

La Química de la vida a escala atómica

Resumen: El conocimiento de la estructura de las macromoléculas es fundamental para la comprensión de los procesos esenciales de la vida. Los trabajos de Roger D. Kornberg, premio Nobel de Química 2006, proporcionan una imagen real de cómo se produce la copia de la información contenida en los genes y de su posterior transmisión a otros lugares de la célula para la síntesis de las proteínas. Roger D. Kornberg ha caracterizado la estructura y la función de las proteínas responsables del proceso de transcripción. La Cristalografía es la única disciplina experimental que permite la obtención de la imagen tridimensional de los enormes complejos macromoleculares a nivel atómico. *In memoriam* de Félix Hernández Cano (1941–2005)

El código universal

En el siglo XVIII Carlos Linneo (1707–1778) consagró su vida a la organización de las formas vivas. Linneo dio el empujón inicial a este colosal esfuerzo, cuando sólo contaba 28 años, con la publicación de su libro sobre la clasificación titulado *Systema Naturae*. En esta obra establecía las bases para la clasificación de los seres vivos a partir de la asignación a cada organismo de dos nombres latinos, el primero, indica el *género* y el segundo, la especie, que a su vez se iban englobando en categorías de mayor rango como la *familia*, el *orden*, la *clase*, el *filum* o el *reino*. De este modo, consiguió agrupar a los seres vivos en función de elementos morfológicos comunes y encontrar orden en la enorme diversidad de los seres vivos.

En realidad, el orden que subyace detrás de toda esa diversidad morfológica, tiene su base a escala molecular y se encuentra regido por unas reglas comunes que unen a todos los seres vivos desde la diminuta bacteria que vive en calderas submarinas a cientos de grados centígrados y a presiones altísimas, hasta el hombre. Este orden tiene dos actores fundamentales, la molécula de ADN, que encierra toda la información genética del organismo, y las proteínas, que son producidas a partir de esa información genética.

En todos los procesos biológicos participan proteínas (muchas de ellas similares entre los distintos reinos de la vida). Las que observamos en la naturaleza han evolucionado, igual que los seres vivos, a través de presión selectiva hasta desarrollar funciones específicas. En el caso de sistemas multicelulares tan complejos como el hombre, las proteínas se han especializado en una gran variedad de funciones hasta llegar a las alrededor de 100.000 proteínas diferentes de nuestro cuerpo

Las propiedades funcionales de las proteínas dependen de su estructura tridimensional. Esta es determinada por las secuencias particulares de aminoácidos de la cadena polipeptídica (típicamente entre 50 y 5.000 aminoácidos) que se pliegan, a partir de cadenas lineales, para generar dominios compactos con estructuras tridimensionales específicas (Figura 1). Los dominios plegados pueden servir como módulos para construir grandes ensamblados, tales como las cápsidas de los virus o las fibras del músculo, o bien pueden proporcionar sitios específicos de unión o catálisis, tal y como ocurre en las



Juan A. Hermoso, Julia Sanz Aparicio, Armando Albert

proteínas que transportan oxígeno o en las enzimas.

Así pues, para comprender la función biológica necesitaríamos deducir o predecir la estructura tridimensional de las proteínas y esto, precisamente, no lo podemos hacer. A pesar de esfuerzos considerables en los últimos 25 años, el problema del plegamiento sigue sin resolverse y es uno de los desafíos intelectuales más apasionantes de la biología molecular. Este problema persiste porque los 20 aminoácidos comunes pueden combinarse en más proteínas que átomos tiene el universo.

Además, hay muchas formas en que pueden generarse dominios estructurales similares a partir de secuencias de aminoácidos diferentes. Por el contrario, la estructura del ADN, construida a partir de 4 únicos nucleótidos que se unen dos a dos, es relativamente simple, regular y predecible. Así, aunque se han extraído algunas reglas generales y, en algunos casos, puede hacerse una predicción aproximada de la forma de una proteína, no es posible predecir de forma precisa y segura la estructura final de una proteína a partir de su secuencia de aminoácidos.

La Cristalografía de rayos X, el método fundamental de la Biología Estructural.

Precisamente éste es el cometido de la Biología Estructural, la determinación de las estructuras tridimensionales de las moléculas biológicas, especialmente proteínas, para obtener las claves de la función de estas moléculas en la célula viva. Para ello existe una serie de técnicas de determinación estructural (Tabla 1) de las que las fundamentales son la cristalografía de rayos X, la resonancia magnética nuclear y la microscopía electrónica. La cristalografía de rayos X es, sin duda, la más importante de todas ellas como lo prueba el hecho de que más del 80% de las estructuras tridimensionales depositadas en el PDB (del inglés *Protein Data Bank*) se han resuelto por este método. La clave del éxito de la cristalografía se basa en cuatro pilares: lo establecido del método, la capacidad de resolver a resolución atómica la estructura tridimensional de proteínas (o agregados de ellas) de cualquier peso molecular, el avance de las técnicas de ingeniería de proteínas (biología molecular) y el enorme desarrollo técnico en la producción y detección de rayos X.

Para comprender lo asentado del método cristalográfico basta recordar que ya en 1912 Max von Laue explicaba cómo la red periódica de un cristal dispersa un haz incidente de rayos X en unas direcciones específicas del espacio, y cómo, con Walter Friedrich y Paul Knipping, descubría el primer

GCMBE, Grupo de Cristalografía Macromolecular y Biología Estructural. Departamento de Cristalografía. Instituto de Química-Física "Rocasolano", CSIC. Serrano 119, 28006 Madrid. C-e: xalbert@iqfr.csic.es

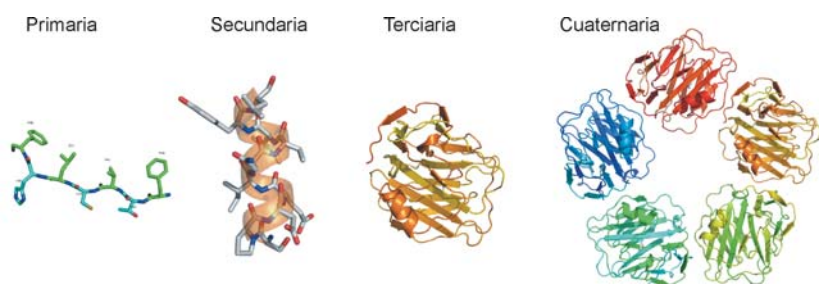


Figura 1. La secuencia de la cadena polipeptídica de una proteína es su estructura primaria. Diferentes regiones de la secuencia forman estructuras regulares tales como hélices alfa o lámina beta, que se determinan estructura secundaria. La estructura terciaria está formada por el empaquetamiento de tales elementos estructurales en una o varias unidades globulares compactas denominadas dominios. La proteína final puede contener varias cadenas polipeptídicas dispuestas en una estructura cuaternaria.

