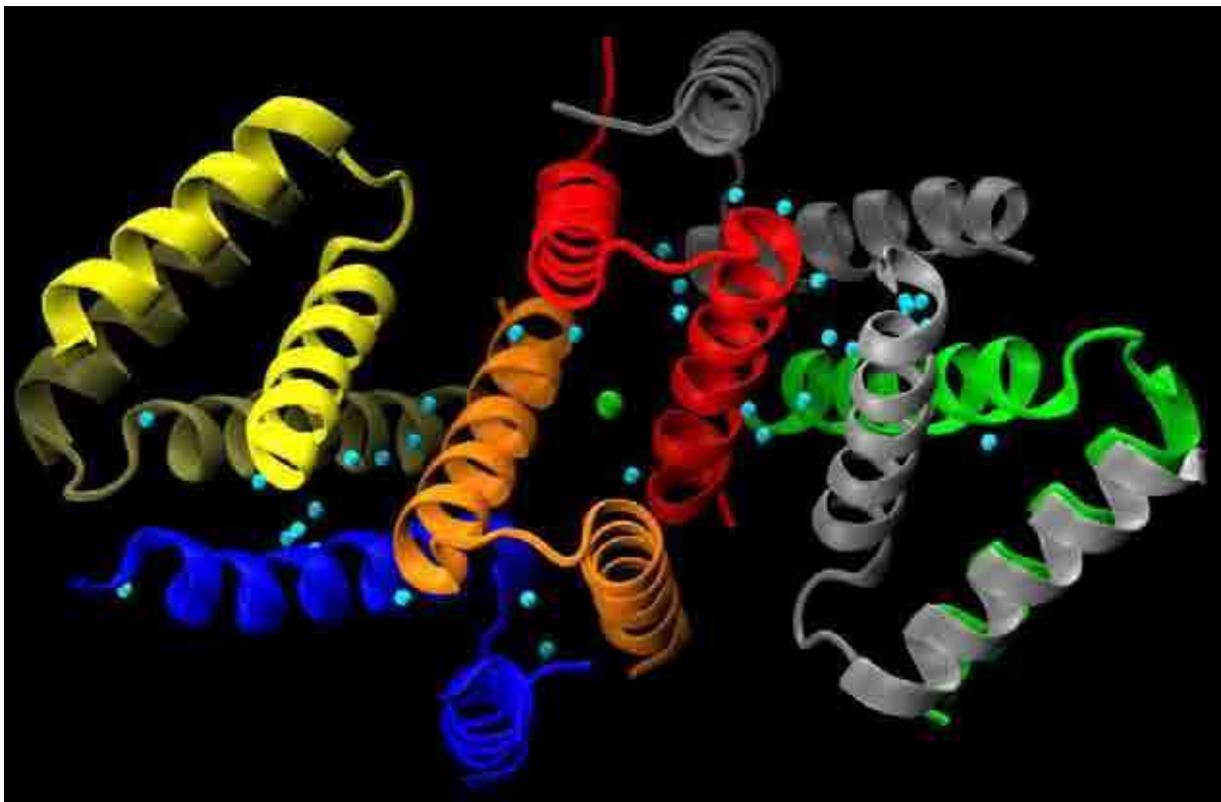


Figura 4.9: Topología general de la estructura cristalina.

Se muestra una vista de la disposición general de los distintos elementos que conforman la unidad asimétrica del cristal. Las ocho cadenas polipeptídicas se muestran en diagrama de cintas, en distintos colores. Los átomos de oxígeno correspondientes a moléculas de agua se muestran como esferas color cian, y el ion cloruro en el centro de la estructura como una esfera verde. Se observa un tetrámero formado por los dímeros A (azul) /D (naranja) y E (amarillo)/F (marrón), y otro formado por los dímeros B (rojo)/C (gris oscuro), y G (gris claro)/H (verde).

Las ocho cadenas polipeptídicas presentan la misma estructura, constituida por un motivo hélice-giro-hélice que se encuentra plegado en forma de X (**Figura 4.9** y **Figura 4.10**). La mayoría de ellas presentaron densidad electrónica bien definida entre los residuos 6 y 49, con ligeras diferencias en los extremos N y C terminales (**Tabla 4.2**) (cabe recordar que la construcción comprende los aminoácidos 1-50 de Bcy1, con una alanina adicional en el extremo N-terminal, que en la numeración utilizada es el aminoácido 1; Bcy1 (1-50) abarca por lo tanto los aminoácidos 2-51). En ninguna de las cadenas se pudieron modelar los primeros aminoácidos del extremo N-terminal debido a la falta de densidad electrónica correspondiente a esa región en los mapas. Esto se debe seguramente al desorden dinámico asociado con la desestructuración intrínseca de esa región de la molécula, lo que genera que las múltiples conformaciones se encuentren promediadas en el tiempo de colección de los datos, y como resultado la densidad electrónica se vuelve muy difusa como para ser interpretada [3]. El mismo fenómeno fue observado en la estructura cristalina del dominio D/D de la subunidad RII α de mamíferos [4].

Las ocho cadenas alinean estructuralmente muy bien entre sí (**Figura 4.10**), con un RMSD promedio de aproximadamente 1,4 Å para átomos de C α equivalentes. Incluso las cadenas laterales muestran conformaciones muy similares en todas las cadenas; las únicas excepciones a la segunda observación son los escasos residuos que preceden a la hélice 1 (Ser5, Leu6, y Lys8 -Pro7 tiene la misma disposición en todas las cadenas-; a partir de Glu9 ya todas las cadenas laterales adoptan conformaciones similares) y los últimos residuos al final de la hélice 2 (Phe47, Leu48, Lys49, Ala50 y Arg51). La variabilidad conformacional de estos residuos no resulta llamativa dado que se encuentran en los extremos de las cadenas, e incluso no en todas las cadenas se observan (ver **Tabla 4.2**). Este análisis demuestra que las ocho cadenas polipeptídicas son estructuralmente equivalentes.

Tabla 4.2: Estructura secundaria de las cadenas polipeptídicas del cristal.

En la segunda columna se indica la numeración de los segmentos que pudieron modelarse exitosamente en cada una de las cadenas por presentar densidad electrónica bien definida. En la tercera y cuarta columna se muestran los aminoácidos correspondientes a las hélices 1 y 2, respectivamente. El código de colores es indicativo de las distintas entidades oligoméricas formadas por ellas: en tonos de verde se representa el tetrámero formado por los dímeros A/D y E/F, y en naranja y rojo, el tetrámero B/C-G/H.

Cadena	Residuos modelados	Hélice 1	Hélice 2
A	6-51 (46)	8-24	29-49
B	5-49 (45)	8-24	29-48
C	5-49 (45)	8-24	29-48
D	7-49 (43)	8-24	29-49
E	9-48 (40)	9-24	29-47
F	8-50 (43)	9-24	29-49
G	6-50 (45)	8-24	29-50
H	6-49 (44)	8-24	29-48

El elevado contenido de estructura α -hélice observado (aproximadamente el 72,5 % de los residuos presentan esta conformación) es consistente con los resultados de CD del Capítulo 3 y con el análisis bioinformático del Capítulo 2. La hélice 1 comprende en promedio los residuos 8-24 (9-23 en base a la predicción bioinformática), y la hélice 2 los residuos 29-49 (28-49 en base a la predicción bioinformática).

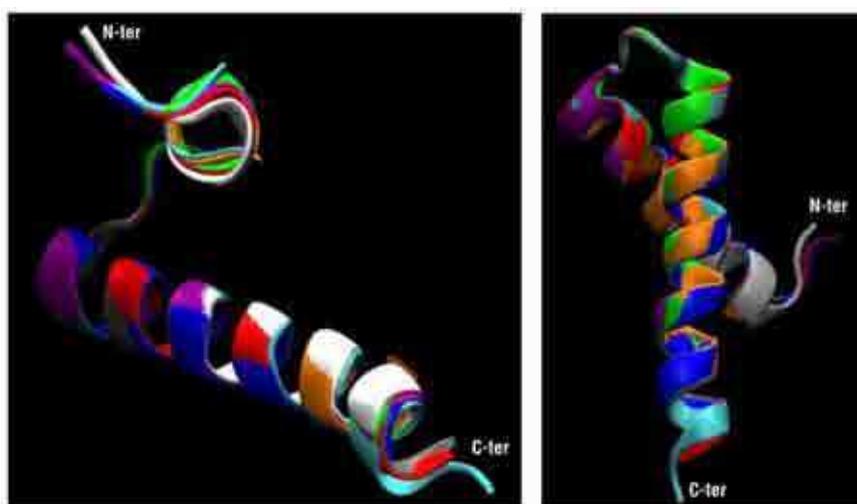


Figura 4.10: Comparación estructural de las cadenas polipeptídicas del cristal.

Alineamiento estructural de las 8 cadenas de proteína que conforman la unidad asimétrica del cristal. Cada cadena se representa como diagrama de cintas en un color distinto; se indican también sus extremos N y C terminales. Se aprecia la superposición de las hélices 1 (N-terminales) y las hélices 2 (C-terminales) de las ocho cadenas.

Una primera inspección revela que las cadenas se agrupan de a pares formando dímeros que presentan el mismo plegamiento que los dominios D/D de mamíferos, lo que se encuentra en concordancia con las predicciones bioinformáticas del Capítulo 2. A su vez, la estructura puede interpretarse en términos de dos tetrámeros equivalentes, con una disposición general de sus subunidades similar a la que se observó para el modelo de la estructura tetramérica construido en el Capítulo 3 (ver **Figura 4.9**). Cada uno de esos dos tetrámeros está compuesto a su vez por dos dímeros, y los cuatro dímeros son también estructuralmente equivalentes, como puede inferirse del hecho de que las 8 cadenas alinean muy bien, como ya se discutió. El ion cloruro se ubica en el centro de la estructura, separando ambos tetrámeros. Es importante remarcar que se acumulan numerosas moléculas de agua en las interfaces entre los dos dímeros que conforman cada tetrámero.

La estructura fue analizada bioinformáticamente utilizando los servidores PDBsum [5] y PDBePISA [6], que proporcionan información acerca de las interfaces que se generan entre las distintas cadenas de un archivo pdb. Ambos servidores identifican claramente la existencia de los dímeros en base a la cantidad de residuos presentes en la interfaz entre ambos monómeros (aproximadamente 17 de cada cadena), el área superficial de la interfaz (aproximadamente 950 Å² para cada cadena) y un elevado número de contactos entre esos residuos. Por otro lado, se observa que también existen áreas interfaciales y un número de contactos significativo entre los dímeros de cada tetrámero, en base a la suma de las contribuciones de cada cadena de un dímero con las dos cadenas del otro dímero que constituyen el tetrámero.

El servidor PDBsum predice una disposición de subunidades consistente con lo que se observa visualmente, es decir, una entidad formada por las cadenas ADEF que está en contacto con la otra entidad, BCGH, ambas conectadas principalmente por el ion cloruro central (ver **Figura 4.9**). Sin embargo, el servidor PDBePISA, que permite identificar complejos potencialmente relevantes en solución en base a cálculos estimativos de su energía libre de disociación (ΔG_{dis}), identifica las entidades diméricas EF y GH, y un tetrámero ABCD, que no resulta evidente en base a la inspección visual, ni consistente con el modelo de la estructura en solución. En este tetrámero, los dos dímeros que lo componen se encuentran conectados a través de una superficie distinta a la de los tetrámeros ADEF y BCGH, con el ion cloruro ubicado en la interfaz entre ambos dímeros (cadenas azul, naranja, roja y gris oscuro en la **Figura 4.9**).

Cabe destacar en este punto que durante la generación del modelo de la estructura tetramérica en solución abordada en el Capítulo 3, se observaron algunos modelos con una disposición de los dímeros similar a la del tetrámero ABCD; sin embargo, estos modelos fueron descartados porque presentaron valores de R_g demasiado elevados en comparación con el dato experimental.

En principio, el análisis bioinformático no necesariamente resulta determinante en cuanto a la composición de subunidades de los potenciales complejos presentes en la estructura. Antes de continuar con el análisis estructural detallado, se procedió a comparar las estructuras de los posibles complejos predichos bioinformáticamente con la información experimental obtenida por SAXS, a efectos de definir sin ambigüedad cuál de esos posibles complejos es representativo de la estructura real en solución.

Comparación de las estructuras cristalina y en solución

Para determinar cuál de los posibles complejos observados en la estructura cristalina es representativo de una especie relevante en solución, se utilizó la información de los experimentos de SAXS del Capítulo 3 (como ya se indicó en la Introducción metodológica de ese capítulo, esta técnica constituye una herramienta de validación de estructuras cristalinas muy poderosa). La estrategia consistió en comparar los perfiles de SAXS teóricos calculados para las distintas estructuras oligoméricas observadas en el cristal con el perfil experimental. Para ello, en primer lugar fue necesario completar los residuos faltantes en la estructura cristalina, dado que todos los aminoácidos contribuyen a la dispersión de rayos X.

Los complejos “candidatos” a representar la estructura real en solución son: los dímeros AD, BC, EF y GH, y los tetrámeros ADEF y BCGH (ver **Tabla 4.2**), que fueron identificados por inspección visual por su similitud estructural con el modelo de la estructura en solución generado en el Capítulo 3, y el tetrámero ABCD, identificado por el servidor PDBePISA como la más probable estructura en solución. Se utilizó el programa MODELLER [7] para completar los residuos faltantes en las estructuras que, como ya se señaló, corresponden al segmento N-terminal desestructurado y al extremo C-terminal de cada cadena. Para cada uno de los posibles complejos, se generaron 20 modelos imponiendo restricciones de simetría para asegurar que dentro de cada complejo todos los segmentos N-terminales presentaran la misma conformación en las distintas cadenas. En principio, dado que estas regiones son

flexibles, esta situación no es representativa de lo que ocurre realmente en solución; sin embargo, la manera más sencilla de analizar la influencia de la conformación de esas regiones sobre el perfil de SAXS y sobre el R_g es imponer estas restricciones. Por otro lado, los modelos generados por NMR de las estructuras en solución de los dominios D/D de las subunidades RII α de mamíferos son siempre simétricos en cuanto a la conformación de este segmento (pdb: 1R2A).

Posteriormente, los 20 modelos de cada grupo (140 modelos en total) fueron enviados al servidor FoXS [8] el cual, como ya se explicó en el Capítulo 3, permite evaluar el perfil teórico de SAXS de una estructura a partir de su archivo pdb, y también proporciona una estimación de su R_g en solución. Se analizó la bondad del ajuste para todos los grupos de modelos en términos del *score* χ^2 proporcionado por el programa, y se compararon los valores de sus R_g con el valor experimental (1,83 nm). En la **Figura 4.11** se presentan los resultados de este análisis.

Como primera observación, es importante destacar que en todos los casos el modelado de los segmentos faltantes mejora sustancialmente el ajuste, siendo la estructura cristalina incompleta el modelo que presentó el peor ajuste a los datos experimentales para todos los grupos. Por otro lado, en lo que respecta a la influencia de la conformación del segmento N-terminal en el valor del R_g , se observó que los modelos que presentan segmentos más extendidos tuvieron valores de R_g más altos, como era esperable. Las diferencias entre las distintas estructuras equivalentes (dímeros AD, CD, EF y GH, y tetrámeros ADEF y BCGH) en cuanto a los valores de R_g obtenidos se deben seguramente a que presentan distinta cantidad de residuos faltantes en la estructura cristalina (ver **Tabla 4.2**). Dado que estos segmentos fueron modelados como *loops* desestructurados, los complejos a los que le faltan más residuos presentan conformaciones más elongadas para estos segmentos, que se traducen en valores de R_g más altos; este es el caso por ejemplo para el tetrámero ADEF en comparación con su equivalente BCGH, o para el dímero EF en comparación con los demás, dado que las cadenas E y F son las más pobres en cuanto a la cantidad de residuos presentes en la estructura cristalina.

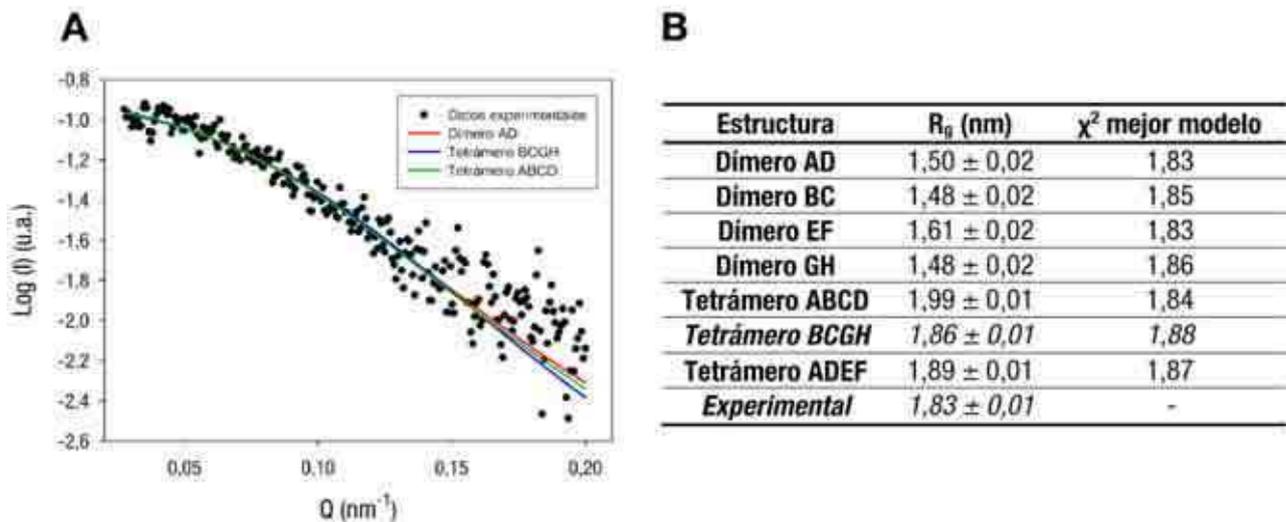


Figura 4.11: Validación por SAXS de los complejos observados en la estructura cristalina.

