

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Biología y Bioquímica Molecular I



**ESTRUCTURA Y MECANISMO DE ACCIÓN DEL
REGULADOR SOS3: RESPUESTA AL ESTRÉS SALINO
EN LA RUTA SOS**

**MEMORIA PRESENTADA PARA OPTAR AL GRADO DE
DOCTOR POR**

María José Sánchez Barrena

Bajo la dirección de los Doctores:

Armando Albert de la Cruz
Martín Martínez Ripoll

Madrid, 2005

ISBN: 84-669-2845-6

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I

ESTRUCTURA Y MECANISMO DE ACCIÓN
DEL REGULADOR SOS3: RESPUESTA AL
ESTRÉS SALINO EN LA RUTA SOS

Esta memoria, presentada como Tesis Doctoral por María José Sánchez Barrena, para optar al Título de Doctor por la Universidad Complutense de Madrid, ha sido realizada en el Grupo de Cristalografía Macromolecular y Biología Estructural del Instituto de Química-Física Rocasolano (CSIC), bajo la dirección del Doctor Armando Albert de la Cruz y del Doctor Martín Martínez Ripoll.

A mis padres. A mi familia. A todas las personas que siempre me han apoyado.

AGRADECIMIENTOS

La realización del trabajo presentado en esta memoria ha sido posible gracias a la ayuda desinteresada de muchas personas que durante estos últimos cuatro años han confiado en mí.

En primer lugar me gustaría dar las gracias a mis directores de tesis, el Dr. Armando Albert de la Cruz y el Dr. Martín Martínez Ripoll, por todo lo que me han enseñado y por lo mucho que he disfrutado trabajando con ellos. Creo que he tenido una suerte inmensa al cruzarme en el camino con personas tan trabajadoras, humildes y honestas, que no sólo me han enseñado ciencia, sino que también se han preocupado de que estuviera “como en casa” desde el primer día, considerando siempre mi opinión y aconsejándome en todo momento.

Esta tesis ha sido realizada gracias al apoyo del Ministerio de Educación y Ciencia, que me ha financiado durante cuatro años mediante una beca del Plan de Formación del Personal Universitario (que a ver si pronto son contratos...). También quiero agradecer a mi tutor, el Catedrático D. José G. Gavilanes, todas las facilidades prestadas para la presentación de esta memoria en la Universidad Complutense de Madrid.

Durante la elaboración de este proyecto, siempre he contado con el respaldo incondicional de todos los miembros del Grupo de Cristalografía Macromolecular y Biología Estructural y del Departamento de Cristalografía del Instituto Rocasolano; cada uno a su manera, han sido partícipes de mi formación científica y de la elaboración de esta memoria. Así quiero dar las gracias al Dr. José Miguel Mancheño, al Dr. Juan Hermoso, a la Dra. Julia Sanz y al Dr. Félix Hernández Cano. También quiero recordar en estas líneas a la Dra. María del Carmen Aprea, al Dr. Pepe Fayos, y a la Dra Lourdes Infantes, con los que también he compartido muchos buenos momentos.

Quiero también dar las gracias a todos mis compañeros de laboratorio del GCMBE, por toda la ayuda prestada y el tiempo y espacio robados durante la realización de esta memoria. A la Dra. Beatriz González, que ya “ha vuelto”, a Tomás, todo un doctor y funcionario ya, a Raquel, que me enseñó y ayudó mucho cuando llegué al GCMBE, a Pablo, siempre tan pausado y comedido, a Inma, que está siempre dispuesta a ayudar a quien lo necesite, a Rafa, el “faseador de Gadolinio”, a Celia, la segunda extremeña en aterrizar en el laboratorio, y a Laura, que aunque del Departamento de Química-Física de Macromoléculas Biológica, también ha pasado largos ratos entre nosotros. Además quiero dar mil gracias a mis compañeras del laboratorio de Biología Molecular, a Martita, la “chimba” por excelencia, que ya es doctora, a Sandra, mi compañera del club “Obsesión” y a Juana Mary, que tantos geles y cultivos prepara para todos. Ellas han sufrido

todas mis angustias con los dichosos rendimientos y me han ayudado y enseñado infinidad de cosas de Biología Molecular y Bioquímica.

Este trabajo ha podido llegar a buen fin gracias también a la ayuda de muchas personas de otros laboratorios como son la Dra. Begoña Monterroso y la Dra. Margarita Menéndez, que vaya lata les he dado con el CD, al Dr. Pepe Saiz, quien ha realizado los experimentos de ultracentrifugación analítica, a la Dra. Rosa Lebrón, por los experimentos de Maldí, a la Dra. Lola Solís, que tantos productos nos ha dejado, por las horas de centrifuga y espectrofotómetro, a la Dra. Marta Martínez, esta zaragozana que tantos consejos de Biología Molecular me ha dado y a la futura doctora María José Aguilar, que me ha puesto al día de la últimas normas de la RAE. También quiero recordar al Dr. Pepe Ramos, a Vera y a Carlitos, del Departamento de Microbiología de la UCO, con los que pasé unas semanas muy agradables aprendiendo a trabajar con levaduras y estrés salino.

Aunque los he dejado para casi el final, quiero mostrar mi más profundo agradecimiento a algunos profesores de la UEX que, de una manera desinteresada, me animaron a venir al GCMBE, en especial a la Dra. Mercedes Tirado, que me enseñó Química-Física de Macromoléculas y que siempre me ha guiado por el buen camino, sin pedir nada a cambio, al Dr. Fernando García Barros, que un día me dijo “haz lo que te gusta y no te conformes con lo que hay aquí” y al Dr. Paco Higes, que me enseñó lo bonita e interesante que es la Cristalografía y que desafortunadamente ya no está entre nosotros. También quiere recordar en estas líneas a tantos compañeros de la UEX con los que he compartido momentos inolvidables.

Por último quiero dar las gracias a mis padres, a mi hermana-amiga-compañera-de-piso Isa, a toda mi familia, a todos mis amigos de Mérida, de Calamonte, de Madrid y de otros lugares, a los que no nombro porque son muchísimos, y finalmente a Juan Ignacio, por ser el mejor compañero que podría imaginar, por estar siempre a mi lado, pase lo que pase. Gracias a todos por vuestra amistad, comprensión y por recordarme siempre qué es lo verdaderamente importante en la vida.

ABREVIATURAS

CD	Dicroísmo circular (<i>Circular dichroism</i>)
CBL	Proteína parecida a la Calcineurina B (<i>Calcineurin B-like</i>)
CIPK	Quinasas que interaccionan con las CBLs (<i>CBL-interacting protein kinases</i>)
DTT	Ditiotreitol
EDTA	Ácido etilendiamino tetra-acético (<i>Ethylenediaminetetraacetic acid</i>)
ESRF	Instalaciones europeas de radiación sincrotrón (<i>European synchrotron radiation facilities</i>)
GST	Glutation-S-transferasa
IPTG	Isopropil-β-D-tiogalactósido
LB	Luria Bertani
MAD	Dispersión anómala múltiple (<i>Multi-wavelength anomalous diffraction</i>)
MIR	Reemplazo isomorfo múltiple (<i>Multiple isomorphous replacement</i>)
MIRAS	Reemplazo isomorfo múltiple con dispersión anómala (<i>Multiple isomorphous replacement with anomalous scattering</i>)
MPD	2-Metil 2,4-pentanodiol
MR	Reemplazo molecular (<i>Molecular replacement</i>)
m/v	Relación masa/volumen
PDB	Banco de datos de estructuras de proteínas (<i>Protein data bank</i>)
PEG 3350	Polietilenglicol 3350 Da
PEG 4K	Polietilenglicol 4 KDa
PKS	Quinasas del tipo a SOS2 (<i>SOS2-like protein kinases</i>)
r.m.s.d.	Desviación cuadrática media (<i>Root mean square deviation</i>)
SAD	Dispersión anómala simple (<i>Single-wavelength anomalous dispersion</i>)
SCaBP	Sensores que unen Ca ²⁺ del tipo a SOS3 (<i>SOS3-like Ca²⁺ sensor/binding proteins</i>)
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato sódico (<i>Polyacrylamide gel electrophoresis in sodium dodecyl sulfate</i>)
SIR	Reemplazo isomorfo simple (<i>Single isomorphous replacement</i>)
SIRAS	Reemplazo isomorfo simple con dispersión anómala (<i>Single isomorphous replacement with anomalous scattering</i>)

SOS	Altamente sensible a la sal (<u>S</u> alt <u>o</u> verly <u>s</u> ensitive)
T _m	Temperatura de desnaturalización (<u>M</u> elting <u>t</u> emperature)
u.a.	<u>U</u> nidades <u>a</u> rbitrarias
UV	<u>U</u> ltraviole <u>t</u> a
v/v	Relación <u>v</u> olumen/ <u>v</u> olumen

ÍNDICE

ÍNDICE

1. INTRODUCCIÓN Y OBJETIVOS	1
1.1.- ESTRÉS SALINO EN PLANTAS	4
1.2.- LA RUTA SALT OVERLY SENSITIVE (SOS) DE <i>Arabidopsis thaliana</i>	7
1.3.- LAS FAMILIAS DE LOS SENSORES DE Ca²⁺ Y DE LAS PROTEÍNAS QUINASAS EN <i>Arabidopsis thaliana</i>	11
1.4.- OBJETIVOS	16
 2. TÉCNICAS EXPERIMENTALES Y RESULTADOS	19
2.1. EXPRESIÓN Y PURIFICACIÓN	21
2.1.1. Introducción.....	21
2.1.2. Expresión y purificación de SOS3.....	23
2.1.2.1. <u>Expresión de SOS3</u>	23
2.1.2.2. <u>Purificación de SOS3</u>	24
2.2. CRISTALIZACIÓN	27
2.2.1.- Introducción.....	27
2.2.1.1. <u>Principios básicos de la cristalización</u>	28
2.2.1.2. <u>Preparación del experimento de cristalización</u>	29
2.2.1.3. <u>Proceso experimental de cristalización</u>	30
2.2.1.4. <u>Métodos de Cristalización</u>	31
2.2.2. Cristalización de SOS3 y sus complejos.....	32
2.3. TÉCNICAS CRISTALOGRÁFICAS Y RESOLUCIÓN ESTRUCTURAL	35
2.3.1. Conceptos generales	35
2.3.2. Resolución estructural de SOS3	45
2.3.2.1. <u>Reemplazo molecular</u>	45
2.3.2.2. <u>Faseado con yoduro</u>	48
2.3.2.3. <u>Construcción y refinamiento de las estructuras</u> <u>de los complejos SOS3 Ca²⁺ y SOS3 Mn²⁺ Ca²⁺</u>	63
2.3.3. Análisis de la estructura de los complejos de SOS3	66

2.4. CARACTERIZACIÓN ESTRUCTURAL DE SOS3 EN SOLUCIÓN	68
2.4.1. Introducción	68
2.4.2. Comportamiento hidrodinámico de SOS3: análisis mediante ultracentrifugación analítica	68
2.4.2.1. Experimentos de equilibrio de sedimentación	69
2.4.2.2. Experimentos de velocidad de sedimentación	71
2.4.2.3. Predicción de propiedades hidrodinámicas a partir de un modelo atómico	72
2.4.3. Análisis estructural de SOS3 mediante espectroscopía de dicroísmo circular	73
2.4.3.1. Experimentos de dicroísmo circular en el UV-lejano y el UV-cercano	73
2.4.3.2. Desnaturalización térmica registrada por CD	75
2.4.4. Análisis de la estructura terciaria de SOS3 mediante espectroscopía de fluorescencia	77
2.4.4.1. Experimentos de emisión de fluorescencia	77
2.4.5. Cromatografía de interacción hidrofóbica	79
2.4.5.1. Experimentos de HIC con SOS3	79
3. DISCUSIÓN	81
3.1. LOS SITIOS DE UNIÓN A METALES DE SOS3	85
3.2. LA ESTRUCTURA DIMÉRICA DE SOS3	91
3.3. AUMENTO DEL CARÁCTER HIDROFÓBICO DE SOS3 POR LA UNIÓN AL Ca²⁺	93
3.4. COMPARACIÓN ESTRUCTURAL DE SOS3 Y CBL2	94
3.5. MECANISMO DE ACTIVACIÓN DE SOS3	98
3.6. MECANISMO MOLECULAR DE SOS3 EN LA RUTA SALT OVERLY SENSITIVE	102
4. CONCLUSIONES	105
5. BIBLIOGRAFÍA	109

6. ANEXO.....	121
----------------------	------------

1. INTRODUCCIÓN Y OBJETIVOS

1. INTRODUCCIÓN Y OBJETIVOS

Las proteínas desempeñan un papel crucial en prácticamente todos los procesos biológicos y pueden considerarse las verdaderas máquinas de la vida. Así por ejemplo, las enzimas son los catalizadores biológicos que controlan todos los procesos químicos en los sistemas vivos, las proteínas se encargan de almacenar y transportar otras moléculas, las hormonas actúan como mensajeros entre diferentes partes del organismo, y los anticuerpos proveen a la célula de inmunidad frente a infecciones. La diversidad de funciones biológicas que desempeñan estas complejas macromoléculas está determinada y es dependiente de la estructura tridimensional. A partir de la información genética, sólo es posible conocer la estructura primaria de las proteínas, la secuencia lineal de aminoácidos que constituyen el esqueleto polipeptídico. Sin embargo, para entender el lenguaje de las proteínas y desentrañar toda la información que contienen, es necesario conocer la estructura tridimensional de las mismas y situarlas en el contexto celular concreto (espacio y tiempo) en el que desempeñan su función.

Las proteínas necesitan relacionarse con su entorno e interactuar con el mismo para llevar a cabo su función. Mediante la interacción con iones, otras moléculas y macromoléculas, las proteínas regulan una gran cantidad de procesos biológicos. La importancia de estas interacciones en biología ha hecho que los procesos de reconocimiento en proteínas sea un área de considerable interés. Estos procesos de reconocimiento tienen como objetivo la transmisión de las señales que se generan en la célula como respuesta a los estímulos a los que está sometida. Las proteínas sufren cambios conformacionales y de estructura cuaternaria para interactuar con sus dianas y transmitir la señal. Todos estos cambios no son más que un reflejo de la naturaleza dinámica de la célula. El análisis estático de la estructura tridimensional de una proteína refleja parcialmente su funcionalidad, por lo que para comprender la complejidad y dinámica de los procesos celulares hay que conocer también cómo interacciona con las moléculas que la rodean y los cambios estructurales que sufre para dicho fin. En estos procesos de reconocimiento y transmisión de señales, las interacciones no covalentes juegan un papel principal, por lo que es necesario poder caracterizar a nivel atómico la naturaleza de dichas interacciones, y así entender en mayor medida las fuerzas que gobiernan todos estos procesos.

1.1.- ESTRÉS SALINO EN PLANTAS

Una característica importante que distingue a las plantas de otros organismos multicelulares complejos es que las plantas tienen un lugar de crecimiento y desarrollo fijo, por lo que han de soportar una multitud de condiciones medioambientales adversas, tales como la sequía, el frío o la salinidad del suelo, que pueden considerarse los más graves de los estreses. El exceso de sales en el suelo inhibe el crecimiento de las plantas y causa grandes pérdidas en la producción agrícola en todo el planeta (Tester & Davenport, 2003; Hasegawa *et al.*, 2000). Cerca del 20% de las áreas cultivadas y del 50% de las tierras de regadío están afectadas por problemas de salinidad (Zhu, 2001).

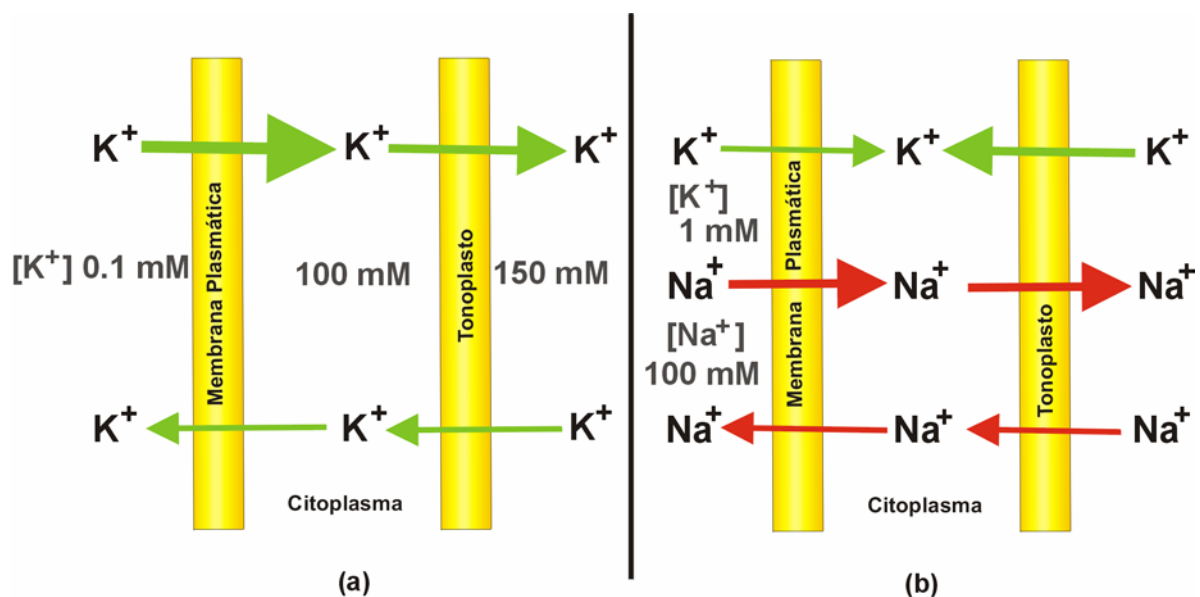


Figura 1.1.1: Homeostasis del Na^+ y del K^+ de una célula en (a) un suelo normal y (b) estresado. En células no estresadas, la concentración de K^+ , $[K^+]$, es muy superior a la del medio, para conseguirlo, el K^+ es asimilado del medio externo y acumulado en el citoplasma y la vacuola; una parte de este K^+ es devuelto al medio para un correcto equilibrio o como consecuencia de un reajuste osmótico. En un suelo estresado, la presencia de altas concentraciones de Na^+ en el medio externo inhibe el flujo de K^+ al interior de la célula y favorece la entrada y acumulación de Na^+ en el citoplasma. La célula pone en marcha mecanismos para la devolución del Na^+ al medio externo o para acumularlo en la vacuola (el grosor de las puntas de flecha indica magnitudes relativas) (Serrano & Rodríguez-Navarro, 2001).

La homeostasis iónica es fundamental para la fisiología de las células. Es necesario regular adecuadamente los flujos de los iones a través de la célula para mantener controlada una concentración baja de iones tóxicos y acumular los iones que son esenciales para la misma. Las plantas emplean un transporte activo primario, mediado por H^+ -ATPasas, y un transporte secundario, mediado por canales o cotransportadores, para mantener altas concentraciones de ión potasio (K^+) y bajas concentraciones de ión sodio (Na^+) en el citosol. La homeostasis intracelular de los iones Na^+ y K^+ es importante para la actividad de muchas enzimas que se

encuentran en el citosol, el mantenimiento del potencial de membrana y la regulación del volumen de la célula (Figura 1.1.1). Por tanto, el mantenimiento de los mecanismos que controlan la homeostasis del Na^+ y K^+ es esencial en condiciones de estrés salino. Entender cómo las plantas se enfrentan a altas concentraciones de Na^+ es de gran importancia para el presente y el futuro de la agricultura en un mundo con suelos cada vez más salinizados. Además, la regulación del transporte iónico mediada por las señales de estrés salino provee un modelo para la comprensión de la regulación general de la homeostasis iónica en las plantas.

La principal consecuencia del estrés salino por Na^+ es que se genera un enorme potencial a través de la membrana plasmática de las células de plantas, que favorece el transporte pasivo de Na^+ al citosol. El Na^+ entra en la célula a través del transportador de alta afinidad de K^+ HKT1 (Rus *et al.*, 2001; Maser *et al.*, 2002) y a través de canales de cationes no selectivos (Amtmann & Sanders, 1999) (Figura 1.1.2). Una cuestión intrigante es entender cómo es posible que a lo largo de la evolución, se hayan mantenido los transportadores de Na^+ , tales como HKT1, dado el efecto tóxico que ocasiona el Na^+ en la célula (Zhu, 2003). El estrés por Na^+ impide la entrada de K^+ a la célula (Hasegawa *et al.*, 2000). Cuando el Na^+ entra en la célula y se acumulan altos niveles del mismo, se hace tóxico para las enzimas que regulan el metabolismo, ocasionando un cese en el crecimiento y desarrollo de la célula, y en última instancia la muerte (Hasegawa *et al.*, 2000). Para combatir la acumulación de Na^+ en el citosol es necesario compartimentarlo en la vacuola o bien expulsarlo de la célula (Figura 1.1.1 y Figura 1.1.2). Esto requiere una regulación de los transportadores de Na^+ del tonoplasto y de la membrana plasmática.

El almacenamiento de Na^+ en la vacuola contribuye no sólo a la disminución de la concentración de Na^+ en el citosol, sino que también contribuye al ajuste osmótico para mantener la entrada de agua desde soluciones salinas. Otros organelos, como los plástidos y mitocondrias, podrían también acumular Na^+ , contribuyendo de este modo a la compartimentación subcelular total. En *Arabidopsis*, la familia de los antiportadores Na^+/H^+ NHX se encarga de dicha función (Blumwald, 2000) (Figura 1.1.2). AtNHX1 y AtNHX2 se encuentran localizados en la membrana vacuolar y sus niveles de transcripción están regulados por la fitohormona ABA (Yokoi *et al.*, 2002). Se ha demostrado que la sobreexpresión de *AtNHX1* en diversas plantas aumenta la tolerancia salina (Apse & Blumwald, 2002).

El papel de la eliminación de Na^+ de la célula no es intuitivo en plantas, ya que el Na^+ expulsado de una presentaría un problema para las células colindantes. Así, el papel del flujo de Na^+ hacia el exterior de la célula debe considerarse en tejidos específicos y en el contexto de la planta, considerada como un todo. En *Arabidopsis*, la expulsión de Na^+ al exterior se lleva a cabo por el antiportador de Na^+/H^+ SOS1, que se localiza en la membrana plasmática (Shi *et al.*, 2000; Qiu *et*

al., 2002; Quintero *et al.*, 2002; Shi *et al.*, 2002). La actividad de SOS1 se ha detectado sólo en plantas estresadas (Qiu *et al.*, 2002). SOS1 es un intercambiador de Na^+/H^+ electroneutral que es específico por Na^+ y no puede transportar Li^+ o K^+ (Qiu *et al.*, 2002, Qiu *et al.*, 2003). La mayor actividad de SOS1 se ha detectado en la raíz y en las células que rodean el tejido vascular de la planta (Shi *et al.*, 2002). El patrón de expresión de SOS1, junto con los resultados obtenidos con mutantes, sugieren que SOS1 posee distintas funciones. Primero, expulsar el Na^+ al entorno de la raíz; segundo, y de manera indirecta, limitar la entrada de Na^+ a la vacuola; y tercero, controlar el transporte a larga distancia de Na^+ entre la raíz y las hojas de la planta. La activación del antiporteador de Na^+/H^+ SOS1 en condiciones de estrés salino está regulado por el complejo formado entre las proteínas SOS2 y SOS3 (Figura 1.1.2) (Qiu *et al.*, 2002; Quintero *et al.*, 2002). A la ruta que controla la expulsión de Na^+ de la célula se le ha denominado SOS, *Salt Overly Sensitive*, de ahí el nombre de las proteínas implicadas.