Tabla 1. Técnicas de determinación estructural en Biología (* Base de Datos de Proteínas (PDB), visitada el 8/11/2006).

	CRISTALOGRAFÍA			RMN	Crio-ME	Scattering en solución	
	Rayos X	Electrones	Neutrones			SAXS	SANS
Estado de la muestra	Cristales pequeños (10^{-3} mm ³)	Cristales 2D (10^{-6} mm ²)	Cristales grandes (1mm ³)	En solución (0,5 ml, ~1mM)	Congelada en hielo vítreo	En solución (0,02 ml, 0,1 mM)	En solución (0,2 ml, 0,1 mM)
Átomos observados	Átomos pesados, C, N, O, S, P, metales 0,5–3 Å,	Función de potencial	Todos los átomos, incluido el H/D	Protones	Resolución no atómica	Resolución no atómica	Resolución no atómica
Resolución	generalmente 2,2 Å	2–10 Å	1,5–10 Å	~1Å (rmsd)	>20 - 30 Å	Forma y distancias	Forma y distancias
Estructuras en bases de datos*	33.826	14	16	5.921	138	14	n.a.
Tamaño límite	Ninguno	Ninguno	<30 KDa	<30-40 KDa	Ninguno	Ninguno	Ninguno
Fortalezas	Alta resolución, grandes ensamblajes	Información directa de las fases, cristales 2D, proteínas de membrana	Enlaces de hidrógeno, hidratación, hidrógenos en sitios de metales	Dinámica, unión, hidrógenos	Grandes ensamblados	Forma, alguna orientación	Orientación forma, variación de contraste

patrón de difracción de rayos X. La aplicación estructural surgió poco después del equipo formado por William y Lawrence Bragg, padre e hijo, que en 1914 resolvieron la estructura del cloruro sódico. La utilización en el caso de las proteínas tuvo que esperar hasta 1957, año en que dos científicos (iniciados bajo la dirección de Sir Lawrence Bragg) Max Perutz y John Kendrew lograron la determinación estructural de la hemoglobina (una proteína encargada del transporte del oxígeno) tras diez largos años de paciencia y perseverancia. Ellos sentaron las bases de la metodología de la cristalografía de proteínas que fue, poco a poco, estableciéndose de modo firme desde entonces (Figura 2).

Además de la metodología cristalográfica, el desarrollo de las técnicas de biología molecular ha sido clave en la explosión de estructuras tridimensionales. Estas técnicas han permitido obtener grandes cantidades de proteínas puras a partir de la inserción de la instrucción genética que codifica a esa proteína, en otros organismos de muy rápido crecimiento como son bacterias y levaduras. Estas técnicas sirven también para facilitar la purificación de la proteína (por ejemplo, mediante la inserción de un fragmento proteico que se adhiera específicamente a una resina) o la resolución estructural mediante la sustitución de aminoácidos naturales por otros modificados que permiten resolver el problema de la fase (sustitución del S del aminoácido metionina por Se con gran capacidad de dispersión anómala).

La Cristalografía de rayos X es, hoy en día, inconcebible

sin el uso de la radiación sincrotrón. Estas fuentes de rayos X son aceleradores de partículas en los que electrones relativistas producen, al quebrar su trayectoria a lo largo del anillo, radiaciones muy intensas que van desde el infrarrojo hasta los rayos X duros. Su alto brillo y colimación, unido a la posibilidad de sintonización de la longitud de onda (imprescindible para las medidas de dispersión anómala) la han convertido en una herramienta absolutamente esencial en la cristalografía de proteínas. Esto lo prueba el hecho de que alrededor del 70% de las estructuras publicadas se han resuelto gracias al uso de esta radiación; un porcentaje que, además, se incrementa cada año. El desarrollo tecnológico se ha producido también en la detección de rayos X, e incluso en la cristalización de las proteínas, donde robots se encargan de la dispensación de las gotas de proteína y precipitante (la sustancia que consigue la formación del cristal proteico) con volúmenes del orden de la centena de nanolitros. Todo esto ha hecho posible que una proteína del tamaño de la hemoglobina determinada por Perutz y Kendrew, sea hoy posible resolverla en unas horas.

El impacto de la Biología Estructural ha sido enorme en campos como la Biología o la Medicina. A finales del siglo XX, el proyecto del Genoma Humano ha producido un enorme cuerpo de información sobre la secuencia del genoma humano. Además, también se conocen los genomas de patógenos humanos desde la tuberculosis a la malaria. Sin embargo, en el siglo XXI necesitaremos información detallada de la estructura, la función y la interacción entre las miles

La Química de la vida a escala atómica

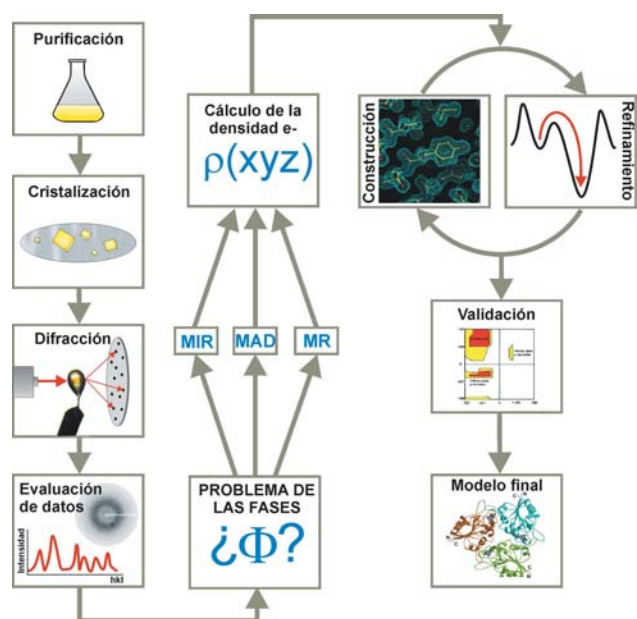


Figura 2. Esquema de las distintas etapas que configuran un estudio estructural mediante Cristalografía de rayos X. La cristalización y el problema de las fases (el experimento de difracción sólo proporciona las amplitudes de las ondas difractadas, pero no sus fases) son los dos cuellos de botella de la resolución estructural. En cristales de proteínas, el problema de la fase puede resolverse mediante tres métodos: (i) introduciendo átomos altamente dispersores, o método del Reemplazo Isomorfo Múltiple (MIR, del inglés, *Multiple Isomorphous Replacement*), (ii) introduciendo átomos dispersores anómalos, o método de Dispersión Anómala Múltiple (MAD, del inglés *Multiwavelength Anomalous Diffraction*) y (iii) mediante el Reemplazo Molecular (MR, del inglés *Molecular Replacement*), haciendo uso de un modelo estructural de una proteína homóloga, previamente determinada.

de proteínas codificadas por estos genomas. Este conocimiento estructural de proteínas clave tanto del hombre como de sus patógenos, es fundamental en el desarrollo de nuevos fármacos, o en la comprensión de las bases moleculares de la enfermedad (por ejemplo, algunas de las enfermedades neurodegenerativas, tales como Alzheimer o Parkinson parecen tener su origen en fallos en el plegamiento de las proteínas) o en la elucidación de los mecanismos de infección de agentes patógenos como bacterias y virus.