(A) Ajuste del mejor modelo de dímero y tetrámeros a los datos experimentales de SAXS. Los valores de χ^2 para el ajuste se muestran en el panel (B). (B) Valores proporcionados por el servidor FoXS para el ajuste del perfil teórico a los datos experimentales para las distintas estructuras. Los valores del R_g corresponden a la media de los 20 modelos generados para cada grupo con su desvío estándar. El χ^2 corresponde al del modelo de cada grupo que mejor ajusta a los datos experimentales (en muchos casos, se encontraron varios modelos con el mismo valor para este parámetro).

Cabe observar que a grandes rasgos todas las estructuras ajustan igualmente bien a los datos experimentales; las diferencias en el parámetro χ^2 entre los distintos grupos resultan muy pequeñas, y por lo tanto no resulta conveniente extraer conclusiones a partir de ese valor. Por el contrario, sí se observan diferencias notables entre los distintos grupos en el R_g ; consecuentemente, se utilizó este parámetro como criterio para evaluar la compatibilidad entre los distintos complejos del cristal y la estructura en solución. Se observa claramente que todos los modelos diméricos presentan valores de R_g significativamente menores que el experimental, en concordancia con los resultados del Capítulo 3, que demostraron que en estas condiciones experimentales la proteína existe exclusivamente como un tetrámero en solución (es importante señalar que las concentraciones utilizadas para la cristalización son similares a las de los experimentos de SAXS). En segundo lugar, la observación más importante de este análisis es que el tetrámero ABCD, predicho por el servidor PDBePISA, presenta un valor de R_g que también es significativamente distinto del valor experimental (1,99 nm y 1,83 nm, respectivamente). En contraposición, el tetrámero BCGH, identificado por

inspección visual por su estructura similar al modelo de la estructura en solución, presenta un valor de R_g muy cercano al experimental (1,86 nm) lo que sugiere fuertemente que, pese a las predicciones bioinformáticas, esta estructura es la que existe en solución.

Para obtener más información que permita reforzar esta suposición, se computaron las propiedades hidrodinámicas de estos dos tetrámeros utilizando el programa HYDROPRO [9], que fue descrito en el Capítulo 3. Los parámetros de los dos tetrámeros resultaron muy similares (no se muestra), a excepción de la D_{max} , que resultó ser bastante más elevada para el tetrámero ABCD, siendo la del tetrámero BCGH más cercana al valor experimental. Por otro lado, se realizaron superposiciones del modelo generado *ab initio* a partir de los datos de SAXS en el Capítulo 3 (no se muestra) con los modelos completos de los dos tetrámeros; si bien el tetrámero BCGH parece adaptarse algo mejor a la forma del contorno de ese modelo, las dos estructuras se encuentran bien contenidas en el mismo, por lo que este criterio tampoco resulta determinante para discriminar entre ambas.

Como comentario final de esta sección, es importante discutir algunos aspectos del cálculo efectuado por el servidor PDBePISA para identificar los complejos potencialmente estables en solución. En primer lugar, hay que mencionar que estos cálculos siempre son estimativos y sirven a manera de comparación, pero no deben interpretarse como valores absolutos realmente confiables para el ΔG_{dis} del complejo. En segundo lugar, para el caso particular del tetrámero predicho ABCD, resulta significativo que de acuerdo a este servidor el ion cloruro central proporciona una contribución energética muy importante a la interacción entre las cadenas B y D (**Figura 4.12**); las cadenas A y C no presentan ningún contacto entre sí. En contraposición, en el tetrámero BCGH las dos cadenas de cada dímero presentan interacciones con las dos cadenas del otro dímero (ver **Figura 4.9**); más aun, el área interfacial resulta mayor para el tetrámero BCGH que para el tetrámero ABCD (1300 Å² para el primero y 700 Å²). Adicionalmente, el hecho de que el tetrámero BCGH no sea identificado como energéticamente favorable por el servidor podría encontrarse en concordancia con la idea desarrollada en el Capítulo 3 acerca de que el equilibrio dímero-tetrámero es dinámico, y que el estado oligomérico se encuentra sujeto a regulación mediada por la concentración de proteína.

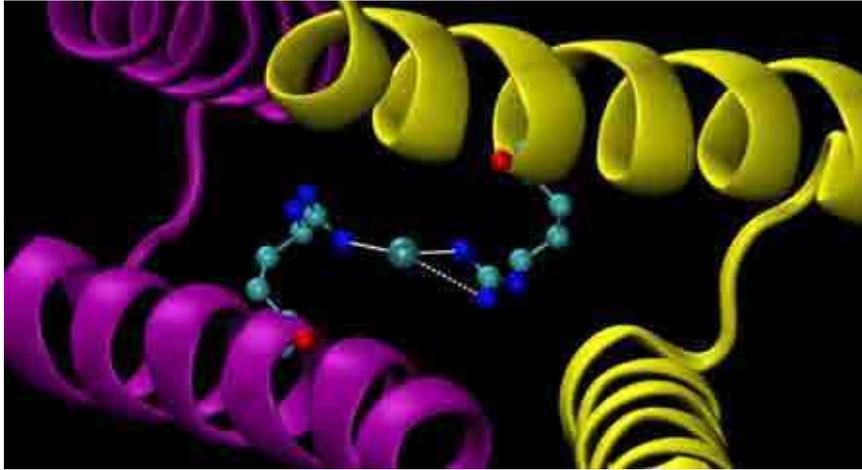


Figura 4.12: Participación del ion cloruro en el tetrámero ABCD.

Los residuos Arg40 ubicados hacia el extremo N-terminal de las hélices 2 de las cadenas B (amarillo) y D (púrpura) establecen interacciones electrostáticas con el ion cloruro central, que aportan contribuciones energéticas importantes en base al cálculo del servidor PDBePISA. Se indican en línea punteada las distancias menores a 4 Å. Si bien no se representa por claridad, las cadenas laterales de los aminoácidos se encuentran bien contenidas en el contorno de densidad electrónica.

Resulta difícil establecer si un complejo mediado por cloruro podría ser fisiológicamente relevante. Se ha demostrado que la concentración citosólica de cloruro en *S. cerevisiae* es baja (alrededor de 0,2 mM) aun cuando se somete a las células a un rango amplio de concentraciones extracelulares, llegando incluso hasta 100 mM, pese a que la membrana plasmática es permeable a cloruro [10]. Más aun, el cloruro en esta concentración resulta tóxico, ya que inhibe la transcripción en extractos libres de células de cepas WT [10]. El medio utilizado para la cristalización contiene una concentración de cloruro 0,4 M, es decir, 2000 veces mayor a la concentración intracelular. Estas observaciones podrían reforzar la idea de que el tetrámero ABCD es un artificio generado por las condiciones de cristalización, y no es representativo de la estructura real en solución.

En concordancia con esta observación, cabe observar que si el tetrámero ABCD fuera realmente la estructura en solución, al disolverse el cristal debería generarse una solución mezcla de tetrámeros y dímeros. La concentración de Bcy1 (1-50) utilizada para obtener cristales fue de 1,5 mg/mL (0,25 mM), que resulta intermedia entre los experimentos de SAXS (1 mM y 0,5 mM) y los de SEC-SLS (100 μM). En ambos experimentos se demostró que las muestras eran monodispersas; por lo tanto, resulta lógico pensar que la muestra utilizada

para obtener los cristales también debía serlo, y en consecuencia resultaría llamativo cristalizar a partir de una solución homogéneamente compuesta por tetrámeros una entidad conformada por un tetrámero y dos dímeros.

En conclusión, el análisis de la sección anterior sugiere fuertemente que la estructura más probable del tetrámero es la que se observó para el modelo de la estructura en solución. Si bien el criterio basado en las diferencias en el R_g , junto con las consideraciones posteriores, parecen proporcionar argumentos concluyentes, para resolver definitivamente la cuestión sería deseable emplear algún abordaje experimental; por ejemplo, podrían mutarse por alanina residuos potencialmente clave para el establecimiento de una u otra estructura, y analizar el estado oligomérico de la proteína resultante por SEC o por SEC-SLS, si fuera necesario.

En las siguientes secciones se describirán en detalle las estructuras de las entidades dimerica y tetramérica, en base a su inspección visual y a la información proporcionada por los servidores mencionados anteriormente.

Análisis de la estructura del dímero

Para desarrollar este análisis, se decidió trabajar con la entidad formada por las cadenas B y C, por ser la que presenta de entre todas las posibilidades la mayor cantidad de residuos, y por estar contenida en el tetrámero BCGH, que es el que se discutirá en la siguiente sección. Se utilizaron los servidores PDBsum y PDBePISA para efectuar el análisis.

Como ya se discutió, la topología de estructura secundaria es similar a la de los dominios D/D de la subunidad RII α de mamíferos. El segmento N-terminal desestructurado comprende los residuos 1-7, la hélice 1 los residuos 8-24, y la hélice 2 los residuos 29-48. El *loop* que conecta las dos hélices (residuos 25-28) está constituido por la secuencia Asn25-Pro26-Ser27-Asp28. La principal diferencia radica en la orientación relativa de las dos cadenas: cuando se superpone una de las cadenas de los dos dominios, la otra queda muy desplazada, sobre todo en lo que respecta a la hélice 1 (**Figura 4.13**).

La interfaz del dímero está conformada por una superficie total de aproximadamente 950 Å² de cada una de las cadenas. Existen unos 20 residuos de cada cadena que aportan contribuciones significativas al área interfacial. Las moléculas de agua quedan completamente excluidas del núcleo central del dímero, que está formado principalmente por aminoácidos

hidrofóbicos. Los principales aminoácidos que contribuyen a estos contactos son: Leu6, Ser10, Glu13, Leu14, Phe17, Gln18, Ile21, Asn25, Pro26, Phe29, Leu30, Ser33, Ala34, Phe37 y Leu41 de las dos cadenas (**Figura 4.14**). Por otro lado, se establecen dos puentes de hidrógeno entre los residuos de Asn38 de ambas cadenas, ubicados aproximadamente en la mitad de la hélice 2: el oxígeno del carbonilo de cada uno de los residuos es aceptor de un puente de hidrógeno del nitrógeno de la amida del otro.

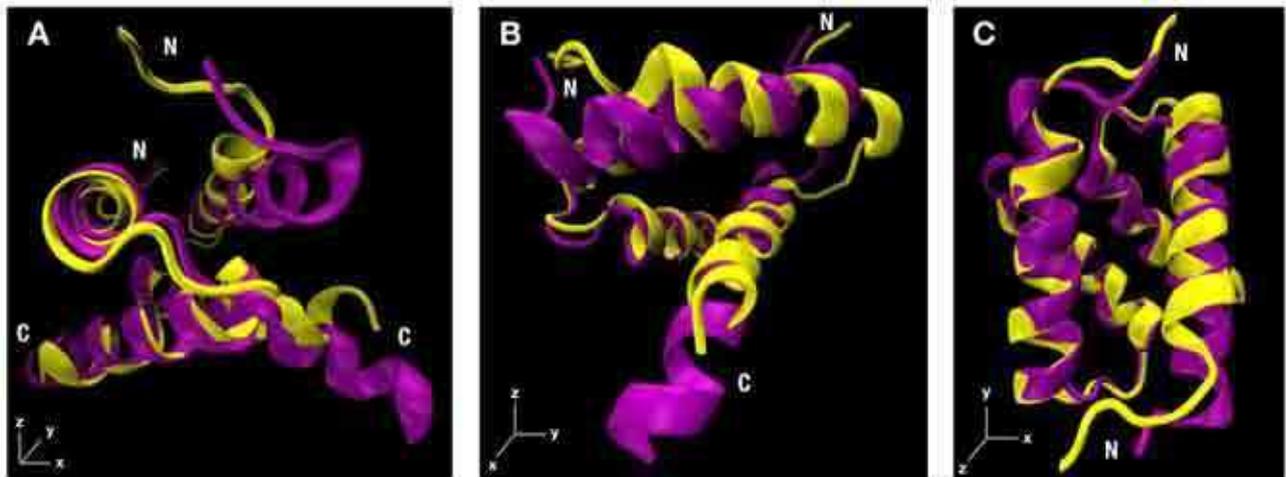


Figura 4.13: Comparación de la estructura global del dímero Bcy1 (1-50) con el dominio D/D de RII α .

Se muestran tres vistas de una superposición de la estructura cristalina del dímero Bcy1 (1-50) (púrpura) con la estructura en solución del dominio D/D de la subunidad RII α (pdb: 1R2A) en diagrama de cintas. Se indican los ejes como referencia, y los extremos N y C terminales de cada cadena. Se superpuso la cadena que se observa a la izquierda en el panel (A). El panel (B) se encuentra rotado respecto del panel (A) 90° alrededor del eje z en sentido horario, mientras que el panel (C) se encuentra rotado respecto del panel (A) 90° alrededor del eje x en sentido antihorario.

Para apreciar la contribución relativa que aportan estos residuos al establecimiento del dímero, en primer lugar se observó cuáles son los más importantes de acuerdo a la cantidad de contactos con la otra cadena en los que participan, y al área interfacial que aportan. Estos residuos son: Phe29 (contactos con 4 residuos distintos y 91 Å²), Leu30 (contactos con 5 residuos distintos y 121 Å²), Phe37 (contactos con 5 residuos distintos 83 Å²) y Leu41 (contactos con 3 residuos distintos y 78 Å²). Por otro lado, se utilizó el módulo *Interface Alanine Scanning* del servidor Robetta [11], que permite identificar computacionalmente residuos potencialmente clave en contactos entre proteínas, mediante una estimación de la

diferencia entre las estabilidades relativas del complejo WT y los complejos que presentan mutados individualmente sus residuos por alanina. Para ello, el programa calcula la variación de energía libre de formación del complejo asociada a esa mutación ($\Delta\Delta G_{\text{bind}}$):

$$\Delta\Delta G_{\text{bind}} = (\Delta G_{\text{complejo}}^{\text{MUT}} - \Delta G_{\text{cadena A}}^{\text{MUT}} - \Delta G_{\text{cadena B}}^{\text{MUT}}) - (\Delta G_{\text{complejo}}^{\text{WT}} - \Delta G_{\text{cadena A}}^{\text{WT}} - \Delta G_{\text{cadena B}}^{\text{WT}})$$

donde MUT y WT representan a las variantes mutante (con un residuo en particular mutado por alanina) y WT de las proteínas del complejo respectivamente, y ΔG son las estabilidades de cada uno de los componentes, calculadas por el programa. Valores positivos y relativamente altos del $\Delta\Delta G_{\text{bind}}$ se interpretan como que ese residuo presenta una contribución apreciable al establecimiento del complejo. De acuerdo a este servidor, los 4 residuos más importantes para el establecimiento del dímero son: Leu14, Phe29, Leu30, Phe37 (de acuerdo al criterio establecido en [11], de que un valor de $\Delta\Delta G_{\text{binding}} > 1,0$ kcal/mol resulta significativo).

alrededor del eje y en sentido horario. En el panel (D) se muestran por comparación los residuos más importantes para el establecimiento del dímero en las subunidades RII α (pdb: 1R2A). (Figura tomada de [12]). En la parte inferior de la figura se muestra un alineamiento de los dominios D/D de Bcy1 y de RII α . Se muestran sombreados en gris los residuos que establecen contactos clave para el dímero en las estructuras de alta resolución de ambas proteínas. La flecha negra indica la posición de la prolina central, y las estrellas negras los residuos que al ser mutados por alanina generan subunidades RII α monoméricas. Las cruces en la parte inferior representan los residuos identificados como más relevantes para la dimerización de Bcy1. La numeración corresponde a la construcción Bcy1 (1-50).

Es interesante notar que de los cinco residuos potencialmente clave para el establecimiento del dímero de Bcy1, los cinco participan de la interfaz de dimerización de las subunidades RII α ; tres se encuentran conservados estrictamente, y dos de éstos generan subunidades R monoméricas en mamíferos cuando son mutados por alanina (**Figura 4.14**).

Análisis de la estructura del tetrámero

Dado que los dos tetrámeros son estructuralmente equivalentes y presentan el mismo conjunto de interacciones, se analizó el tetrámero BCGH, que es el que presenta mayor cantidad de aminoácidos modelados, utilizando los servidores PDBsum y PDBePISA, y por inspección visual de la estructura. La topología general del tetrámero se muestra en la **Figura 4.15**.

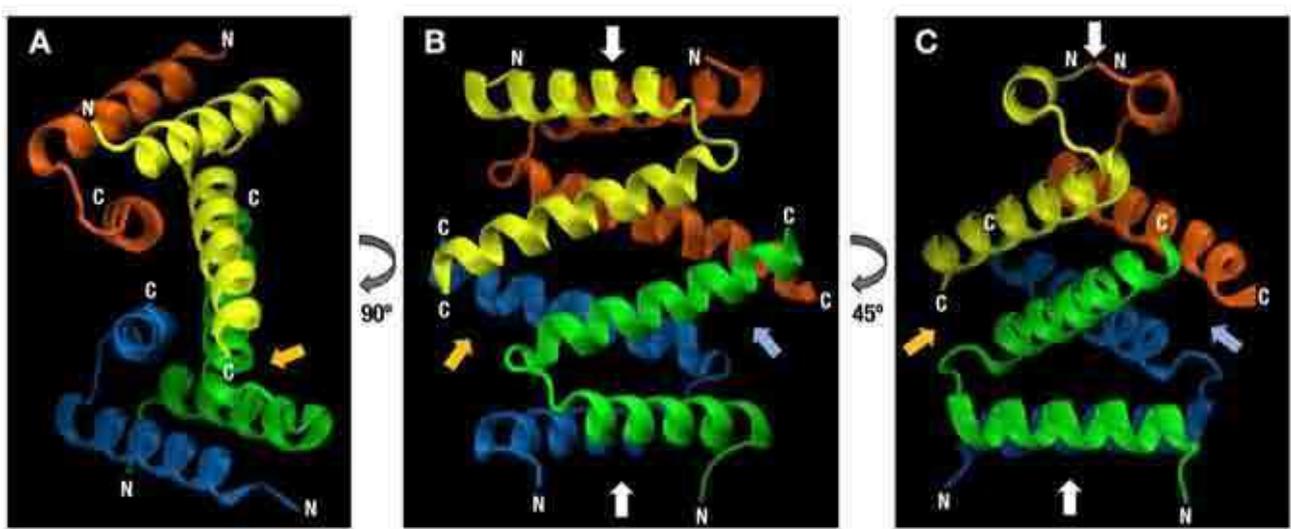


Figura 4.15: Topología general del tetrámero Bcy1 (1-50).