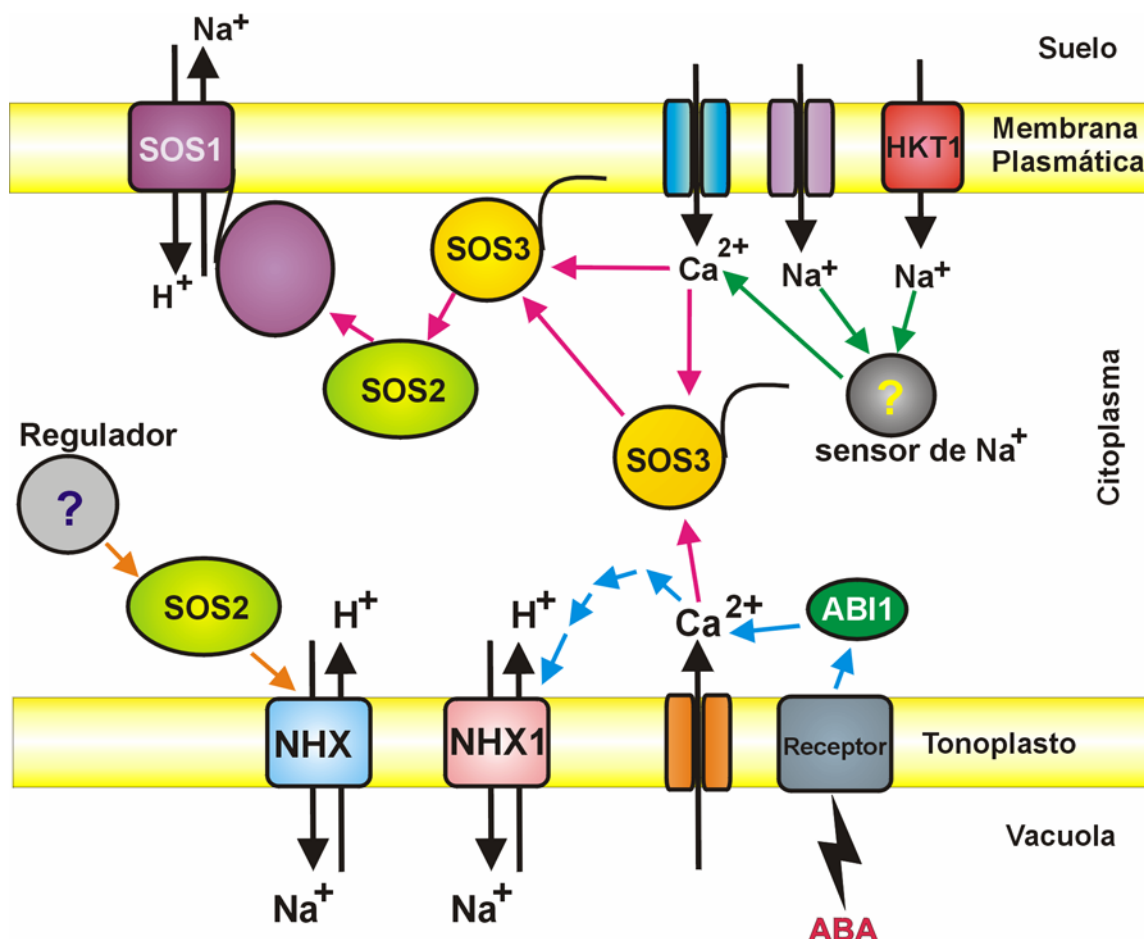


Figura 1.1.2: Esquema de la maquinaria encargada de controlar la homeostasis del Na^+ cuando la planta se encuentra en condiciones de estrés salino.

El estudio molecular de todos los sistemas involucrados en el control de la toxicidad del Na^+ tiene una gran importancia por la aplicación biotecnológica que subyace. La generación de plantas transgénicas en las que se introduzcan, por ejemplo, sistemas como el de la ruta SOS o los antiportadores NHX, va a permitir el crecimiento y desarrollo de estas plantas en terrenos salinos, por lo que solucionará en gran medida los problemas actuales de la agricultura mundial, motivados por la sobreexplotación de los suelos y el cambio climático.

1.2.- LA RUTA SALT OVERLY SENSITIVE (SOS) DE *Arabidopsis thaliana*

Mediante estudios genéticos y bioquímicos se ha identificado la ruta *Salt Overly Sensitive* (SOS) para la homeostasis del Na^+ y la tolerancia salina en *Arabidopsis* (Zhu, 2000; Zhu, 2003). En esta ruta, SOS3 es una proteína que une el ión calcio (Ca^{2+}) y es capaz de sentir la señal del Ca^{2+} que se genera en el citosol como respuesta al estrés salino. SOS3 presenta similitud secuencial con la subunidad reguladora de la Calcineurina, una fosfatasa del tipo 2B, y los sensores neuronales de calcio de origen animal (Liu & Zhu, 1998; Guo *et al.*, 2001). Sin embargo, *Arabidopsis* no posee ninguna proteína fosfatasa del tipo 2B y el sensor de Ca^{2+} SOS3 activa a una proteína quinasa, SOS2, y no a una fosfatasa. SOS3 es una pequeña proteína miristoilada que no presenta actividad enzimática por sí misma; la unión del Ca^{2+} , la miristoilación y los extremos N y C-terminal son requeridas para la función de SOS3 en la tolerancia salina (Ishitani *et al.*, 2000) (Figura 1.2.1). Se ha propuesto que la miristoilación es importante para el reclutamiento de SOS2 en la membrana plasmática, donde se encuentra el antiportador SOS1 (Figura 1.1.2) (Ishitani *et al.*, 2000).

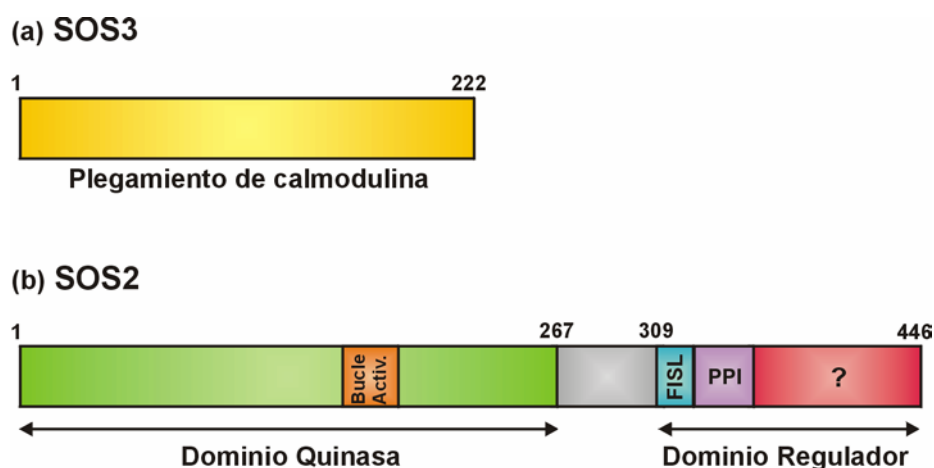


Figura 1.2.1: Esquema donde se compara el tamaño de (a) SOS3 con el de (b) SOS2.

SOS2 es una proteína quinasa del tipo Ser/Thr que también participa en la tolerancia salina de *Arabidopsis* (Liu *et al.*, 2000). El análisis del doble mutante *sos3/sos2* ha demostrado que SOS3 y

SOS2 actúan en la misma ruta (Figura 1.1.2) (Halfter *et al.*, 2000). SOS2 es una proteína que pertenece a la familia de las RD quinastas (Nolen *et al.*, 2004) y que posee un dominio quinasa N-terminal similar al encontrado en las proteínas SNF1 (*Suc nofermenting1*) y AMPK (*AMP-activated kinase*) (Hardie, 1999), y un dominio regulador C-terminal de plegamiento desconocido, que se compone de un motivo denominado FISL, un dominio PPI (*protein phosphatase interaction*), para la interacción con fosfatasas del tipo 2C (tales como ABI1 y ABI2), y uno o varios motivos en el extremo C-terminal con función desconocida (Figura 1.2.1(b)) (Gong *et al.*, 2004). Se ha demostrado que SOS3 interacciona físicamente con SOS2 mediante ensayos de doble híbrido en levadura e *in vitro* (Halfter *et al.*, 2000). Además, se sabe que en presencia de Ca^{2+} , SOS3 activa la capacidad de la quinasa SOS2 para fosforilar sustratos (Halfter *et al.*, 2000). La unión de SOS3 a SOS2 está mediada por un motivo de veintidós aminoácidos del dominio regulador de SOS2, denominado FISL, que autoinhibe la actividad quinasa mediante la interacción intramolecular entre los dominios quinasa y regulador (Guo *et al.*, 2001) (Figura 1.2.2 y Figura 1.2.3). Mediante estudio *in vivo* e *in vitro*, se ha propuesto un mecanismo para la activación de SOS2 (Figura 1.2.3).

In vitro y en ausencia de SOS3, la quinasa puede ser activada bien mediante la mutación Thr¹⁶⁸→Asp (Figura 1.2.2 y Figura 1.2.3(a)) en el bucle de activación, que mimetiza la fosforilación por otra quinasa, bien mediante la supresión del motivo autoinhibitorio FISL, o bien por la combinación de ambos, que produce un efecto sinérgico en la actividad quinasa de SOS2 (Guo *et al.*, 2001; Gong *et al.*, 2002a). El cambio de dos residuos adicionales del bucle de activación, Ser¹⁵⁶ y Tyr¹⁷⁵ a Asp también activa la quinasa (Figura 1.2.2), sugiriendo que la fosforilación múltiple debe ser importante para la actividad de SOS2 (Gong *et al.*, 2002b). El bucle de activación es un motivo estructural que regula la actividad de este tipo de quinastas y está implicado en la interacción con el cofactor, Mg^{2+} o Mn^{2+} , el correcto posicionamiento del ATP y la interacción con el sustrato (Nolen *et al.*, 2004). Por tanto, una distorsión de la conformación del bucle de activación impide una correcta orientación del ATP, de los residuos catalíticos y de la proteína sustrato. La fosforilación del bucle de activación es el mecanismo más común para que éste adquiera la conformación adecuada y se lleve a cabo la reacción.

In vivo, se ha propuesto que la quinasa puede ser activada mediante la unión del complejo SOS3 Ca^{2+} al motivo FISL, mediante fosforilación del residuo Thr¹⁶⁸ del bucle de activación o mediante una combinación de ambos (Figura 1.2.2 y Figura 1.2.3(b)). Además, se ha propuesto que la fosforilación/desfosforilación de SOS2 por una quinasa/fosfatasa puede jugar un papel muy importante en la regulación de SOS2 *in vivo*. Aunque aún no se conoce la quinasa/fosfatasa, se ha propuesto que las fosfatasas del tipo 2C, como por ejemplo ABI2 que se unen a SOS2, desfosforilen el bucle de activación (Ohta *et al.*, 2003).

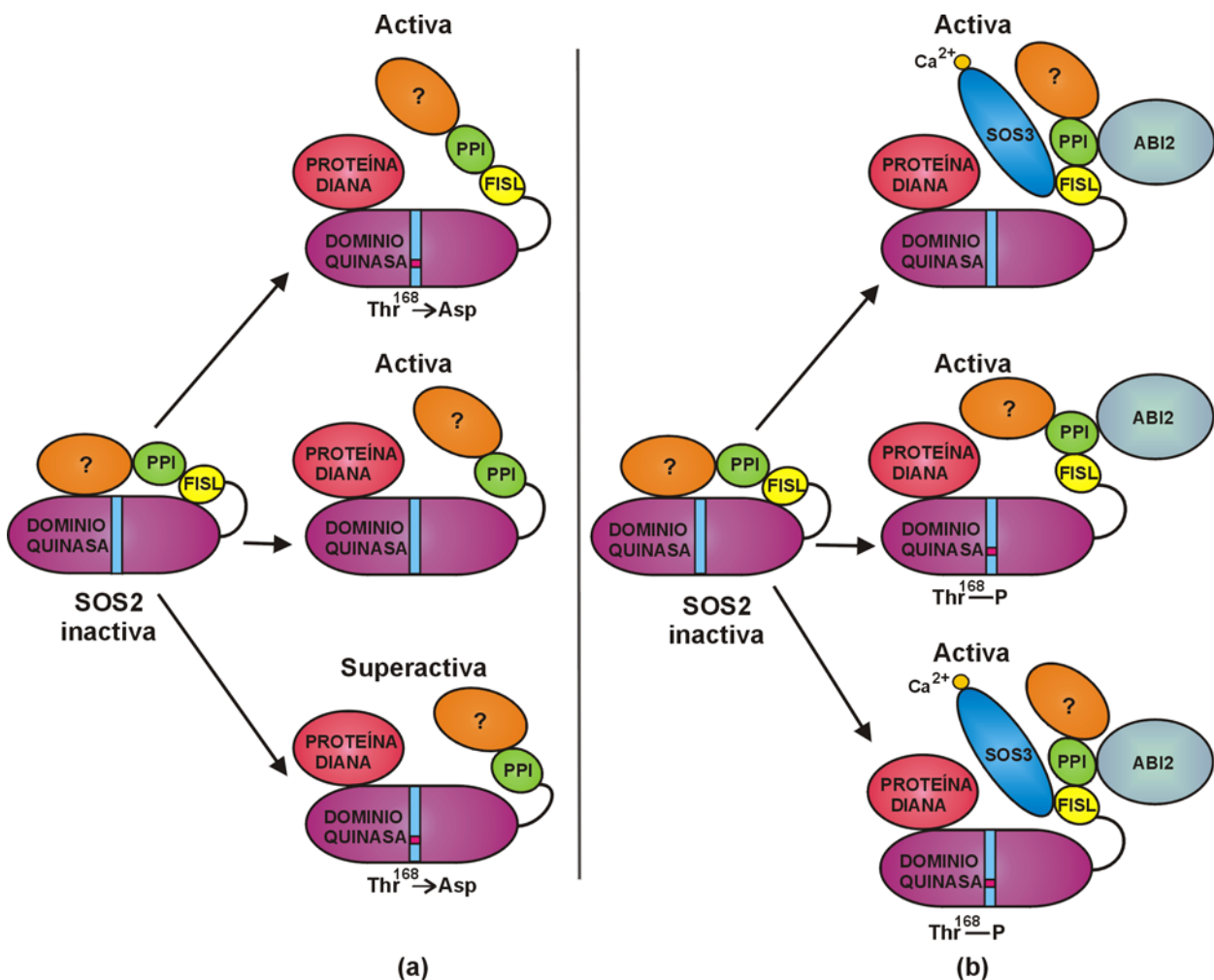


Figura 1.2.3: Mecanismo de activación de SOS2 (a) *in vitro* y (b) el propuesto *in vivo*. El extremo C-terminal del dominio regulador de función desconocida aparece en naranja. El bucle de activación del dominio quinasa se muestra en azul claro. El dominio de interacción con fosfatasa del tipo C, como por ejemplo la proteína ABI2 (en azul), se muestra en verde y con la etiqueta PPI.

En condiciones de estrés, las proteínas SOS2 y SOS3 son necesarias para la activación del antiportador de Na⁺/H⁺ SOS1 (Figura 1.1.2). (Qiu *et al.*, 2002). SOS1 está compuesta por un dominio formado por doce hélices transmembrana, y un dominio C-terminal (de más de setecientos aminoácidos) que reside en el citoplasma (Shi *et al.*, 2000). Estudios realizados en levadura han proporcionado evidencias adicionales de la interacción de SOS2 y SOS3 y la regulación de SOS1. La coexpresión de SOS1, SOS2 y SOS3 aumenta dramáticamente la tolerancia salina en levadura en un mutante que carece de transportadores de Na⁺ endógenos (Quintero *et al.*, 2002). El complejo SOS2-SOS3 fosforila y activa al antiportador SOS1 expresado en levadura, aumentando la exclusión del Na⁺ y aumentando la tolerancia salina (Quintero *et al.*, 2002). Estos resultados ponen de manifiesto la funcionalidad de estas tres proteínas en la homeostasis iónica y la tolerancia a la sal en *Arabidopsis*. Además, estudios recientes han demostrado que la sobreexpresión de SOS1 o de un mutante de SOS2

constitutivamente activo mejora la tolerancia salina en *Arabidopsis* (Shi *et al.*, 2003; Guo *et al.*, 2004), sugiriendo que la sobreexpresión conjunta de SOS1, SOS2 y SOS3 debe aumentar en gran medida la tolerancia a la sal en plantas.

Finalmente, es interesante remarcar que se ha demostrado que existe coordinación entre los antiportadores de Na^+/H^+ localizados en el tonoplasto y en la membrana plasmática de *Arabidopsis*. Así, aunque NHX1 está regulado por la hormona ABA, otros antiportadores del tipo NHX presentan una actividad dependiente de la quinasa SOS2, pero no de SOS1 o SOS3 (Qiu *et al.*, 2004). Sin embargo, se desconoce qué proteína activa a SOS2 en esta nueva ruta y, además, la regulación de NHX no implica cambios en la fosforilación del transportador (Qiu *et al.*, 2004). Estos resultados muestran que la quinasa SOS2 juega un papel muy importante en el mantenimiento de la homeostasis del Na^+ y es un componente crítico en la maquinaria implicada en la tolerancia salina en plantas.

1.3.- LAS FAMILIAS DE LOS SENSORES DE Ca^{2+} Y DE LAS PROTEÍNAS QUINASAS EN *Arabidopsis thaliana*

Las plantas experimentan cambios de temperatura, luz y humedad, excesos y deficiencias de minerales y ataques de organismos herbívoros y patógenos durante el ciclo de vida. Todos estos estímulos provenientes del exterior provocan señales específicas que alteran su desarrollo y crecimiento. Los segundos mensajeros son pequeñas moléculas que bien atraviesan la membrana o bien se generan en la misma, para transformar una señal primaria que es recibida en la membrana celular desde el medio externo.

El ión calcio (Ca^{2+}) actúa como segundo mensajero en las rutas de señalización inducidas por procesos fisiológicos y medioambientales en las plantas (Trewavas & Malho, 1998; Sanders *et al.*, 2002; White & Broadley, 2003). De esta forma, los sucesos que ocurren en el exterior son percibidos como oscilaciones en la concentración citosólica del Ca^{2+} , y en la especificidad de la señal influyen tanto el cambio en la concentración, como la cinética y compartimentación de la misma (Harmon *et al.* 2000; Scrase-Field & Knight, 2003). En la célula, los niveles de Ca^{2+} están fuertemente regulados, y pequeños cambios transitorios en la concentración intracelular de Ca^{2+} proveen de información a la célula para la posterior regulación de la actividad enzimática y de la expresión de genes como respuesta a un determinado estímulo. Por tanto, las señales provenientes del exterior son transformadas a oscilaciones en la señal del Ca^{2+} , que posteriormente tienen que ser registradas por diferentes sensores y que mediante la unión de dicho catión regulan distintas rutas de señalización (Snedden & Fromm, 1998; Zielinski, 1998). Por tanto, además de la especificidad de la señal del Ca^{2+} , el tipo de sensor que responde a la señal

determina un segundo nivel de regulación. Teniendo en cuenta que un mismo sensor puede interactuar con diversas dianas, el potencial de regulación específica de distintas rutas se magnifica indudablemente.

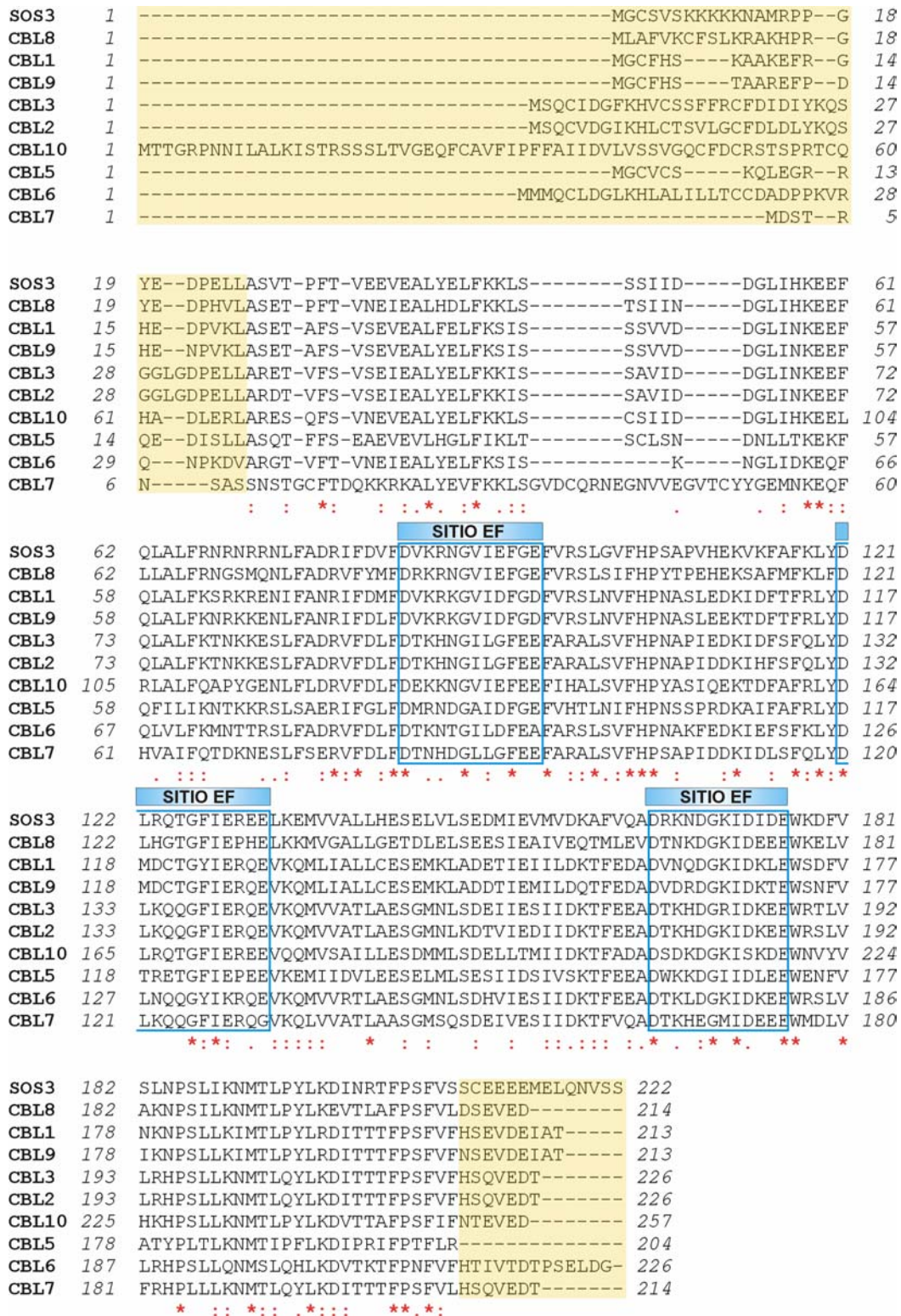


Figura 1.3.1: Alineamiento secuencial de las CBLs de *A. thaliana* (Pearson & Lipman, 1998). Los símbolos *, ; y . indican los residuos idénticos, conservados y semiconservados, respectivamente. Los recuadros azules muestran los sitios EF de unión a Ca²⁺. Con fondo amarillo se muestran los extremos N y C-terminal.