Del ADN al ribosoma: la transcripción a nivel molecular

Además de las aplicaciones médicas, la Biología Estructural ha desvelado procesos fundamentales de la vida y sin duda uno de los más importantes es la codificación del mensaje genético y su "traducción" en la síntesis de las proteínas (volvemos pues, sobre nuestros actores, el ADN y las proteínas). La estructura de la doble hélice del ADN (Figura 3), determinada por Watson y Crick en 1953 a partir de los patrones de difracción obtenidos por Maurice Wilkins y Rosalind Franklin, encierra en cada triplete de nucleótidos la información para la producción de un aminoácido. De este modo, en esas largas secuencias del ADN se encierra la información para la producción de todas las proteínas del organismo. Desde el descubrimiento del código genético por Crick quedó claro cómo a partir de 4 nucleótidos se podrían codi-

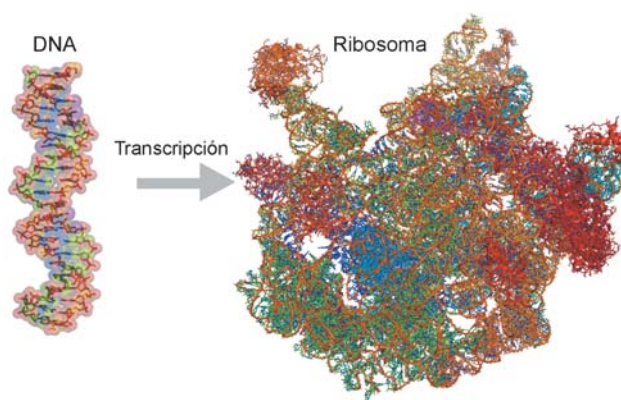


Figura 3. La transcripción del ADN es el primer proceso de la expresión genética. Durante la transcripción genética, las secuencias de ADN son copiadas a ARN mediante una enzima llamada ARN-polimerasa. La transcripción produce ARN mensajero como primer paso de la síntesis de proteínas que se lleva a cabo en los ribosomas.

ficar los 20 aminoácidos esenciales; quedaba aún un importante paso por comprender: ¿cómo esa información codificada en el ADN del núcleo puede llegar hasta los ribosomas (los gigantes agregados moleculares donde se sintetizan las proteínas) que se encuentran en el citoplasma?

La respuesta a esta cuestión se encuentra, en gran medida, en los trabajos estructurales de Roger Kornberg que ha conseguido desentrañar, mediante una importantísima serie de estudios cristalográficos, este proceso fundamental para la vida. Trabajos que le han valido la concesión del premio Nobel de Química de 2006.

La transcripción es un proceso fundamental en el desarrollo celular, en el cual la información genética almacenada en el ADN se activa a través de la síntesis de ARN mensajero (mARN) complementario, mediante la intervención de unas enzimas denominadas polimerasas. Los ribosomas traducen posteriormente este mARN sintetizando proteínas, que son las verdaderas responsables del control de todos los procesos que transcurren en la célula y por tanto de la regulación del funcionamiento general del organismo.

Durante algún tiempo se creía que el proceso de transcripción en las células eucariotas era similar al descrito en bacterias, que carecen de núcleo. Sin embargo, ahora se sabe que las células eucariotas presentan el ADN cromosómico unido a ciertas proteínas muy conservadas evolutivamente (histonas), y empaquetado formando la cromatina, en base a unas unidades conocidas como nucleosomas. Además, la maquinaria molecular implicada en la transcripción de las células eucariotas es mucho más compleja que en las células bacterianas, presentando niveles adicionales de regulación. En particular, contienen tres formas diferentes de ARN polimerasa,^[1,2] siendo la denominada ARN-polimerasa II esencial en todo el proceso, ya que transcribe todos los genes que codifican la síntesis de proteínas. Esta enzima es capaz de reconocer el sitio de inicio en un gen, separar las dos hebras del ADN, copiar una hebra en el ARN utilizando unidades de ribonucleósido-trifosfato como bloques y, finalmente, unir de nuevo las dos hebras mientras se trasloca a lo largo de la cadena de ADN. Para realizar esta tarea, la ARN-polimerasa II necesita la asistencia de cinco factores de transcripción, denominados TFIIB, D, E, F y H.

Los primeros estudios sobre el proceso de transcripción en células eucariotas datan de los años 50, cuando Weiss y

Gladstone hallaron la actividad de la ARN-polimerasa en el núcleo de células hepáticas de rata.^[3] Sin embargo, y dada la mayor facilidad encontrada en la purificación de enzimas procedentes de cultivos bacterianos, la investigación se centró en la regulación de la transcripción en células procariotas. Así, en 1965, Jacques Monod, André Lwoff y Francois Jacob obtuvieron, el premio Nobel de Medicina, entre otras cosas, por su descripción de cómo transcurre la transcripción en bacterias, demostrando que la ARN-polimerasa necesita la asistencia de lo que fue denominado factor sigma, que se une a la enzima y es capaz de reconocer una secuencia concreta de ADN, el promotor, indicando así el principio y el final de la información a copiar.