Se muestran tres vistas en diagrama de cintas del tetrámero Bcy1 (1-50). Las cuatro cadenas se representan en colores distintos por claridad: azul (B) y verde (C) (dímero 1), y naranja (G) y amarillo (H) (dímero 2). Se indica el ángulo de rotación entre los distintos paneles, y los extremos N y C terminales de cada cadena. La flecha blanca señala la posición de las potenciales superficies de unión a AKAPs. Las flechas naranja y azul indican la posición aproximada de las interacciones que se detallan en la **Figura 4.16 (C)** y **(D)**, respectivamente.

El tetrámero está formado por dos dímeros estructuralmente equivalentes, que interactúan mediante las caras accesibles al solvente (las que no contribuyen al núcleo de dimerización) de las hélices 2 (las hélices C-terminales) de los cuatro protómeros. El núcleo de tetramerización está constituido por un manojo formado por estas cuatro hélices; las dos hélices 2 de cada dímero se encuentran dispuestas en forma de X una respecto de la otra, y cada una de ellas está alineada de forma antiparalela con la hélice correspondiente del protómero opuesto (es decir, el extremo N-terminal de una contra el extremo C-terminal de la opuesta). La estructura es altamente simétrica; por ejemplo, existe un eje C_2 que atraviesa las dos superficies potenciales de unión a AKAPs (indicadas con flechas blancas en la **Figura 4.15**). Estas superficies están expuestas al solvente, lo que resulta consistente con su potencial participación en la formación de complejos macromoleculares; están dispuestas formando un ángulo de casi 90° entre ellas (**Figura 4.15 (C)**). Más aun, como consecuencia de la estequiometría del complejo existen dos de estas superficies por cada entidad, lo que representa una diferencia remarcable con los dominios D/D de las subunidades R de mamíferos.

Resulta interesante remarcar que los extremos C-terminales de cada una de las cuatro cadenas presentan una disposición tal que hay suficiente espacio como para que el *linker* se ensamble a continuación (ver por ejemplo la **Figura 4.15 (B)**). Debido a la cercanía de los extremos C-terminales de cadenas adyacentes, esta observación sugiere que al menos las regiones N-terminales de dos *linkers* contiguos podrían encontrarse muy cercanas en el espacio, lo que se encuentra en concordancia con la hipótesis formulada en el Capítulo 2 de que la desaparición de los péptidos correspondientes al segmento 58-80 por entrecruzamiento químico podría ser evidencia de la cercanía espacial entre estos segmentos; cabe recordar que esta región contiene el *cluster II* de serinas fosforilables, que podría ser un elemento regulatorio.

La estructura observada es similar al modelo generado en el Capítulo 3 en base a la información obtenida de los experimentos en solución, con la diferencia de que las entidades diméricas se encuentran mejor ensambladas en la estructura cristalina. Esta disposición espacial sugiere un modelo para el tetrámero completo de Bcy1: el núcleo formado por el dominio D/D tetramérico se encontraría en el centro de la estructura, y a partir de él se extenderían de forma radial hacia fuera los *linkers* de las cuatro cadenas, con los dominios de unión a cAMP ubicados en la periferia.

La estructura puede analizarse en términos de dos categorías según la forma en que interactúan las cadenas de a pares: existen dos pares de cadenas cuyas hélices 2 se encuentran alineadas de forma paralela y que por lo tanto comparten una cara entera (B-G y C-H), mientras que los otros dos pares de cadenas se disponen en forma de X e interactúan por lo tanto de manera cruzada, mediante sus extremos N-terminales (B-H y C-G) (ver **Figura 4.15**). El primer tipo de interfaz presenta una superficie de interacción algo mayor, y también mayor cantidad de residuos involucrados en la misma (**Tabla 4.3**).

El área inaccesible al solvente en la interfaz ubicada entre los dos dímeros que conforman el tetrámero es de aproximadamente 1300 Å². Se observan moléculas de agua en la interfaz, algunas de las cuales podrían desempeñar alguna función estructural, como se discutirá más adelante. Existen unos 11 residuos de cada cadena que aportan contribuciones significativas al área interfacial. En concordancia con el modelo de la estructura en solución presentado en el Capítulo 2, y en contraste con el núcleo hidrofóbico de dimerización de los dímeros individuales, la interfaz del tetrámero está dominada por residuos polares y cargados, que median interacciones electrostáticas y puentes de hidrógeno (ver **Tabla 4.3**).

Tabla 4.3: Resumen de las interacciones que sostienen el tetrámero.

Se muestran el área superficial excluida del solvente, los residuos involucrados, y los tipos de interacciones para las distintas interfaces que constituyen el tetrámero. En la tercera fila se muestran en negrita los residuos que participan de las dos interfaces. La cantidad de interacciones de cada tipo indicadas corresponden a una de las dos interfaces simétricamente relacionadas, por lo que para obtener el total de interacciones los números de las filas 4, 5 y 6 deben multiplicarse por dos.

Interfaz	B-G/C-H	B-H/C-G
Área interfacial (Å ²)	350	300
Residuos de cada cadena involucrados (cantidad)	Asp28, Leu30, Gln31, Asn35, Asn38 , Glu42, Arg45 , Lys49 (8)	Asn38 , Leu41, Gln44, Arg45 , Leu48 (5)
Puentes de hidrógeno	3	2
Puentes salinos	2	0
Moléculas de agua estructurales	3	0

Existen un total de 8 residuos cargados negativamente, 8 residuos cargados positivamente, 16 residuos polares no cargados, y 12 residuos de leucina que aportan interacciones relevantes para el establecimiento del tetrámero. Adicionalmente, se observan tres moléculas de agua que están involucradas en puentes de hidrógeno entre cadenas adyacentes. Estos elementos se encuentran dispuestos de forma tal que los 12 residuos hidrofóbicos se encuentran sumergidos en la interfaz del tetrámero (**Figura 4.16**, paneles **(A)** y **(B)**). Existe además una cavidad central en la que se acomodan las moléculas de agua que resultan estructuralmente relevantes. El tetrámero se encuentra sostenido por un total de 10 puentes de hidrógeno, al menos 4 puentes salinos (dado que para algunos pares de residuos, por ejemplo Arg-Asp, puede identificarse más de una interacción) y al menos 6 puentes de hidrógeno mediados por agua (dado que es posible que existan más moléculas de agua en la cavidad central que no se observaron en el modelo).

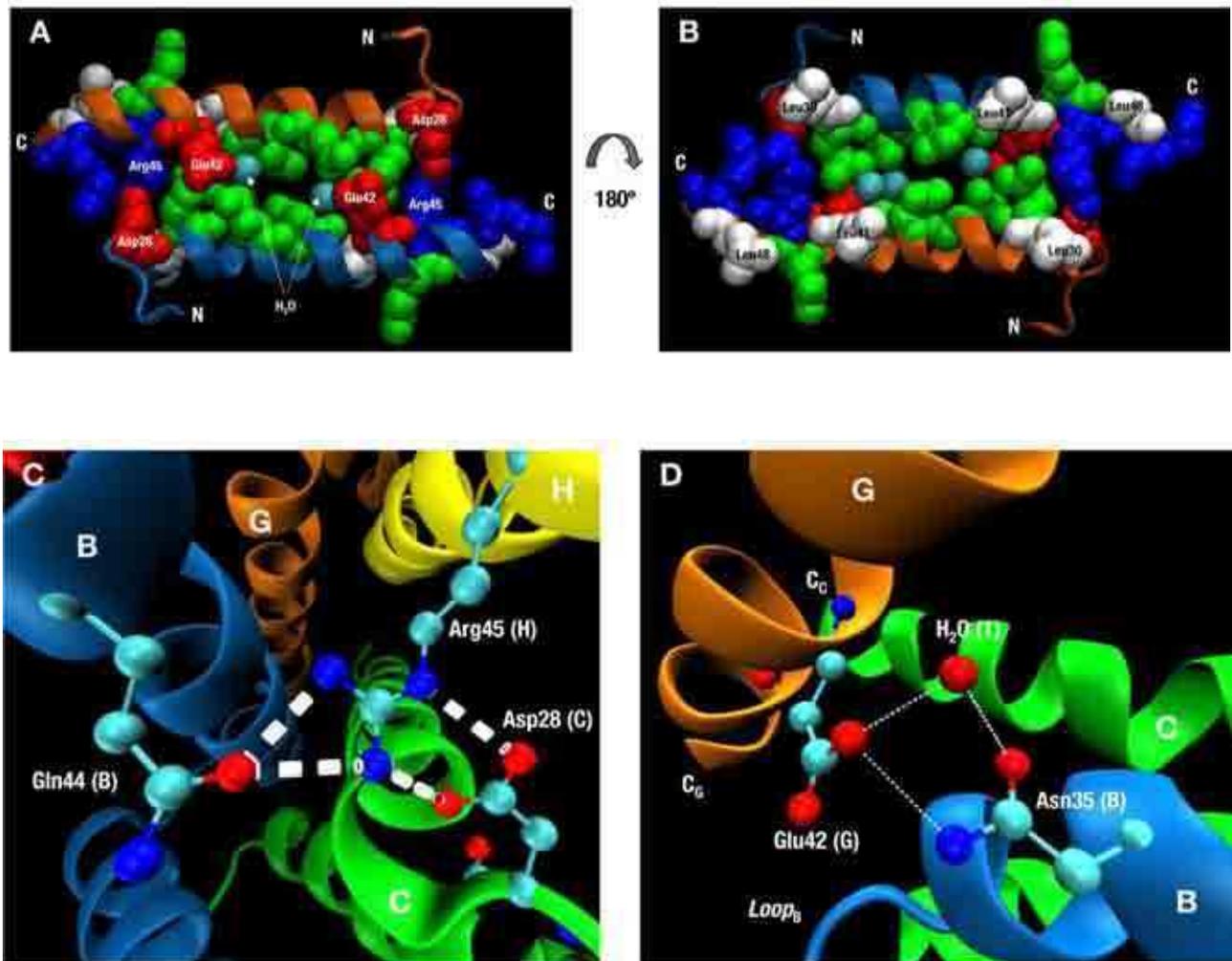


Figura 4.16: Residuos que establecen contactos importantes en el tetrámero Bcy1 (1-50).

En la parte superior se muestran dos vistas de la interfaz formada por las cadenas B (azul) y G (naranja) (segunda columna en la **Tabla 4.3**). Se muestran solamente las hélices 2 de cada una de las cadenas (residuos 25-51), dado que son las que establecen los contactos. Las hélices se representan en diagrama de cintas, y las cadenas laterales de los aminoácidos como esferas de Van der Waals. El código de colores para los aminoácidos es: azul (básico), rojo (ácido), verde (polar, no cargado) y blanco (hidrofóbico). Las moléculas de agua se representan como esferas turquesas. Se indican los extremos N y C terminales de cada una de las cadenas (donde el extremo N-terminal corresponde al aminoácido 25). Se indica también la ubicación de los residuos más importantes para el establecimiento del tetrámero. En el panel (A) se muestra la superficie expuesta al solvente, mientras que en el panel (B) se muestra la superficie sumergida en la interfaz del tetrámero, que interactúa con la superficie simétricamente relacionada formada por las cadenas C y H. En los paneles inferiores se muestran en detalle algunas de las interacciones más relevantes para la tetramerización. (C) Contacto entre tres de las cuatro cadenas del tetrámero mediado por el residuo Arg45. (D) Contacto entre las hélices 2 apareadas de las cadenas B y G mediado por una molécula de agua estructural. En ambos paneles, las líneas punteadas representan probables interacciones de puente hidrógeno y/o puente salino. Si bien no se muestran por simplicidad, los residuos resaltados se encuentran bien contenidos

en los contornos de densidad electrónica de los mapas. La ubicación de estos residuos en el contexto de la estructura global puede apreciarse en la **Figura 4.15**, señalada con flechas naranja y azul.

Para apreciar la contribución relativa que aportan los distintos residuos a la estabilidad del tetrámero, en primer lugar se observó cuáles son los más importantes de acuerdo a la cantidad de contactos en los que participan y al área interfacial que aportan. Estos residuos son: Asp28 (participa en dos puentes hidrógeno y puentes salinos; aporta 50 Å²), Glu42 (participa en un puente hidrógeno y aporta 58 Å²) y Arg45 (participa en dos puentes hidrógeno y puentes salinos; aporta 85 Å²). Por otro lado, los residuos hidrofóbicos Leu41 y Leu48, aunque naturalmente no participan del mismo tipo de interacciones que los anteriores, aportan áreas interfaciales significativas (45 y 90 Å² respectivamente). Todos estos residuos se muestran en la **Figura 4.16**.

Con el objetivo de identificar los residuos potencialmente clave para el establecimiento del complejo, nuevamente se utilizó el módulo *Interface Alanine Scanning* del servidor Robetta. Los residuos que presentaron un *score* significativo son: Asp28, Asn35, Gln44, Arg45 y Leu48. De estos residuos, el que presentó un score particularmente elevado fue Arg45; este residuo parece ser central en el establecimiento de la estructura del tetrámero en base al análisis anterior; reforzando esta idea, cabe destacar además que participa de los dos tipos de interfaces definidas en la **Tabla 4.3**.

En los paneles inferiores de la **Figura 4.16** se ilustran algunas de las interacciones más relevantes establecidas por estos residuos. Los residuos Asp28 (ubicado en el *loop* entre las dos hélices), Gln44 y Arg45 (ubicados hacia el final de la hélice 2) de tres cadenas distintas están involucrados en una importante red de puentes de hidrógeno y puentes salinos (**Figura 4.16 (C)**). Resulta destacable también que Arg45 de la cadena H (en este ejemplo) media un contacto entre dos monómeros de un mismo dímero, las cadenas B y C, que no existiría de no formarse el tetrámero. Es importante destacar que, por la simetría de la estructura, en cada “punta” del tetrámero formada por los extremos C-terminales de dos cadenas de dímeros opuestos (ver por ejemplo la **Figura 4.15 (B)**), existe un par de esta tríada de residuos; de esta manera, existen en total cuatro de estos *clusters* en el tetrámero, dos en cada extremo, que sirven a modo de “cierre”. Es probable que estos elementos constituyan las interacciones más importantes que justifican la existencia del tetrámero.

En la **Figura 4.16 (B)** se muestra otra interacción importante que está mediada por una molécula de agua que posiblemente desempeñe un rol estructural. Reforzando esta idea, se observó que en distintas posiciones equivalentes del cristal, en los dos tetrámeros, el modelo contiene otras moléculas de agua en una disposición muy similar.

Análisis de las secuencias de los dominios D/D de diversos organismos

Considerando que la estructura tetramérica de una subunidad R de la PKA descrita en este trabajo se encuentra en fuerte discordancia con el concepto previamente establecido de que todas estas proteínas son dimericas, y que los únicos modelos para los que se cuenta con información estructural de los dominios D/D son mamíferos y *S. cerevisiae*, se llevó a cabo un análisis de secuencias para determinar cuáles son los determinantes aminoacídicos que justifican las diferencias en la estructura cuaternaria, y especular acerca de la posible existencia de subunidades R tetraméricas en otros organismos. A pesar de que ya se ha publicado un estudio filogenético de los dominios de unión a cAMP [13], no existe ningún trabajo sistemático sobre los dominios D/D.

En primer lugar, se analizaron las secuencias de los dominios D/D de hongos. En el trabajo de Tesis Doctoral de Jimena Rinaldi, llevado a cabo en nuestro laboratorio, se encontró mediante un análisis bioinformático de 40 secuencias de subunidades R fúngicas que existe un grupo de hongos, pertenecientes al *subphylum* Pezizomycotina, que no presentan en su extremo N-terminal un dominio D/D, y por lo tanto se espera que sus subunidades R sean monoméricas, al igual que las proteínas de *Dictyostelium discoideum* y *Paramecium*, para las cuales el estado oligomérico ya ha sido estudiado [14, 15].

En primer lugar, se decidió ampliar esa colección para lograr mayor representatividad de los distintos grupos taxonómicos del reino *Fungi*. Por lo tanto, se construyó una colección de 99 secuencias de subunidades R de hongos, que fueron obtenidas de la base de datos UniProt utilizando la herramienta BLAST. Estas secuencias fueron agrupadas en base a la clasificación taxonómica de los hongos [16] en los siguientes grupos: Pezizomycotina (46 secuencias), Saccharomycotina (25 secuencias), Basidiomycota (14 secuencias), Mucoromycotina (10 secuencias), Taphrinomycotina (2 secuencias), Blastocladiomycota (1 secuencia) y Chytridiomycota (1 secuencia). La mayoría de los hongos presentan una única isoforma de la subunidad R de la PKA, a excepción de los miembros del *subphylum* Mucoromycotina [17].

El análisis bioinformático de estas secuencias consistió en la construcción de un alineamiento múltiple y en la predicción de los elementos de estructura secundaria de todas ellas (no se muestra). Como era esperable, se encontraron tres regiones distintas en base a su grado de conservación: los dominios de unión a cAMP, que se encuentran altamente conservados entre todos los hongos; la región del *linker* que, por el contrario, presenta baja e incluso regiones de nula conservación (siendo el *linker II* más conservado que el *linker I*), con una gran proporción de residuos de prolina y glicina. Las predicciones sugieren que esta región se encuentra desestructurada, consistentemente con la falta de conservación y con la escasa tendencia de los residuos de glicina y prolina a participar en elementos de estructura secundaria; el IS es el único segmento dentro de esta región que presenta alta conservación.