Existen dos tipos de sensores, aquellos que cuando unen Ca^{2+} , cambian de conformación y modulan su propia función o actividad mediante interacciones intramoleculares (*sensor responders*), o aquellos que no poseen actividad enzimática y tras la unión del Ca^{2+} cambian de conformación para interactuar con sus correspondientes dianas, por ejemplo, quinasas o fosfatasas (*sensor relays*) (Sanders *et al.*, 2002). En plantas, los sensores del tipo “*responders*” mejor caracterizados son las proteínas quinasas dependientes de Ca^{2+} (CDPK) (Harmon *et al.*, 2000) y entre los “*relays*” se encuentran las calmodulinas. SOS3 pertenece a este segundo tipo de sensores, ya que no tiene actividad enzimática y se ha demostrado que regula la actividad de la quinasa SOS2 (Halfter *et al.*, 2000).

Desde el aislamiento inicial y caracterización del sensor de Ca^{2+} SOS3 (Liu & Zhu, 1998) y de su diana SOS2 (Liu *et al.*, 2000), se han identificado otros nueve sensores y veinticuatro quinasas que conforman dos nuevas familias de proteínas en *Arabidopsis* (Kudla *et al.*, 1999; Shi *et al.*, 1999; Kim *et al.*, 2000; Albrecht *et al.*, 2001; Luan *et al.*, 2002; Kolukisaoglu *et al.*, 2004, Gong *et al.*, 2004). La familia de los sensores de Ca^{2+} a la que pertenece SOS3 se ha denominado S CaBP (*SOS3-like Ca²⁺ sensor/binding proteins*) (Guo *et al.*, 2001) o CBL (*Calcineurin B-like*) (Kudla *et al.*, 1999). A la familia de SOS2 se le ha dado el nombre de PKS (*SOS2-like protein kinases*) (Halfter *et al.*, 2000) o CIPK (*CBL-interacting protein kinases*) (Shi *et al.*, 1999). Estas dos familias de proteínas solo se han encontrado en plantas superiores, por lo que se ha sugerido que desempeñan una función específica en las mismas.

La familia de proteínas CBL está bastante conservada en tamaño. La mayoría de los genes predicen cadenas polipeptídicas de 23.5 a 26 KDa en *Arabidopsis*. La única excepción es CBL10, que presenta una extensión en el extremo N-terminal, que aunque no corresponde a ningún motivo o secuencia conocida, es de esperar que tenga relevancia funcional (Figura 1.3.1). La identidad secuencial de las diferentes CBLs varía entre un 29 y un 92% (33-93% de similitud). Las proteínas más parecidas son CBL1 y CBL9, y CBL2 y CBL3 que presentan una identidad secuencial del 89 y 92%, respectivamente. Incluso las proteínas más divergentes, CBL5 y CBL10, exhiben un 29% de identidad en la secuencia. Las diez proteínas presentan tres manos EF para la unión del Ca^{2+} , aunque el grado de conservación de las mismas con respecto al que presentan las calmodulinas es muy variado. El espaciado entre las manos EF es igual, de manera que las pequeñas variaciones en tamaño únicamente afectan a los extremos N y C-terminal. Estos resultados también han sido observados en las CBLs de arroz, sugiriendo que las CBLs de plantas deben de tener una estructura tridimensional conservada, mostrando por tanto un modo de acción similar. Esto último se refleja en que las CBLs interactúan con las CIPKs a través del mismo motivo estructural de las últimas, el motivo FISL o NAF (Halfter *et al.*, 2000; Guo *et al.*, 2001; Albrecht *et al.*, 2001).

La familia de proteínas CIPK también presentan un elevado grado de conservación en *Arabidopsis thaliana* y en *Oriza sativa* (Kolukisaoglu *et al.*, 2004). En *Arabidopsis*, estas proteínas presentan una identidad secuencial que está en el rango del 33-78%. El dominio quinasa N-terminal está mucho más conservado (51-90% de identidad y 63-96% de similitud secuencial) que el dominio regulador C-terminal (24-58% de identidad y 36-69% de similitud secuencial) (Figura 1.2.2). Una excepción en el dominio regulador es el motivo FISL, que se encuentra altamente conservado en toda la familia de proteínas y que por tanto, define a estas quinasas como dianas de las CBLs en la señalización por Ca^{2+} . Debido al elevado grado de conservación, es de esperar que las CIPKs presenten la misma funcionalidad que SOS2 (Figura 1.2.2).

El elevado número de CBLs y CIPKs sugiere la existencia de una red compleja de posibles interacciones, que asegura una descodificación de la señal del Ca^{2+} diferenciada y que redundante en una correcta respuesta de la célula a los estímulos percibidos. Entre los factores que regulan esta red hay que destacar:

- **La diferente afinidad por el Ca^{2+} de las CBLs.** En animales, es conocido que los sensores de Ca^{2+} , como los sensores neuronales (NCS) difieren en la afinidad por dicho catión (Burgoyne & Weiss, 2001). Esto implica que una familia de sensores puede actuar en un rango grande de concentraciones de Ca^{2+} y cada miembro descodificar señales de Ca^{2+} muy específicas. Sin embargo, este tipo de estudios no ha sido desarrollado con las CBLs. Un nivel de regulación adicional del Ca^{2+} consistiría en la dependencia del catión para la interacción sensor-quinasa. Por ejemplo, SOS3 interacciona con SOS2 en ausencia de Ca^{2+} , pero éste es necesario para la activación de la quinasa; sin embargo, la pareja CBL1/CIPK1 solo interacciona en presencia de Ca^{2+} .
- **La diferente especificidad de las CBLs por las CIPKs.** Aunque es posible un número elevado de combinaciones CBL/CIPK, se ha demostrado que existen preferencias en la interacción de estas proteínas. Pocas CBLs interaccionan con múltiples CIPKs, mientras que la mayoría de las CBLs únicamente interaccionan con unas pocas (alrededor de cuatro) CIPKs; también ocurre que distintas CBLs interaccionan con una misma CIPK (Batistic & Kudla, 2004; Kim *et al.*, 2000; Guo *et al.*, 2001). Ello induce a pensar que habrá sensores, como SOS3, que solo funcionarán ante un estímulo, pero habrá otros, como CBL1, que distribuyan las señales provenientes de distintos estímulos a distintas quinasas (Figura 1.3.2). Sin embargo, estos estudios se han realizado mediante ensayos de doble híbrido en levadura, y aún queda por confirmar si estas parejas también interaccionan *in vivo* y por estudiar las bases estructurales de las interacciones.

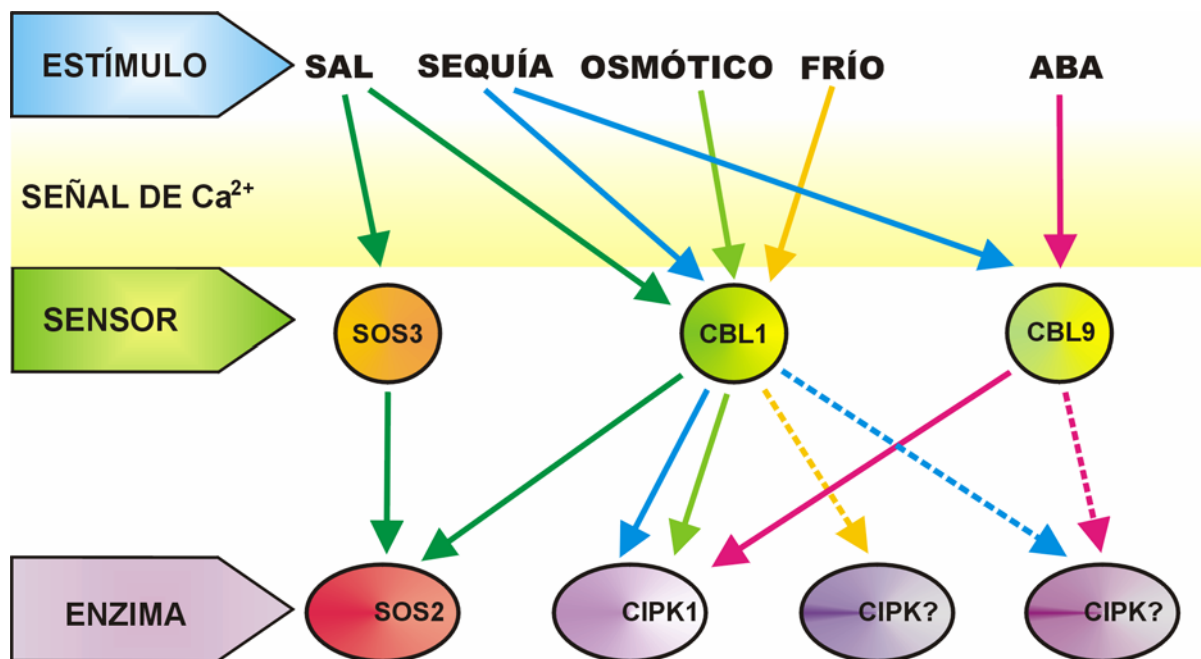


Figura 1.3.2: Rutas de señalización por estrés reguladas a través de la red de interacciones CBL/CIPK. Los diferentes colores de la flechas indican la especificidad de las respuestas a los distintos estreses. Las líneas discontinuas indican que la interacción con la quinasa no se ha demostrado.

- **La localización temporal y espacial de las CBLs y las CIPKs en el contexto celular y de la planta.** Algunas CBLs presentan sitios de miristoilación, como SOS3, y se han localizado solo en la membrana plasmática y otras, tanto en la membrana plasmática como en la vacuolar (Batistic & Kudla, 2004). Esta localización específica de las CBLs va a asegurar la necesaria separación espacial de procesos de señalización que responden a diferentes estímulos y que tienen que ser procesados al mismo tiempo. En contra de lo que ocurre con las CBLs, las CIPKs no presentan en su secuencia señales que determinen su localización (Kolukisaoglu *et al.*, 2004). Por tanto, su localización dependerá de la proteína CBL con la que interaccionen. Acerca de la localización de las parejas de proteínas en la planta, los experimentos de expresión realizados con algunas CBLs y CIPKs muestran patrones de expresión específicos según el tejido y el estado de desarrollo de la planta, de manera que se podrán esperar: (i) distintas funciones para un mismo complejo CIPK/CBL, según el órgano en donde se exprese, y (ii) una misma función para distintos complejos que actúen en el mismo tejido. Esto significa que la red de interacciones va a poder ser lo suficientemente dinámica para responder a los cambios ambientales en cualquier estadio de desarrollo de la planta (Batistic & Kudla, 2004).

- **La especificidad de las CIPKs por su sustrato.** Hasta ahora solo es conocido el sustrato de SOS2, el antiporteador Na^+/H^+ SOS1.

Hasta la fecha, únicamente se ha podido establecer la función fisiológica de unas pocas CBLs y CIPKs. El complejo mejor caracterizado es el formado por SOS3 y SOS2, que como ya se ha mencionado, regula la homeostasis del Na^+ durante el estrés salino en las raíces (Liu & Zhu, 1998; Halfter *et al.*, 2000; Liu *et al.*, 2000). También es conocido que CBL1 está implicada en la respuesta al frío, a la sequía, al estrés salino (en toda la planta, no como SOS3) y al estrés osmótico (Figura 1.3.2) (Albrecht *et al.*, 2003; Cheong *et al.*, 2003). Para la regulación del estrés salino interaccionaría con SOS2, y para la regulación del estrés osmótico y la sequía con CIPK1. Por otro lado, se ha visto que CBL1 y CBL9, sensores muy conservados, tienen funciones opuestas en la regulación de la sequía, activándose CBL1 ante la escasez de agua. Además, CBL9 tiene otra función dependiente de ABA de manera que regula la síntesis y señalización de la misma mediante la interacción con una quinasa que aún se desconoce (CIPK?) (Figura 1.3.2). CBL1 parece también interaccionar con esta quinasa desconocida, ya que su actividad genera la expresión de genes inducidos por ABA, de una manera análoga a como ocurre cuando se activa CBL9. Esto quiere decir que CBL1 y CBL9 tiene funciones solapantes y compiten por las mismas quinasas en ciertas rutas.

1.4.- OBJETIVOS

El exceso de sal en los suelos inhibe el crecimiento de las plantas y causa grandes pérdidas en la producción agrícola mundial. La ruta *Salt Overly Sensitive* (SOS) juega un papel determinante en la tolerancia salina en plantas y el mantenimiento de concentraciones bajas de Na^+ en el citosol, de ahí que el establecimiento de las bases estructurales de dicho proceso sea de gran importancia para entender la función molecular del sistema y para poder en un futuro darle una aplicación biotecnológica.

Como se ha descrito previamente, el sensor de Ca^{2+} SOS3 es una proteína clave en la ruta SOS, ya que se encarga de iniciar la respuesta para la tolerancia salina. SOS3 es capaz de sentir la señal citosólica del Ca^{2+} provocada por el estrés salino, para posteriormente activar a la quinasa SOS2. Por tanto, los objetivos fundamentales de este trabajo se centran en **entender cómo SOS3 une el Ca^{2+} y regula la actividad de SOS2**, para comprender en última instancia **cómo funciona específicamente mediando la señal proveniente del estrés salino** y no de otros tipos de estreses que pueden también provocar cambios en la concentración citosólica de Ca^{2+} .

Con estos objetivos se plantea abordar el estudio estructural de SOS3. Además, la caracterización estructural de SOS3 y su comparación con la proteína CBL2, la única estructura conocida de las CIPKs/CBLs, permite poner de manifiesto los determinantes estructurales asociados a la función de cada una de ellas y sirve para caracterizar en mayor medida esta familia de proteínas, de las que hasta ahora, poco se conoce de su función molecular. Asimismo, y desde un punto de vista metodológico, se pretende sistematizar y comprobar la utilidad del anión yoduro como faseador de estructuras cristalinas, empleando distintas estrategias de resolución cristalográfica.

Concretamente, los objetivos que se proponen en esta tesis pueden sintetizarse en:

- Determinar la estructura tridimensional de SOS3:
 - Expresión y purificación.
 - Cristalización.
 - Recogida y análisis de datos de difracción de rayos X.
 - Resolución estructural.
 - Análisis de la estructura.
 - Determinación del modo de unión a Ca^{2+} .
 - Comparación estructural con sus homólogos estructurales.

- Caracterizar la unión a Ca^{2+} y otros iones metálicos mediante técnicas en solución, tales como:
 - Dicroísmo circular.
 - Desnaturalizaciones registradas por dicroísmo circular.
 - Emisión de fluorescencia de triptófano.
 - Cromatografía de interacciones hidrofóbicas.

- Estudiar del estado de asociación:
 - Experimentos de equilibrio y velocidad de sedimentación por ultracentrifugación analítica.

- Proponer un mecanismo de respuesta al Ca^{2+} e interacción con la quinasa SOS2.

2. TÉCNICAS EXPERIMENTALES Y RESULTADOS

2. TÉCNICAS EXPERIMENTALES Y RESULTADOS

2.1. EXPRESIÓN Y PURIFICACIÓN

2.1.1. Introducción

Los cristales de proteína, como los de cualquier compuesto, están caracterizados microscópicamente por la agrupación de una población química y conformacionalmente única, en este caso de moléculas de proteína, según un modelo de repetición periódica. A priori no se conoce cómo va a cristalizar la macromolécula, por lo que como regla general, las características necesarias para abordar el estudio cristalográfico de una proteína son la cantidad, la pureza y la homogeneidad conformacional de la misma. Esto explica por qué en el pasado las primeras proteínas que se estudiaron por técnicas biofísicas y cristalográficas fueron aquellas que eran más abundantes, estables y fáciles de aislar.

El problema de la cantidad se ha resuelto con la aparición de métodos de ingeniería genética que hacen posible clonar y sobreexpresar genes en células bacterianas o eucariotas. La producción masiva de una proteína concreta puede llegar a ser de un cuarto del total expresado, pero a altas concentraciones intracelulares las proteínas pueden formar agregados, cuerpos de inclusión o desnaturalizarse. Por tanto, los niveles de producción deben ser optimizados para obtener la mayor proporción de la proteína deseada, funcionalmente activa y en su conformación natural. Además, con el uso de ciertos vectores es posible expresar proteínas de fusión con cola de histidinas o GST, con lo que se facilita la posterior detección y purificación de las proteínas diana. Para el estudio de complejos proteicos se ha elaborado una estrategia muy interesante basada en la coexpresión de múltiples genes diana en *E. coli*, habiéndose diseñado para ello vectores específicos que son capaces de coexpresar varios genes. Con esta estrategia se consiguen rendimientos óptimos de expresión de cada gen, se fomenta la actividad y solubilidad de las proteínas expresadas y se protege cada subunidad de posibles degradaciones.

La purificación de una proteína tiene como objetivo final obtener el material deseado con el máximo rendimiento, utilizando el proceso más corto, y a ser posible, más económico. Aunque suele considerarse simplemente un mero trámite en cristalografía, en este proceso a veces es posible extraer información acerca de propiedades de la macromolécula que pueden ser útiles para estudios ulteriores. Para desarrollar un buen protocolo de purificación es recomendable conocer previamente qué cantidad (del orden de miligramos) y pureza (al menos del 95%) se necesita, las propiedades físicas de la macromolécula, información sobre el sistema de expresión

de la misma y el intervalo de condiciones compatibles con la estabilidad de la muestra. Inicialmente se parte de un protocolo estándar, que ha de ser optimizado y posteriormente escalado a las cantidades deseadas. La purificación de muestras biológicas crudas se puede dividir en tres fases cromatográficas diferentes: captura, purificación intermedia y pulido (Figura 2.1.1).

- **La fase de captura** consiste en concentrar la muestra y en eliminar el grueso de contaminantes. Para ello se usan columnas cromatográficas con gran capacidad de carga y buenas propiedades de flujo.
- **La fase de purificación intermedia** puede abarcar varios pasos y consiste en separar los contaminantes más importantes, tales como otras proteínas y ácidos nucleicos, de la muestra parcialmente pura. En este punto es importante usar columnas de máxima resolución, ya que los contaminantes que quedan tienen propiedades cromatográficas similares a las de la proteína diana. Esto se lleva a cabo combinando técnicas de distinta selectividad que se usan en modo de alta resolución (uso de gradientes de elución, tamaño de bola más pequeño, etc.). El flujo y la capacidad de carga tienen que controlarse para que la resolución sea máxima.
- **La fase de pulido** sirve para alcanzar la pureza final de la muestra y para eliminar las trazas de contaminantes que incluso pueden ser variantes estructurales, tales como agregados, de la misma proteína. Con ninguna técnica cromatográfica se alcanzan rendimientos del 100%, por lo que el rendimiento final dependerá del número de pasos incluidos.

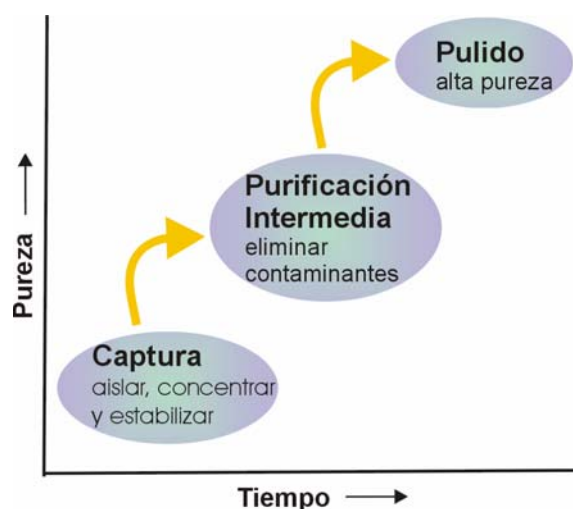


Figura 2.1.1: Secuencia de la estrategia de purificación en tres fases.

Una vez purificada, para abordar el trabajo cristalográfico es necesario concentrar la muestra y conseguir que ésta sea homogénea. Para medir la homogeneidad (en cuanto a tamaño) de una solución proteica se puede acudir a la técnica de ultracentrifugación analítica. Mediante un experimento de equilibrio de sedimentación, se puede extraer información de la dependencia de la masa molecular de la proteína con la concentración. Una proteína es homogénea cuando únicamente existe una especie molecular en solución, y por tanto, no hay ningún equilibrio de asociación. La monodispersidad de una solución proteica también es posible estudiarla mediante cromatografía de penetrabilidad o dispersión dinámica de luz. En el caso de que la solución presente distintos agregados moleculares, habrá que buscar los ligandos adecuados para estabilizar alguna de las especies moleculares, o bien separar los unos de los otros si no existe un equilibrio entre las diferentes especies.

2.1.2. Expresión y purificación de SOS3

2.1.2.1. Expresión de SOS3

El gen *SOS3* de *Arabidopsis thaliana* fue clonado en el plásmido de expresión pET14b (Novagen, Madison, WI, USA) para obtener pET-*His*-*SOS3*. El plásmido se transformó en células de *Escherichia coli* pertenecientes a la línea BL21(DE3) para la expresión de la proteína con cola de histidinas (Ishitani et al., 2000). Las cepas de *E.coli* con la construcción pET-*His*-*SOS3* se crecieron a 37 °C durante toda la noche, para así disponer de un cultivo primario, en medio LB (10 g bacto-tryptona, 5 g extracto de levadura y 5 g de NaCl por litro de solución) al que se le añadió ampicilina hasta una concentración final de 50 µg/ml. Alícuotas de 5 ml de este cultivo se subcultivaron en 500 ml de LB fresco con ampicilina y se dejaron crecer hasta una densidad óptica (600 nm) $DO_{600}=0.6-0.8$, situación en la que las células se encontraban en fase exponencial de crecimiento. La inducción de la expresión de la proteína recombinante se inició con la adición de IPTG hasta una concentración final en el medio de 0.3 mM. Después de tres horas más de incubación se recogieron las células inducidas mediante centrifugación (10 min, 8300 g).

La comprobación de que la inducción se había llevado a cabo con éxito se realizó mediante electroforesis, cargando muestras de antes y de después de inducir la expresión en un gel SDS-PAGE.