A pesar de la intensa labor realizada en los años 70, los científicos no fueron capaces de identificar en las células eucariotas moléculas análogas al factor sigma bacteriano. Al contrario de lo que sucedía con la ARN-polimerasa de bacterias, las enzimas purificadas a partir de células eucariotas no eran capaces de transcribir ADN, y estaba claro que se necesitaba un sistema enzimático capaz de transcribir específicamente ADN exógeno *in vitro*. En 1979, se reportó un extracto de cultivos de células de tejido humano que, junto con ARN-polimerasa II, fue capaz de iniciar específicamente la transcripción de un promotor vírico.^[4] El análisis de ese extracto mostró la existencia de una serie de componentes, que fueron referidos como factores generales de transcripción (GTF) porque están involucrados en virtualmente todos los genes.^[5] Además, diversos estudios condujeron a la identificación de elementos *enhancer*, secuencias concretas de ADN que no tienen actividad promotora por sí mismas, pero que pueden estimular la transcripción ya que unen proteínas activadoras de genes, que a su vez controlan el proceso.^[6]

La contribución de Kornberg

Roger D. Kornberg inició sus estudios sobre la transcripción en su estancia post-doctoral en el Laboratorio de Investigación Médica (MRC) en Cambridge, trabajando con Francis Crick y Aaron Klug en la estructura de la cromatina. Estudios previos utilizando difracción de rayos X habían demostrado que la cromatina está compuesta por la repetición de unidades de una longitud en torno a 100 Å. Hoy se sabe que las cuatro histonas *core* o nucleares, forman un octámero alrededor del cual se enrolla el ADN, en una longitud variable en función del organismo. Kornberg y Thomas, en 1974, observaron por primera vez la existencia de un tetrámero de las histonas H3 y H4,^[7] tras lo cual el mismo Kornberg propuso la formación de un octámero como unidad básica de la cromatina, el nucleosoma, que a su vez contenía un fragmento de ADN de 200 pares de bases.^[8]

De vuelta a Stanford, continuó sus estudios centrándose en el desarrollo de un nuevo sistema de transcripción *in vitro* utilizando células de levadura, concretamente de *Saccharomyces cerevisiae*,^[9,10] más fáciles de manipular que las procedentes de mamífero. A pesar de ello, el grupo de Kornberg invirtió diez años de trabajo en optimizar el sistema hasta que fue capaz de obtener ARN-polimerasa y los factores de transcripción en la forma adecuada. Esto le permitió identificar y aislar por primera vez el complejo mediador, *Med*, compuesto por 20 proteínas diferentes,^[11,12,13] e imprescindible para la transcripción de la información en las células eucariotas. En

bacterias, los represores y activadores de la transcripción se acoplan directamente a la ARN-polimerasa afectando su unión al promotor. En eucariotas, por el contrario, la cromatina y el complejo mediador constituyen niveles adicionales de regulación entre los factores de transcripción específicos de un gen y la ARN-polimerasa II, posibilitando una regulación mucho más sofisticada. Así, la gran complejidad de los organismos eucarióticos multicelulares se debe a la actuación concertada de sustancias-señal específicas de cada tejido, secuencias *enhancer* en el ADN y el complejo mediador, lo cual posibilita la transcripción de ciertos genes en cada tejido y, por tanto, la diferenciación celular.

Con este hallazgo, quedaban definidos los tres componentes esenciales de la regulación y transcripción génica en eucariotas: los factores generales de transcripción, el complejo mediador y la ARN-polimerasa II. Sin embargo, quedaba por entender el papel de cada uno de ellos en el proceso, lo cual pasaba necesariamente por una meticulosa caracterización de su interacción a nivel molecular. Con casi 60 subunidades distintas y un peso molecular de 3 millones de daltons la compleja maquinaria de la transcripción eucariota constituía un reto formidable para abordar su estudio estructural. Desde el principio, Kornberg decidió que la pieza clave era la ARN-polimerasa II, molécula compuesta a su vez de 12 subunidades, con un peso de 0,5 millones de daltons, inestable y muy difícil de aislar. Combinando técnicas de microscopía electrónica y cristalografía de rayos X pudo, por fin, resolver el problema después de 20 años de duros trabajos bioquímicos en la expresión y purificación de las distintas proteínas, para lo cual fue decisivo el sistema de transcripción *in vitro* que le permitió, además, producir la cantidad de muestra imprescindible para la obtención de cristales. Por otra parte, desarrolló un sistema de formación de cristales bidimensionales de proteína sobre capas lipídicas para los estudios de microscopía electrónica.^[14]

En 2001, dio a conocer en *Science* la primera estructura de la ARN-polimerasa II, que constaba de 10 subunidades, a una resolución inicial de 2,8 Å.^[15] El mismo año^[16] publicó la estructura del factor de elongación de la ARN-polimerasa II con los correspondientes fragmentos de ADN-molde y ARN-copia en el momento de la transcripción. Los científicos tenían por primera vez una verdadera imagen molecular de cómo transcurre el proceso de la transcripción (Figura 4).

Desde entonces, el grupo de Kornberg ha publicado una docena más de estructuras cristalinas de la ARN-polimerasa II formando distintos complejos funcionales con ADN, ARN, nucleótidos y otras proteínas implicadas en el proceso. En 2004, presentó un modelo del complejo iniciador de la transcripción, que incluye todos los factores generales de transcripción, en el que combina los resultados obtenidos por cristalografía y microscopía electrónica^[17] (Figura 5).

En uno de sus últimos trabajos^[18] ha revelado la estructura por microscopía electrónica de una parte del complejo mediador que contiene siete subunidades, con un peso molecular de 223 KDa. Mediante ensayos bioquímicos muestra que la interacción de la ARN-polimerasa II con el complejo mediador requiere la asistencia del factor general de transcripción TFIIF. El paso siguiente, deberá ser el estudio por difracción de rayos X de la ARN-polimerasa II, en complejos que incluyan elementos del complejo mediador, lo cual dará la imagen definitiva de la maquinaria completa responsable de la regulación transcripcional.

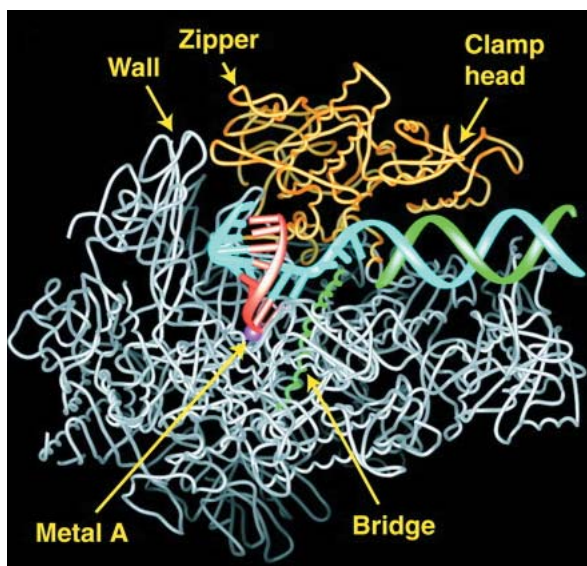


Figura 4. Primera imagen molecular del proceso de la transcripción obtenida por Kronberg en 2001.^[17] estructura de la ARN-polimerasa II, representada en color blanco, un fragmento del ADN molde, en color azul, y un fragmento del ARN formado, representado en color rojo.