Finalmente, en concordancia con la Tesis Doctoral de Jimena Rinaldi, se observaron dos grandes grupos de secuencias en cuanto a las propiedades del extremo N-terminal: un grupo al que se le predice un motivo hélice-giro-hélice, que corresponde a un fragmento de aproximadamente 50 aminoácidos con algunos residuos característicos de los dominios D/D de mamíferos, y un grupo que no presenta estos elementos. De acuerdo con estas observaciones, el primer grupo (hongos de Basidiomycota, Blastocladiomycota, Chytridiomycota, Mucoromycotina, Saccharomycotina, y 10 miembros de Pezizomycotina) presenta un dominio D/D, mientras que el segundo grupo (la mayor parte de los hongos de Pezizomycotina) no lo presenta. En base a este análisis, un total de 63 proteínas se predicen diméricas, mientras que las 36 restantes se espera que sean monoméricas. Para el análisis siguiente se utilizaron solamente las primeras.

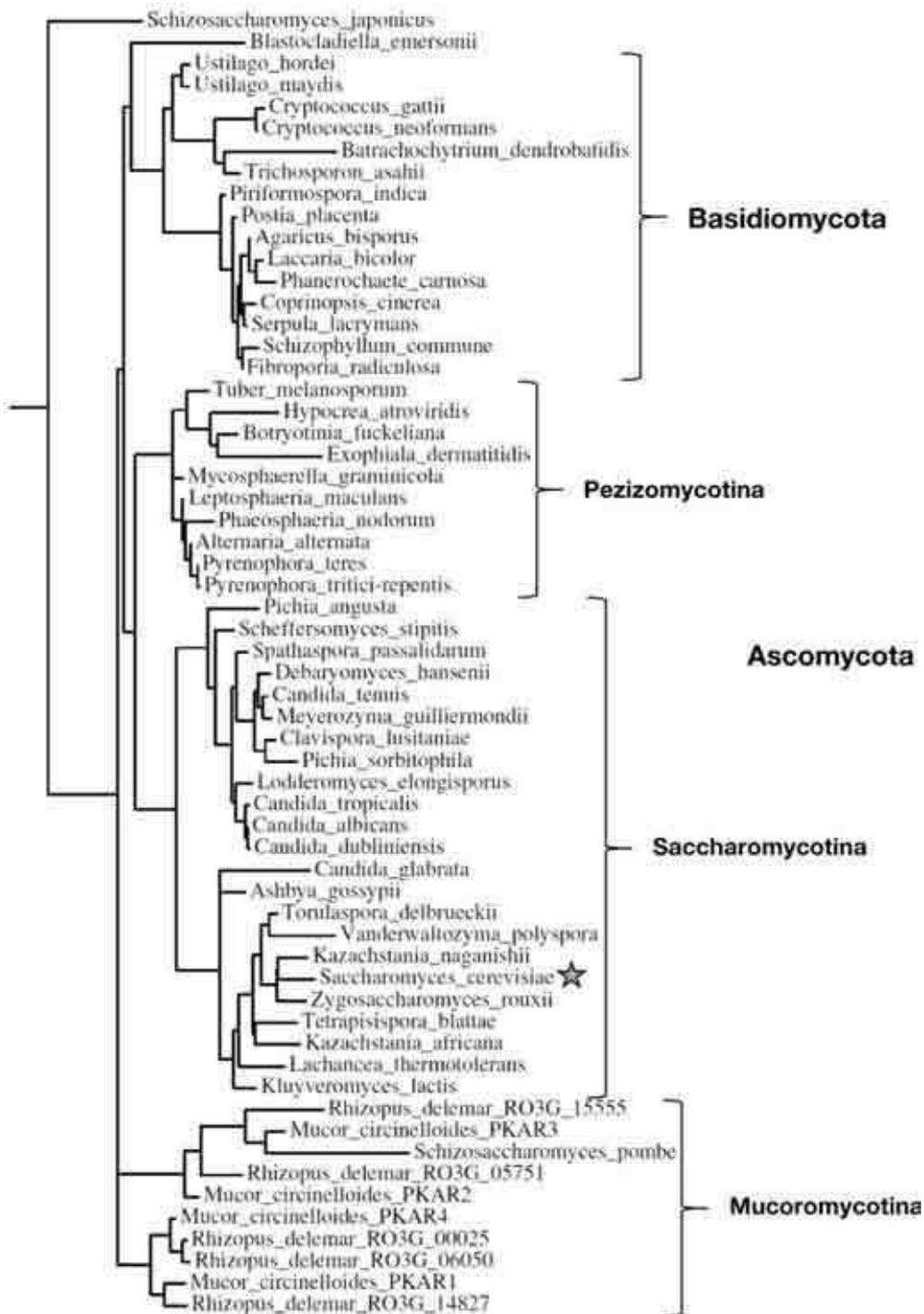
En segundo lugar, se procedió a analizar la presencia de los determinantes que median la tetramerización de Bcy1 en distintos grupos de organismos. Para ello, se amplió la colección de secuencias de hongos con representantes de mamíferos (7 subunidades RI α , 6 subunidades RI β , 7 subunidades RII α , 7 subunidades RII β) e insectos; estos últimos presentan solamente dos isoformas, una más similar a las RI y la otra más similar a las RII de mamíferos. Se colectaron 11 secuencias de tipo RI y 11 secuencias de tipo RII.

Se predijo la estructura secundaria para todas estas secuencias; como era esperado, todas presentan un motivo hélice-giro-hélice en su extremo N-terminal. A continuación, se construyó un alineamiento múltiple correspondiente a la región de estas dos hélices (residuos

8-51 de Bcy1 (1-50), o 7-50 de Bcy1) y en base a él un árbol filogenético a efectos de agrupar las secuencias que presentan características similares (**Figura 4.17**).

El árbol consta de dos grupos principales que divergen de la rama principal: un grupo que contiene a los dominios D/D de mamíferos e insectos, y otro grupo que contiene a los de hongos. Por lo tanto, los dominios D/D de metazoos se encuentran más relacionados entre sí que con los de hongos, aun considerando la diversidad de isoformas. Por otro lado, dentro de cada grupo se aprecia una clara jerarquía: los dominios D/D de las subunidades RI se dividen claramente de los de las subunidades RII, y a su vez, dentro de cada grupo los de insectos conforman un subgrupo, las isoformas α de mamíferos otro, y las isoformas β de mamíferos otro.

Dentro del grupo de las proteínas de hongos, se observa que salvo cuatro organismos (*Schizosaccharomyces pombe*, *Schyzosaccharomyces japonicus*, *Blastocladiella emersonii*, y *Batrachochytrium dendrobatidis*, que son organismos para los que aun existe controversia en cuanto a su clasificación taxonómica), el ordenamiento de sus dominios D/D refleja claramente la clasificación taxonómica: el grupo monofilético *Dikarya*, que consta de los *phyla* Basidiomycota y Ascomycota, diverge de los linajes fúngicos basales (grupo Mucoromycotina). A su vez, dentro del *phylum* Ascomycota, las secuencias de los dominios D/D se agrupan según la división de los organismos en sus correspondientes *supbhyla*: Pezizomycotina y Saccharomycotina.



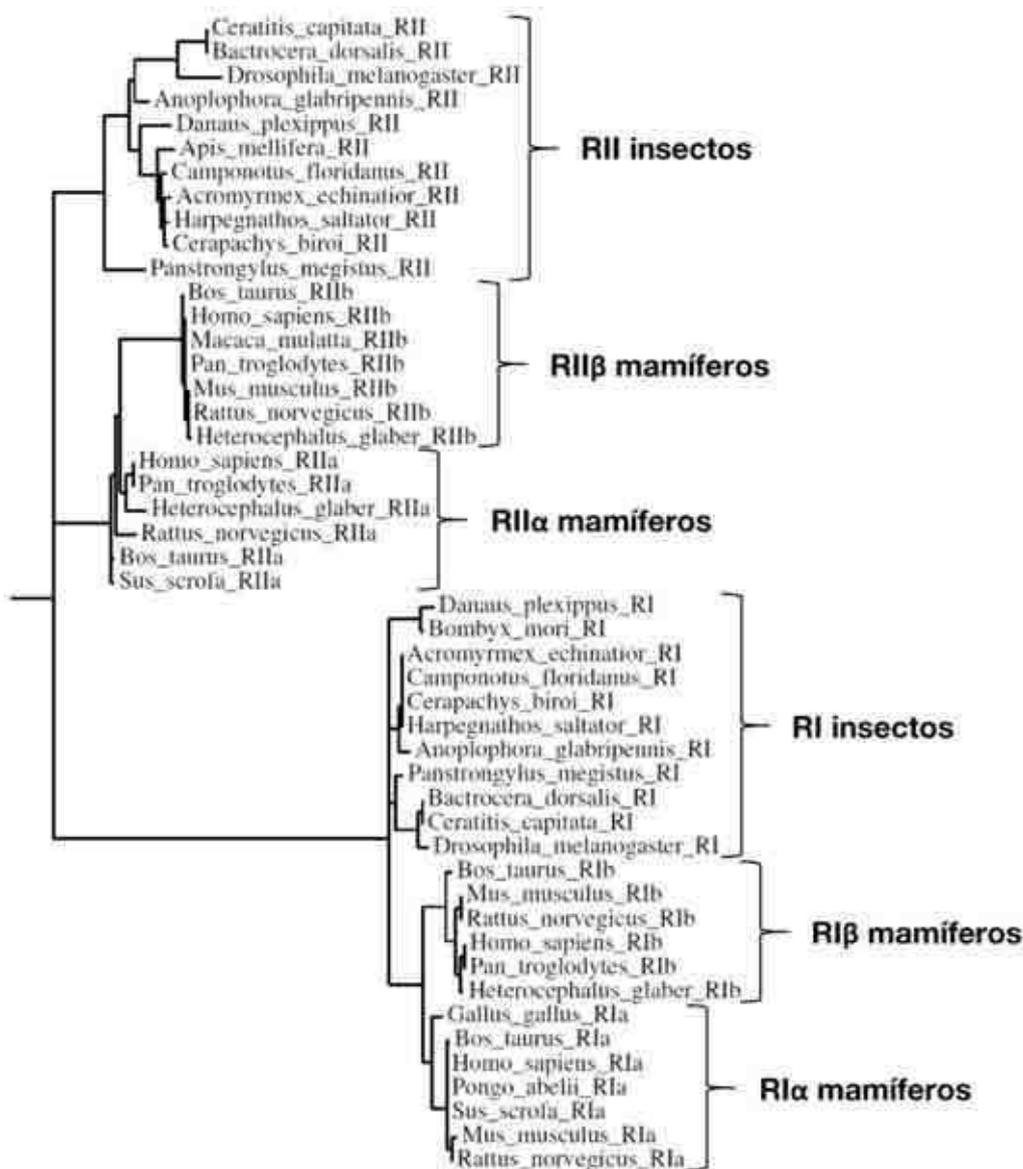


Figura 4.17: Árbol filogenético de los dominios D/D de las subunidades R.

Se muestran por separado las dos ramas divergentes del árbol: los dominios D/D de las subunidades R de mamíferos e insectos (panel inferior), y los de hongos (panel superior). Se indican a un lado los distintos grupos de organismos. La ubicación del dominio D/D de Bcy1 se indica con una estrella.

A continuación, las distintas secuencias se agruparon de acuerdo a los grupos que se evidenciaron en el análisis filogenético (ver **Figura 4.17**), y se construyó un “logo” para cada uno de los grupos (**Figura 4.18**). Este se analizó con el objetivo de identificar los posibles determinantes de la tetramerización, con la evidencia obtenida del análisis de la estructura cristalina.

Figura 4.18: Comparación de dominios D/D de distintos organismos.

Se muestran los logos correspondientes a los distintos grupos identificados en el árbol filogenético de la **Figura 4.17**. La numeración de las secuencias corresponde a la de Bcy1 (1-50). En la parte superior se muestran las dos α -hélices del dominio D/D de Bcy1. Las columnas sombreadas representan posiciones de interés: en naranja se muestran las únicas dos posiciones estrictamente conservadas entre todas las secuencias analizadas: el residuo de prolina que separa las dos hélices, y la fenilalanina que cuando es mutada por alanina impide la dimerización de las subunidades RII α de mamíferos. La columna roja indica la posición en la que se “quiebra” el alineamiento, como consecuencia de la inserción de un aminoácido en las secuencias de hongos. Las columnas azules indican residuos importantes en el establecimiento de la estructura tetramérica de Bcy1: Asp28, Gln31, Leu 41 y Arg45. Las flechas negras señalan la ubicación de estos dos últimos residuos en las subunidades RII. Finalmente, la columna verde representa una característica distintiva de casi todos los dominios D/D de hongos: un residuo de cisteína muy conservado que no está presente en las subunidades R de mamíferos e insectos.

En primer lugar, se observa que ciertas posiciones muy características de los dominios D/D de mamíferos están conservadas en todos los grupos; tal es el caso de la prolina que separa las dos hélices (Pro26 en Bcy1 (1-50)) y de Phe37 en Bcy1 (1-50) que, como ya se señaló, cuando es mutada por alanina en la subunidad RII α de mamíferos, la proteína resultante es incapaz de dimerizar. Debido a su elevada conservación (son los únicos dos residuos que presentan conservación estricta entre todos los grupos), estos dos residuos se tomaron como referencia para el análisis posterior del alineamiento.

Por otro lado, se analizó la conservación de residuos clave para la tetramerización de Bcy1. En primer lugar, se observa que Asp28 se encuentra estrictamente conservado en hongos y en las subunidades RII de mamíferos e insectos, pero no en las subunidades RI. Por otro lado, Gln31 se encuentra conservada en todos los hongos, pero no en mamíferos (de todos modos, este residuo no se encuentra entre los más relevantes para el establecimiento del tetrámero, pero aun así resulta notable su conservación estricta en todos los hongos).

El aspecto más notable del alineamiento consiste en que, hasta Phe37, las secuencias de todos los grupos se encuentran bien alineadas; sin embargo, dos residuos más adelante, las subunidades R de hongos presentan una inserción que tiene como consecuencia un desplazamiento en el registro helicoidal de sus hélices 2 respecto de las proteínas de mamíferos e insectos. En concordancia con esta observación, la posición en la que ocurre esta inserción presenta mucha variabilidad en las secuencias de hongos (columna roja en la **Figura 4.18**). Resulta notable que este desplazamiento repercute en la ubicación en el alineamiento

del residuo más importante para el establecimiento del tetrámero de Bcy1: Arg45. Se observa que en las subunidades RII de mamíferos e insectos, que presentan el residuo Asp28 que es el otro componente clave del tetrámero, esta arginina está presente pero desplazado en una posición respecto del de Bcy1 (flecha negra en la **Figura 4.18**). Lo mismo ocurre con Leu41.

Para analizar las posibles implicancias estructurales de estas observaciones, se analizaron comparativamente las estructuras cristalinas de los dominios D/D de RII α y de Bcy1, mediante un alineamiento estructural (**Figura 4.19**).

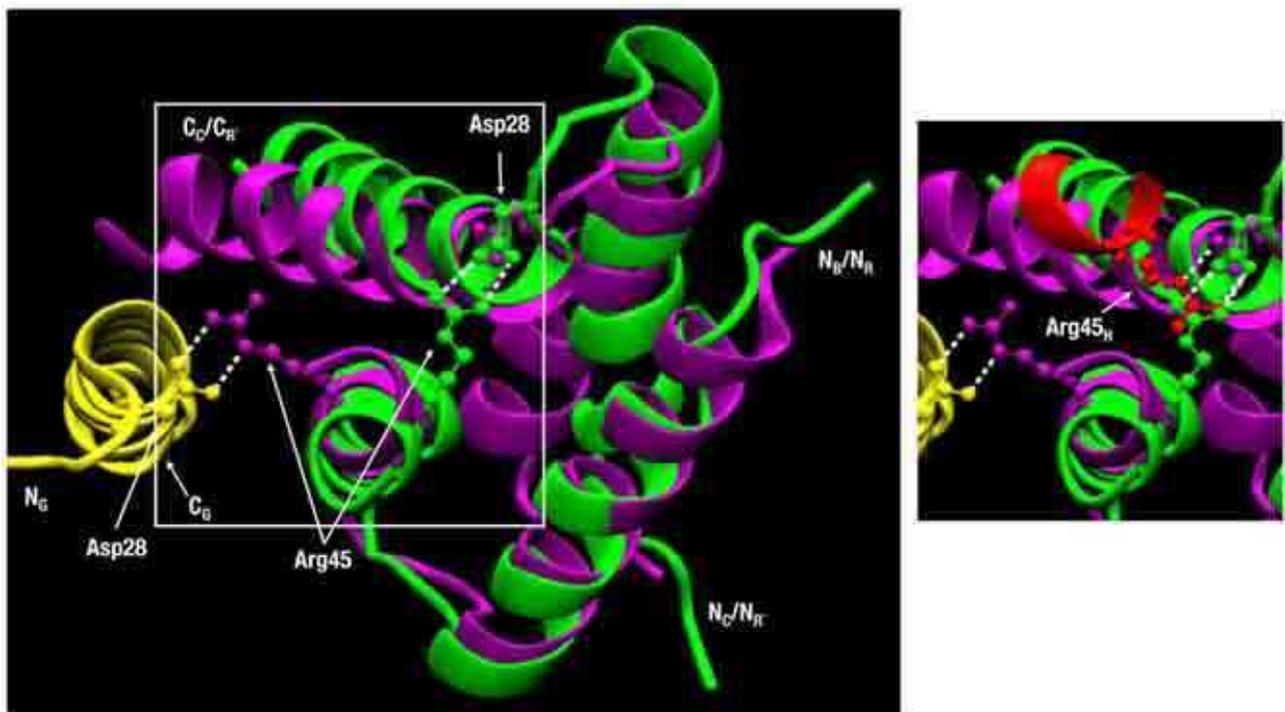


Figura 4.19: Comparación estructural de los dominios D/D de Bcy1 y RII α .