2.1.2.2. Purificación de SOS3

Cromatografía de afinidad a níquel

La proteína recombinante SOS3 se expresó con una cola de histidinas con el fin de purificarla mediante una cromatografía de afinidad a níquel, ya que los anillos de imidazol de las histidinas poseen una alta afinidad al catión Ni^{2+} que se encuentra covalentemente unido a la matriz de la columna. Este tipo de cromatografías de selectividad tan elevada facilita la purificación de las proteínas, pudiéndose conseguir altos grados de pureza. Con este propósito, se resuspendieron las células en tampón 1 (20 mM fosfato de sodio pH 7.4, 0.5 M NaCl, 5 mM imidazol, 0.05% (m/v) NaN_3) y se lisaron por sonicación. Después se centrifugó la muestra a 4°C (30 min, 34500 g) y se recogió el sobrenadante resultante, que se filtró (diámetro de poro 0.45 μm ; Millipore Corporation, Bedford, MA, USA) para posteriormente aplicarlo a una columna de níquel-sefarosa (volumen de 5 ml; Amersham Biosciences Limited, UK) que previamente estaba equilibrada en tampón 1. A continuación se lavó la columna con cinco volúmenes de tampón 1 para limpiarla de las proteínas que no se unían, seguido de un lavado con tampón 20 mM fosfato de sodio pH 7.4, 0.5 M NaCl, 30 mM imidazol, 0.05% (m/v) NaN_3 , para eluir las proteínas que se unían débil o inespecíficamente. Los pasos de la purificación se verificaron mediante una electroforesis en un gel SDS-PAGE (Figura 2.1.2(a)).

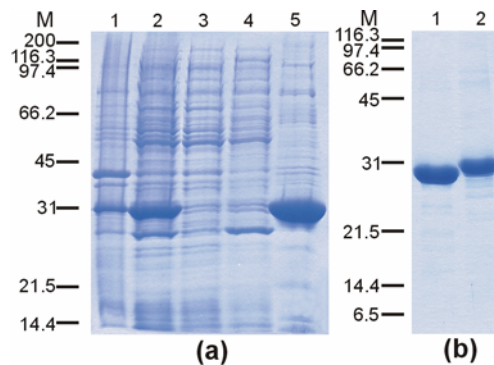


Figura 2.1.2: (a) Expresión y purificación de SOS3 por cromatografía de afinidad a níquel. Las proteínas se analizaron con un gel SDS-PAGE al 12%. Las líneas 1 y 2 muestran las fracciones precipitada y soluble después de la lisis celular. La línea 3 muestra la fracción proteica que no se une a la columna de afinidad. La línea 4, fracción de proteína eluida con tampón 30 mM imidazol. Línea 5, fracción eluida con 500 mM de imidazol. (b) Corte en columna de la proteína con cola de histidinas. El análisis se realizó con un gel SDS-PAGE al 15%. Línea 1, proteína cortada eluida tras la digestión. Línea 2, proteína no cortada eluida con tampón 500 mM imidazol. Las muestras se cargaron en condiciones reductoras. Los marcadores de peso molecular (M, en KDa; BioRad Laboratorios Inc., USA) aparecen indicados. Los geles se tificaron con azul de Coomassie.

Digestión con trombina

Antes de eluir His-SOS3 de la columna de afinidad se hizo un corte de la cola de histidinas en columna. El extremo que lleva las histidinas suele encontrarse bastante expuesto al solvente y consiste en una sucesión de seis cargas positivas y a veces genera problemas en la cristalización de la proteína porque puede ocasionar desorden y repulsiones intermoleculares. Con este fin se reequilibró la columna con tampón 1 y se aplicó la proteasa trombina (Amersham Biosciences Limited, UK) a la columna (7.5 unidades de trombina por mg de His-SOS3). La digestión se incubó a 20°C durante una noche. La proteína SOS3 cortada se eluyó con tampón 1 y se comprobó la eficiencia del corte por electroforesis en gel SDS-PAGE (Figura 2.1.2(b)) y mediante espectrometría de MALDI-TOF. En la Figura 2.1.3 se observa cómo la diferencia de masas entre His-SOS3 (27876.2 Da) Da y SOS3 (25973.65 Da) era de 1902.55 Da, valor que coincidía con la masa de la cola de histidinas esperada 1882.1 Da, si se considera el error del aparato en la medida de las masas.

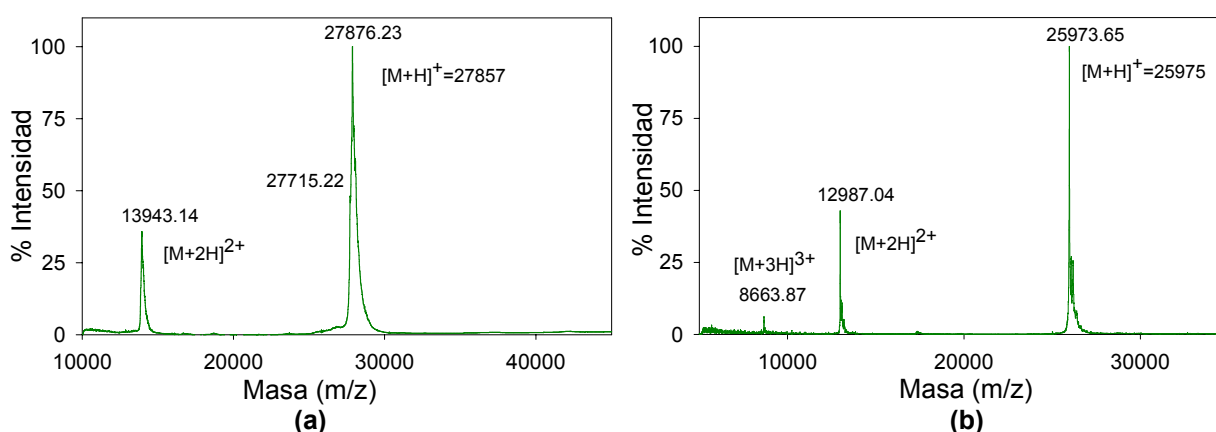


Figura 2.1.3: Espectro MALDI-TOF de His-SOS3 (a) y SOS3 (b).

Intercambio aniónico

La pureza de SOS3 fue mejorada mediante una cromatografía de intercambio iónico con una columna Resource Q (volumen 6 ml; Amersham Biosciences Limited, UK). Primero se cambió el tampón en el que se encontraba SOS3 con una columna PD-10 (Amersham Biosciences Limited, UK) para dejarla en tampón 2 (20 mM Tris-HCl pH 7.5, 0.05 M NaCl, 0.05% NaN₃). Una vez preequilibrada la columna, cargada la muestra y lavada con cinco volúmenes de tampón 2 se aplicó un gradiente de 0.05 M a 0.5 M de NaCl. La proteína eluyó a 0.1 M NaCl (Figura 2.1.4). El

segundo pico que se recogió (0.5 M NaCl) también correspondía a SOS3, pero debía de ser una variante estructural menos estable, quizá mal plegada, que precipitaba al intentar concentrarla.

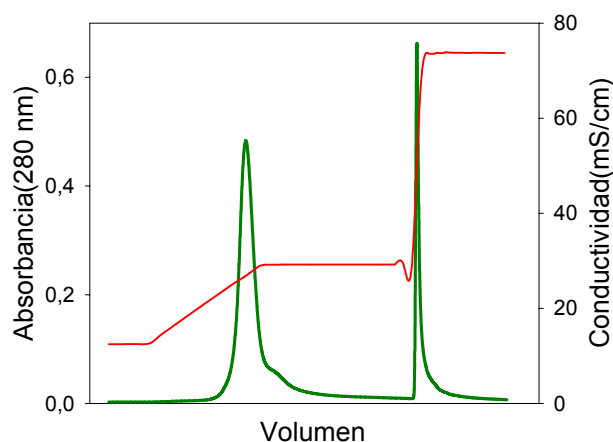


Figura 2.1.4: Cromatografía de intercambio aniónico de SOS3. La línea roja representa la conductividad de la solución. La línea verde representa la absorbancia registrada a 280 nm.

Finalmente, la muestra de SOS3 pura se dejó en tampón 2 con la ayuda de una columna PD-10 (Amersham Biosciences Limited, UK) y se concentró a 20 mg/ml con un concentrador de proteínas Amicon YM-10 (Millipore Corporation, Bedford, MA, USA). La concentración final de SOS3 se determinó espectrofotométricamente usando un coeficiente de absorción molar de $10930 \text{ M}^{-1}\text{cm}^{-1}$ a 280 nm (<http://us.expasy.org>). La pureza final de la muestra fue examinada mediante electroforesis en gel SDS-PAGE y por espectrometría de masas (Figura 2.1.3), donde se obtuvo una masa, (M+H) = 25975 Da, que era coincidente con el valor esperado, 25974.7 Da.

Para corroborar la homogeneidad de la muestra se realizó un experimento de equilibrio de sedimentación en función de la concentración de SOS3, haciendo uso de una ultracentrífuga analítica Beckman Optima XL-A de seis canales (paso óptico de 12 mm) y un rotor AnTi50. Los resultados mostraron que SOS3 se encuentra en equilibrio entre la forma monomérica y dimérica (Apartado 2.4.2). La heterogeneidad que presentaba la muestra se contrastó mediante la realización de una cromatografía de penetrabilidad con una columna Superdex 200 (Amersham Biosciences Limited, UK). Una vez equilibrada la columna con dos volúmenes de tampón 2, se cargaron 3 ml de muestra a $1.92 \cdot 10^{-4} \text{ M}$. Durante la cromatografía se recogieron dos picos. A pesar de que éstos salieron a unos tiempos teóricos correspondientes a proteínas de mayor masa molecular, se asumió que el primero y mayoritario correspondía al dímero y el segundo, que era aproximadamente un 5% del primero, al monómero. Las fracciones recogidas se examinaron mediante electroforesis en gel SDS-PAGE para corroborar que los picos recogidos correspondían a SOS3. Este tipo de cromatografía es muy sensible a factores como la forma de la

macromolécula, por lo que el desajuste entre el tiempo de elución y la masa molecular de la proteína podría deberse a que la forma del monómero y del dímero de SOS3 no es globular.

2.2. CRISTALIZACIÓN

2.2.1.- Introducción

La determinación de la estructura tridimensional de cualquier molécula por difracción de rayos X depende, en primer lugar, del éxito en la obtención de monocristales adecuados para someterlos a difracción.

Un cristal se puede considerar como una red imaginaria de unidades elementales idénticas, denominadas cada una de ellas celdilla unidad. La repetición por translación en las tres direcciones del espacio de este elemento estructural genera el volumen del cristal. Si el cristal está formado por celdillas unidad, la celdilla a su vez presenta en su interior una unidad mínima, denominada unidad asimétrica. La aplicación de ciertas operaciones de simetría a la unidad asimétrica genera la celdilla unidad, que está definida por tres ejes (a, b, c) y tres ángulos interaxiales (α, β, γ) (Figura 2.2.1) que se usan como sistema de referencia para las coordenadas atómicas.

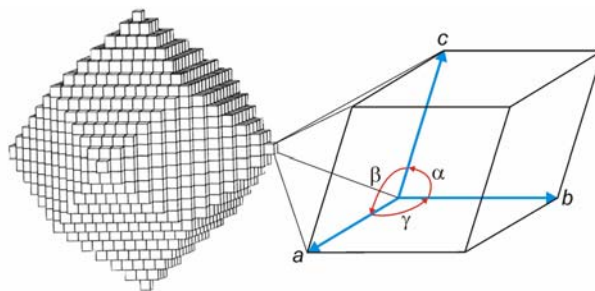


Figura 2.2.1: Celdilla unidad (derecha). El apilamiento de celdillas forma el cristal (izquierda). Imagen cedida por los autores, Dr. Martín Martínez-Ripoll y Dr. Juan A. Hermoso y publicada en el libro “Estructura de Proteínas” (Gómez-Moreno *et al.*, 2003).

Los cristales de proteínas difieren mucho de los cristales convencionales de moléculas pequeñas (como los compuestos orgánicos en general) o sales inorgánicas (como la sal común). Las características de los primeros están condicionadas por la naturaleza propia de estas macromoléculas. Así, debido al tamaño de las proteínas y a su esfera de solvatación, el volumen de la celdilla unidad suele ser bastante grande. Además, como las macromoléculas se

empaquetan con baja densidad, al ordenarse en el cristal dejan grandes canales constituidos por moléculas de agua desordenadas y el contenido en solvente suele ser elevado (50-70%) (Mathews, 1968). Por todo ello, las interacciones entre moléculas son muy pequeñas, de manera que los cristales son muy frágiles y, en ocasiones, suelen presentar un cierto grado de desorden.

2.2.1.1. Principios básicos de la cristalización

Para producir cristales de cualquier compuesto, las moléculas en solución, en presencia de agentes precipitantes suaves, son conducidas a un estado de sobresaturación, estado termodinámicamente inestable, que posteriormente deriva en la aparición de una fase cristalina, de manera que las moléculas son llevadas a un equilibrio entre las fases sólida y líquida (Figura 2.2.2) (Ducruix & Giegé, 1992).

El proceso de cristalización se puede dividir en dos fases: la fase de nucleación (formación de los primeros agregados ordenados) y la fase de crecimiento del cristal (en donde el conjunto de interacciones atractivas entre las moléculas son máximas, y las repulsivas mínimas).

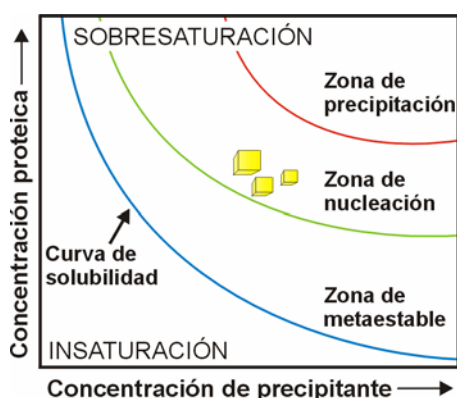


Figura 2.2.2: Descripción esquemática de un diagrama bidimensional de solubilidad, en el que se ilustra el cambio de concentración de proteína con la concentración de agente precipitante (Ducruix & Giegé, 1992).

Durante este proceso hay que superar una barrera energética, como ocurre en cualquier reacción química (Figura 2.2.3). En el caso de las proteínas, la superación excesiva del punto de saturación, sobre todo si sucede rápidamente, conlleva con gran frecuencia la precipitación de proteína en forma de agregados amorfos, no ordenados, que no son adecuados para los estudios de difracción. Por el contrario, si la sobresaturación es leve se alcanza un estado metaestable que no da lugar a núcleos cristalinos (a menos que se le aporte energía al sistema) pero que sería la zona ideal para el crecimiento de cristales sin nucleación de otros nuevos (Figura 2.2.2). Por tanto, la formación de cristales requiere alcanzar lentamente un grado limitado de sobresaturación, para que se formen pocos núcleos, de manera que la masa proteica se invierta en el crecimiento de los cristales y así éstos adquieran un buen tamaño.

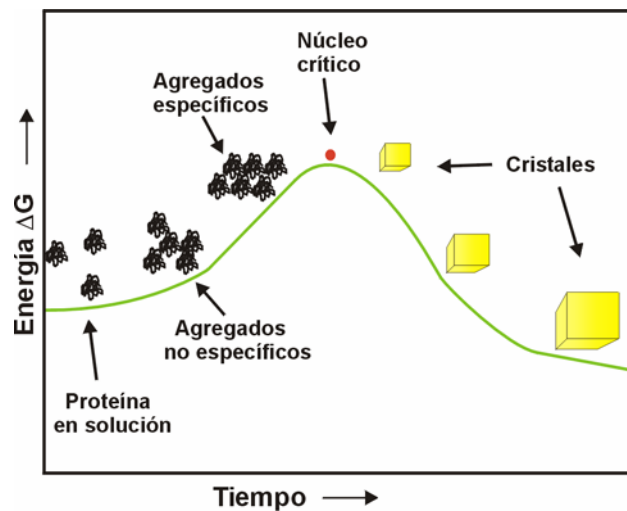


Figura 2.2.3: Diagrama de energía de la cristalización. Imagen extraída de la página web <http://www-structmed.cimr.cam.ac.uk/Course/Crystals>.

2.2.1.2. Preparación del experimento de cristalización

Antes de comenzar a cristalizar la proteína es conveniente controlar ciertos aspectos:

- ¿Está la proteína pura?

Comprobar por electroforesis y espectrometría de masas.

- ¿Está bien plegada?

Comprobar la actividad de la proteína, si es el caso, y hacer un espectro de dicroísmo circular para ver si presenta estructura secundaria.

- ¿Hace tiempo que se purificó?

Con el tiempo, las proteínas se degradan y generan heterogeneidad.

- ¿Es monodispersa?

Corroborar mediante cromatografía de penetrabilidad, experimentos de equilibrio de sedimentación y ensayos de dispersión dinámica de luz.

- ¿Necesita estar la proteína en condiciones reductoras?

Si tiene cisteínas, usar β -mercaptoetanol o DTT en la solución proteica.

- ¿Es la proteína estable a temperatura ambiente?
- ¿Se ha cristalizado alguna vez alguna proteína similar?
- ¿Necesita la proteína la adición de algún compuesto para permanecer estable en solución?

En ocasiones las proteínas unen ligandos que estabilizan ciertas regiones móviles de la proteína, que pueden favorecer las interacciones intramoleculares reduciendo la flexibilidad, cambiar las propiedades de superficie de la proteína e incluso causar cambios conformacionales. En estos

casos, antes de comenzar con la cristalización sería aconsejable adicionar dicho ligando o agente estabilizante a la solución proteica.

- ¿Se dispone de diferentes construcciones de la proteína?

A veces, las proteínas están formadas por distintos dominios, o poseen extremos móviles que confieren flexibilidad, características que no favorecen la cristalización. En estos casos, es aconsejable hacer uso de las técnicas de biología molecular para producir los dominios por separado o la proteína sin los extremos móviles, eliminando así el carácter flexible de la macromolécula.

- ¿Es posible trabajar con una proteína homóloga de otro organismo?

A veces, una simple mutación puntual es suficiente para que una proteína cristalice. Trabajar con especies diferentes es la manera más fácil de disponer de una colección de mutaciones puntuales que no afecten a la funcionalidad, pero que confieran mayor estabilidad a la proteína. Los organismos termófilos son un buen ejemplo, ya que sus proteínas tienen la característica general de ser muy estables.

2.2.1.3. Proceso experimental de cristalización

La cristalización es un proceso que depende de muchos parámetros químico-físicos que aún hoy en día no se han llegado a racionalizar y controlar totalmente, con lo que se puede considerar que este proceso es fundamentalmente empírico.

La estrategia que se sigue para alcanzar la sobresaturación consiste en empezar con una concentración relativamente elevada de proteína (alrededor de 10 mg/ml), y aumentar la concentración por eliminación gradual de agua, introduciendo además ciertas variables que afecten a la solubilidad de la proteína, como son:

- Naturaleza y concentración de diferentes agentes precipitantes, como por ejemplo polímeros (los más utilizados son los polietilenglicoles), sales, compuestos orgánicos volátiles y no volátiles,... etc.
- pH.
- Temperatura.
- Naturaleza y concentración de aditivos, como por ejemplo sales, ligandos comunes de proteínas,... etc.

Esta búsqueda de las condiciones de cristalización se desarrolla en dos pasos, primero se realiza una búsqueda de condiciones, seguida, en segundo lugar, de una optimización sistemática de las condiciones iniciales.

- **Búsqueda de condiciones iniciales:** Debido a que el número de factores que controlar es excesivo, es imposible realizar una búsqueda sistemática en la que se exploren todos los factores posibles. Por este motivo se han diseñado estrategias de muestreo, que utilizan matrices que exploran el espacio de cristalización. La matriz más utilizada actualmente fue desarrollada por Jancarín y Kim (1991) y consiste en el rastreo de condiciones que han dado en el pasado resultados favorables. Sin embargo también existen programas que generan matrices que exploran al azar el espacio de cristalización, basándose únicamente en incompatibilidades de reactivos.
- **Optimización:** En primer lugar se hace un rastreo alrededor de las condiciones iniciales, variando la concentración de precipitante, de proteína, la temperatura o el método de cristalización. Una vez refinadas estas condiciones, se prueban ciertos aditivos, como compuestos orgánicos, cationes, aniones, detergentes,... etc. Además, también se recurre a las técnicas de macro y microsembrado y al crecimiento de cristales en geles de agarosa.

2.2.1.4. Métodos de Cristalización

Existen muchos métodos de cristalización: la difusión de vapor, la diálisis, la contradifusión a través de interfase o el denominado *microbatch*.

La difusión de vapor es la técnica más ampliamente utilizada. Una pequeña gota que contiene la solución de proteína y la de cristalización, se equilibra contra un reservorio que contiene la solución de cristalización a una concentración mayor que en la gota. El equilibrio se consigue por difusión de las especies volátiles (agua o solvente orgánico), hasta que la presión de vapor en la gota se iguala con la del reservorio. Aunque existen diversas modalidades de esta técnica, la más utilizada es la de gota colgante ([Figura 2.2.4\(a\)](#)).

El método de contradifusión se utiliza para la optimización de cristales a partir de condiciones de cristalización conocidas. Como se muestra en la [Figura 2.2.4\(b\)](#), la solución de proteína, en contacto con la solución de cristalización a través de una interfase líquido-líquido, se equilibra por difusión de la segunda, que es más densa, a través de la primera. El experimento se realiza en un capilar de pequeño diámetro (0.3 mm), con lo que se minimiza el transporte de masa por convección, y solo se produce transporte por difusión, con lo que, en principio, existe un aporte de proteína al cristal más regular, que debe redundar en un crecimiento cristalino más ordenado y en cristales con menores problemas mecánicos. Los cristales aparecen a lo largo del capilar debido al gradiente de sobresaturación que se genera, obteniéndose al principio precipitados amorfos y, a lo largo del capilar, cristales creciendo separados unos de otros (Gavira *et al*, 2002). Una ventaja

adicional de esta técnica es que, en general, no es necesario manipular los cristales para recoger datos de difracción.

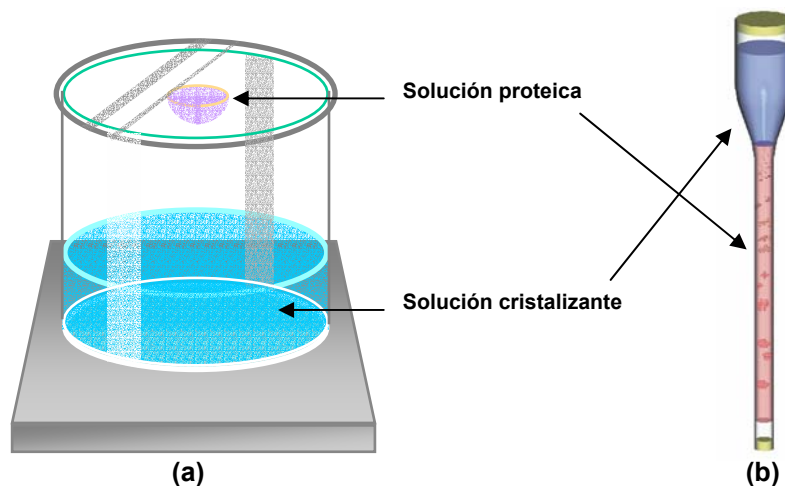


Figura 2.2.4: Esquema de los montajes experimentales de cristalización mediante (a) la técnica de difusión de vapor en gota colgante y (b) la de contradifusión en capilar (Gavira *et al.*, 2002).