El uso de la cristalografía para la obtención de imágenes de las moléculas aisladas es una técnica común en la actualidad. Sin embargo, la captura de un proceso dinámico en la forma en que transcurre es una tarea que trasciende del mero uso de la cristalografía. Lo verdaderamente notable del trabajo de Kornberg es el uso combinado de la cristalografía más metódica con un sólido conocimiento bioquímico, lo que le ha permitido un profundo control del proceso que persigue caracterizar. Por todo ello, la Academia sueca le ha concedido el galardón y resalta lo más notable de su trabajo desde el año 2000 como *detailed crystallographic pictures describing the transcription apparatus in full action*.

La Biocristalografía en el Instituto de Química-Física "Rocasolano"

En la estructura molecular está escrita su reactividad futura y pasada, su funcionalidad y su relación con el medio. El esfuerzo por obtener esa información ha guiado siempre la carrera de los cristalógrafos que, desde mediados del siglo XX, han trabajado en el Instituto de Química Física "Rocasolano" del CSIC. Esta idea es aplicable a sistemas de pequeño tamaño como moléculas orgánicas que empaquetan en un cristal no caprichosamente; a enzimas en los cuales se persigue entender cómo la molécula reconoce, une, desestabiliza un sustrato, fomenta la reactividad y, posteriormente, desestabiliza los productos; a sistemas complejos formados por proteínas que deben interactuar con otras débilmente para transmitir un estímulo puntual en un determinado tiempo; o a toxinas de membrana, que son proteínas capaces de reestructurarse completamente para realizar su función tanto en ambientes hidrofílicos como hidrofóbicos. Toda esta riqueza informativa está contenida en la estructura. Para desentrañarla, la cristalografía es una herramienta esencial, pero muchas veces además, es necesario desarrollar nuevas estrategias metodológicas tanto de resolución como de análisis o recurrir a diversas técnicas complementarias.

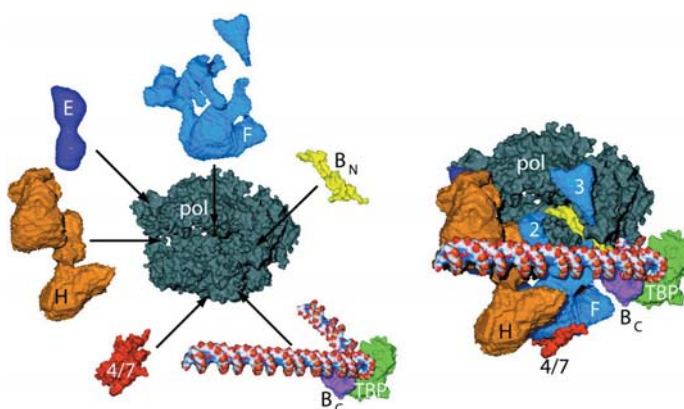


Figura 5. Modelo del complejo iniciador de la transcripción propuesto por Kronberg.^[18] Las estructuras obtenidas por difracción de rayos X y microscopía electrónica, a la izquierda, fueron ensambladas en un complejo de iniciación completo, que contiene la ARN-polimerasa (Pol), los cinco factores generales de transcripción y el fragmento de ADN.

En la actualidad, el objetivo del Grupo de Cristalografía Macromolecular y Biología Estructural (GCMBE) es interpretar distintos fenómenos biológicos mediante estudios estructurales a escala atómica, molecular y supramolecular. La idea es entender los procesos esenciales de la vida y utilizar este conocimiento para resolver problemas de interés biotecnológico o biomédico. Una característica importante de nuestra investigación es que está llevada a cabo por investigadores formados en distintos campos de la ciencia y siempre en estrecha colaboración. Esto permite la realización de proyectos "frontera" como los relacionados con la estructura y función de proteínas, siempre situados en la interfase de la química, la física y la biología. Así, nuestro laboratorio combina y desarrolla técnicas de biología molecular y química de proteínas, técnicas de cristalización de alto rendimiento, técnicas cristalográficas y análisis de bases de datos, sacando la máxima ventaja al hecho de contar con la experiencia acumulada desde hace muchos años y de la continua incorporación de jóvenes investigadores formados en los mejores laboratorios del mundo. Además, nuestra posición dentro de Instituto de Química-Física "Rocasolano", nos sitúa en un entorno privilegiado para interactuar con otros investigadores responsables de técnicas químico-físicas como la resonancia magnética nuclear, las técnicas calorimétricas y la espectroscopía de masas.

La investigación del grupo se organiza en torno a distintas líneas de trabajo, todas ellas interrelacionadas, y que incluyen estudios sobre enzimología estructural, estructura y enfermedad, interacción lípido-proteína y sistemas modulares.

Enzimología estructural

La enzimología se ocupa del estudio de procesos químicos catalizados por enzimas. Dentro de ella, la enzimología estructural, se apoya en datos estructurales para realizar este estudio. La estrategia consiste en abandonar el concepto estático de los estudios estructurales para darles una orientación más dinámica tratando de extraer información de los distintos intermedios de reacción y de la forma en la que la enzima ayuda a que la reacción progrese. La información química obtenida es utilizada para entender y resolver problemas biotecnológicos, como por ejemplo, los estudios realiza-

dos sobre la generación racional de enzimas más estables y con propiedades catalíticas mejoradas^[19] o el diseño de nuevos fármacos contra la depresión.^[20-21]

Una línea importante de nuestra investigación trata sobre las enzimas implicadas en el ciclo de la metionina (Figura 6). Este ciclo comprende un conjunto de reacciones esenciales para el funcionamiento celular y que regulan los niveles de metionina/homocisteína en el organismo. En el primer paso, metionina y ATP producen S-adenosilmetionina (SAM) en una reacción singular en dos etapas catalizada por la metionina adenosiltransferasa (MAT). La importancia de esta enzima radica en el papel esencial de SAM como principal donador de metilos en las numerosas reacciones de transmetilación que se producen en la célula, y por tanto se encuentra implicado en la síntesis y modificación de una gran variedad de metabolitos esenciales, tales como fosfolípidos, ADN, o proteínas, jugando un papel clave en la regulación de sus rutas metabólicas. En consecuencia, el mal funcionamiento en la síntesis y control de los niveles de SAM contribuye a la aparición de multitud de patologías, por lo que el control de la actividad MAT es de vital importancia para el normal desarrollo y funcionamiento celular. Por su parte, la betaína-homocisteína metiltransferasa (BHMT) regenera metionina a partir de homocisteína, por lo que es una enzima clave en el mantenimiento de los niveles plasmáticos de homocisteína, lo cual se ha demostrado como un claro factor de riesgo en el desarrollo de enfermedades cardiovasculares. La estructura tridimensional de estas dos enzimas y de sus complejos^[22-23]