En el panel de la izquierda se muestra una superposición estructural del dímero BC de Bcy1 (1-50) (cintas púrpura), con el dominio D/D de la subunidad RII α de mamíferos (pdb: 2IZY) (cintas verdes). En amarillo se muestra la hélice 2 de la cadena G del tetrámero BCGH de Bcy1 (1-50); la cadena H se omite por claridad. Se indican los extremos N y C terminales de cada una de las cadenas. Se muestran algunos aminoácidos característicos con el mismo código de colores que las cadenas. Se observa que el residuo Asp28 de la cadena R' del dominio D/D de RII α forma un doble puente salino/puente hidrógeno con Arg45 de la cadena opuesta. El residuo Arg45 de la cadena B de Bcy1 (1-50) se encuentra desplazado, formando un doble puente salino/puente hidrógeno con Asp28 de la cadena G. En el panel de la derecha, se muestra como *inset* una representación de los residuos 44-46 de la cadena H de Bcy1 (1-50). Se observa que Arg45 de esa cadena forma un doble puente salino/puente hidrógeno con Asp28 de la cadena C.

La observación de las estructuras cristalinas confirma la predicción realizada en base al análisis de las secuencias. En el dominio D/D de la subunidad RII α de mamíferos, existe un doble puente salino/puente hidrógeno intercatenario entre Arg45, ubicada hacia el final de la hélice 2 de un protómero, y Asp28, ubicado en el *loop* del protómero opuesto. En cambio, para la cadena B de Bcy1 (1-50), pese a que su hélice 2 alinea estructuralmente bastante bien con la de RII α , se observa que Arg45 se encuentra desplazada unos 90° respecto de la posición que adopta en el dominio D/D de RII α (**Figura 4.19**, panel de la izquierda); esto corresponde al desplazamiento esperado por la inserción de un aminoácido. Como consecuencia, pese a que el protómero opuesto del dímero (C) presenta el residuo de Asp28 en la misma posición que RII α , no es posible establecer el mismo contacto.

Más aun, resulta notable que el doble puente salino/puente hidrógeno Arg45-Asp28 sí existe, pero se establece no con el residuo Asp28 de la cadena C, sino con el de la cadena G, que forma parte del otro dímero de los dos que constituyen el tetramero. Esta interacción resulta perfectamente simétrica, dado que el residuo de Asp28 de la cadena C, que como ya se señaló se encuentra bien alineado con el de la proteína de mamíferos, establece el mismo tipo de contacto con Arg45 de la cadena H (**Figura 4.19**, panel de la derecha).

Estas observaciones sugieren fuertemente que el principal elemento responsable de la tetramerización es la inserción de un residuo en la posición que antecede a Arg45. El análisis del logo permite observar que en todos los grupos de hongos esta inserción se encuentra presente, y los residuos Asp28 y Arg45 están estrictamente conservados. Una hipótesis de trabajo interesante es que el estado oligomérico descrito en este trabajo para Bcy1 sea extensible a las subunidades R de todos los hongos.

Como comentario final, otra característica distintiva de las secuencias de hongos es la presencia de un residuo de cisteína muy conservado hacia el comienzo de la hélice 2 (Cys33). Si se observa la estructura cristalina de Bcy1 (1-50) (que presenta una serina en esta posición), se puede apreciar que estos dos residuos se encuentran muy bien posicionados en el espacio como para establecer un puente disulfuro intercatenario (no se muestra), lo que podría representar otra característica distintiva de los dominios D/D de las subunidades R de hongos.

CONCLUSIONES Y PERSPECTIVAS

En este capítulo se describe la primera estructura de alta resolución de un dominio D/D de la subunidad R de la PKA de un organismo no mamífero. En base a este análisis, fue posible determinar cuáles son los elementos que definen la estructura tetramérica de Bcy1 (1-50). Por otro lado, mediante el análisis de las secuencias se identificaron las principales diferencias entre esta estructura y la de los dominios D/D de mamíferos, que determinan que los estados oligoméricos de estas proteínas en solución sean distintos. Una observación importante es que, mientras que en el establecimiento del dímero participan tanto las hélices 1 como las hélices 2 de los dos protómeros, la estructura del tetrámero es consecuencia solamente de la participación de los residuos de la hélice 2 que no forman parte del núcleo de dimerización. La estructura sugiere un modelo para el tetrámero de Bcy1 WT con el dominio D/D ubicado en el centro, y los *linkers* irradiando a partir de los extremos C-terminales de las hélices 2 de las cuatro cadenas; en este modelo, los dominios de unión a cAMP se ubicarían en la periferia, donde tendrían cierta movilidad como consecuencia de la flexibilidad conformacional del *linker*.

Un posible punto de interés que surge de los resultados de este capítulo es obtener mutantes de Bcy1 en su *locus* cromosómico que sean diméricas, y analizar las consecuencias fisiológicas de perturbar el estado de oligomerización. Un buen candidato para lograr esto es el residuo Arg45, que demostró ser de importancia central en el establecimiento del tetrámero.

MATERIALES Y MÉTODOS

Obtención de cristales

Los cristales que se utilizaron para resolver la estructura se obtuvieron utilizando PEG 4000 como agente precipitante. Para ello, se tomaron como núcleos pequeños cristales obtenidos a partir de una solución de PEG 4000 30%, Tris 0,1 M, MgCl₂ 0,2 M, pH 8,5, y se sembraron (*microseeding*) en una gota de idéntica composición pero con una concentración de PEG 4000 24%. Para inocular los cristales se utilizó un bigote de gato. Estos cristales se montaron en el equipo de difracción de rayos X, para lo cual fueron pasados previamente por una solución crioprotectora conteniendo PEG 4000 30%, glicerol 20%, Tris 0,1 M, MgCl₂ 0,2 M, pH 8,5.

Colección de datos

Los datos se colectaron a 108 K utilizando un equipo Cryostream Series 700, Oxford Cryosystems, con radiación de la línea de Cu K_α (1,5418 Å). Se utilizó una fuente de ánodo rotatorio de cobre MicroMax007-HF (Rigaku), espejos Varimax-HF (Rigaku) y un detector del tipo *Image plate* Mar345 (Mar Research). Se colectaron 200 *frames* de 15 minutos de exposición cada uno.

Procesamiento de los datos y resolución de la estructura

Los datos se procesaron utilizando el paquete XDS. Para la resolución de la estructura, se intentó en primera instancia utilizar como sondas de búsqueda la estructura cristalina del dominio D/D de la subunidad RII α (pdb: 2IZX, 2IZY). Estos intentos fracasaron debido a que las dos estructuras resultaron ser, finalmente, bastante diferentes, y a la dificultad intrínseca asociada a la resolución de estructuras con gran cantidad de copias por unidad asimétrica.

Para sortear esta dificultad, se generaron modelos *ab initio* como sondas de búsqueda para reemplazo molecular, utilizando el programa Rosetta. El modelado se llevó a cabo asumiendo un dímero simétrico, y se generaron 30000 modelos en total. Se tomaron los primeros 1000 de acuerdo al *score* proporcionado por el programa, y fueron agrupados en diferentes *clusters*. Los primeros 20 modelos de cada uno de los 20 mejores clusters se

probaron como sondas utilizando el programa Phaser. De esta manera logró resolverse la estructura, que fue posteriormente refinada utilizando el programa Buster.

Herramientas bioinformáticas

Para el análisis bioinformático de las secuencias y de las estructuras se utilizaron distintos programas que funcionan *online*. Para el análisis de la estructura cristalina se utilizaron los servidores PDBePISA [6] y PDBsum [5]. Para generar los alineamientos de secuencias se utilizó el programa Clustal Omega [18]. Para generar los árboles filogenéticos se utilizó el programa Phylogeny.fr [19], proporcionando como *input* los alineamientos generados con Clustal Omega. Los logos para los alineamientos de cada familia se obtuvieron utilizando el programa WebLogo [20]. Las predicciones de estructura secundaria se realizaron utilizando el servidor Phyre [21]. Finalmente, la visualización de las estructuras y la generación de las imágenes se llevó a cabo utilizando el programa VMD [22].

La colección de secuencias generada se muestra en la siguiente tabla:

Organismo	UniProt ID	Dominio D/D	Clasificación
Mucor circinelloides PKAR1	A1YZ32	Sí	Mucormycotina
Mucor circinelloides PKAR2	C1K6Q3	Sí	Mucoromycotina
Mucor circinelloides PKAR3	H6U7H4	Sí	Mucormycotina
Mucor circinelloides PKAR4	H6U7H5	Sí	Mucormycotina
Rhizopus delemar	RO3G 05751*	Sí	Mucormycotina
Rhizopus delemar	RO3G 14827*	Sí	Mucormycotina
Rhizopus delemar	RO3G 00025*	Sí	Mucormycotina
Rhizopus delemar	RO3G 15555*	Sí	Mucormycotina
Rhizopus delemar	RO3G 16579*	Sí	Mucormycotina
Rhizopus delemar	RO3G 06050*	Sí	Mucormycotina
Batrachochytrium dendrobatidis	F4NTV7	Sí	Chytridiomycota
Blastocladiella emersonii	P31320	Sí	Blastocladiomycota
Agaricus bisporus	K9I457	Sí	Basidiomycota
Coprinopsis cinérea	A8NXY5	Sí	Basidiomycota
Cryptococcus gattii	E6QY35	Sí	Basidiomycota
Cryptococcus neoformans	Q9HFV9	Sí	Basidiomycota
Fibroporia radiculosa	J4GU90	Sí	Basidiomycota
Laccaria bicolor	B0CYR4	Sí	Basidiomycota
Phanerochaete carmosa	K5WJ63	Sí	Basidiomycota
Piriformospora indica	G4TW49	Sí	Basidiomycota
Postia placenta	B8PEC1	Sí	Basidiomycota
Schizophyllum commune	Q33DX2	Sí	Basidiomycota
Serpula lacrymans	F8PNR9	Sí	Basidiomycota

Capítulo 4 Estructura cristalina del dominio D/D de Bcy1

<i>Trichosporon asahii</i>	J5QEQ0	Sí	Basidiomycota
<i>Ustilago hordei</i>	I2FNH9	Sí	Basidiomycota
<i>Ustilago maydis</i>	P49605	Sí	Basidiomycota
<i>Schizosaccharomyces japonicus</i>	B6K056	Sí	Taphrinomycotina
<i>Schizosaccharomyces pombe</i>	P36600	Sí	Taphrinomycotina
<i>Ashbya gossypii</i>	Q75AM2	Sí	Saccharomycotina
<i>Candida albicans</i>	Q9HEW1	Sí	Saccharomycotina
<i>Candida dubliniensis</i>	B9WAG0	Sí	Saccharomycotina
<i>Candida glabrata</i>	Q6FQL6	Sí	Saccharomycotina
<i>Candida tenuis</i>	G3AYE5	Sí	Saccharomycotina
<i>Candida tropicalis</i>	C5M619	Sí	Saccharomycotina
<i>Clavispora lusitanae</i>	C4YC33	Sí	Saccharomycotina
<i>Debaryomyces hansenii</i>	Q6BZG7	Sí	Saccharomycotina
<i>Kazachstania africana</i>	H2AMU5	Sí	Saccharomycotina
<i>Kazachstania naganishii</i>	J7RPN6	Sí	Saccharomycotina
<i>Kluyveromyces lactis</i>	Q6CPK7	Sí	Saccharomycotina
<i>Lachancea thermotolerans</i>	C5DBK8	Sí	Saccharomycotina
<i>Lodderomyces elongisporus</i>	A5DUX9	Sí	Saccharomycotina
<i>Meyerozyma guilliermondii</i>	A5DNT5	Sí	Saccharomycotina
<i>Pichia angusta</i>	E7R4G7	Sí	Saccharomycotina
<i>Pichia pastoris</i>	F2QW39	Sí	Saccharomycotina
<i>Pichia sorbitophila</i>	G8YB63	Sí	Saccharomycotina
<i>Saccharomyces cerevisiae</i>	P07278	Sí	Saccharomycotina
<i>Scheffersomyces stipitis</i>	A3LN26	Sí	Saccharomycotina
<i>Spathaspora passalidarum</i>	G3AK93	Sí	Saccharomycotina
<i>Tetrapisispora blattae</i>	I2H3M4	Sí	Saccharomycotina
<i>Torulaspora delbrueckii</i>	G8ZZQ6	Sí	Saccharomycotina
<i>Vanderwaltozyma polyspora</i>	A7TNN2	Sí	Saccharomycotina
<i>Yarrowia lipolytica</i>	F2XG10	Sí	Saccharomycotina
<i>Zygosaccharomyces rouxii</i>	C5DWM8	Sí	Saccharomycotina
<i>Ajellomyces capsulata</i>	C0NGJ1	No	Pezizomycotina
<i>Ajellomyces dermatitidis</i>	C5GF89	No	Pezizomycotina
<i>Alternaria alternata</i>	J7G926	Sí	Pezizomycotina
<i>Arthroderma gypseum</i>	E5R0T4	No	Pezizomycotina
<i>Arthroderma otae</i>	C5FV92	No	Pezizomycotina
<i>Aspergillus clavatus</i>	A1CKA1	No	Pezizomycotina
<i>Aspergillus kawachii</i>	G7XCM5	No	Pezizomycotina
<i>Aspergillus niger</i>	Q9C196	No	Pezizomycotina
<i>Aspergillus oryzae</i>	Q5H7P2	No	Pezizomycotina
<i>Aspergillus terreus</i>	Q0CP60	No	Pezizomycotina
<i>Beauveria bassiana</i>	J4KQ85	No	Pezizomycotina
<i>Blumeria graminis</i>	Q9HEP7	No	Pezizomycotina
<i>Botryotinia fuckeliana</i>	C0H5W5	Sí	Pezizomycotina
<i>Chaetomium thermophilum</i>	G0SDM9	No	Pezizomycotina
<i>Coccidioides immitis</i>	I9XN00	No	Pezizomycotina
<i>Colletotrichum gloeosporioides</i>	Q0PXU3	No	Pezizomycotina

Capítulo 4 Estructura cristalina del dominio D/D de Bcy1

<i>Colletotrichum orbiculare</i>	Q9C1C2	No	Pezizomycotina
<i>Emericella nidulans</i>	O59922	No	Pezizomycotina
<i>Exophiala dermatitidis</i>	H6C4K9	Sí	Pezizomycotina
<i>Gibberella zeae</i>	I1RZQ6	No	Pezizomycotina
<i>Grosmannia clavigera</i>	F0XSY0	No	Pezizomycotina
<i>Hypocrea atroviridis</i>	Q86ZN7	Sí	Pezizomycotina
<i>Leptosphaeria maculans</i>	E5ADF7	Sí	Pezizomycotina
<i>Magnaporthe oryzae</i>	O14448	No	Pezizomycotina
<i>Metarhizium acridum</i>	E9DXG2	No	Pezizomycotina
<i>Metarhizium anisopliae</i>	E9EWM7	No	Pezizomycotina
<i>Mycosphaerella graminicola</i>	Q1KTF0	Sí	Pezizomycotina
<i>Nectria haematococca</i>	C7YRC1	No	Pezizomycotina
<i>Neosartorya fischeri</i>	A1D768	No	Pezizomycotina
<i>Neosartorya fumigata</i>	Q96UX3	No	Pezizomycotina
<i>Neurospora crassa</i>	Q01386	No	Pezizomycotina
<i>Neurospora tetrasperma</i>	F8MF61	No	Pezizomycotina
<i>Paracoccidioides brasiliensis</i>	C0S191	No	Pezizomycotina
<i>Penicillium chrysogenum</i>	B6H413	No	Pezizomycotina
<i>Penicillium marneffeii</i>	B6Q719	No	Pezizomycotina
<i>Phaeosphaeria nodorum</i>	Q0V3C1	Sí	Pezizomycotina
<i>Podospora anserina</i>	B2B4N9	No	Pezizomycotina
<i>Pyrenophora teres</i>	E3RJD4	Sí	Pezizomycotina
<i>Pyrenophora tritici-repentis</i>	B2VYL4	Sí	Pezizomycotina
<i>Sclerotinia sclerotiorum</i>	A7EYX0	No	Pezizomycotina
<i>Talaromyces stipitatus</i>	B8LUQ6	No	Pezizomycotina
<i>Thielavia heterothallica</i>	G2Q7A9	No	Pezizomycotina
<i>Thielavia terrestris</i>	G2QQL7	No	Pezizomycotina
<i>Trichophyton rubrum</i>	F2SDV1	No	Pezizomycotina
<i>Trichophyton tonsurans</i>	F2S0I8	No	Pezizomycotina
<i>Tuber melanosporum</i>	D5GGF7	Sí	Pezizomycotina
<i>Bos taurus</i>	P00514	R1a	Mamífero
<i>Bos taurus</i>	Q17QF5	R1b	Mamífero
<i>Bos taurus</i>	P00515	R11a	Mamífero
<i>Bos taurus</i>	P31322	R11b	Mamífero
<i>Rattus norvegicus</i>	P09456	R1a	Mamífero
<i>Rattus norvegicus</i>	P81377	R1b	Mamífero
<i>Rattus norvegicus</i>	P12368	R11a	Mamífero
<i>Rattus norvegicus</i>	P12369	R11b	Mamífero
<i>Homo sapiens</i>	P10644	R1a	Mamífero
<i>Homo sapiens</i>	P31321	R1b	Mamífero
<i>Homo sapiens</i>	P13861	R11a	Mamífero
<i>Homo sapiens</i>	P31323	R11b	Mamífero
<i>Mus musculus</i>	Q9DBC7	R1a	Mamífero
<i>Mus musculus</i>	P12849	R1b	Mamífero
<i>Mus musculus</i>	P12367	R11a	Mamífero
<i>Mus musculus</i>	P31324	R11b	Mamífero