2.2.2. Cristalización de SOS3 y sus complejos

Las condiciones preliminares de cristalización se establecieron con el método de difusión de vapor en gota colgante a 293 K (método y temperatura que se utilizaron en los experimentos que se enumerarán, a no ser que se indique lo contrario) y usando las soluciones de cristalización “Crystal Screen I”, “Crystal Screen II”, “Crystal Screen Lite” e “Index” (Hampton Research, CA, USA), que están basadas en la matriz de rastreo de condiciones ideada por Jancarin y Kim (1991). Así, se mezclaron 1 μl de SOS3 pura con 1 μl de solución de cristalización y las gotas se pusieron en equilibrio con 500 μl de solución de cristalización.

Se obtuvieron microcristales en forma de aguja con la condición 40 del “Crystal Screen Lite” (Hampton Research, CA, USA) (Tabla 2.2.1.a). Sin embargo, estos cristales, que aparecían en 3 ó 4 días no evolucionaban y eran demasiado pequeños para someterlos a difracción.

	Muestra	Precipitante	Aditivo	*Relación
a	SOS3	10% (v/v) isopropanol, 0.1 M tampón citrato pH 5.6,	-	1:1
b	SOS3-Mn	20% (m/v) PEG 4K, 30% (v/v) MPD, 0.1 M tampón citrato pH 4.9	-	2:1
c	SOS3	20% (m/v) PEG 4K, 30% (v/v) MPD, 0.1 M tampón citrato pH 5.2	sacarosa 30% (m/v) y etanol 30% (v/v)	2:1:0.75:0.75
d	SOS3-Ca	24% (w/v) MPD, 18% (w/v) PEG 4K, 0.1 M tampón citrato sódico pH 4.8	Nal 1 M	2:1:0.75
e	SOS3-Mn-Ca	24% (w/v) MPD, 16% (w/v) PEG 4K, 0.1 M tampón citrato sódico pH 4.8	Nal 1 M	2:1:0.75

* Relación indica los microlitros de proteína, precipitante y aditivos que se pusieron en la gota.

Tabla 2.2.1: Condiciones de cristalización de SOS3 con el método de gota colgante.

Con el fin de controlar mejor el sistema de cristalización y optimizar los cristales se sustituyó el isopropanol por MPD, precipitante orgánico similar pero no volátil, y se adicionó PEG 4K para estudiar el efecto de la inclusión de un precipitante de cadena larga en las condiciones de cristalización (Huang *et al*, 1999) (Tabla 2.2.1.b). Además, a la solución proteica (10 mg/ml) se le adicionó $MnCl_2$ hasta una concentración de 3.85 mM (relación molar proteína: Mn^{2+} 1:10), ya que se había visto previamente, por experimentos de desnaturalización térmica registrada por dicroísmo circular, que el Mn^{2+} estabilizaba a SOS3 (Apartado 2.4.3.2). De esta manera, al cabo de una semana, se obtuvieron agrupaciones de cristales en forma de aguja y de placa (Figura 2.2.5(a)). Sin embargo, estos cristales tampoco eran aptos para recoger datos de difracción.

El uso de colecciones de pequeñas moléculas o aditivos, que se añaden a las condiciones de cristalización, es una estrategia adecuada para la mejora de las condiciones de cristalización. Así, se probó la colección de aditivos de la casa comercial Hampton Research (CA, USA) alrededor de las condiciones en que se obtuvieron las placas. Con la utilización de sacarosa como aditivo y SOS3 (10.7 mg/ml) se consiguió otro tipo de cristales en forma de barra (Tabla 2.2.1.c). La calidad de estos cristales mejoró al adicionar etanol a la gota (Figura 2.2.5(b)). Sin embargo, no se consiguieron datos de difracción a una resolución mayor de 3.2 Å.

En ocasiones, debido a la fragilidad de los cristales de proteína, su manipulación, o la propia tensión superficial de la gota, producen una destrucción interna del orden cristalino y por tanto de la difracción. En estos casos, el uso del método de cristalización de contradifusión en capilar es adecuado. Para ello, se mezclaron 4 µl de SOS3 (15 mg/ml) con 1.2 µl de sacarosa (30% (m/v) y

1.2 μl de etanol (30% (v/v)). El volumen de solución madre utilizado duplicaba al de la mezcla proteica. Se dejó el capilar en posición horizontal y al cabo de dos semanas se obtuvieron cristales que se estudiaron por difracción sin sacarlos del capilar. Se alcanzaba una resolución similar a la obtenida con los cristales crecidos por difusión de vapor en gota colgante, por tanto, la mala calidad de los mismos no se debía a su manipulación, por lo que se asumió la existencia de un gran desorden interno.

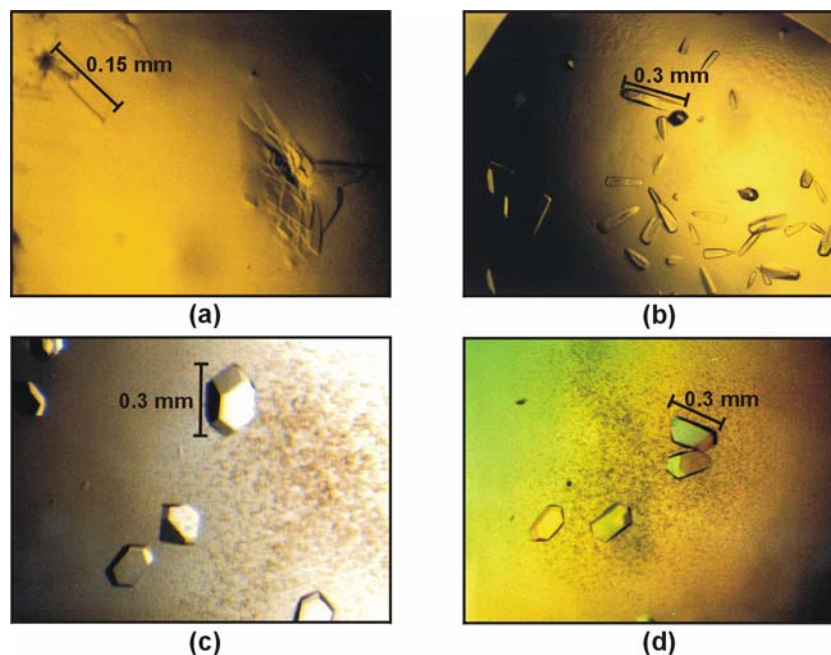


Figura 2.2.5: Imágenes de cristales de SOS3 en diferentes condiciones de cristalización, indicando el tamaño de los mismos.

Finalmente se consiguieron cristales de SOS3 en complejo con Ca^{2+} (Figura 2.2.5(c)), y con Ca^{2+} y Mn^{2+} (Figura 2.2.5(d)). Con la idea de partir de una muestra lo más homogénea química y conformacionalmente, y para obtener mejores cristales, se sometió la muestra a una incubación durante 1 h con una resina quelante (Hampton Research, CA, USA), que une preferentemente iones metálicos divalentes (Dunn *et al.*, 1980). A continuación se añadió controladamente CaCl_2 en una proporción SOS3(10 mg/ml): CaCl_2 1:10 y por otro lado CaCl_2 y MnCl_2 en una relación 1:10:10. Con esta estrategia los cristales mejoraron, y más aún con la adición de yoduro como aditivo en las condiciones de cristalización (Tabla 2.2.1.d y e) y, como se verá posteriormente, la calidad de los mismos fue suficiente para permitir la resolución estructural.

2.3. TÉCNICAS CRISTALOGRÁFICAS Y RESOLUCIÓN ESTRUCTURAL

2.3.1. Conceptos generales

La biología molecular tiene como fin último entender los procesos biológicos en términos de la física y la química de las macromoléculas que participan en ellos. Una de las diferencias esenciales entre la química de los sistemas vivos y los inertes es la gran complejidad estructural de las macromoléculas biológicas. Por tanto, la química de la vida no se podrá desvelar sin antes conocer a resolución atómica, o cercana a la atómica, la estructura de las macromoléculas biológicas, especialmente las proteínas (Branden & Tooze, 1991), que son las verdaderas máquinas de la vida. Así, entender la estructura de cada proteína es entender en qué consiste la vida (Max Perutz, Premio Nobel de Química).

La Cristalografía de rayos X es una de las técnicas físicas más potentes para la obtención de información estructural de la materia ordenada. El estudio estructural realizado proporciona una fotografía tridimensional, a escala atómica, del material cristalizado y, aplicado a las proteínas, proporciona información de la estructura tridimensional, plegamiento y también detalles de un gran significado físico-químico, tales como distancias, ángulos de enlace, estado térmico vibracional de los átomos o empaquetamiento cristalino, información clave para entender la función biológica.

La técnica alternativa para la determinación estructural de moléculas a alta resolución es la Resonancia Magnética Nuclear (RMN). Sin embargo, esta técnica está limitada por el tamaño molecular. La Cristalografía de rayos X, por el contrario, no presenta dicha limitación y se aplica al estudio de la estructura tridimensional de proteínas, virus, ácidos nucleicos, complejos nucleoproteicos y complejos proteína-proteína. Por otro lado, la microscopía electrónica se ha convertido recientemente en una técnica complementaria para grandes agregados moleculares, aunque con un grado de resolución mucho menor.

Concepto de difracción

La difracción de rayos X es el fenómeno físico fundamental a través del cual se manifiesta la interacción de los rayos X con la materia cristalina. Los rayos X, al incidir sobre un cristal, interaccionan con los electrones de los átomos que componen el cristal, haciéndolos vibrar acopladamente con las variaciones periódicas de su campo eléctrico. Los electrones oscilantes se convierten, de esta manera, en focos de nueva radiación X que se emite en cualquier dirección, fenómeno que se denomina “dispersión”. Estos rayos X dispersados por los electrones interfieren entre sí, pudiéndose dar interferencias destructivas. Sin embargo, como los átomos en un cristal

están ordenados de un modo regular y periódico, también ocurre que en determinadas direcciones, las ondas interfieren constructivamente, reforzándose, dando lugar al fenómeno de dispersión cooperativa o difracción.

El modo en que la Cristalografía de rayos X llega al conocimiento de la estructura cristalina de cualquier material puede comprenderse mejor si se compara con el símil que representa la observación de un objeto a través de un microscopio óptico. En éste, la radiación dispersada por el objeto en estudio se recombina de nuevo, a través de un sistema de lentes, para dar lugar a la imagen aumentada del objeto dispersor. Sin embargo, no existe ninguna lente que sea capaz de focalizar los rayos X dispersados por los átomos. Para ello, las técnicas cristalográficas hacen uso de cálculos matemáticos denominados “síntesis de Fourier”, con los que se consigue “focalizar” los rayos X.

Toma de datos de difracción

En los experimentos de difracción se hace incidir un haz monocromático de rayos X (con una longitud de onda alrededor de 1 \AA) sobre un monocristal de la muestra en estudio.

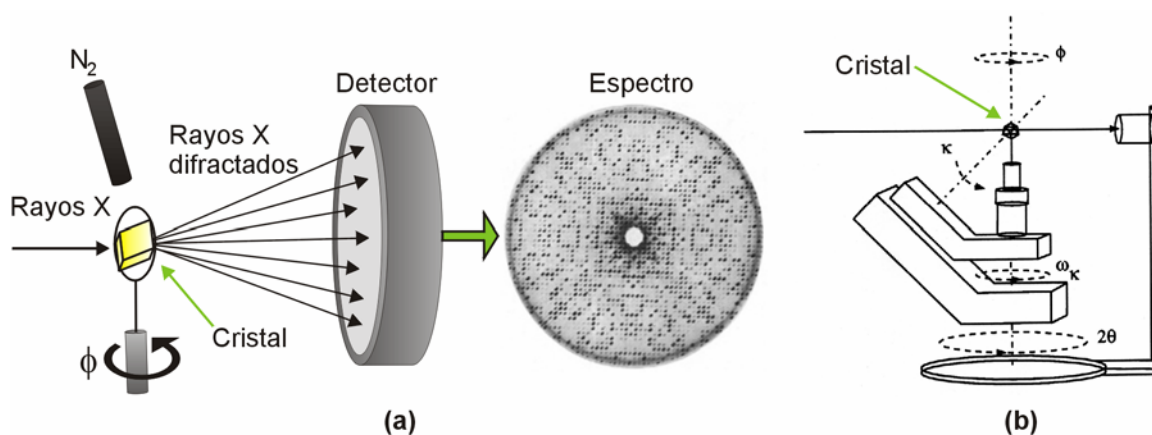


Figura 2.3.1: (a) Esquema de un experimento de difracción con geometría Phi. (b) Esquema de un difractómetro con geometría Kappa.

El espectro de difracción de rayos X contiene un patrón de intensidades correspondientes a los haces difractados por el cristal y la recogida de datos de difracción consiste en la medida de dichas intensidades y su disposición espacial respecto al sistema de referencia del laboratorio. El diseño experimental implica la toma de imágenes sucesivas, en un detector plano situado en la perpendicular del haz de rayos X, para distintas posiciones angulares del cristal, hasta obtener un

conjunto de datos completo. Esto se consigue mediante oscilaciones simples del cristal (geometría Phi) o bien mediante una combinación de giros más complejos (geometría Kappa) (Figura 2.3.1).

La estrategia de recogida de los datos de difracción y la elección de la fuente de radiación dependerá del tipo de datos de difracción que sean necesarios para poder abordar la resolución estructural, pudiéndose realizar en un difractor con fuente convencional o en una estación de radiación sincrotrón:

- Las características técnicas de un difractor con fuente convencional le confieren una gran versatilidad y un bajo coste a la hora de diseñar y realizar una estrategia de recogida de datos (Figura 2.3.2(a)).
- Tanto en una fuente convencional, como de sincrotrón, es posible la implantación de la geometría Kappa, clave para poder orientar el cristal y diseñar estrategias para obtener conjuntos de datos de difracción completos y/o muy redundantes (Figura 2.3.1(b)).
- Un sincrotrón es un acelerador de partículas en donde se genera radiación de diferentes longitudes de onda, entre ellas rayos X (2.3.2(b)). La radiación X que se produce es muy “brillante”, es decir muy intensa y concentrada, característica muy importante para el caso de cristales de proteína, donde el fenómeno de difracción es débil. Otra ventaja del sincrotrón frente a las fuentes convencionales es que la radiación producida es sintonizable. Esto permite recoger datos a diferentes longitudes de onda que, como se verá posteriormente, puede ser muy útil a la hora de resolver las estructuras.

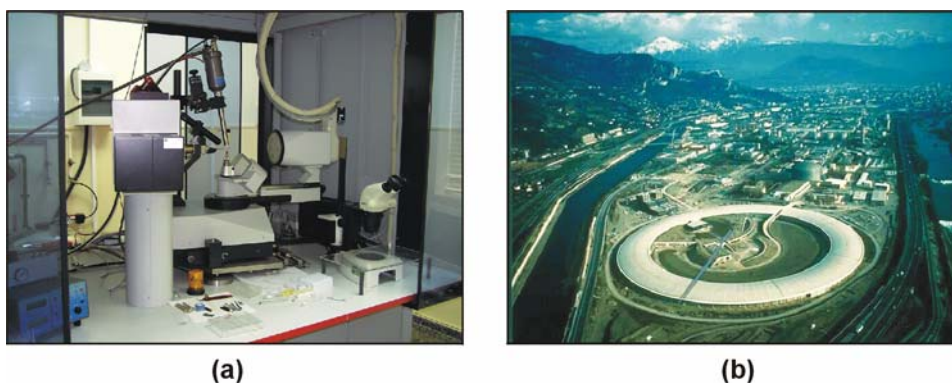


Figura 2.3.2: (a) Imagen un difractor con fuente convencional. (b) Imagen de una estación de radiación sincrotrón.

Los cristales de proteína, debido a las características especiales que presentan, sufren un gran deterioro al ser expuestos a la radiación. Así, la longitud de onda penetrante y la intensidad del haz, unidas a prolongados tiempos de exposición, provocan fenómenos de absorción que hacen que la muestra se caliente, generándose radicales libres en las moléculas de proteína y solvente, que, con el tiempo, dan lugar a daños irreparables en el cristal que redundan en la pérdida de orden interno. Por este motivo es conveniente enfriar el cristal durante el proceso de difracción

(Garman & Schneider, 1997), mediante un flujo laminar de nitrógeno frío (120-100 K) (Figura 2.3.1(a)). Para hacer uso de la criocristalografía es necesario previamente bañar el cristal en una solución crioprotectora (glicerol, PEGs, etilenglicol, sacarosa, etc.) (McFerrin & Snell, 2002) para que las moléculas de agua no estructurales se reemplacen por las crioprotectoras, impidiendo así la formación de hielo y la consiguiente ruptura del cristal.

Resolución estructural

Como se ha mencionado anteriormente, un espectro de difracción de rayos X consiste en un patrón de intensidades. La información estructural está concentrada en dicho patrón de difracción (Figura 2.3.3) y en forma de los denominados factores de estructura, asociados a cada intensidad, cuya fase se pierde en el experimento de difracción.

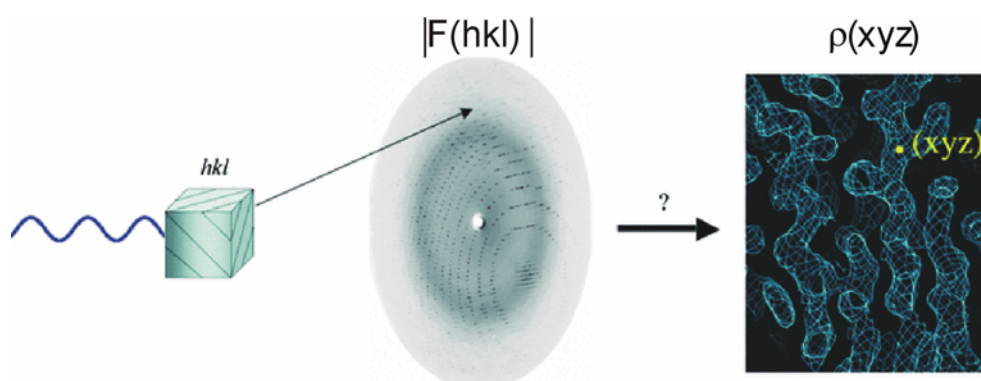


Figura 2.3.3: Proceso de resolución estructural.

Haciendo uso de la teoría de interferencias de ondas es fácil demostrar que la geometría del espectro da cuenta de la celdilla unidad y la simetría de las intensidades informa de la simetría cristalina. Además, se puede establecer una relación directa entre la estructura cristalina y el experimento, mediante la transformada de Fourier (Figura 2.3.4).

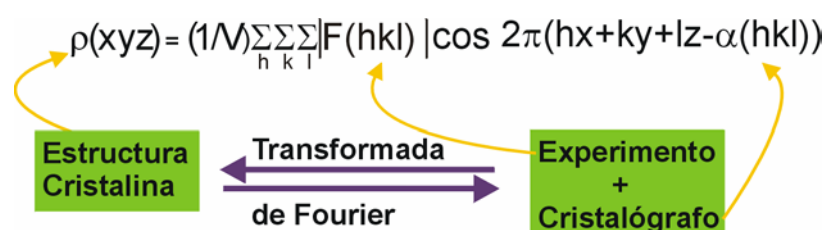


Figura 2.3.4: Función de densidad electrónica. V es el volumen de la celdilla unidad y $\alpha(hkl)$ es la fase asociada al módulo del factor de estructura, $|F(hkl)|$. Es interesante resaltar el hecho de que la densidad electrónica en un punto viene determinada por la información contenida en todos los factores de estructura, es decir, por la información de todo el espectro. Inversamente, el valor de un factor de estructura (módulo y fase) viene determinado por las posiciones de todos los átomos que forman la estructura, $\rho(xyz)$.

Esta expresión (Figura 2.3.4) relaciona la densidad electrónica, $\rho(xyz)$, en cada punto (xyz) del cristal, es decir la estructura, y el experimento de difracción, donde $|F(hkl)|$ es la raíz cuadrada de la intensidad de cada punto (hkl) del espectro de difracción. Para poder resolver la ecuación es necesario disponer no solo de los módulos, $|F(hkl)|$, sino también de las fases, $\alpha(hkl)$, asociadas a los factores de estructura, pero cuya información se pierde durante el experimento. Esto constituye el denominado “problema de la fase” y el trabajo del cristalógrafo consiste en averiguar el valor de $\alpha(hkl)$. Mediante diferentes metodologías que se expondrán a continuación, es posible obtener las fases, pudiéndose calcular el mapa de densidad electrónica y, por tanto, la estructura.

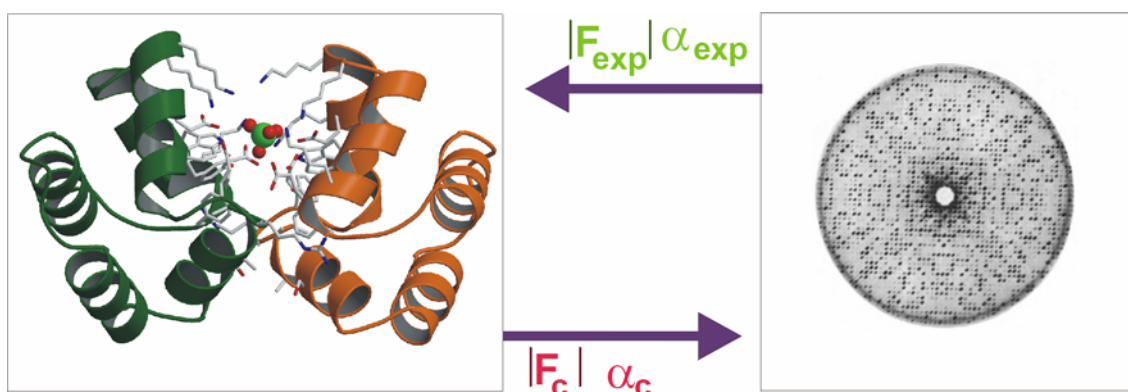


Figura 2.3.5: Relación entre la estructura y un espectro de difracción.

Metodologías para recuperar las fases

Conocida la amplitud ($|F_{exp}(hkl)|$) no es posible derivar, de manera directa, el valor de la fase ($\alpha_{exp}(hkl)$), únicamente conociendo la estructura molecular o la densidad electrónica es posible saber los valores de las fases ($\alpha_c(hkl)$). Por tanto, es necesario disponer previamente de algún conocimiento parcial de la densidad electrónica o de la estructura para calcular unas fases iniciales, $\alpha_c(hkl)$, que junto con las amplitudes experimentales, $|F_{exp}(hkl)|$, permitan calcular una función de densidad electrónica aproximada (Figura 2.3.5). Este es el fundamento de los denominados “métodos de faseado” (Tabla 2.3.1). Los programas que se suelen utilizar para resolver el problema de la fase se muestran en la Tabla 2.3.2.

Método	Conocimiento previo
Reemplazo molecular	Modelo homólogo
Reemplazo isomorfo	Subestructura de átomos pesados
Dispersión anómala	Subestructura de átomos dispersores anómalos

Tabla 2.3.1: Métodos de faseado.

El problema de la fase para cristales formados por moléculas de tamaño pequeño y medio ha sido resuelto mediante los llamados *métodos directos* y *métodos de Patterson*. Los primeros están basados en cálculos estadísticos y fueron desarrollados por varios autores a lo largo del siglo XX, entre los que cabe mencionar a Jerome Karle y Herbert Hauptmann, que compartieron el Premio Nobel de Química en 1985. Los segundos son métodos vectoriales basados en relaciones interatómicas, y para cuyo cálculo se omiten las fases de los factores de estructura.