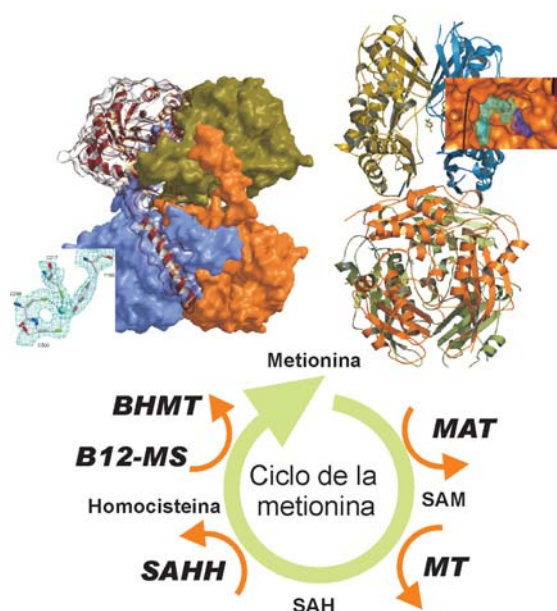


Figura 6. El ciclo de la metionina comprende un conjunto de reacciones esenciales para el funcionamiento celular y que regulan los niveles de metionina/homocisteína en el organismo, por lo que su mal funcionamiento produce serias patologías, hepáticas y cardiovasculares principalmente. En el primer paso, metionina y ATP producen S-adenosilmetionina (SAM) catalizada por la metionina adenosiltransferasa (MAT). SAM juega un papel esencial como principal donador de metilos en las numerosas reacciones de transmetilación que se producen en la célula. En el último paso del ciclo, se produce metionina de nuevo mediante la intervención alternativa de dos enzimas, una de ellas, la betaína-homocisteína metiltransferasa (BHMT). MAT y BHMT de mamíferos son enzimas oligoméricas formadas por cuatro subunidades.

ha sido fundamental para entender los mecanismos catalíticos a escala molecular.

Interacción proteína lípido

La Biología Estructural de las proteínas que interactúan con lípidos es, hoy en día, el reto más importante que tiene la Biocristalografía. Las proteínas de membrana suponen el ~25% de los genomas secuenciados y alrededor del 50% son posibles dianas farmacológicas; sin embargo, de las ~40.000 estructuras disponibles en el *Protein Data Bank*, sólo ~80 corresponden a proteínas de membrana. El estudio de estos sistemas no es simple. Implica el desarrollo de nuevas metodologías experimentales y, por tanto, es costoso en tiempo y recursos. En el GCMBE, estamos realizando un serio esfuerzo para situarnos entre los grupos implicados en estos sistemas. En este sentido, el estudio de toxinas y de lipasas es, particularmente, interesante ya que son estables tanto en solución como en una interfase lipídica.

Las toxinas formadoras de poros en membrana (PFP) muestran la importante propiedad de existir en dos estados bien definidos: uno, en forma de monómero soluble y, otro, en forma de poros oligoméricos capaces de lisar las células tras la interacción con su membrana. Dentro de este área, hemos llevado a cabo la determinación de la estructura monomérica soluble de la Sticolisina II (StnII) (20 KDa) así como su cristalización 2D en presencia de monocapas lipídicas y su posterior análisis por microscopía electrónica.^[24] Estos datos han permitido proponer un novedoso mecanismo de formación de poros (Figura 7). Además, hemos determinado la estructura nativa de la lectina formadora de poros de *Laetiporus sulphureus* (LSL) de 35 KDa.^[25] La proteína presenta dos dominios, uno con actividad de reconocimiento de carbohidratos (actividad lectina) y otro de unión a membranas. Este estudio ha permitido indicar cuáles son las regiones claves en el reconocimiento de carbohidratos y su relación con el proceso de inserción en membrana.

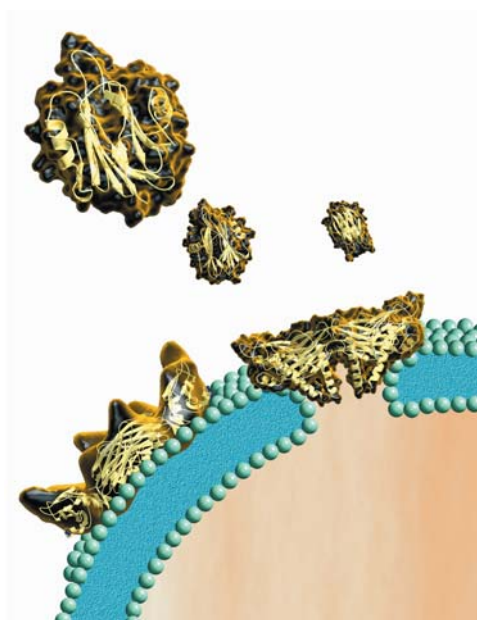


Figura 7. Etapas del mecanismo propuesto de formación de poros en actinoporinas: el monómero se une a la membrana, se agrupa en un estado tetramérico (preporo) y, finalmente, se forma el poro funcional.

Biocristalografía de sistemas complejos

Los procesos esenciales para la vida raramente son desempeñados por moléculas aisladas. Casi siempre, implican redes de proteínas que interaccionan y actúan coordinadamente para desarrollar una función compleja. Además, estas redes funcionan en un determinado compartimento celular (periferia celular, membrana, citosol, retículo...). Las proteínas han desarrollado estrategias que garantizan su correcta localización y relación con otras macromoléculas y que aseguran el desarrollo de un proceso biológico concreto. Éstas incluyen, la presencia en su estructura de diferentes dominios funcionales, por ejemplo, una proteína con un dominio de unión al ADN y otro de anclaje a la membrana que estabiliza el ADN en la periferia de la membrana.^[26] La determinación estructural de estas proteínas "multidominios" y de complejos intramoleculares sitúa a la Cristalografía más cerca de lo "vivo" y permite obtener una imagen más completa del problema biológico.