Capítulo 4 Estructura cristalina del dominio D/D de Bcy1

<i>Gallus gallus</i>	Q5ZM91	RIa	Mamífero
<i>Pongo abelii</i>	Q5REL1	RIa	Mamífero
<i>Sus scrofa</i>	P07802	RIa	Mamífero
<i>Heterocephalus glaber</i>	G5AZ14	RIb	Mamífero
<i>Pan troglodytes</i>	K7B4J4	RIb	Mamífero
<i>Heterocephalus glaber</i>	G5C9G1	RIIa	Mamífero
<i>Pan troglodytes</i>	G2HJH1	RIIa	Mamífero
<i>Sus scrofa</i>	C1PIG4	RIIa	Mamífero
<i>Macaca mulatta</i>	I0FTB7	RIIb	Mamífero
<i>Heterocephalus glaber</i>	G5C2D5	RIIb	Mamífero
<i>Pan troglodytes</i>	K7C6K9	RIIb	Mamífero
<i>Acromyrmex echinatio</i>	F4WV86	RI	Insecto
<i>Anoplophora glabripennis</i>	V5GNL5	RI	Insecto
<i>Bactrocera dorsalis</i>	A0A034VHL6	RI	Insecto
<i>Bombyx mori</i>	A5HMG0	RI	Insecto
<i>Camponotus floridanus</i>	E2AEL2	RI	Insecto
<i>Cerapachys biroi</i>	A0A026WQN0	RI	Insecto
<i>Ceratitis capitata</i>	W8C2E9	RI	Insecto
<i>Danaus plexippus</i>	G6CTL4	RI	Insecto
<i>Drosophila melanogaster</i>	P16905	RI	Insecto
<i>Harpegnathos saltator</i>	E2C7R5	RI	Insecto
<i>Panstrongylus megistus</i>	A0A069DTM3	RI	Insecto
<i>Acromyrmex echinatio</i>	F4WZN8	RII	Insecto
<i>Anoplophora glabripennis</i>	V5GYF5	RII	Insecto
<i>Bactrocera dorsalis</i>	A0A034V9R4	RII	Insecto
<i>Camponotus floridanus</i>	E2ACB7	RII	Insecto
<i>Cerapachys biroi</i>	A0A026WYU0	RII	Insecto
<i>Ceratitis capitata</i>	W8B5W7	RII	Insecto
<i>Danaus plexippus</i>	G6DRS2	RII	Insecto
<i>Drosophila melanogaster</i>	P81900	RII	Insecto
<i>Harpegnathos saltator</i>	E2BQT5	RII	Insecto
<i>Panstrongylus megistus</i>	A0A069DTE6	RII	Insecto
<i>Apis mellifera carnica</i>	Q6ZXJ1	RII	Insecto

REFERENCIAS

1. Rhodes G: **Crystallography Made Crystal Clear (Third Edition)**. *Academic Press* 2006.
2. Drenth J: **Principles of Protein X-Ray Crystallography, Third Edition**. *Springer* 2007.
3. Wlodawer A, Minor W, Dauter Z, Jaskolski M: **Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures**. *The FEBS journal* 2008, **275**(1):1-21.
4. Gold MG, Lygren B, Dokurno P, Hoshi N, McConnachie G, Tasken K, Carlson CR, Scott JD, Barford D: **Molecular basis of AKAP specificity for PKA regulatory subunits**. *Mol Cell* 2006, **24**(3):383-395.
5. de Beer TA, Berka K, Thornton JM, Laskowski RA: **PDBsum additions**. *Nucleic Acids Res* 2014, **42**(Database issue):D292-296.
6. Krissinel E, Henrick K: **Inference of macromolecular assemblies from crystalline state**. *J Mol Biol* 2007, **372**(3):774-797.
7. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller**. *Current protocols in bioinformatics / editorial board, Andreas D Baxevanis [et al]* 2006, **Chapter 5**:Unit 5.6.
8. Schneidman-Duhovny D, Hammel M, Sali A: **FoXS: a web server for rapid computation and fitting of SAXS profiles**. *Nucleic Acids Res* 2010, **38**(Web Server issue):W540-544.
9. Garcia De La Torre J, Huertas ML, Carrasco B: **Calculation of hydrodynamic properties of globular proteins from their atomic-level structure**. *Biophysical journal* 2000, **78**(2):719-730.
10. Jennings ML, Cui J: **Chloride homeostasis in *Saccharomyces cerevisiae*: high affinity influx, V-ATPase-dependent sequestration, and identification of a candidate Cl⁻ sensor**. *The Journal of general physiology* 2008, **131**(4):379-391.
11. Kortemme T, Baker D: **A simple physical model for binding energy hot spots in protein-protein complexes**. *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**(22):14116-14121.
12. Banky P, Roy M, Newlon MG, Morikis D, Haste NM, Taylor SS, Jennings PA: **Related protein-protein interaction modules present drastically different surface topographies despite a conserved helical platform**. *J Mol Biol* 2003, **330**(5):1117-1129.
13. Canaves JM, Taylor SS: **Classification and phylogenetic analysis of the cAMP-dependent protein kinase regulatory subunit family**. *J Mol Evol* 2002, **54**(1):17-29.
14. Gunzburg J, Veron M: **A cAMP-dependent protein kinase is present in differentiating *Dictyostelium discoideum* cells**. *The EMBO journal* 1982, **1**(9):1063-1068.
15. Carlson GL, Nelson DL: **The 44-kDa regulatory subunit of the *Paramecium* cAMP-dependent protein kinase lacks a dimerization domain and may have a unique autophosphorylation site sequence**. *The Journal of eukaryotic microbiology* 1996, **43**(4):347-356.
16. Wang H, Xu Z, Gao L, Hao B: **A fungal phylogeny based on 82 complete genomes using the composition vector method**. *BMC evolutionary biology* 2009, **9**:195.
17. Ocampo J, Fernandez Nunez L, Silva F, Pereyra E, Moreno S, Garre V, Rossi S: **A subunit of protein kinase a regulates growth and differentiation in the fungus *Mucor circinelloides***. *Eukaryotic cell* 2009, **8**(7):933-944.
18. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R: **Analysis Tool Web Services from the EMBL-EBI**. *Nucleic Acids Res* 2013, **41**(Web Server issue):W597-600.
19. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M *et al*: **Phylogeny.fr: robust phylogenetic analysis for the non-specialist**. *Nucleic Acids Res* 2008, **36**(Web Server issue):W465-469.
20. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: a sequence logo generator**. *Genome research* 2004, **14**(6):1188-1190.
21. Kelley LA, Sternberg MJ: **Protein structure prediction on the Web: a case study using the Phyre server**. *Nat Protoc* 2009, **4**(3):363-371.
22. Humphrey W, Dalke A, Schulten K: **VMD: visual molecular dynamics**. *Journal of molecular graphics* 1996, **14**(1):33-38, 27-38.

CAPÍTULO 5

INTERACCIÓN DEL DOMINIO D/D DE BCY1 CON EL PÉPTIDO IRA2

“ Andábamos sin buscarnos
pero sabiendo que
andábamos para
encontrarnos.”

Julio Cortázar, “ Rayuela ”

ANTECEDENTES

1) En mamíferos, los dominios D/D son responsables de la interacción con las proteínas de anclaje de la familia de las AKAPs, que se establece entre la superficie expuesta al solvente formada por las hélices 1 de los dos protómeros, de carácter hidrofóbico, y la cara no polar de un α -hélice anfipática de unos 20 aminoácidos que está presente en todas las AKAPs (dominio AKB). Los residuos hidrofóbicos de los dos componentes del complejo son cruciales para la interacción.

2) Hasta la fecha, no se han identificado en *S. cerevisiae* (ni tampoco en hongos en general) proteínas que cumplan con las características de las AKAPs. Sin embargo, se han descrito proteínas que son responsables de la localización subcelular de Bcy1 bajo ciertas condiciones [1], así como también un grupo de proteínas que interactúan con la región 1-85 de Bcy1 a través de péptidos que parecen presentar características similares a las de las AKAPs [2]. Una de estas proteínas, Ira2, es una candidata interesante a ser clasificada como posible AKAP, debido a que: 1) colocaliza con Bcy1 bajo ciertas condiciones [2], y 2) participa de la misma vía de señalización que la PKA. Sin embargo, en contraste con las interacciones clásicas D/D-AKB, el complejo Ira2-Bcy1 parece involucrar principalmente residuos cargados positivamente del péptido, mientras que la mutación por alanina de sus aminoácidos hidrofóbicos no perjudica la interacción.

OBJETIVOS Y ABORDAJES UTILIZADOS

Para completar la caracterización del dominio D/D de Bcy1 desde un punto de vista más funcional, se decidió evaluar su capacidad de interactuar con proteínas utilizando como modelo el péptido Ira2. Para ello, se plantearon los siguientes objetivos para este capítulo:

1) Evaluar la capacidad del péptido Ira2 de plegarse en forma de una α -hélice anfipática, y de esta manera decidir si se asemeja estructuralmente a los dominios AKB de las AKAPs. Para ello se utilizó la técnica de CD.

2) Estudiar la capacidad del péptido de interactuar en solución con Bcy1 WT mediante técnicas de espectroscopia de fluorescencia, y determinar la constante de afinidad para

Capítulo 5 Interacción del dominio D/D de Bcy1 con el péptido Ira2

compararla con las reportadas para las interacciones D/D-AKB de mamíferos. Para ello se utilizó una variante del péptido con un residuo de cisteína adicional que permite su marcación con un fluoróforo. Cabe destacar que en el trabajo ya mencionado, desarrollado en nuestro laboratorio, la interacción de Bcy1 con Ira2 completa se demostró mediante experimentos de *pull down*, mientras que los ensayos que involucraron al péptido se realizaron en fase heterogénea, utilizando *arrays*.

3) Determinar la capacidad de Bcy1 (1-50) de interactuar con el péptido Ira2, mediante técnicas de fluorescencia y de CD, y obtener la constante de afinidad de la interacción para compararla con la de Bcy1 WT.

4) En caso de verificarse esa interacción, hipotetizar acerca de cuáles podrían ser las superficies de Bcy1 (1-50) que se encuentran involucradas en la misma, para comparar con el sistema de mamíferos. Para ello, se cuenta con la información disponible de los *arrays* de péptidos, que permitieron determinar cuáles son los residuos de Ira2 que son clave para la interacción, y con la estructura cristalina de Bcy1 (1-50).

Hasta el momento, los resultados correspondientes a estos objetivos no resultan del todo satisfactorios, por lo que esta línea de trabajo se encuentra aun en desarrollo. En la siguiente sección integrada de resultados, métodos y discusión, se presenta una síntesis de los experimentos preliminares correspondientes a los objetivos planteados anteriormente, y se especulan algunas posibles explicaciones del comportamiento observado.

RESULTADOS, MÉTODOS Y DISCUSIÓN

Para los experimentos de este capítulo se decidió trabajar con un péptido de Ira2 de 20 aminoácidos (CTTTEVANNIIQKILAKIRSF L) que corresponde a los aminoácidos 2511-2530 de la proteína. Esta secuencia es suficiente para garantizar la interacción con Bcy1 WT [2]. El péptido se sintetizó con un residuo de cisteína adicional en su extremo N-terminal para poder derivatizarlo con un fluoróforo; se eligió marcar en este extremo debido a que los tres residuos que son clave para la interacción con Bcy1 (Lys2521, Lys2525 y Arg 2527) se encuentran en el extremo C-terminal del péptido. El péptido fue sintetizado por la compañía GenBiotech (<http://www.genbiotech.com.ar/>).

Estructura secundaria del péptido Ira2 en solución

Para estudiar la estructura del péptido en solución, en primer lugar se trató con iodoacetamida (que es un reactivo alquilante de cisteínas), para evitar su dimerización mediante la formación de un puente disulfuro intercatenario. Posteriormente, el péptido derivatizado fue purificado por cromatografía en fase reversa utilizando una fase fija C18 en un equipo HPLC.

Con el objetivo de determinar si el péptido puede plegarse en forma de una α -hélice anfipática en solución, se registraron espectros de CD. Dado que los péptidos tan cortos rara vez adoptan su estructura nativa en solución, dado que suelen requerir de la interacción con otras regiones de la proteína o con proteínas interactoras, se realizó una curva de agregado de trifluoroetanol (TFE), que es un reactivo utilizado comúnmente para estabilizar estructuras helicoidales de péptidos que presentan cierta tendencia intrínseca a adoptar esa estructura. En la **Figura 5.1 (A)** se muestran los resultados obtenidos. Se observa que aun en ausencia de TFE el péptido tiene cierta tendencia a plegarse en forma de α -hélice (curva naranja claro), lo cual resulta llamativo para un péptido de tan solo 20 aminoácidos. Esta hélice termina de consolidarse por agregado de cantidades crecientes de TFE (curva rojo oscuro). Por otro lado, en la **Figura 5.1 (B)** se muestra una representación de “*Helical Wheel*” del péptido. Se observa claramente que la hélice es anfipática, con un momento hidrofóbico muy pronunciado en la dirección de la cara no polar, constituida principalmente por residuos alifáticos; por el contrario, la otra cara está compuesta por residuos polares.

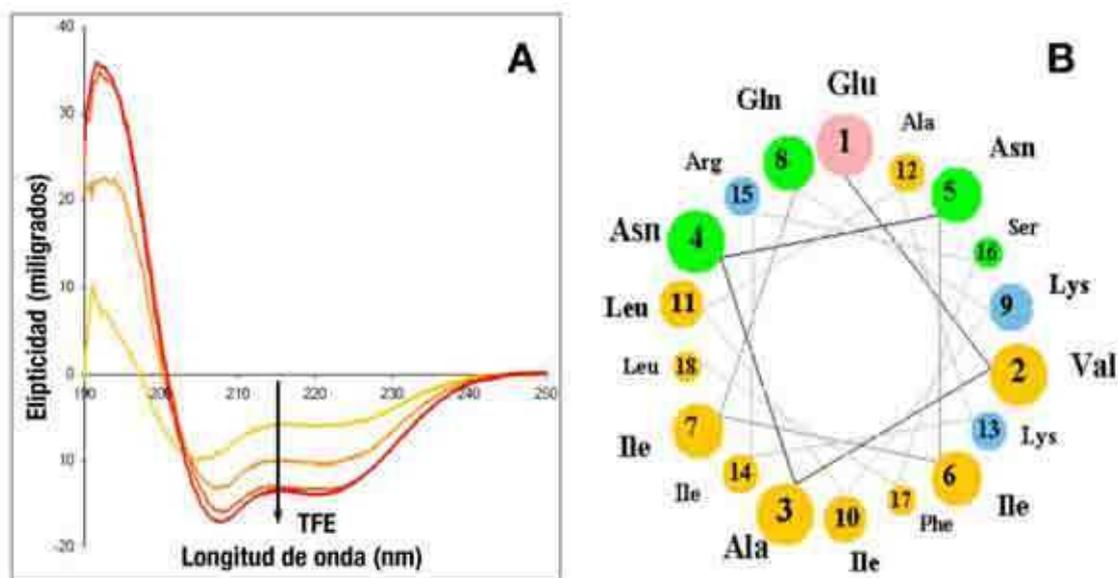


Figura 5.1: Estructura secundaria del péptido Ira2.

(A) Espectros de CD en el UV lejano de una solución 4 μ M de péptido Ira2 en buffer Tris-HCl 10 mM pH 6,8; 10 mM NaCl; 1 mM DTT (línea naranja claro), y de soluciones de igual concentración de péptido y cantidades crecientes de TFE (10%-30%; naranja-rojo). (B) Representación de “*Helical Wheel*” de los residuos 3-20 del péptido. El código de colores es: naranja (hidrofóbico), verde (polar, no cargado), azul (básico), rosa (ácido); la flecha indica la dirección del momento hidrofóbico.

Los resultados de esta sección demuestran que el péptido Ira2 tiene la capacidad de plegarse en forma de una α -hélice anfipática en solución, lo que es una característica típica de los dominios AKB de todas las AKAPs de mamíferos. Por otro lado, los residuos de la cara hidrofóbica son, a excepción de una Phe, todos alifáticos; esto también se encuentra en concordancia con las características conocidas para las AKAPs de mamíferos, dado que, como ya se discutió en la Introducción, la interacción con los dominios D/D está mediada por este tipo de residuos. Finalmente, se observa que los tres residuos básicos que son determinantes para la interacción del péptido Ira2 con Bcy1 WT se ubican a ambos costados de la cara hidrofóbica.

Interacción entre el péptido Ira2 y las proteínas Bcy1 WT y Bcy1 (1-50)

Para llevar a cabo este análisis, en primer lugar se procedió a la marcación del péptido con un fluoróforo adecuado para estudiar la interacción por técnicas de fluorescencia. Se

utilizó el reactivo IAEDANS (**Figura 5.2**) que fue elegido por: 1) reaccionar selectivamente con residuos de cisteína, 2) ser sensible al entorno, lo hace muy apropiado para espectroscopía de fluorescencia, y 3) presentar un tiempo de vida medio apropiado para mediciones de anisotropía.

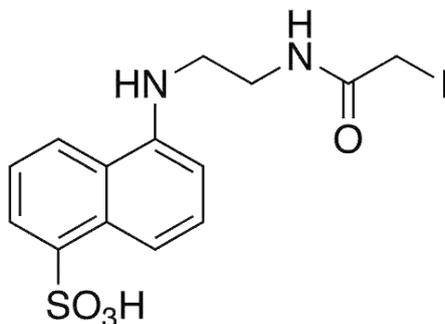


Figura 5.2: Estructura química del IAEDANS.

El compuesto reacciona mediante un ataque nucleofílico del azufre desprotonado del residuo de cisteína sobre el carbono en posición α al carbonilo del reactivo, que resulta un buen electrófilo por estar acompañado del yoduro terminal.

Previo a la reacción, el péptido purificado por HPLC fue tratado con el agente reductor TCEP para evitar el bloqueo de los grupos de cisteína por dimerización; este reactivo tiene la característica de no interferir con la marcación, a diferencia de otros agentes reductores más convencionales como el DTT o el β -mercaptoetanol. A continuación, se llevó a cabo la reacción de marcación en buffer PBS a temperatura ambiente, en la oscuridad, utilizando un exceso molar de 5 veces de reactivo respecto del péptido. Posteriormente, el péptido marcado fue purificado del exceso de reactivo y del péptido sin marcar por HPLC, e identificado por espectrometría de masa y por su espectro de fluorescencia. La **Figura 5.3** muestra estos resultados.