Programa	Utilidad	Método
MOLREP (Vagin & Teplyakov, 1997)	Reemplazo molecular	Patterson
AMoRe (Navaza, 1994)	Reemplazo molecular	Patterson
Shake-and-Bake (Miller <i>et al</i> , 1994)	Búsqueda de átomos pesados	Directos
SHELXS (Sheldrick, 1991)	Búsqueda de átomos pesados	Patterson o directos
RANTAN (Yao, 1981)	Búsqueda de átomos pesados	Directos
SHELXD (Usón & Sheldrick, 1999; Schneider & Sheldrick, 2002)	Búsqueda de átomos pesados (SIR(AS), MIR(AS), MAD, SAD)	Directos y Patterson
SOLVE (Terwillinger & Berendzen, 1999)	Búsqueda de átomos pesados y faseado (MAD, SAD)	Patterson
SHARP (de La Fortelle & Bricogne, 1997)	Refinamiento de posiciones de átomos pesados y faseado (SIR(AS), MIR(AS), MAD)	Directos
SHELXE (Debreczeni <i>et al</i> , 2003)	Faseado (SIR(AS), MIR(AS), MAD, SAD)	Directos

Tabla 2.3.2: Programas de faseado.

En cristales de proteínas, el problema de la fase puede resolverse mediante tres métodos: (i) mediante el *Reemplazo Molecular* (MR, del inglés *Molecular Replacement*), haciendo uso de un modelo estructural de una proteína homóloga, previamente determinada, (ii) introduciendo átomos altamente dispersores, o método del *Reemplazo Isomorfo Múltiple* (MIR, del inglés, *Multiple Isomorphous Replacement*) o (iii) introduciendo átomos dispersores anómalos, o método de *Dispersión Anómala Múltiple* (MAD, del inglés *Multi-wavelength Anomalous Diffraction*). Con cualquiera de estos tres métodos podemos calcular unas fases iniciales, $\alpha_c(hkl)$ que, junto con las amplitudes experimentales, $|F_{exp}(hkl)|$, nos permitirán calcular una función de densidad electrónica (ecuación de la [Figura 2.3.4](#)) aproximada, sobre la cual podremos construir nuestro modelo estructural.

Reemplazo molecular (MR)

Cuando se dispone de un modelo estructural de una proteína, con secuencia homóloga a la que queremos resolver, es posible utilizarlo como punto de partida en el cálculo de las fases iniciales. Esta técnica fue descrita inicialmente por Michael Rossmann y David Blow (Rossmann & Blow, 1962). Para ello se requiere que el modelo conocido tenga una identidad secuencial con la

proteína en estudio, de al menos un 25%, y una r.m.s.d. (desviación cuadrática media) entre las posiciones de los átomos $C\alpha$ del modelo y de la nueva estructura final menor de 2.0 Å. El propósito de esta metodología consiste en rotar, $[R]$, y trasladar, T , el modelo de estructura conocida, A (Figura 2.3.6), dentro de la celdilla unidad del cristal de la proteína de estructura desconocida, B (Figura 2.3.6). Matemáticamente se busca que la diferencia entre el patrón de difracción teórico de este híbrido y el patrón de difracción del cristal de la proteína de estructura desconocida sea mínima. (Figura 2.3.6). Cuando se encuentra la mejor orientación y traslación para el modelo, se calculan las fases utilizando las coordenadas del mismo y las amplitudes medidas experimentalmente, para obtener una función de densidad electrónica aproximada de la molécula en estudio.

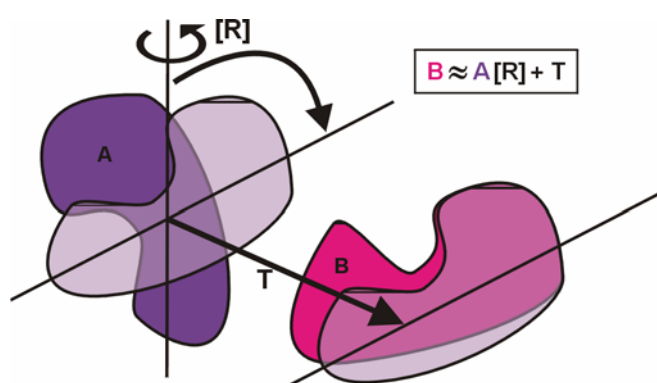


Figura 2.3.6: Esquema del proceso de reemplazo molecular.

Reemplazo isomorfo (MIR)

El empleo de esta metodología en proteínas fue aplicada por primera vez por Max Perutz y John Kendrew (Perutz, 1956; Kendrew *et al*, 1958). Consiste en incluir un número limitado de átomos pesados, en posiciones ordenadas del cristal de proteína en estudio. La inclusión del átomo pesado (Pt, Au, Hg, Se, I, Br, etc.) puede ser mediante inmersión del cristal en una solución del mismo o mediante cocrystalización. Si el cristal del derivado de átomo pesado es isomorfo con respecto al de proteína, es decir tienen la misma celdilla y la misma posición de las moléculas en la celdilla, es posible obtener la subestructura de los átomos pesados por comparación de los espectros de difracción de la proteína nativa y del derivado mediante métodos directos o de Patterson. En un caso ideal de isomorfismo perfecto, y si se dispone de más de un derivado de átomo pesado, el problema de la fase tiene una solución vectorial de forma que, conocida la subestructura, es posible calcular las fases del cristal de la proteína nativa (Blundell & Johnson, 1976). En cualquier caso, la falta de isomorfismo, el escaso poder dispersor de los átomos pesados frente al resto de la estructura y las dificultades en encontrar derivados diferentes, hacen

que las fases obtenidas sean solo aproximadas. Cuando se dispone de un único derivado la metodología de reemplazo isomorfo se denomina SIR, y es necesario recurrir a fuentes adicionales para la obtención de fases (ver sección “Dispersión anómala”). Si se dispone de varios derivados la metodología se denomina MIR.

Dispersión anómala (MAD)

Además de la difracción, la interacción de los rayos X con la materia cristalina lleva asociada fenómenos de absorción. La absorción hace que la radiación dispersada sufra un desplazamiento en fase y amplitud que se denomina dispersión anómala (Blundell & Johnson, 1976). Los átomos con un elevado número de electrones, los más pesados, son los que contribuyen a la dispersión anómala, pudiéndose considerar este fenómeno despreciable en los átomos ligeros, como el H, C, N y O que componen las macromoléculas biológicas. La inclusión de dispersores anómalos también se suele realizar, como en el caso del reemplazo isomorfo, mediante cocrystalización o inmersión en soluciones que contengan dispersores anómalos. El efecto de la dispersión anómala es observable experimentalmente cuando se comparan las intensidades de parejas de reflexiones simétricas (pares de Friedel), que deberían ser iguales, pero que no lo son ($I(hkl) \neq I(-h-k-l)$), dando lugar a las diferencias anómalas o de Bijvoet. Como en las espectroscopías, el fenómeno es dependiente de la longitud de onda de la radiación incidente. Las diferencias anómalas pueden usarse, como las diferencias isomorfas, para localizar los dispersores anómalos mediante métodos directos o de Patterson. Un formulismo vectorial parecido al mencionado para el caso de SIR/MIR permite asignar unas fases aproximadas.

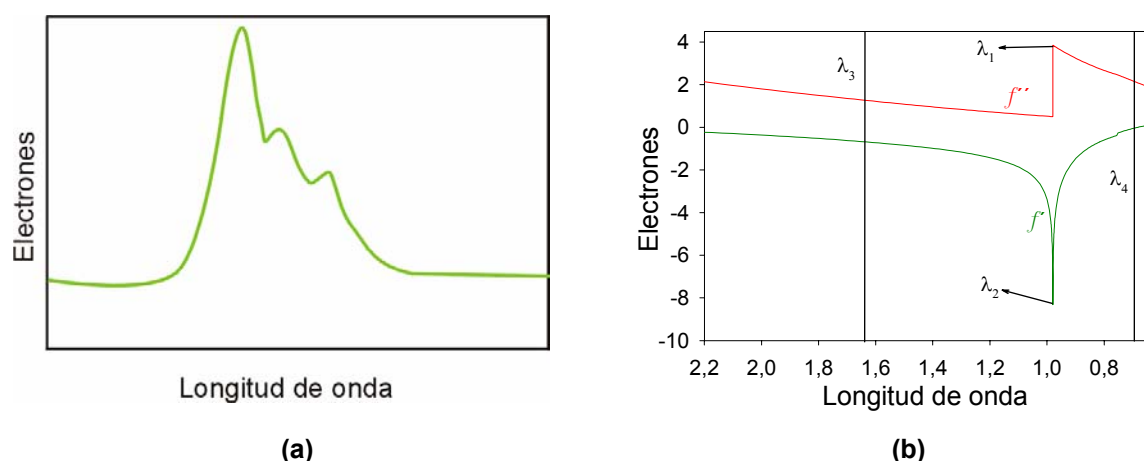


Figura 2.3.7: (a) Curva de absorción (f'') determinada mediante fluorescencia. (b) Variación de la dispersión anómala para un átomo dispersor. La curva es la suma de la componente dispersiva (f') y la componente de absorción (f''). La longitud de onda está expresada en Å.

El diseño experimental suele requerir el uso de la radiación sintonizable sincrotrón e implica la inclusión de dispersores anómalos y la selección de la longitud de onda a la que se recogerán los espectros. Al menos es necesario obtener datos a una longitud de onda cercana al máximo de absorción (SAD) para resolver el problema de las fases. Existe también la posibilidad de recoger datos de difracción a diferentes longitudes de onda, denominándose MAD el método de obtención de fases. Para ello, es necesario determinar experimentalmente la curva real de absorción mediante un espectro de fluorescencia en el sincrotrón, ya que el ambiente que rodea al dispersor anómalo puede afectar al fenómeno de absorción (Figura 2.3.7(a)). En este caso, los datos se recogen generalmente a tres longitudes de onda: (i) λ_1 , en el pico de absorción, (ii) λ_2 , en el punto de inflexión de la curva de absorción y (iii) λ_3 , longitud de onda remota donde la dispersión anómala es mínima (también es posible recoger a otra longitud de onda remota, λ_4) (Figura 2.3.7(b)). De esta manera, con un único cristal que contenga dispersores anómalos es posible resolver la estructura. La dispersión anómala también puede romper la ambigüedad de fase que aparece en el método de SIR; al método que utiliza las diferencias isomorfas y anómalas se le conoce como SIRAS. Por analogía, el uso de varios derivados de átomo pesado isomorfos, combinado con el empleo de la información anómala, se denomina MIRAS.

Metodologías para mejorar las fases

En ocasiones, los valores de las fases obtenidas inicialmente son solo aproximados, y en estas circunstancias, los mapas de densidad electrónica que se calculan a partir de éstas no son completamente interpretables. Por tanto, han de ser mejoradas haciendo uso de distintos métodos de modificación de densidad, que están basados en algún tipo de conocimiento estructural previo (Read & Kleywegt, 2001; Read, 2003).

Las técnicas de modificación de densidad más empleadas son: el aplanado de solvente (*solvent flattening*), el promediado por simetría no cristalográfica (*non-crystallographic averaging*), el ajuste a histograma (*histogram matching*) y la extensión de fases (*phase extension*).

- **El aplanado de solvente** parte del conocimiento de la fracción que ocupan solvente y proteína en el cristal. Asume que la densidad electrónica no puede ser negativa y fija unos valores promedios para la densidad de la proteína y del solvente.
- **La técnica de promediado por simetría no cristalográfica** es también una técnica muy potente y consiste en imponer equivalencia en los valores de densidad electrónica cuando existen varias copias en la unidad asimétrica relacionadas por simetría no cristalográfica. La densidad correspondiente a cada copia se promedia de manera que aumenta la

relación señal/ruido. El promediado es más efectivo cuanto mayor sea el número de copias independientes en la unidad asimétrica y cuanto mayor sea el contenido de solvente.

- **La técnica de ajuste a histograma** es una mejora de la técnica de aplanado de solvente. Asume que la densidad electrónica de la región proteica sigue una distribución de densidad electrónica que no es constante.
- **La técnica de extensión de fases** se emplea para propagar la información que se tiene de las fases desde baja a alta resolución. De esta manera, si se tienen las fases para un conjunto de datos a baja resolución, se puede hacer uso de éstas sobre un conjunto de datos a alta resolución para obtener fases más allá de la resolución inicialmente obtenida. En este caso se asume que la distribución de densidad electrónica existente en los datos a baja resolución se va a conservar para los datos de alta resolución.

La mejora de fases es un proceso cíclico que termina cuando se alcanza la convergencia y que conlleva la aplicación alternada de dos procesos: la modificación del mapa y combinación de fases, de manera que con un mapa mejorado se recalculan fases, que luego se vuelven a utilizar para obtener un nuevo mapa mejorado. Todas estas técnicas pueden llevarse a la práctica con programas como DM (Cowtan & Main, 1998), SOLOMON (Leslie, 1987a; Abrahams & Leslie, 1996), RESOLVE (Terwillinger, 2000) o CNS (Brünger *et al.*, 1998).

Trazado de la cadena polipeptídica y refinamiento

Una vez que se dispone de un mapa de densidad electrónica interpretable, se procede al trazado de la cadena polipeptídica, proceso que se puede automatizar si se dispone de datos a alta resolución (mejor de 2.5 Å) (Perrakis *et al.*, 1999). Si el mapa no es de la suficiente calidad, el trazado tiene que realizarse manualmente. El modelo estructural inicial no suele explicar demasiado bien las observaciones experimentales (desacuerdo entre $|F_o|$ y $|F_c|$, mostrando un elevado factor de desacuerdo R (Ecuación 2.3.1). Para mejorar el modelo (minimizar R hasta un 20-25%) se realizan procedimientos de refinamiento matemático, seguidos del cálculo de mapas de densidad electrónica y de una reconstrucción manual, necesaria cuando existen errores serios en el modelo que el refinamiento matemático no puede corregir.

$$R(\%) = 100 \left(\frac{\sum_{hkl} |F_o| - k |F_c|}{\sum_{hkl} |F_o|} \right)$$

Ecuación 2.3.1: Factor de desacuerdo.

Los parámetros atómicos que se refinan son las coordenadas (x,y,z) de la posición de cada átomo, el factor térmico B isotrópico, que da cuenta de la vibración térmica o movilidad del átomo, y la ocupación. Esto quiere decir que en una macromolécula, el número de parámetros por refinar es muy superior al número de datos experimentales, lo que obliga a introducir en el refinamiento una serie de restricciones geométricas (longitudes y ángulos de enlace, interacciones de van der Waals, ángulos de torsión, etc.), que se recogen en diccionarios.

Existen distintos tipos de refinamiento que se aplican dependiendo del estado en que se encuentre el modelo. En las primeras etapas del refinamiento se aplican los métodos de templado simulado (*simulated annealing*) y de cuerpo rígido. Posteriormente se aplican métodos probabilísticos de máxima verosimilitud (*maximum likelihood*). Un valor bajo del factor de desacuerdo R no corresponde necesariamente a un buen modelo final. Para solucionar este problema, se ha introducido el método estadístico de validación cruzada (Kleywegt & Brünger, 1996). Así, los datos de difracción se dividen en dos grupos: un grupo que contiene el 90-95% de los datos y que se utiliza en el proceso de refinamiento (*working set*), y un grupo que contiene el 10-5% restante de datos de difracción (*test set*), que no se usan en el refinamiento. De este último conjunto de datos se calcula otro factor de desacuerdo R denominado R -free. El R -free mide el grado en que el modelo predice los datos de difracción incluidos en el conjunto denominado *test set*, que fueron excluidos del proceso de refinamiento y modelado. Por tanto, cambios en el modelo que no mejoren la capacidad del mismo para predecir este 10% de datos de difracción harán que el R -free no baje. Por tanto, el R -free está correlacionado con la exactitud del modelo atómico. Los programas que se suelen utilizar para refinar las estructuras son REFMAC (Murshudov *et al.*, 1997) y CNS (Brünger *et al.*, 1998) y para la reconstrucción manual O (Jones *et al.*, 1991)

El modelo final puede ser considerado aceptable cuando presenta unos buenos valores para el factor de desacuerdo ($R \approx 20$ -25% y R -free $< 30\%$), se ajusta al mapa de densidad electrónica y presenta la mayoría de torsiones de enlace dentro de los valores permitidos (Ramachandran *et al.*, 1963).

2.3.2. Resolución estructural de SOS3

2.3.2.1. Reemplazo molecular

Cuando se abordó la resolución estructural de SOS3, no se conocía la estructura de ningún miembro de la familia de las SCaBPs/CBLs. Así que se realizó una búsqueda de estructuras homólogas a SOS3 con el servidor 3D-PSSM (*Protein Fold Recognition*) (<http://www.sbg.bio.ic.ac.uk/~3dpssm>) para intentar resolver la estructura por reemplazo

molecular. Las dos proteínas que presentaban mayor homología con SOS3 eran la frecuencia humana o sensor neuronal de calcio 1 (NCS1; código PDB 1g8i; Bourne *et al.*, 2001) y la calcineurina B humana (CnB; código PDB 1aui; Kissinger *et al.*, 1995), que presentaban unos porcentajes de identidad secuencial de 26.6 y 19.8%, respectivamente (Altschul *et al.*, 1990).

Toma de datos de difracción

Se recogieron datos de difracción de un cristal de SOS3 crecido con la técnica de difusión de vapor en gota colgante en presencia de los aditivos sacarosa y etanol ([Apartado 2.2.2](#), [Tabla 2.2.1](#), [Figura 2.2.5\(b\)](#)). El cristal se montó en un lazo de nylon y se congeló a 120 K con una corriente de nitrógeno. No fue necesario el uso de ninguna solución crioprotectora porque los cristales crecían en una solución madre que ya poseía propiedades crioprotectoras. Los datos de difracción se midieron haciendo uso de la radiación $\text{CuK}\alpha$, en un equipo Kappa CCD-2000 de Bruker-Nonius con espejos Montel ([Figura 2.3.8\(a\)](#)).

Los datos de difracción se procesaron con HKL2000 (Otwinowski & Minor, 1997). Las estadísticas del procesamiento de datos se detallan en la [Tabla 2.3.3](#). El porcentaje de solvente en la celdilla unidad era de 63.6% (Mathews, 1968; Kantardjieff & Rupp, 2003), lo que suponía la existencia de 8 moléculas de SOS3 por celdilla, o lo que es lo mismo, una molécula de SOS3 por unidad asimétrica ($V_M = 3.38 \text{ \AA}^3/\text{Da}$) (Anderson & Hovmöller, 2000).

SOS3	
Grupo espacial	$I4_1$
Celdilla unidad (Å)	$a=b=93.65, c=80.08$
Rango de resolución(Å)	30.0-3.2
Observaciones	18957
Reflexiones únicas	5736
Completitud (%)	98.9 (99.5)
Multiplicidad	3.3 (3.3)
*R_{merge} (%)	0.09 (0.42)
$I/\sigma(I)$	15.3 (5.2)
Contenido de solvente (%)	63.6

* $R_{\text{merge}} = \sum_{hkl} |I_{hkl} - \langle I_{hkl} \rangle| / \sum_{hkl} I_{hkl}$ para todo el rango de resolución y la concha de resolución 3.20–3.31 Å. Los valores entre paréntesis corresponden a los datos de la concha de mayor resolución: 3.20–3.31 Å.

Tabla 2.3.3: Estadísticas del procesamiento de datos de difracción.

En busca de una solución de reemplazo molecular

Se realizaron múltiples intentos para tratar de encontrar una solución de reemplazo molecular al problema de la fase de SOS3. Todos los intentos se realizaron con los modelos estructurales de las dos proteínas cuya homología era más elevada, NCS1 y CnB, y con modelos generados a partir de éstas como modelos de poli-alaninas o con los bucles eliminados. Finalmente se encontró una solución contrastada con el programa AMoRe (Navaza, 1994), tomando como modelo la estructura del NCS1. En la [Tabla 2.3.4](#) se especifican los parámetros que se fijaron y las estadísticas (correlación entre los espectros calculado y el observado y el Factor-*R*) de la solución obtenida.

	SOS3
Grupo Espacial	I4 ₁
Distancia entre centros de masa (Å)	25
Resolución	15-3.5
Coefficiente de Correlación (CC) (%)	32.1
Factor-<i>R</i> (%)	50.3

Tabla 2.3.4: Solución de reemplazo molecular donde se muestran los parámetros clave para la obtención de la solución y las estadísticas que suelen dar cuenta de una buena solución.

El mapa de densidad electrónica calculado era extremadamente ruidoso, no era predictivo y estaba claramente sesgado por el modelo del NCS1, es decir, las fases calculadas no eran de la suficiente calidad como para poder obtener un buen mapa.

Mientras se intentaba resolver la estructura de SOS3, apareció el modelo tridimensional de SCaBP1/CBL2 (código PDB 1uhn; Nagae *et al.*, 2003), que había sido resuelto mediante cristalografía de rayos X. Este modelo poseía una identidad secuencial con SOS3 del 50.4% (Altschul *et al.*, 1990). Sorprendentemente, a pesar de la elevada identidad secuencial, no pudo encontrarse una buena solución de reemplazo molecular cuando se empleaba la estructura de SCaBP1/CBL2 como modelo. En principio, este resultado sugiere que deben de existir importantes diferencias estructurales entre SOS3 y SCaBP1/CBL2 para que el reemplazo molecular fallara.

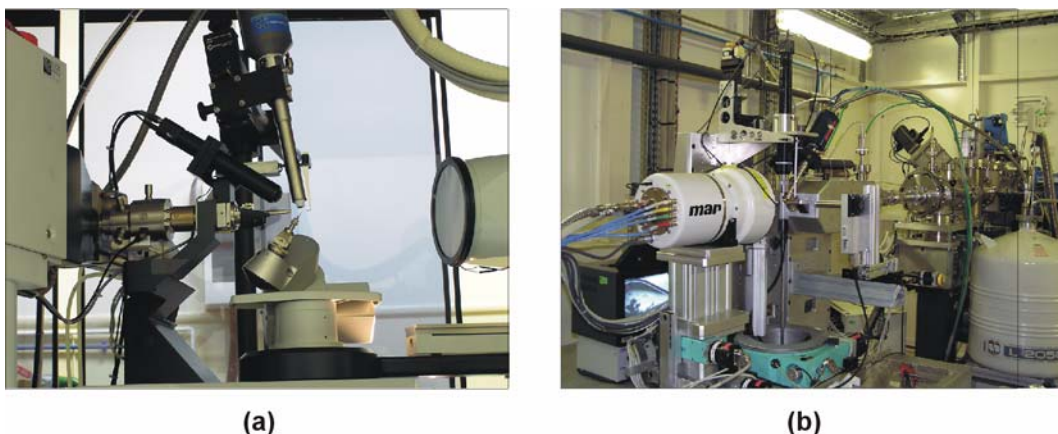


Figura 2.3.8: (a) Imagen de un equipo de difracción Kappa CCD-2000 de Bruker-Nonius. (b) Cabaña de difracción de la línea BM16 del ESRF.