Las *Choline Binding Proteins* (CBPs) son un grupo de proteínas de superficie presentes en neumococos y sus bacteriófagos. Todas ellas son modulares, de modo que uno de los módulos es responsable de la actividad funcional y el otro del anclaje a la pared del neumococo vía reconocimiento de las moléculas de colina presentes en los ácidos teicóicos. El interés de esta línea de trabajo es doble, por un lado estas proteínas son fundamentales en los mecanismos de virulencia y en la interacción patógeno/hospedador y por otro tienen importantes implicaciones en biomedicina en el desarrollo de nuevos agentes terapéuticos en la lucha contra la resistencia a antibióticos. Como primer paso determinamos la estructura de la endolisina del fago Cp-1 (Cpl-1) de 39 KDa.^[27] Esta estructura, la primera de una CBP, nos ha mostrado el mecanismo de hidrólisis de la pared bacteriana por el fago, así como la interacción entre módulos que estabiliza la proteína. Esta estructura abre una nueva vía en la lucha contra el neumococo, al demostrarse su actividad *in vivo* como agente *enzimático*. Posteriormente, hemos determinado, la estructura de la fosforilcolin esterasa Pce de *Streptococcus pneumoniae* (606 aa) a resolución atómica^[28] (Figura 8). La estructura completa de esta enzima nos ha permitido postular un mecanismo de hidrólisis *in vivo* de la fosforilcolina presente en la superficie bacteriana. Además, ha mostrado las profundas

implicaciones que esta proteína tiene en el mecanismo de patogénesis de neumococo. Por un lado, la localización externa de Pce facilitaría la eliminación del mensajero PAF (factor de activación plaquetario) del torrente sanguíneo, bloqueando la señal de alerta ante la inflamación. Por otro lado, la activación de la enzima Pce en determinados momentos permitiría la reducción del número de las moléculas de fosforilcolina evitando el reconocimiento por parte de las proteínas del sistema inmune humano.

Las células son capaces de adoptar decisiones complejas a partir de diferentes estímulos externos. Toda toma de decisión requiere la codificación y transmisión de información desde la membrana celular hasta el núcleo celular donde se dispara una respuesta concreta. Este flujo de información está mediado por redes de proteínas que interaccionan entre sí y que son capaces de reconocer a otras pequeñas moléculas o iones. Las proteínas quinasas presentan en su estructura dominios discretos que sirven como andamiajes en su relación con otras proteínas de forma que regulan su función. La estructura de la quinasa SOS2 muestra este tipo de organización, la estructura del complejo entre SOS2 y la proteína reguladora y de anclaje a membrana SOS3 proporciona una imagen directa del mecanismo de regulación^[29] (Figura 9).

En definitiva, la estructura de las proteínas encierra la clave para comprender los mecanismos esenciales de la vida. A pesar de los más de 30 premios Nobel directamente relacionados con la Cristalografía, desde la estructura del cloruro sódico en 1914 a los trabajos de Kornberg, el nivel de complejidad ha ido aumentando. Sin embargo, cada día nos sorprenden nuevas estructuras que nos indican que todavía nos encontramos en los inicios de esta fascinante exploración de la Química de la vida.

Roger D. Kornberg nació en San Luis (EE UU) en 1947. Es licenciado en Químicas por la Universidad de Harvard y doctor por la Universidad de Stanford. Como otros investigadores realizó una estancia postdoctoral en Cambridge en el Laboratorio de Investigación Médica (MRC). En la actualidad, ostenta la cátedra George A. Winzer de Biología Estructural en la Escuela de Medicina de la Universidad de Stanford. (Figura 10)

Como su padre, Arthur Kornberg, que obtuvo el premio Nobel en Fisiología y Medicina en 1959, ha centrado sus estudios en la investigación genética. El uso de técnicas cristalo-

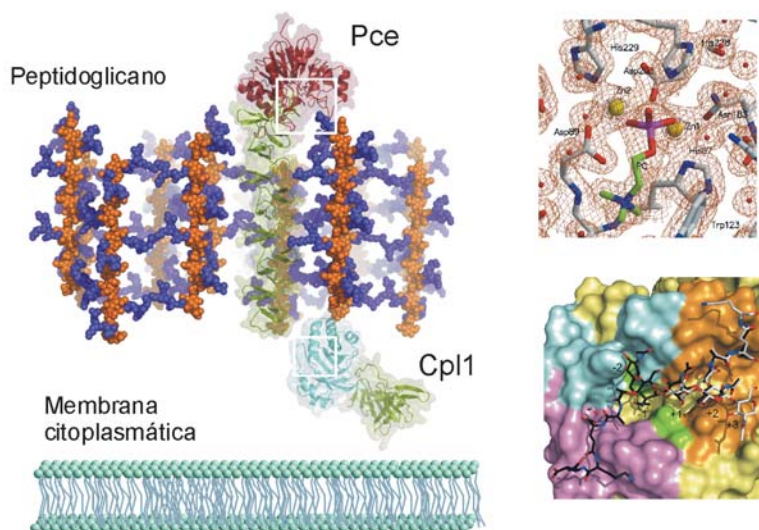


Figure 8. Estructura tridimensional de fosforilcolin esterasa (Pce) y Cpl-1, unidas a la pared celular del neumoco. Pce hidroliza las moléculas de fosforilcolina de la envuelta externa del peptidoglicano. Cpl-1 hidroliza el peptidoglicano para liberar a la progenie fágica. Se indican con un recuadro el sitio activo de cada enzima y se amplía a la derecha. El mecanismo hidrolítico de Pce se basa en el centro binuclear de Zn. La imagen inferior derecha muestra la glicosidasa Cpl-1 con los análogos de la pared bacteriana en el sitio activo y en los canales de reconocimiento del sustrato.

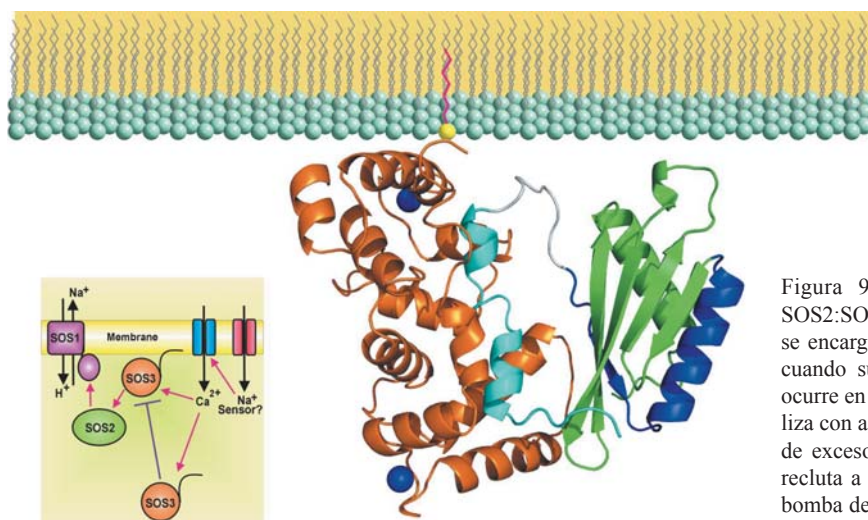


Figura 9. Estructura cristalográfica del complejo SOS2:SOS3. Las proteínas que componen la ruta SOS, se encargan de bombear sodio hacia fuera de la célula cuando su concentración es demasiado elevada. Esto ocurre en situaciones de sequía o cuando el riego se realiza con agua salinizada. SOS3 traduce la señal primaria de exceso de sodio y se coloca en la membrana. Allí recluta a SOS2, el activador de la bomba. SOS1 es la bomba de sodio.