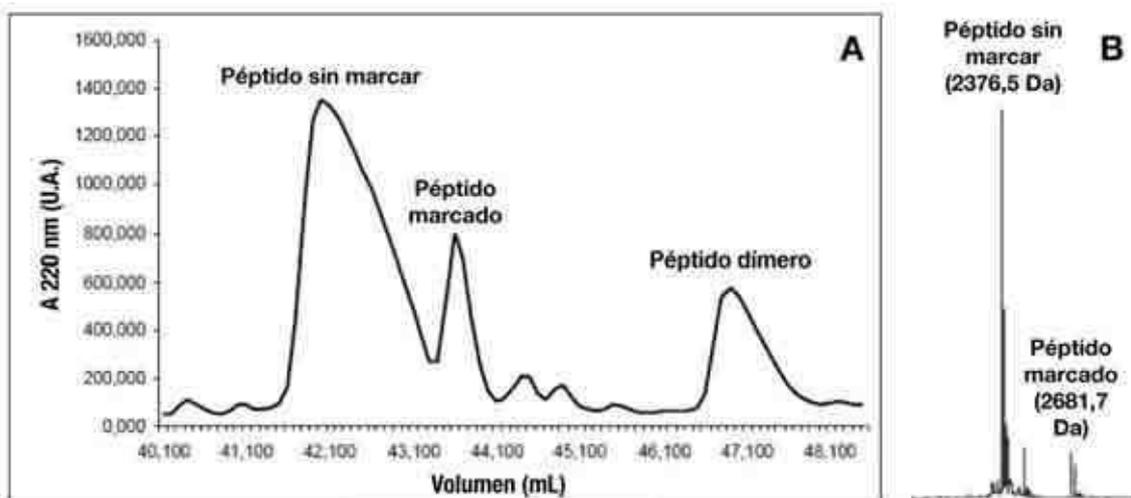


Figura 5.3: Marcación y purificación del péptido Ira2.

(A) Cromatograma registrado para la purificación del péptido Ira2 marcado con IAEDANS; se utilizó una columna de fase reversa C18 y un gradiente lineal de 100% solvente A a 100% solvente B en 70 minutos (solvente A: agua, ácido trifluoroacético 0,05%; solvente B: acetonitrilo, ácido trifluoroacético 0,05%) con un flujo de 1 mL/min. Se indican los picos correspondientes al exceso de péptido sin marcar, al péptido marcado, y a péptido sin marcar que se encontró dimérico. El exceso de reactivo eluye a volúmenes mucho menores. (B) Espectro de masa del pico correspondiente al péptido marcado en el cromatograma del panel (A). La MM esperada para el péptido sin marcar es 2375,9 Da. La MM del IAEDANS sin el yoduro (que se pierde en la reacción de marcación) es 307,3 Da. La diferencia de masa entre los dos picos es de 305 Da.

A continuación, se analizó el comportamiento del espectro de fluorescencia del péptido marcado como consecuencia del agregado de Bcy1 WT purificada. Se observó que frente a agregados sucesivos de proteína ocurre *quenching* de la fluorescencia del péptido, lo que puede ser indicativo de interacción (Figura 5.4 (A)). Por lo tanto, se intentó obtener una curva de unión por titulación de una solución de péptido marcado de concentración conocida con cantidades crecientes de Bcy1 WT. Si bien el experimento se repitió en numerosas oportunidades bajo distintas condiciones experimentales, nunca se pudo llegar a saturación porque para los últimos puntos de la curva se observó agregación de la muestra. No se pudo determinar la causa de este fenómeno.

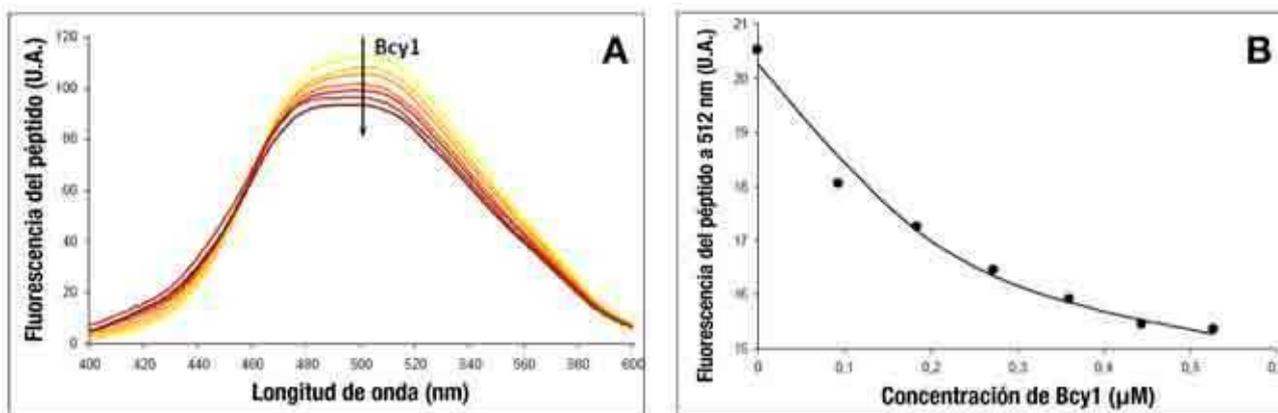


Figura 5.4: Unión del péptido Ira2 marcado a Bcy1 WT.

(A) Espectros de fluorescencia del péptido Ira2 marcado registrados en ausencia de Bcy1 (línea amarilla) y luego de agregados sucesivos de Bcy1. Los espectros se registraron a 25° C en buffer PBS con una longitud de onda de excitación de 370 nm. (B) Curva de unión de Bcy1 al péptido marcado (los datos no corresponden a los espectros del panel (A)). Se representó la señal de fluorescencia del péptido marcado en el máximo del espectro de emisión en función del agregado de Bcy1. La concentración del péptido fue de 0,45 μM (determinada por absorbancia de acuerdo al espectro de absorción del IAEDANS). Los puntos corresponden a los datos experimentales, mientras que la línea sólida representa el ajuste de acuerdo a un modelo de unión 1:1.

Pese a las dificultades experimentales mencionadas, se observa que la curva de unión presenta el aspecto esperado para este tipo de experimentos hasta el punto en que la muestra comenzó a agregar. Aunque no se logró llegar a saturación, se ajustaron los datos según un modelo de unión 1:1, y se estimó a partir del ajuste una K_d para el complejo, que resultó ser del orden de las decenas de nM. Aunque este valor no puede ser considerado con seguridad como la K_d , es probable que la estimación aproximada del orden de magnitud de la afinidad no esté errada. Este valor es del mismo orden que los reportados para los complejos D/D-AKB de mamíferos.

Para verificar que el cambio espectral obedece efectivamente a la interacción entre Bcy1 WT y el péptido, una vez completada la titulación se intentó desplazar el péptido marcado unido mediante agregado de péptido sin marcar proveniente del *stock*. De esta manera, se descubrió que éste viene impurificado con algún componente que presenta fluorescencia en el rango espectral utilizado, y por lo tanto no se pudo verificar este punto. Para continuar con

estos estudios será necesario purificar el péptido sin marcar de estas impurezas para validar los resultados.

A continuación, se intentó realizar el mismo tipo de estudio utilizando Bcy1 (1-50). En contraste con lo observado para Bcy1 WT, se observó un incremento de la fluorescencia del péptido como consecuencia del agregado de la proteína (no se muestra). Luego de realizar distintos tipos de controles, se concluyó que el efecto se debe a que el péptido, al encontrarse en bajas concentraciones y contener residuos cargados, se adhiere a las paredes de la cubeta, y que el agregado de Bcy1 (1-50) lo desplaza a la solución, incrementando la fluorescencia. Se intentaron registrar espectros en presencia de BSA o de *tween* como *carriers* para intentar minimizar estos efectos, pero no se obtuvieron resultados que permitan confirmar o descartar la interacción. También se intentaron mediciones de anisotropía de fluorescencia, pero no se observó evidencia de interacción.

Considerando estas dificultades experimentales, y aprovechando que se conoce el comportamiento de Bcy1 (1-50) frente a la desnaturalización térmica (ver Capítulo 3), se decidió evaluar la formación de un complejo con el péptido sin marcar mediante CD. Este abordaje suele ser bastante utilizado, y se rige por la hipótesis de que la formación de un complejo con el péptido debería proporcionar alguna estabilidad adicional a Bcy1 (1-50) frente a la desnaturalización térmica, que se vería reflejada en un incremento de la T_m . Se registraron entonces por CD curvas de desnaturalización térmica de Bcy1 (1-50) en presencia y ausencia de un exceso de péptido Ira2 sin marcar. La **Figura 5.5** muestra los resultados obtenidos.

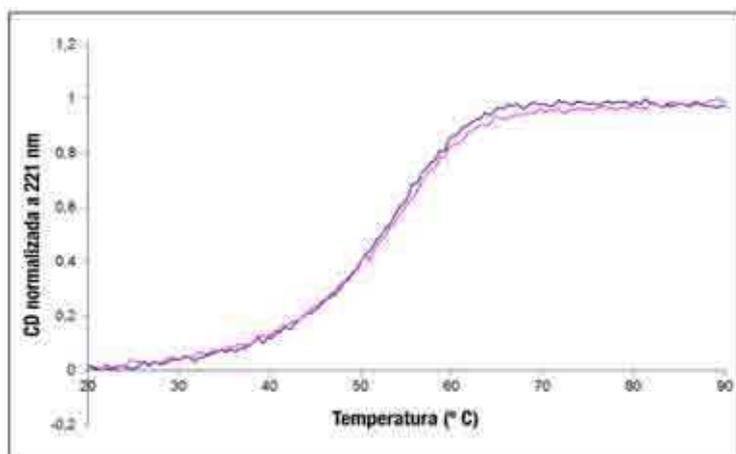


Figura 5.5: Curvas de desnaturalización térmica de Bcy1 (1-50) en presencia y ausencia de péptido.

Se registró la elipticidad a 221 nm en función de la temperatura para una muestra 30 μM de Bcy1 (1-50) en presencia (línea rosa) o ausencia (línea azul) de un exceso de péptido Ira2 derivatizado con iodoacetamida. Las curvas se registraron en buffer PBS, y la señal fue corregida por la contribución del péptido a la señal de CD considerando los espectros nativo y desnaturalizado del péptido aislado. La señal se normalizó de acuerdo al mínimo valor de CD (correspondiente a la temperatura ambiente) para poder comparar ambas curvas.

Se observa un leve incremento (0,4^o C) en la T_m de Bcy1 (1-50) como consecuencia del agregado de un exceso de péptido Ira2. Sin embargo, esta diferencia resulta demasiado pequeña como para ser considerada significativa. Estos resultados no confirman que haya interacción entre ambos, pero tampoco se puede descartar en base a estos experimentos. Sería necesario repetir estos ensayos utilizando otras concentraciones de proteína y péptido, y otras condiciones de buffer, para descartar que la interacción no sea dependiente de la fuerza iónica, por ejemplo.

Análisis de las posibles superficies de Interacción de Bcy1 (1-50)

Si bien, como ya se señaló, en base a los resultados de este capítulo no se puede confirmar que el péptido Ira2 interactúe con Bcy1 (1-50), numerosos antecedentes avalan la hipótesis de que esa interacción debe existir. En primer lugar, como ya se señaló, se ha demostrado que la mutante Bcy1 $\Delta(1-85)$, a la que le falta el dominio D/D, no es capaz de interactuar [2]. Por otro lado, también se ha demostrado que la región 1-48 de Bcy1 es responsable de la localización subcelular de la proteína bajo ciertas condiciones de crecimiento [1]. Adicionalmente, los resultados del análisis bioinformático del Capítulo 2 sugieren que a partir del aminoácido 51 comienza la región del *linker*, que es desestructurada; esta observación no descarta que esa región participe de la interacción, dado que muchas veces las regiones intrínsecamente desestructuradas de las proteínas están involucradas en la formación de complejos, y adquieren estructura como consecuencia de la interacción. Sin embargo, el hecho de que Bcy1 (1-50) sea un dominio bien constituido, junto con el concepto de que en el sistema de mamíferos el dominio D/D es el responsable de la interacción con los dominios AKB de las AKAPs, y las similitudes estructurales que existen entre el péptido Ira2 y los dominios AKB de las AKAPs de mamíferos, sugieren fuertemente que la interacción podría darse a través de esta región. Por lo tanto, se decidió analizar la superficie del tetrámero Bcy1

(1-50) en base a su estructura cristalina para identificar potenciales sitios de interacción. En la **Figura 5.6** se muestran los resultados de este análisis.

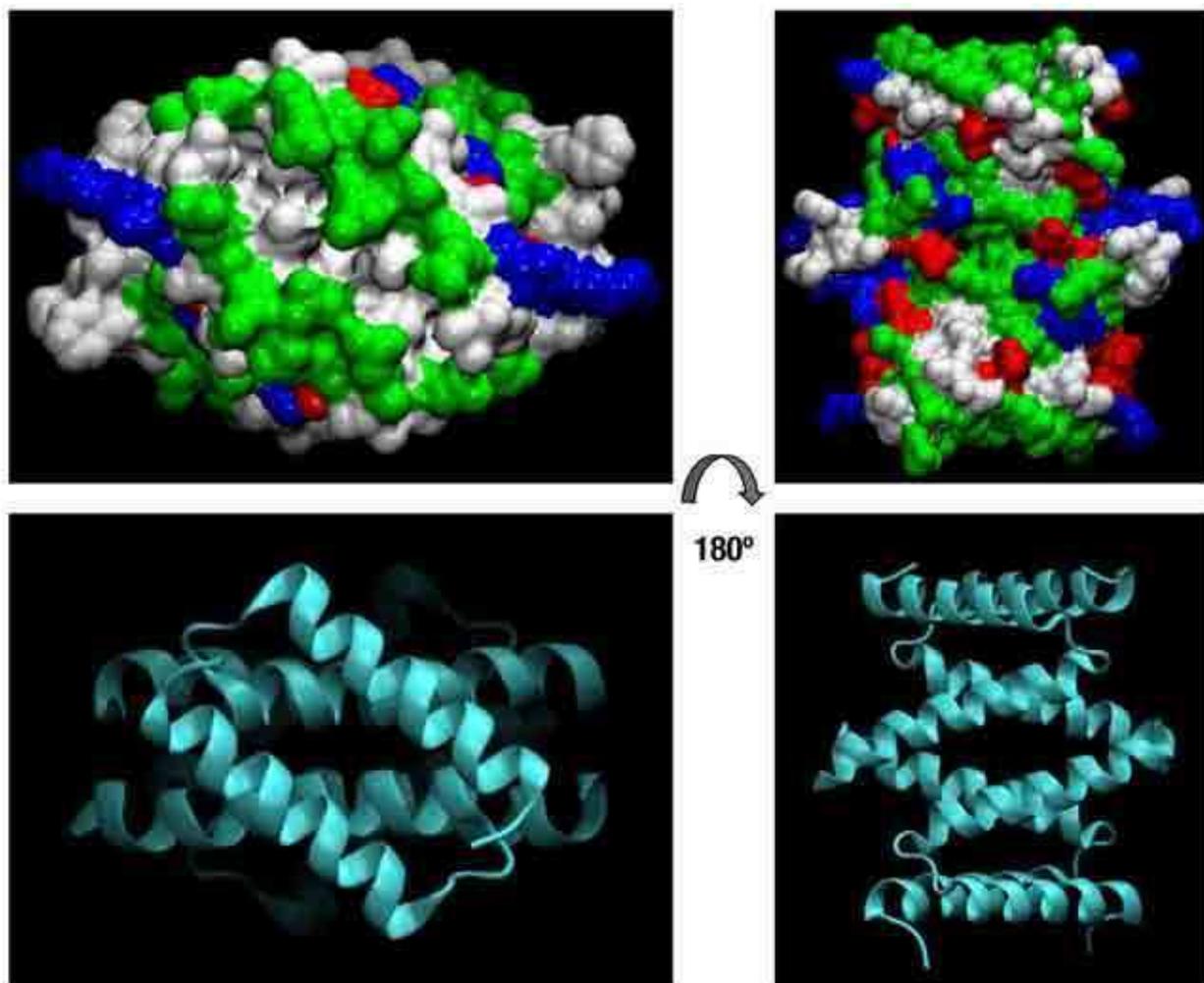


Figura 5.6: Características de la superficie del tetramero Bcy1 (1-50).

Se muestran dos vistas superficiales de la estructura de Bcy1 (1-50). En los paneles de la izquierda se muestra la superficie de unión a AKAPs “clásica” del sistema de mamíferos, mientras que en los paneles de la derecha se muestra la superficie que se constituye por los residuos expuestos de las hélices 2 de las dos cadenas de dímeros distintos en el tetramero; esta superficie es única de Bcy1 (1-50). En la parte superior se muestra la representación de la superficie, con los aminoácidos según el código de colores: blanco (hidrofóbico), verde (polar, no cargado), azul (básico), rojo (ácido). Como referencia, en el panel inferior se muestra el diagrama de cintas correspondiente a cada una de las imágenes del panel superior.

La superficie de unión a AKAPs “clásica” presenta características muy similares a la de los dominios D/D de mamíferos: se observa un surco hidrofóbico central, que en la subunidad RII α es responsable de la interacción con los residuos alifáticos de los dominios AKB de las AKAPs. Por otro lado, como consecuencia de la estructura tetramérica de Bcy1 (1-50), se generan nuevas posibles superficies de interacción, que son únicas de esta proteína. Como se observa en los paneles de la derecha de la **Figura 5.6**, esta superficie presenta características muy distintas a la anterior, ya que se encuentra enriquecida en residuos polares y cargados.

Como ya se señaló, en un trabajo de nuestro laboratorio se demostró que la mutación por alanina de los residuos del péptido Ira2 Lys2521, Lys2525 y Arg 2527, compromete seriamente la interacción con Bcy1. Por otro lado, en ese mismo trabajo se demostró que un péptido derivado de la AKAP-7 de mamíferos, que no interactúa con Bcy1, presenta cierta afinidad por la misma cuando se mutan por alanina algunos de los residuos que constituyen la cara hidrofóbica de la α -hélice anfipática [2]. En base a estas observaciones, en este trabajo se postuló que la interacción Ira2-Bcy1 podría estar dominada por interacciones electrostáticas.