2.3.2.2. Faseado con yoduro

En vista de que la técnica de reemplazo molecular no era adecuada para resolver la estructura de SOS3, se intentó fasear con yoduro, ya que se disponía de cristales crecidos en presencia de NaI.

El uso de haluros como átomos pesados para el faseado de estructuras macromoleculares biológicas ha abierto un nuevo camino para la solución del problema de las fases (Dauter *et al.*, 2000; Dauter & Dauter, 2001; Dauter *et al.*, 2001). Hasta hace una década, el método más común para obtener derivados de átomos pesados consistía en sumergir el cristal en soluciones de átomo pesado a muy baja concentración (del orden de mM) durante un tiempo que va desde minutos hasta días. Un buen derivado se definía como aquel que producía un bajo número de sitios con elevada ocupación y orden, y además era isomorfo con el cristal de la proteína nativa. En los últimos años este panorama ha cambiado, la posibilidad de congelar los cristales ha permitido sumergir los cristales en soluciones de átomo pesado más concentradas y durante menor tiempo. Además, se han desarrollado programas que son capaces de resolver subestructuras compuestas de un elevado número de átomos con ocupancia y orden variados. Todos estos avances han permitido el uso del yoduro como anión faseador. Sus características le convierten en una opción ventajosa a la hora de solucionar el problema de las fases, como ocurre con otros candidatos estrella como el Se o incluso el Br, y es que nos encontramos ante uno de los candidatos más versátiles para el faseado de estructuras de proteínas.

El empleo del yoduro como átomo pesado tiene muchas ventajas, que se pueden resumir en que no es tóxico, no produce alteraciones significativas en la estructura de la proteína ya que reemplaza moléculas de agua, es muy soluble, es barato, no altera las condiciones de cristalización y fundamentalmente, puede emplearse para resolver las estructuras sin necesidad

de acudir a la radiación sincrotrón. A continuación se detallan de manera más pormenorizada las particularidades de su uso como átomo pesado y sus características:

- Las sales más empleadas para la inclusión de yoduro son NaI y KI, pero además son posibles las mezclas KI/I₂, que generan aniones I₃⁻ y I₅⁻, también útiles para fasear, aunque tienen el inconveniente de que son insolubles en PEGs (Evans & Bricogne, 2003). La sustitución del anión en la red cristalina se puede hacer, tanto por inmersión del cristal durante unos 10 segundos en soluciones crioprotectoras que llevan la sal en una concentración que varía entre 0.25 y 1 M, como mediante cocrystalización, para lo que se emplean concentraciones del orden de 0.2 M. Las proteínas suelen cristalizarse en presencia de NaCl, sal muy parecida desde el punto de vista químico-físico al NaI, por lo que a veces es posible reproducir el crecimiento de cristales en soluciones madres que además incluyan NaI.
- El anión I⁻ interacciona mediante enlaces de hidrógeno con los grupos amida de la cadena principal de Trp, Tyr, Phe, Leu, Gly, Arg y Asn o de las cadenas laterales de Asn o Gln, mediante formación de pares iónicos con residuos como Arg o Lys, mediante puentes de hidrógeno con los grupos hidroxilos como los de las cadenas laterales de los residuos Ser o Thr o con el nitrógeno del anillo indólico del Trp (Dauter & Dauter, 2001). Cuanto mayor sea la superficie molecular de una proteína, mayor será el número de I⁻ que se unan.
- El anión I⁻ difunde rápidamente por los canales que forman los cristales de proteína. Por tanto, en principio siempre será mejor obtener el derivado por cocrystalización que por inmersión, ya que los iones sustituidos se encontrarán únicamente en posiciones de equilibrio, teniendo una ocupación y un orden mayor, por lo que el cálculo de la subestructura será menos complejo.
- Las técnicas de faseado que se aplican a los derivados de yoduro son diversas, MIR, MIRAS, SIR, SIRAS y SAD. La única técnica que no se puede aplicar es MAD, ya que el eje de absorción se encuentra a una longitud de onda fuera del rango disponible en el sincrotrón. Sin embargo la señal anómala a la longitud de onda de la radiación CuK α es alta, por lo que es posible fasear habiendo recogido los datos de difracción en un difractómetro con fuente convencional. En este caso es muy aconsejable orientar el cristal si se dispone de un aparato con geometría Kappa, para así recoger datos con una elevada redundancia y completitud y poder incluso recoger los pares de Friedel en una misma imagen, minimizando así los errores en la medida de las intensidades. Si por el contrario

se tiene acceso a la radiación sincrotrón, es posible medir fácilmente a una longitud de onda elevada, para hacer la absorción máxima.

Toma de datos de difracción

Se recogieron datos de difracción en la línea española BM16 del ESRF (*European Synchrotron Radiation Facilities*) de Grenoble (Figura 2.3.8(b)) de los cristales de SOS3 en complejo con Mn^{2+} y Ca^{2+} y de SOS3 en complejo con Ca^{2+} (Apartado 2.2, Tabla 2.2.1.d y .e), que habían crecido en presencia de NaI. Los cristales se midieron a unas longitudes de onda de 1.75 Å donde $f'' = 8.39$ electrones, y a 0.92 Å donde $f'' = 2.85$ electrones (f'' es la componente imaginaria del factor de dispersión atómica). Los datos se procesaron con MOSFLM (Leslie, 1987b)) y programas del paquete CCP4 (Bailey, 1994). Las estadísticas que se muestran proceden del escalado conjunto de los pares de Friedel. Los datos relativos a las recogidas y a las estadísticas del procesado se muestran en la Tabla 2.3.5.

	Complejo SOS3 Ca^{2+}		Complejo SOS3 Mn^{2+} Ca^{2+}	
Longitud de onda (Å)	0.92	1.75	0.92	1.75
Grupo espacial	$I4_1$		$I4_1$	
Celdilla unidad (Å)	91.6	91.6	85.4	91.9
Resolución (Å)	62.0-2.75	45.6-2.8	46.0-3.0	49.9-3.0
$I/\sigma(I)$	4.5(1.8)	2.6(1.4)	4.5(1.5)	2.6(1.9)
* R_{merge} (%)	5.7(41.8)	11.7(47.5)	8.6(51.2)	10.6(39.7)
Complejitud (%)	100(100)	100(100)	97.7(97.7)	96.7(96.7)
Multiplicidad	3.7(3.7)	7.5(7.4)	4.5(4.5)	3.9(3.8)
Solvente(%)	64.3		64.7	

* $R_{merge} = \frac{\sum_{hkl} |I_{hkl} - \langle I_{hkl} \rangle|}{\sum_{hkl} I_{hkl}}$. Los valores entre paréntesis corresponden a los datos de la concha de mayor resolución: 2.90-2.75 y 2.95-2.80 Å (para el complejo SOS3 Ca^{2+} y longitud de onda 0.92 y 1.75, respectivamente), y 3.16-3.00 Å (para el complejo SOS3 Mn^{2+} Ca^{2+}).

Tabla 2.3.5: Recogida de datos y estadísticas de procesado.

Cálculo de la subestructura y faseado

El cálculo de la subestructura se realizó de dos modos, por un lado calculando un mapa anómalo con las fases que se obtuvieron mediante reemplazo molecular y localizando los dispersores anómalos y, por otro, *ab initio*, con el programa SHELXD (Usón & Sheldrick, 1999; Schneider & Sheldrick, 2002). La transformada de Fourier en la cual los valores de los módulos de los factores

de estructura se sustituyen por las diferencias anómalas entre los pares de Friedel (dano) informa sobre las posiciones de los átomos pesados.

El cálculo de mapas anómalos es una buena manera de encontrar dispersores anómalos cuando se dispone de un conjunto de fases aproximadas. Estos mapas dan una impresión extremadamente optimista de la calidad de los datos, debido a que las fases son más importantes que las amplitudes a la hora de determinar la calidad de un mapa. Aunque las fases calculadas mediante reemplazo molecular no eran lo suficientemente aproximadas para obtener un buen mapa de densidad electrónica de SOS₃, permitieron localizar las posiciones de los dispersores anómalos (Figura 2.3.9(a)) y con ellos fasear la estructura. De esta manera se identificaron hasta 7 posiciones de iones yoduro. Con esta subestructura y las diferencias anómalas se consiguió fasear mediante la técnica de SAD (Tabla 2.3.6). Tras realizar los cálculos de mejora de fases (SOLOMON (Leslie, 1987a; Abrahams & Leslie, 1996) y DM (Cowtan & Main, 1998)), se obtuvieron unos mapas de densidad electrónica que ya eran de la suficiente calidad como para construir el modelo de SOS₃ Ca²⁺ (Figura 2.3.9(c)). Además se intentó fasear comparando los datos recogidos a las longitudes de onda de 1.75 y 0.92 Å y calculando las diferencias de las diferencias anómalas (MAD). Las estadísticas de faseado eran ligeramente mejores (Tabla 2.3.6), pero tras el aplanado de solvente no se observaba ninguna mejoría en la calidad del mapa de densidad electrónica.

	SAD	seudo-MAD
Rango de resolución	45.6 - 2.8	45.6 - 2.8
R-Cullis	0.68	0.68
Poder de faseado	1.65	1.73
Figura de mérito (FOM)	0.30	0.36

Tabla 2.3.6: Estadísticas que se emplean para caracterizar la calidad del faseado determinado por SHARP (de la Fortelle & Bricogne, 1997).

Por otro lado, se obtuvo la subestructura de yoduro *ab initio*, haciendo uso únicamente de las diferencias anómalas (Figura 2.3.9(b)) con el programa SHELXD (Usón & Sheldrick, 1999; Schneider & Sheldrick, 2002). Se llegó a la misma solución, pero esta vez se faseó y se aplanó el solvente con el programa SHELXE (Debreczeni *et al.*, 2003). Las estadísticas del faseado se muestran en la Tabla 2.3.7.

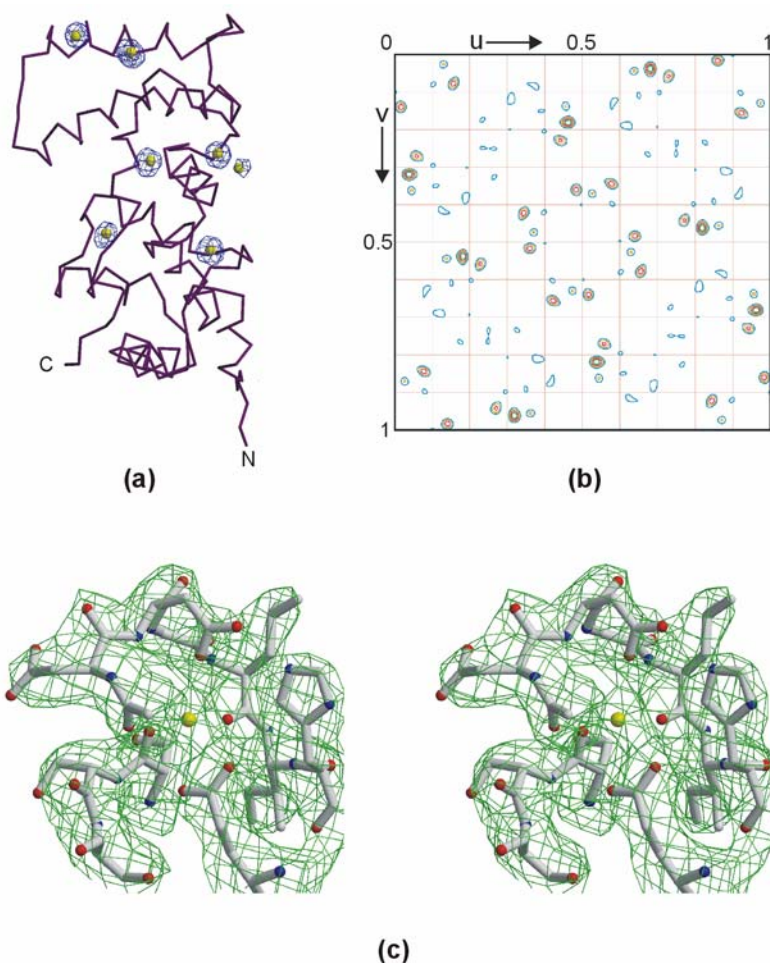


Figura 2.3.9: (a) Sección del mapa de diferencias anómalas (en azul) contorneado a 4σ . Los iones yoduro se representan con bolas amarillas. (b) Como ilustración, se muestra una sección del mapa de Patterson anómalo correspondiente a los datos de SOS3 medidos a $\lambda=1.75$ Å. Los picos más altos están relacionados con las posiciones de los iones yoduro, destacados frente al ruido (cada contorno corresponde a 1σ). (c) Vista estereográfica de una sección del mapa de densidad electrónica experimental (contorneado a 1σ) en el sitio de unión a Ca^{2+} EF1 de SOS3.

	SHELXD		SHELXE
Resolución (Å)	45.6 – 3.5	Resolución (Å)	45.6 – 3.5
CC(total)/CC(débil)	42.29/19.84	Contraste	1
PAT-FOM	24.03	Conectividad	0.92

Tabla 2.3.7: Estadísticas del faseado con SHELXD y SHELXE.

Análisis comparativo de dos casos prácticos de faseado con yoduro

El empleo de sales de yoduro para la resolución estructural de proteínas es reciente (Dauter *et al.*, 2000; Dauter *et al.*, 2001; Dauter & Dauter, 2001) y existen pocos estudios sistemáticos en la literatura en los que se analicen las variables experimentales que conducen a la resolución de estos problemas estructurales. Por este motivo, es muy interesante la realización de un estudio sistemático de estas variables. Para ello se ha tomado como caso práctico de estudio la resolución estructural de SOS3 y de la proteína AS-48 (Sánchez-Barrena *et al.*, 2003). La estructura de AS-48 ha sido resuelta por la doctoranda como trabajo de iniciación a la investigación en el Departamento de Cristalografía del Instituto de Química-Física Rocasolano. En la [Tabla 2.3.8](#) se muestran las estadísticas del procesado de datos de los cristales nativo y derivado de yodo de AS-48.

	Cristal Nativo	Cristal derivado
Grupo Espacial	I222	
Celdilla unidad (Å)	79.47, 83.40, 99.83	79.79, 83.65, 99.43
Resolución (Å)	25.24(1.46)	33.90(2.40)
I/σ(I)	0.30	0.36
R_{merge} (%)	6.4(28.7)	12.3(21.5)
Complejidad (%)	94.6(94.6)	100(100)
Multiplicidad	4.0 (3.0)	9.9 (9.9)

* $R_{merge} = \frac{\sum_{hkl} |I_{hkl} - \langle I_{hkl} \rangle|}{\sum_{hkl} I_{hkl}}$. Los valores entre paréntesis corresponden a los datos de la concha de mayor resolución: 1.54-1.46 Å (para el cristal nativo) y 1.73-1.64 Å (para el derivado de NaI).

Tabla 2.3.8: Estadísticas del procesado de datos de AS-48.

Las diferencias fundamentales entre ambos estudios se detallan a continuación:

- La inclusión del yoduro en los cristales de AS-48 se realizó mediante inmersión en una solución crioprotectora 1 M NaI, mientras que en el caso de SOS3 se realizó mediante cocrystalización. Esto redundaba en el nivel de sustitución del yoduro en el cristal. Para el derivado de AS-48 se encontraron 15 I⁻, mientras que para SOS3 solo 7, siendo la masa molecular en la unidad asimétrica comparable en ambos casos.

- La recogida de datos del derivado de AS-48 (AS-48i) se realizó en un difractómetro convencional usando radiación $\text{CuK}\alpha$; por el contrario, los datos de difracción de SOS3 se recogieron en el ESRF (Grenoble) a una longitud de onda de 1.75 Å.
- La resolución estructural de AS-48 se realizó mediante la metodología de SIRAS (haciendo uso de las diferencias anómalas e isomorfas, ya que se disponía de cristales de proteína nativa, y de un derivado de yoduro, y éstos eran comparables). Sin embargo, la resolución de SOS3 se realizó mediante SAD (solo se emplearon las diferencias anómalas).

Consideraciones generales

Ya se han mencionado anteriormente las ventajas del empleo de las sales de yoduro como anión faseador. Entre ellas destacan el hecho de que las sales de yoduro suelen producir derivados isomorfos y que dicho anión puede utilizarse como dispersor anómalo. Sin embargo, el relativamente bajo número de electrones que presenta el yoduro respecto a otros dispersores convencionales (Pt, Au, Sm, Gd) hacen que las diferencias anómalas e isomorfas sean menores. Por tanto, es fundamental extremar el cuidado en la recogida de datos y en el escalado de los mismos. En este sentido, el empleo de la geometría Kappa para la recogida de datos es muy aconsejable porque es posible obtener datos de difracción muy redundantes y orientar el cristal, pudiéndose recoger en una misma imagen los pares de Friedel y de esta manera minimizar las diferencias entre ellos debidas a errores en la medida. Por otro lado, es también importante realizar un escalado de los datos en donde se promedien los pares de Friedel de manera independiente; de esta manera es posible disponer de una mejor estimación de los mismos, factor importante ya que las diferencias anómalas suelen ser pequeñas y por tanto están sometidas a elevados errores.

Para tener éxito en el faseado con yoduro es necesario realizar un análisis cuidadoso de las señales anómalas e isomorfas, con el fin de saber si son significativas frente a los errores experimentales que llevan asociadas. Delimitar los valores de resolución entre los que la señal es significativa es de particular importancia para los métodos de resolución de las subestructuras de átomo pesado, sobre todo los basados en métodos directos. Los últimos se fundamentan en factores de estructura normalizados (E), y los $\langle |E| \rangle$ son constantes para una concha de resolución, de manera que se les da más peso a los datos a alta resolución, por lo que se debe asegurar de la calidad de éstos y trabajar a una resolución en donde la relación señal/ruido sea adecuada. De la misma manera, los métodos directos son notoriamente sensibles a la presencia de incluso un pequeño número de reflexiones mal medidas, por lo que es importante eliminar

estas reflexiones. Esto es especialmente problemático en el caso de las diferencias, que pueden ser especialmente ruidosas.

El análisis de la existencia de señal anómala e isomorfa, el grado de sustitución del átomo pesado (en este caso yoduro) y la calidad de la señal se pueden estimar con los programas de CCP4 (Bailey, 1994) SCALA (Evans, 1993) y SCALEIT (Howell & Smith, 1992) y también con XPREP (Sheldrick *et al.*, 2001).

Los parámetros que se van a analizar son: la variación del *R-iso* y *R-ano* con la resolución, la dependencia de la señal anómala e isomorfa con el valor de $|F|$ y la resolución, la relación señal/ruido para las señales anómala e isomorfa. Una vez analizados se va a estudiar cómo se puede obtener la subestructura y cómo se puede fasear de una manera sencilla y automatizada.

Variación del *R-iso* y *R-ano* con la resolución

El parámetro *R-sym* (expresado en tanto por ciento o en tanto por uno) refleja el acuerdo entre reflexiones que debido a la simetría deberían ser iguales. Por tanto, valores altos de *R-sym* se relacionan con errores en la medida de las observaciones (si el grupo espacial está bien asignado) y su análisis frente a la resolución se utiliza para estimar la bondad de los datos de difracción.

El análisis comparado del *R-sym* frente al acuerdo observado entre reflexiones afectadas por la dispersión anómala, *R-ano*, que da cuenta de las diferencias anómalas ($\Delta F^{\pm} = |F_{PH(+)}| - |F_{PH(-)}|$), indica si la señal anómala es significativa. El subíndice PH se refiere al cristal de proteína (P) con el átomo pesado, *heavy* (H). En la [Figura 2.3.10\(a\)](#) se muestra una representación del *R-sym* y el *R-ano* para los datos de SOS3 y del derivado de yoduro de AS-48 (AS-48i) frente a la resolución. Como se puede observar para ambos casos, a una resolución superior a 4.5 Å la señal anómala, *R-ano*, está muy por debajo de la variabilidad en la medida de las reflexiones motivadas por errores experimentales, y por tanto estas diferencias a alta resolución no serían significativas. Sin embargo, a partir de aproximadamente 4.5 Å de resolución *R-ano* > *R-sym*, indicando buena calidad de los datos anómalos. Por otro lado, el análisis comparado del *R-ano* de AS-48 y SOS3 indica que este parámetro no es sensible a la diferente sustitución y longitud de onda a la que se midieron los cristales.

El parámetro *R-merge* refleja el error experimental asociado a la medida de una reflexión concreta. Por otro lado, el parámetro *R-iso* da cuenta de de las diferencias entre el cristal nativo y el derivado motivadas por la sustitución, es decir, informa de las diferencias isomorfas ($\Delta F = |F_{PH}| - |F_P|$). En la [Figura 2.3.10\(b\)](#) se representa el *R-iso* y el *R-merge* frente a la resolución

para los datos de AS-48. Las diferencias son muchos más marcadas que para el caso de los datos anómalos y son significativas frente a la variabilidad experimental. Los valores de R -iso tan elevados dan cuenta de una elevada sustitución.

Para poder realizar este estudio con AS-48 primero hubo que poner en la misma escala relativa los datos del derivado y los del cristal nativo, procedimiento que se llevó a cabo con el programa FSHCALC (Kraut *et al.*, 1962). Posteriormente, las diferencias isomorfas (R -iso) se calcularon con el programa SCALEIT (Howell & Smith, 1992). Los valores de R -merge, R -sym y R -ano se obtuvieron con el programa SCALA (Evans, 1993).

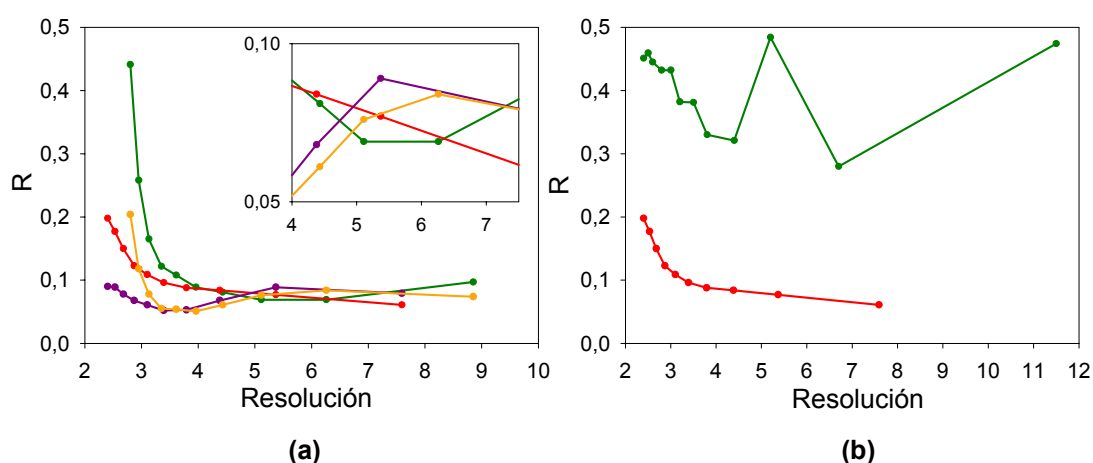


Figura 2.3.10: (a) Representación del R -sym y R -ano vs resolución (en Å) para los datos de SOS3 (en verde y naranja, respectivamente) y del derivado de AS-48 (en rojo y morado, respectivamente). (b) Representación del R -iso y R -merge vs resolución (en Å) de AS-48i (en verde y rojo, respectivamente).