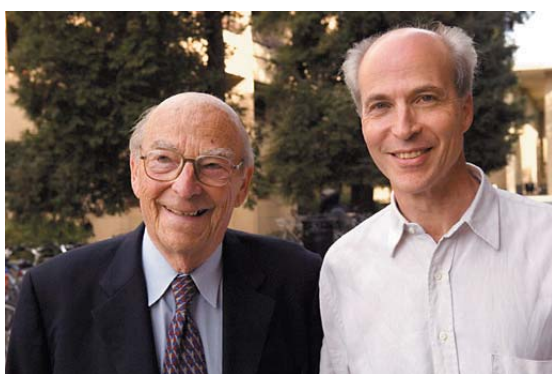


Figura 10. Roger D. Kornberg (derecha), premio Nobel de Química de 2006, con su padre, Arthur Kornberg, premio Nobel de Medicina y Fisiología de 1959.

gráficas ha marcado la diferencia con otros investigadores y ha sido fundamental en el esclarecimiento del proceso de transcripción de los genes.

Referencias

- [1] R. G. Roeder, W. J. Rutter, *Nature* **1969**, *224*, 234–237.
- [2] C. Keding, M. Gniazdowski, J. L. Mandel, F. Gissinger, Q. Chambon, *Biochem. Biophys. Res. Commun.* **1970**, *38*, 165–171.
- [3] S. B. Weiss, L. A. Gladstone, *J. Am. Chem. Soc.* **1959**, *81*, 4118–4119.
- [4] P. A. Weil, D. S. Luse, J. Segall, R. G. Roeder, *Cell* **1979**, *18*, 469–484.
- [5] T. Matsui, J. Segall, P. A. Weil, R. G. Roeder, *J. Biol. Chem.* **1980**, *255*, 11992–11996.
- [6] P. Moreau, R. Hen, B. Wasyluk, R. Everett, M. P. Gaub, P. Chambon, *Nucleic Acid Res.* **1981**, *9*, 6047–6068.
- [7] R. D. Kornberg, J. O. Thomas, *Science* **1974**, *184*, 865–868.
- [8] R. D. Kornberg, *Science* **1974**, *184*, 868–871.
- [9] N. F. Lue, R. D. Kornberg, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8839–8843.
- [10] M. H. Sayre, H. Tschonochner, R. D. Kornberg, *J. Biol. Chem.* **1992**, *267*, 23376–23382.
- [11] R. J. Kelleher III, P. M. Flanagan, R. D. Kornberg, *Cell* **1990**, *61*, 1209–1215.
- [12] P. M. Flanagan, R. J. Kelleher III, M. H. Sayre, H. Tschonochner, R. D. Kornberg, *Nature* **1991**, *350*, 436–438.
- [13] Y. J. Kim, S. Björklund, Y. Li, M. H. Sayre, R. D. Kornberg, *Cell* **1994**, *77*, 599–608.
- [14] E. E. Uzgiris, R. D. Kornberg, *Nature* **1983**, *301*, 125–129.
- [15] P. Cramer, D. A. Bushnell, R. D. Kornberg, *Science* **2001**, *292*, 1863–1876.
- [16] A. L. Gnat, P. Cramer, J. Fu, D. A. Bushnell, R. D. Kornberg, *Science* **2001**, *292*, 1876–1882.
- [17] D. A. Bushnell, K. D. Westover, R. E. Davis, R. D. Kornberg, *Science* **2004**, *303*, 983–988.
- [18] Y. Takagi, G. Calero, H. Komori, J. A. Brown, A. H. Ehrenberg, A. Hudmon, F. J. Asturias, R. D. Kornberg, *Mol. Cell.* **2006**, *23*, 355–364.
- [19] J. Sanz-Aparicio, J. A. Hermoso, M. Martínez-Ripoll, J. L. Lequerica, J. Polaina, *J. Mol. Biol.* **1998**, *275*, 491–502.
- [20] A. Albert, L. Yenush, M. R. Gil-Mascarell, P. L. Rodriguez, S. Patel, M. Martínez-Ripoll, T. L. Blundell, R. Serrano, *J. Mol. Biol.* **2000**, *295*, 927–938.
- [21] S. Patel, M. Martínez-Ripoll, T. L. Blundell, A. Albert, *J. Mol. Biol.* **2002**, *320*, 1087–1094.
- [22] B. González, M. A. Pajares, J. A. Hermoso, L. Alvarez, F. Garrido, J. R. Sufrin, J. Sanz-Aparicio, *J. Mol. Biol.* **2000**, *300*, 363–375.
- [23] B. González, M. A. Pajares, M. Martínez-Ripoll, T. L. Blundell, J. Sanz-Aparicio, *J. Mol. Biol.* **2004**, *338*, 771–782.
- [24] J. M. Mancheño, J. Martín-Benito, M. Martínez-Ripoll, J. G. Gavilanes, J. A. Hermoso, *Structure* **2003**, *11*, 1319–1328.
- [25] J. M. Mancheño, H. Tateno, I. J. Goldstein, M. Martínez-Ripoll, J. A. Hermoso, *J. Biol. Chem.* **2005**, *280*, 17251–17259.
- [26] A. Albert, D. Muñoz-Espin, M. Jiménez, J. L. Asensio, J. A. Hermoso, M. Salas M, W. J. Meijer, *J. Biol. Chem.* **2005**, *280*, 42486–42489.
- [27] J. A. Hermoso, B. Monterroso, A. Albert, B. Galán, O. Ahrazem, P. García, M. Martínez-Ripoll, J. L. García, M. Menéndez, *Structure* **2003**, *11*, 1239–1249.
- [28] J. A. Hermoso, L. Lagartera, A. Gonzalez, M. Stelter, P. Garcia, M. Martínez-Ripoll, J. L. Garcia, M. Menéndez, *Nat. Struct. Mol. Biol.* **2005**, *12*, 533–538.
- [29] M. Sánchez-Barena, M. Martínez-Ripoll, J. Zhu, A. Albert, *J. Mol. Biol.* **2005**, *345*, 1253–264.