Si Bcy1 (1-50) interactuara con el péptido Ira2 a través de la superficie “clásica” de unión a AKAPs, los residuos hidrofóbicos de la α -hélice anfipática (**ver Figura 5.1 (B)**) podrían encajar en el surco hidrofóbico que se observa en el panel superior izquierdo de la **Figura 5.6**, de manera análoga a la interacción D/D RII α -AKB. Sin embargo, no parece claro que los residuos básicos del péptido (que se ubican a ambos lados de la cara hidrofóbica) pudieran establecer interacciones electrostáticas con residuos negativamente cargados, ya que estos no se observan en esa superficie de Bcy1 (1-50). En esta línea, cabe recordar que se ha propuesto que la fosforilación del *cluster I* de serinas de Bcy1 (Ser3 y Ser4, localizadas en la región desestructurada N-terminal, y Ser9, localizada al inicio de la hélice 1) podría estar involucrada en la regulación de la interacción con alguna proteína responsable de la localización subcelular de Bcy1, de identidad aun desconocida [1]. Si este fuera el caso de Ira2, esas serinas fosforiladas, que sí se encuentran en las inmediaciones de esta superficie, podrían establecer interacciones electrostáticas con los residuos básicos del péptido. De ser así, esto podría justificar que no se haya observado unión entre Bcy1 (1-50) y el péptido en ninguno de los experimentos, dado que en la proteína recombinante obtenida de bacterias esos residuos no se encuentran fosforilados.

Capítulo 5 Interacción del dominio D/D de Bcy1 con el péptido Ira2

Por otro lado, existe también la posibilidad de que la interacción ocurra a través de la superficie que se muestra en el panel de la derecha de la **Figura 5.6**, que involucra a los dos dímeros que constituyen el tetrámero. Como ya se señaló, esa superficie se encuentra enriquecida en residuos polares y cargados; en particular, se aprecian numerosos residuos ácidos que podrían mediar interacciones electrostáticas con los residuos básicos del péptido. Como se discutirá en las Consideraciones finales, a menudo la oligomerización es un mecanismo que permite la generación de superficies de interacción con proteínas.

CONCLUSIONES Y PERSPECTIVAS

Los resultados preliminares presentados en este capítulo sugieren (aunque no demuestran) que, en concordancia con lo que se encuentra reportado, Bcy1 WT interactúa con el péptido Ira2 en solución. Sin embargo, aun no se ha podido determinar el valor de la constante de afinidad por los problemas experimentales que ya fueron mencionados. Para validar que el *quenching* de la fluorescencia del péptido se debe a la interacción con Bcy1, es necesario realizar ensayos de competencia por agregado del péptido sin marcar.

Por otro lado, no se obtuvo evidencia de que Bcy1 1-50 interactúe por sí misma con el péptido. Esto podría deberse a distintos motivos: 1) es posible que los experimentos de desnaturalización térmica no resulten adecuados para detectar la formación del complejo; 2) Bcy1 (1-50) podría no ser suficiente para garantizar la interacción, y requerirse por lo tanto de regiones presentes en el segmento 51-85; 3) podrían requerirse modificaciones post-traduccionales, como las fosforilaciones ya propuestas, para que ocurra la interacción. Próximamente se repetirán los ensayos de fluorescencia utilizando reactivos que permiten tapizar el interior de las cubetas para evitar el pegado del péptido a las paredes, para evaluar la primera hipótesis. De no observarse interacción, se generarán construcciones con los residuos de serina del *cluster I* mutados por aspártico para emular la fosforilación. Finalmente, se está trabajando en nuestro laboratorio (Tesis Doctoral de Enzo Tofolón, dirigida por la Dra. Silvia Moreno y codirigida por la Dra. Silvia Rossi) en la obtención de la construcción Bcy1 (1-85), para evaluar la segunda hipótesis.

REFERENCIAS

1. Griffioen G, Branduardi P, Ballarini A, Anghileri P, Norbeck J, Baroni MD, Ruis H: **Nucleocytoplasmic distribution of budding yeast protein kinase A regulatory subunit Bcy1 requires Zds1 and is regulated by Yak1-dependent phosphorylation of its targeting domain.** *Molecular and cellular biology* 2001, **21**(2):511-523.
2. Galello F, Moreno S, Rossi S: **Interacting proteins of protein kinase A regulatory subunit in *Saccharomyces cerevisiae*.** *J Proteomics* 2014, **109C**:261-275.

CONSIDERACIONES FINALES

“ Find all your need in your mind
If you take the time ”

Dream Theater, “ Take the time ”

CONCLUSIONES

En este trabajo se presenta la caracterización estructural del dominio D/D de la subunidad R de la PKA de *Saccharomyces cerevisiae*, Bcy1. Este es el primer estudio de este tipo que se lleva a cabo en un organismo no mamífero.

El análisis bioinformático y la estructura cristalina demostraron que la estructura secundaria, los residuos hidrofóbicos que constituyen el núcleo de dimerización, y la arquitectura general de los dominios D/D diméricos de mamíferos están presentes en Bcy1. Pese a que presenta algunas características que son típicas de los dominios D/D de las subunidades RII α , como por ejemplo la estructura del segmento N-terminal que precede a la hélice 1, el dominio D/D de Bcy1 contiene elementos distintivos, como la extensión adicional de la hélice 2, que presenta 1,5 giros más que el de la subunidad RII α . Esta observación es consistente con las diferencias observadas en cuanto a la orientación relativa de los dominios de unión a cAMP, que como se mencionó en la Introducción, en Bcy1 es muy distinta a la de las dos subunidades R de mamíferos [1]. En conjunto, estas observaciones demuestran que Bcy1 es una subunidad R con características estructurales únicas.

En concordancia con esta idea, se encontró una diferencia notable en cuanto a la estructura cuaternaria: inesperadamente, el dominio D/D de Bcy1 es un tetrámero en solución. También se demostró por experimentos de SLS y de entrecruzamiento químico la existencia de una forma de Bcy1 WT con esta misma estructura cuaternaria, que no había sido reportada hasta el momento. Esta forma oligomérica parece ser dinámica en solución, de manera dependiente de la concentración, como lo sugieren los experimentos de SEC y de CD de diluciones sucesivas del dominio D/D.

El análisis de las secuencias obtenidas de distintos organismos, en conjunto con las estructuras cristalinas, sugiere que la principal causa de esta diferencia en los estados oligoméricos de las proteínas de levaduras y mamíferos es una inserción de un aminoácido que existe hacia el final de la hélice 2 del dominio D/D de Bcy1. Esta inserción posiciona al residuo Arg45 en conformaciones sustancialmente distintas en los dominios D/D de Bcy1 y RII α , lo que a su vez tiene un impacto importante en la estructura cuaternaria. Esta observación está relacionada con las diferencias de estructura secundaria que se aprecian en esta región; cabe destacar que, cuando se alinearon las secuencias al inicio del Capítulo 2, antes de conocer la estructura cristalina, para optimizar el alineamiento fue necesario introducir un *gap* en los dominios D/D de mamíferos en esa misma posición.

Muchas veces, las asociaciones transitorias entre proteínas se encuentran dominadas por interacciones entre residuos polares o cargados, en contraste con las asociaciones permanentes que son típicamente hidrofóbicas [2]. En concordancia con esta observación, y con la idea de que la estructura cuaternaria de Bcy1 es dinámica en solución, la estructura cristalina reveló que, si bien cada dímero está sostenido por el núcleo hidrofóbico clásico de dimerización, las interacciones entre los dímeros que conforman el tetrámero son predominantemente de tipo puente salino y puente de hidrógeno.

Se ha propuesto que los organismos han evolucionado para mantener la oligomerización de sus proteínas debido a que esto les confiere ventajas funcionales [3]. Por ejemplo, la localización de un sitio de unión con un ligando en la interfaz entre los componentes de un complejo proporciona un mecanismo de regulación a través del estado de oligomerización, que es dependiente de la concentración: cuando la concentración es lo suficientemente alta como para formar el oligómero, el sitio de unión existe, mientras que a bajas concentraciones el complejo se separa en sus componentes individuales, y por lo tanto el sitio de unión se desarma. Además de este mecanismo regulatorio, el ensamblado de oligómeros puede aumentar el número de proteínas interactoras, ya que ofrece nuevas superficies de interacción que no están presentes en las subunidades individuales. [3].

Como ya se discutió en la Introducción, en la PKA la principal función de la dimerización de la subunidad R de mamíferos es proporcionar la superficie de interacción para las AKAPs. En otros organismos se han descrito holoenzimas diméricas RC que son completamente funcionales en cuanto a su actividad catalítica dependiente de cAMP, ya sea que existan así naturalmente debido a que sus subunidades R no tienen dominio D/D [4], o bien que fueran generadas artificialmente por cortes con proteasas en la región del *linker* [5]. Estas observaciones sugieren que en este sistema la dimerización se encuentra fuertemente vinculada a la interacción con otras proteínas. Efectivamente, en el caso de *S. cerevisiae*, la mutante Bcy1 $\Delta(1-85)$ es incapaz de establecer oligómeros; en concordancia, esa mutante tampoco es capaz de interactuar con Ira2 [6].

En vista de lo discutido en los párrafos anteriores, ampliando el concepto de “dimerización” al de “oligomerización”, se puede especular que la tetramerización de Bcy1 está relacionada con la generación de nuevas superficies de interacción con otras proteínas. Como se aprecia en la estructura cristalina, por cada tetrámero existen dos superficies clásicas

de unión a AKAPs, que son predominantemente hidrofóbicas, similares a las de los dominios D/D de mamíferos. Sin embargo, la interacción de Bcy1 con Ira2 está mediada por determinantes diferentes que los de la interacción clásica de los dominios D/D de mamíferos con los dominios AKB de las AKAPs: mientras que la segunda es de naturaleza hidrofóbica, para la primera resultan cruciales tres aminoácidos del péptido cargados positivamente [6]. En el modelo del tetrámero se aprecian sendas superficies expuestas a ambos lados del complejo, que podrían proporcionar sitios de interacción novedosos. Más aun, estas superficies están enriquecidas en aminoácidos polares y cargados (varios de ellos negativamente); es posible que Ira2 interactúe con Bcy1 a través de estas superficies adicionales.

El concepto de que la estructura cuaternaria de Bcy1 es dinámica, junto con la idea de que la tetramerización está relacionada con el establecimiento de superficies de interacción con otras proteínas, sugiere un rol regulatorio para la oligomerización de Bcy1, que a su vez podría estar regulada por la concentración intracelular de la proteína. En concordancia, se observó en este trabajo que una cepa de *S. cerevisiae* que sobreexpresa el fragmento Bcy1 (1-50), que podría tener la capacidad de alterar *in vivo* el estado de oligomerización de Bcy1 endógena, genera efectos fenotípicos apreciables, asociados a una actividad PKA que se encuentra atenuada respecto de la cepa WT. Esto podría estar asociado a la alteración de la localización subcelular de la holoenzima, o bien a la perturbación de algún mecanismo de regulación alostérico entre las subunidades R del tetrámero. Siguiendo la línea de la segunda hipótesis, la estructura cristalina reportada en este trabajo sugiere un modelo para el tetrámero de Bcy1 WT en el que los cuatro extremos N-terminales de los protómeros se agrupan en el centro, conformando el dominio D/D tetramérico, con los *linkers* irradiando hacia la periferia de forma concéntrica. Esto podría proporcionar mecanismos adicionales de comunicación entre los dominios de unión a cAMP de las cuatro subunidades del tetrámero, de forma análoga, aunque distinta, a lo que ocurre en las holoenzimas tetraméricas de mamíferos, que fueron discutidas en la Introducción.

Pese a que aun debe estudiarse la existencia de la estructura tetramérica *in vivo* (cabe aclarar que tampoco ha sido verificado en mamíferos que las subunidades R sean constitutivamente diméricas *in vivo*, y sin embargo siempre se consideran de esa manera), la evidencia presentada en este trabajo sugiere que, en contraste con la idea clásica de que las únicas formas posibles de la subunidad R son monómeros y dímeros (dependiendo de si tienen o no tienen dominio D/D) existe una mayor diversidad de estados oligoméricos. Esto

amplía notablemente el panorama en cuanto a la diversidad estructural que presentan estas proteínas en los diferentes organismos.

PERSPECTIVAS FUTURAS

A continuación se enumeran algunos puntos de interés que surgen a partir de los resultados de este trabajo y que pueden servir como disparadores de futuras líneas de investigación, y se proponen posibles abordajes experimentales para algunas de ellas:

1) Análisis del estado oligomérico de Bcy1 *in vivo*. Esto podría abordarse en primera instancia por entrecruzamiento químico intracelular, aprovechando que el EGS atraviesa las membranas, y posterior análisis de las distintas formas oligoméricas por *western blot*. En un segundo nivel, podrían utilizarse técnicas de microscopía de fluorescencia.

2) Análisis del posible rol fisiológico de la tetramerización. Para completar el panorama del análisis fenotípico presentado en el Capítulo 3, aprovechando el conocimiento que se tiene acerca de la estructura cristalina de Bcy1 (1-50), podrían generarse mutantes puntuales en el *locus* cromosómico de Bcy1 que generen proteínas diméricas y no tetraméricas, y evaluar los fenotipos asociados a estas mutaciones. Un buen candidato para este abordaje es el residuo Arg45.

3) Análisis del estado de oligomerización de la subunidad R de otros hongos, para verificar las predicciones, basadas en el análisis de secuencias, de que podrían ser tetraméricas al igual que Bcy1.

4) Estudios en solución de la estructura de la subunidad R completa. Siguiendo el enfoque utilizado para la subunidad R de mamíferos, conociendo ya las estructuras de alta resolución de los dominios D/D y de unión a cAMP, para completar el panorama estructural de Bcy1 podrían utilizarse técnicas como SAXS. Efectivamente, este trabajo está siendo llevado a cabo en nuestro laboratorio.

5) Análisis de la estructura cuaternaria de la holoenzima PKA de *S. cerevisiae*. La estructura tetramérica de Bcy1 presentada en este trabajo sugiere un interrogante clave, relacionado con el estado de oligomerización de la holoenzima. ¿Es posible que exista una holoenzima PKA octamérica, R₄C₄?

6) Estudio de la interacción de Bcy1 con Ira2 y otras proteínas interactoras. Este es un punto clave, que fue abordado de manera preliminar en este trabajo y para el cual no se obtuvieron resultados concluyentes, pero que potencialmente resulta de gran interés. Para ello, se continuará con los ensayos de interacción entre el péptido Ira2 y Bcy1, a efectos de elucidar si basta con el dominio D/D para garantizar la interacción o si se requieren segmentos adicionales, si las fosforilaciones del *cluster I* de serina desempeñan algún papel, etc. Eventualmente, de confirmarse la interacción, sería de gran interés mapear la superficie de Bcy1 (1-50) responsable de la misma, utilizando técnicas como intercambio hidrógeno-deuterio acoplado a espectrometría de masa y, en última instancia, la cristalización y resolución de la estructura del complejo. Finalmente, sería interesante ampliar el análisis a otras proteínas, como por ejemplo Hsp60.

7) Estudio estructural y funcional de la construcción Bcy1 (1-85). Este fragmento resulta de gran interés, dado que los resultados de este trabajo sugieren que el fragmento 51-85 podría estar involucrado en numerosos aspectos. Por ejemplo, resulta importante elucidar si el *linker I* desempeña algún papel en la oligomerización, si el *cluster II* de serinas fosforilables participa en la misma o bien en la interacción con otras proteínas, etc. Este punto también está siendo abordado actualmente en nuestro laboratorio.

8) Análisis del papel que desempeña el residuo de cisteína conservado en las subunidades R de hongos. Como se describió en el Capítulo 4, existe un residuo de cisteína conservado en casi todos los hongos hacia la mitad de la hélice 2. Se hipotetizó que este residuo podría estar involucrado en la formación de un puente disulfuro intercatenario. Sería interesante verificar esta hipótesis en aquellos hongos que lo presentan, y analizar qué rol fisiológico podría desempeñar esa modificación. Este punto resulta de interés dado que sería una característica distintiva de los dominios D/D de hongos, y además se ha reportado para la subunidad RI α de mamíferos que el puente disulfuro característico de estas subunidades desempeña un rol regulatorio de la activación de la holoenzima y de la interacción con las AKAPs que es dependiente del estado redox de la célula [7].

REFERENCIAS

1. Rinaldi J, Wu J, Yang J, Ralston CY, Sankaran B, Moreno S, Taylor SS: **Structure of yeast regulatory subunit: a glimpse into the evolution of PKA signaling.** *Structure (London, England : 1993)* 2010, **18**(11):1471-1482.
2. Jones S: **Computational and structural characterisation of protein associations.** *Advances in experimental medicine and biology* 2012, **747**:42-54.
3. Matthews JM, Sunde M: **Dimers, oligomers, everywhere.** *Advances in experimental medicine and biology* 2012, **747**:1-18.
4. Gunzburg J, Veron M: **A cAMP-dependent protein kinase is present in differentiating Dictyostelium discoideum cells.** *The EMBO journal* 1982, **1**(9):1063-1068.
5. Paveto C, Passeron S, Corbin JD, Moreno S: **Two different intrachain cAMP sites in the cAMP-dependent protein kinase of the dimorphic fungus Mucor rouxii.** *European journal of biochemistry / FEBS* 1989, **179**(2):429-434.
6. Galello F, Moreno S, Rossi S: **Interacting proteins of protein kinase A regulatory subunit in Saccharomyces cerevisiae.** *J Proteomics* 2014, **109C**:261-275.
7. Brennan JP, Bardswell SC, Burgoyne JR, Fuller W, Schroder E, Wait R, Begum S, Kentish JC, Eaton P: **Oxidant-induced activation of type I protein kinase A is mediated by RI subunit interprotein disulfide bond formation.** *J Biol Chem* 2006, **281**(31):21827-21836.