Dependencia de la señal anómala e isomorfa con el valor de $|F|$ y la resolución

Tanto la señal anómala como la isomorfa deben ser independientes de la resolución y del valor de $|F|$. Por tanto, el análisis del valor promedio de las diferencias isomorfas dividido por el valor de $|F|$, ($\langle |diso| \rangle / |F|$), y el valor promedio de las diferencias anómalas en valor absoluto dividido por $|F|$, ($\langle |dano| \rangle / |F|$), frente a la resolución indican la calidad y la magnitud de dichas señales (Howell & Smith, 1992).

La [Figura 2.3.11\(a\)](#) muestra la variación de estos parámetros con la resolución tanto para los datos de SOS3 como para los de AS48i. El análisis de los resultados indica que:

- La calidad de los datos es buena en todos los casos ya que las diferencias anómalas e isomorfas son constantes con la resolución, exceptuando los límites de alta resolución.

- En el caso de SOS3, la señal anómala de los datos recogidos en el sincrotrón a 1.75 Å (línea verde), es comparable con la señal de los datos recogidos en un difractor con fuente convencional Kappa-CCD de Bruker-Nonius (línea naranja). En el difractor convencional fue posible recoger los pares de Friedel en la misma imagen y la redundancia de los datos era de 15. De esta observación se puede deducir que a pesar de que f'' es casi el doble a la longitud de onda de 1.75 frente a 1.54 Å, es más importante realizar recogidas de datos cuidadosas, en donde se minimizan errores en la medida.
- En el caso de AS-48, la magnitud y por tanto la calidad de la señal isomorfa (línea azul) es superior a la de la señal anómala (línea morada). Un indicador del error en la medida de la señal anómala puede obtenerse a través del cálculo de la variación de la señal anómala con la resolución para los datos de la proteína nativa (línea roja).

Los parámetros necesarios para realizar este análisis se calcularon con el programa SCALEIT (Howell & Smith, 1992).

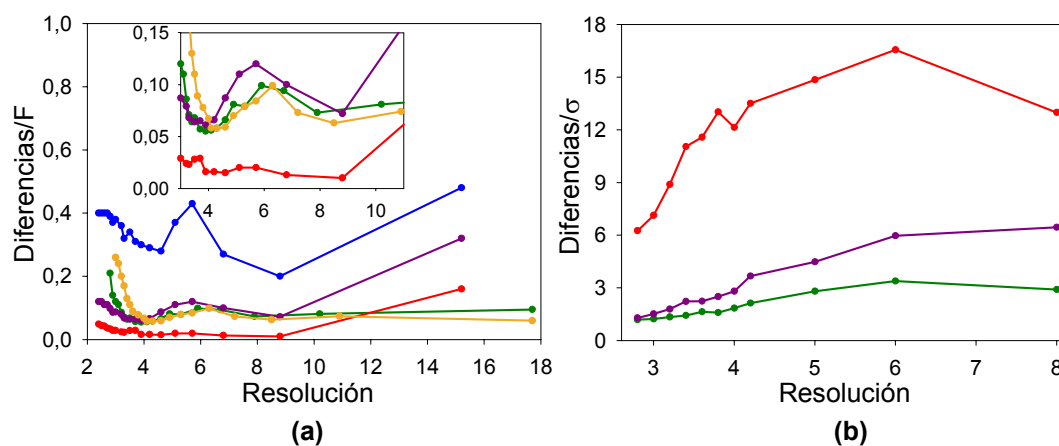


Figura 2.3.11: (a) Variación frente a la resolución de $\langle |diso| \rangle / |F|$, para los datos de AS-48 (en azul), y $\langle |dano| \rangle / |F|$, para los datos de SOS3 recogidos a una longitud de onda de 1.75 Å (en verde), de SOS3 recogidos en un difractor Kappa-CCD 2000 (en naranja), de AS-48i (en morado) y de AS48nat (en rojo). (b) Variación frente a la resolución de $diso/\sigma(diso)$, para los datos de AS-48 (en azul), y $dano/\sigma(dano)$, para SOS3 (en verde) y AS-48 (en morado), con la resolución.

Relación señal/ruido para las señales anómala e isomorfa

La relación entre las diferencias isomorfas (diso) y anómalas (dano) frente al error asociado a los valores de dichos parámetros (σ) muestra también la bondad de los datos, la calidad y magnitud de las señales. Por tanto, la relación señal/ruido de estas diferencias ($diso/\sigma(diso)$ y $dano/\sigma(dano)$)

frente a la resolución es un buen criterio para limitar la alta resolución, donde las intensidades correspondientes a las reflexiones son más débiles y llevan asociadas grandes errores.

En la [Figura 2.3.11\(b\)](#) se presenta el análisis de la relación señal/ruido ($d_{iso}/\sigma(d_{iso})$ y $d_{ano}/\sigma(d_{iso})$) para SOS3 y AS-48. La relación señal/ruido disminuye cuando la resolución aumenta. En nuestra experiencia, un buen criterio de corte sería quedarse con aquellos datos que tengan un valor señal/ruido de al menos 1.5. En relación con estos criterios, la señal isomorfa de AS-48 es significativa en todo el rango de resolución y la señal anómala hasta una resolución de 3.0 Å. En el caso de SOS3 la señal anómala es buena hasta 3.5 Å.

El programa XPREP (Sheldrick *et al.*, 2001) es capaz de hacer un análisis bastante fiable de la relación señal/ruido de las diferencias en función de la resolución y utiliza el programa DIFFE (Blessing *et al.*, 1999) para analizar y rechazar las reflexiones deficientemente medidas y así preparar los datos para la búsqueda de la subestructura.

Cálculo de la subestructura

Una vez analizada la calidad y magnitud de la señal, hay que resolver la subestructura. En la actualidad, los programas Shake-and-Bake (SnB) y SHELXD han introducido un algoritmo muy potente para resolver subestructuras a resolución moderada compuestas por un número elevado de sitios. Los programas convencionales que utilizaban métodos directos no eran capaces de solucionar este tipo de subestructuras. Este algoritmo consiste en la alternancia repetida e incondicional de un ciclo de refinamiento de fases en el espacio recíproco (*Shaking*) seguido de un ciclo de mejora de fases en el espacio real mediante la aplicación de restricciones geométricas (*Baking*). La eficiencia de estos nuevos métodos directos puede ser mejorada considerablemente usando subestructuras iniciales que no sean simplemente aleatorias, como hace SHELXD, que parte de subestructuras aleatorias de prueba evolucionadas, que además son consistentes con la función de Patterson. Tanto SnB como SHELXD computan figuras de mérito que permiten al usuario juzgar la calidad de las soluciones que se obtienen, para poder decidir si la solución es correcta (Weeks *et al.*, 2003). El programa SHELXD es muy efectivo y robusto a la hora de encontrar la subestructura, sin embargo, hay que tener en cuenta ciertos parámetros críticos para este fin: la resolución a la que truncan los datos, que se evalúa muy bien con XPREP o SHELXC, el número de sitios que buscar, que debe de ser al menos un 20% del número real, y en el caso de inmersión del cristal en una solución de átomo pesado, permitir que la búsqueda sea posible en posiciones especiales. Una vez encontrada la subestructura, hay que obtener las fases, y a pesar de los progresos alcanzados en el faseado automático, el tiempo necesario para obtener un mapa interpretable es mucho mayor que el empleado por SHELXD para resolver la subestructura, que

puede ser de segundos o minutos. Así, hasta hace poco tiempo, existía la necesidad de encontrar un modo simple, rápido y robusto de encontrar las fases. Con este fin ha aparecido el programa SHELXE, que aunque no es un programa demasiado eficaz para la mejora de fases, es capaz de dar información acerca de si la estructura se puede resolver con los datos recogidos, puede establecer el enantiómero correcto, e incluso a veces calcula mapas que son de la suficiente calidad para trazar la cadena polipeptídica. Sin embargo este programa sólo funciona correctamente en el caso de que se tengan datos de elevada calidad, el contenido de solvente sea alto (más del 60%) o cuando la resolución de los datos del cristal nativo sea muy alta ($\geq 1.5 \text{ \AA}$).

A continuación se va a mostrar el análisis de las soluciones que se obtienen con SHELXD y SHELXE para los casos de AS-48 y SOS3 cuando el programa se ejecuta de manera automática. La alta resolución se cortó tal y como el programa XPREP sugería, y se buscaron 15 sitios para AS-48 y 7 para SOS3, ya que éstos fueron el número de yoduro que se habían utilizado para fasear. En el caso de AS-48 la solución de la subestructura era buena tanto si se calculaba mediante la técnica de SIRAS, SAD o SIR; como ya se ha dicho la señal anómala también fue suficiente para resolver la estructura de SOS3. Los parámetros resultantes de la soluciones que se van a evaluar son: la ocupación de los sitios y las figuras de mérito.

Ocupación de los sitios

La ocupación de los sitios indica el grado de sustitución del yoduro en el cristal de proteína. En la [Figura 2.3.12](#) se muestra el grado de ocupación de los sitios en AS-48 y SOS3. El análisis de estos datos indica que la inmersión de los cristales de proteína en una solución concentrada de NaI 1 M (como es el caso de AS-48) produce una mayor sustitución que la cocrystalización a una concentración de la sal de 0.2 M (como es el caso de SOS3). El patrón de ocupación que siguen los yoduros en el caso de AS-48 es típico de las inmersiones, hay muchos sitios y con ocupación muy diversa. Sin embargo, la ocupación de los últimos sitios en SOS3 es bastante baja para tratarse de una cocrystalización, y no existe ningún salto brusco en ocupación que distinga las soluciones buenas del ruido, aunque los siete primeros sitios son correctos, ya que se consiguió fasear con estos mismos. Por otro lado, si se compara SIRAS, SAD y SIR, ambas soluciones son parecidas, pero con SIRAS la ocupación de los últimos sitios es mayor; esto se debe a que como se está combinando información (anómala e isomorfa), el refinamiento de las ocupaciones va mejor.

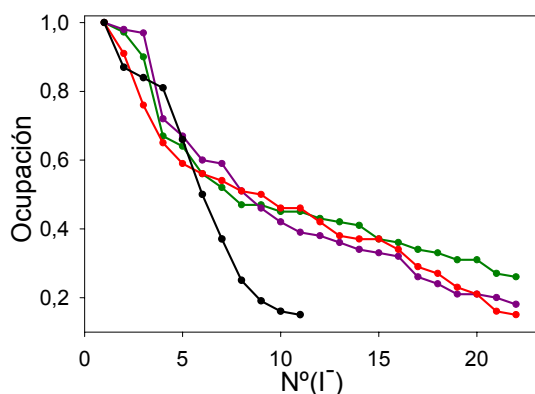


Figura 2.3.12: Representación de la ocupación de los iones I^- con el número de sitios encontrados para las soluciones de SAD de SOS3 (en negro) y de SIRAS, SAD y SIR de AS-48 (en verde, morado y rojo, respectivamente).

Figuras de Mérito

Los programas de cálculo de subestructuras generan una batería de soluciones cuya bondad se encuentra caracterizada mediante figuras de mérito. Si se comparan las figuras de mérito de las soluciones correctas de AS-48 y SOS3 frente a todas las posibles soluciones que SHELXD genera, se comprueba que siempre están contrastadas. Este contraste es mayor para la solución de SIRAS, y mejor el de la solución proveniente de SAD que de SIR (estos datos no se muestran). En nuestra experiencia, las figuras de mérito que dan cuenta de una buena solución son:

- La correlación entre los módulos de los factores de estructura normalizados observados ($|E_o|$) y los calculados ($|E_c|$). Esta figura de mérito que se calcula con todos los datos (CC), y solo con los valores de E más pequeños (CC-débil). Una CC del 35% y una CC-débil de 15% suelen ser indicios de una buena solución.
- La función mínima ($R(\varphi)$): Es una figura de mérito altamente fiable. Se espera que tenga un valor mínimo cuando las fases calculadas sean iguales a las reales. Una buena solución es la que tiene menor valor.
- R cristalográfico (R -crist): mide las diferencias entre los factores de estructura normalizados observados y calculados. Es también una figura de mérito altamente fiable y tiene un valor pequeño para la solución correcta.
- *PATFOM*: Es el valor de la función mínima de Patterson para un número determinado de átomos y mide el acuerdo entre el mapa de Patterson observado y el calculado con la subestructura obtenida. Aunque el valor absoluto depende de la estructura en cuestión, las soluciones correctas suelen tener un valor más alto.

	Método	CC	CC-débil	R-crist	Función mínima	PATFOM
AS-48	SIRAS	51,37	41,35	0.31	0.44	7.07
	SAD	35.92	20.77	0.32	0.53	6.57
	SIR	27,99	11,46	0.37	0.55	6.91
SOS3	SAD	42,29	19,84	0.27	0.39	24.03

Tabla 2.3.9: Estadísticas de las soluciones encontradas con SHELXD.

Los valores de las distintas figuras de mérito para las soluciones de SOS3 y AS-48 se presentan en la [Tabla 2.3.9](#). Para la búsqueda de la subestructura de AS-48 puede verse cómo las figuras de mérito son mejores empleando la técnica de SIRAS respecto a SAD y SIR. Parece que la información anómala es más potente que la isomorfa para encontrar la subestructura, ya que las figuras de mérito son mejores. Como era de esperar, la conjunción de información (SIRAS) permite obtener la solución con figuras de mérito más fiables, da cuenta de una mejor estimación de la posición de los átomos pesados, y además encuentra la solución en un número menor de ciclos de búsqueda. Si se compara la solución de SAD de SOS3 con la de AS-48, las figuras de mérito son mejores para la primera, probablemente porque la subestructura es menos compleja y porque los datos se recogieron a una longitud de onda donde el yoduro es mejor dispersor anómalo.

Obtención del mapa de densidad electrónica: faseado

Una manera rápida de ver si la solución encontrada para la subestructura es correcta es intentar calcular las fases con SHELXE. Las premisas que se asumen en SHELXE para simplificar el faseado mediante SAD, SIR, SIRAS y MAD son:

- que las coordenadas y ocupación de los átomos pesados son las correctas y no necesitan ser refinadas. Los factores térmicos, B, se ignoran,
- que existe solo un tipo de dispersor anómalo, y
- que no se utilizan distribuciones de probabilidad complicadas.

Todo esto significa que a veces es posible que se esté ante una solución de la subestructura correcta pero no lo suficientemente bien estimada para calcular las fases. Esto implicaría que SHELXE no calcularía bien el mapa de densidad electrónica y que habría que utilizar un programa de faseado más complejo como SHARP (de la Fortelle & Bricogne, 1997). Las estadísticas que utiliza SHELXE en el faseado son:

- Correlación del mapa (*CC-mapa*).
- Conectividad del mapa. Valores elevados de estos dos parámetros son indicios de una buena solución.

Como se trabaja con factores de estructura, no hay ningún modo de determinar el enantiómero correcto, por lo que el faseado se realiza teniendo en cuenta los dos enantiómeros. Un indicio de que la solución de la subestructura es correcta y de que el faseado con SHELXE está yendo bien es que la correlación y la conectividad de uno de los enantiómeros mejoran con los ciclos de faseado, mientras que en el otro enantiómero se quedan estancadas; además, las buenas soluciones alcanzan un valor elevado de estos parámetros en pocos ciclos.

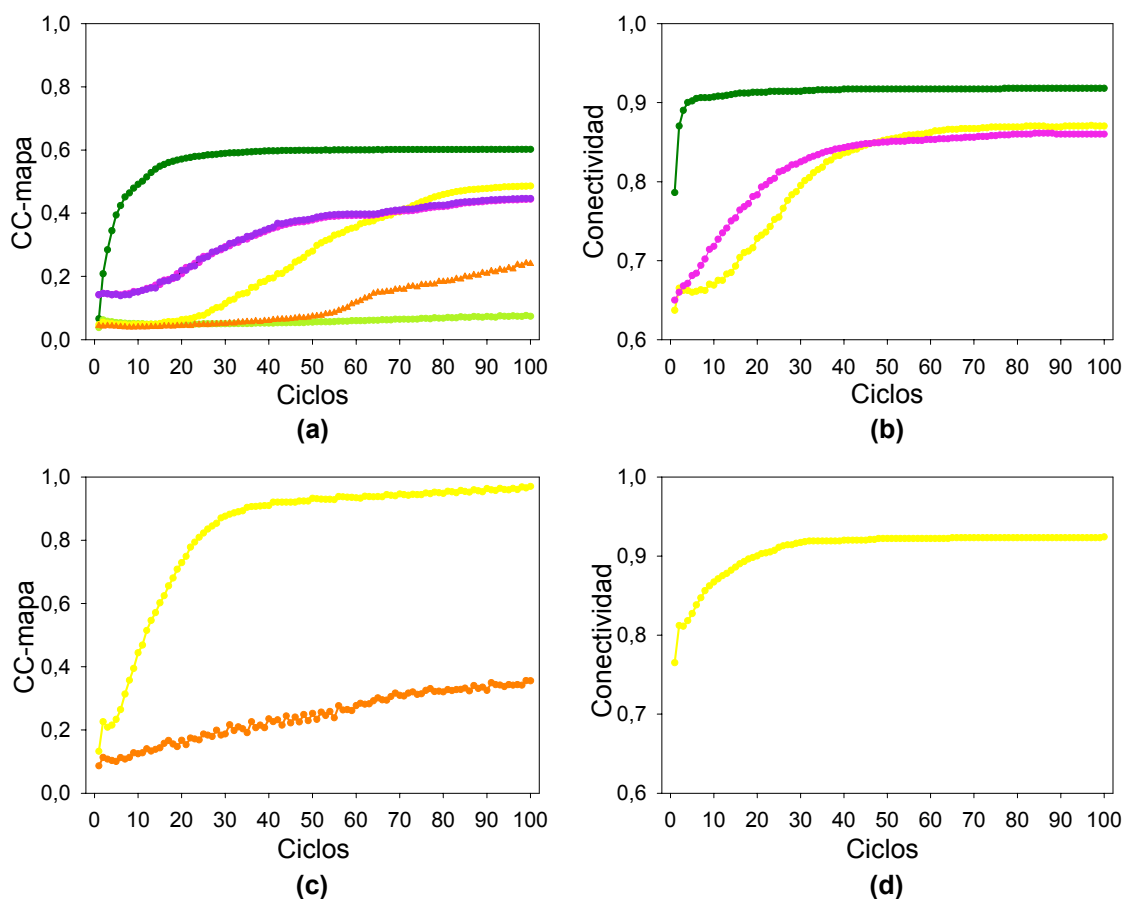


Figura 2.3.13: Representación de la correlación (CC-mapa) y conectividad del mapa con el número de ciclos para las soluciones de SAD de SOS3, y SIRAS, SAD y SIR de AS-48, obtenidos con el programa SHELXE. (a) CC-mapa para las soluciones de AS-48; los colores verde oscuro y claro corresponde a la solución por SIRAS y a la de su enantiomérica, en amarillo y naranja la solución por SAD y su enantiomérica; en rosa y violeta la solución por SIR y su enantiomérica. (b) Conectividad para las soluciones de AS-48 por SIRAS (en verde), SAD (en amarillo) y SIR (en rosa). (c) CC-mapa de la solución de SOS3 por SAD (en amarillo) y su enantiomérica (en naranja). (d) Conectividad de la solución por SAD de SOS3.

En la [Figura 2.3.13\(a\)](#) y en la [Figura 2.3.13\(b\)](#) se muestran la correlación y la conectividad del mapa para las soluciones de SIRAS, SAD y SIR frente al número de ciclos de faseado, respectivamente. El faseado mediante SIRAS permite diferenciar entre los dos enantiómeros (la correlación de uno es mejor que la del otro) y el mapa calculado sería de suficiente calidad como

para trazar la cadena polipeptídica (la conectividad es del 90%). Sin embargo, faseando con SAD, aunque se diferencian los enantiómeros, síntoma de que la solución es correcta, el mapa no es de buena calidad. Esto quiere decir que la subestructura debería de ser refinada para tener una mejor estimación de las fases. La solución mediante SIR es aún peor y ni siquiera se distingue un enantiómero del otro; la correlación es igual. Otro indicio de que el faseado ha ido bien es que si se analizan las tendencias de la curva de la correlación y de la conectividad, frente al número de ciclos de faseado, cuando la solución es correcta se alcanza rápidamente un valor asintótico de correlación y conectividad.

En la [Figura 2.3.13\(c\)](#) y en la [Figura 2.3.13\(d\)](#) se muestran la correlación y conectividad del mapa obtenidos mediante SAD para los datos de SOS3. Con el enantiómero correcto se calcula un mapa de densidad electrónica totalmente interpretable. Si se compara la correlación del mapa de AS-48 (SIRAS) y de SOS3 (SAD) se observa que es mejor la correlación en el segundo caso, al igual que el mapa. Probablemente esto se deba a que el porcentaje de solvente en la celdilla de AS-48 es del 54% mientras que en la de SOS3 es del 64%; además es más sencillo el sistema de SOS3 (1 molécula por unidad asimétrica) que el de AS-48 (4 moléculas por unidad asimétrica).

2.3.2.3. Construcción y refinamiento de las estructuras de los complejos SOS3 Ca²⁺ y SOS3 Mn²⁺ Ca²⁺

Una vez se hubo obtenido un buen mapa experimental de densidad electrónica para los datos del complejo SOS3 Ca²⁺, se procedió a la construcción de los modelos. Los datos de difracción que se utilizaron para el refinamiento de las estructuras fueron los recogidos a una longitud de onda de 0.92 Å ([Tabla 2.3.5](#)), ya que estos datos eran los que alcanzaban mayor resolución.

El trazado de la cadena polipeptídica del complejo SOS3 Ca²⁺ se realizó manualmente con el programa O versión 8.0 (Jones *et al.*, 1991), aunque los extremos N y C-terminal de la cadena y el bucle A67-A75 no estaban definidos. La construcción de la cadena polipeptídica no fue difícil ya que la densidad electrónica de la mayoría de las cadenas laterales estaba bien definida. El refinamiento del modelo se realizó con los programas CNS (Brünger *et al.*, 1998) y REFMAC (Murshudov *et al.*, 1997). Inicialmente se realizó un refinamiento de cuerpo rígido con REFMAC (Murshudov *et al.*, 1997) tomando 5 cuerpos, constituidos cada uno de ellos por los tramos que iban de una mano EF (*EF hand*) a otra, incluyendo cadenas laterales y tomando como límites de resolución 20-3.5 Å. Con el modelo resultante se realizó una dinámica molecular de templado simulado con CNS (Brünger *et al.*, 1998) (se llegó hasta 2500 K, límites de resolución 20-3.5 Å,) para aliviar las tensiones que pudieran existir. El modelo resultante fue sometido a un refinamiento restringido con REFMAC (Murshudov *et al.*, 1997) y se calcularon mapas $2F_o-F_c$ y F_o-F_c . Con la

ayuda de estos mapas y mediante un proceso cíclico de modelado y refinamiento se terminó de construir la cadena y se añadieron 4 iones Ca^{2+} y los 7 I⁻. Se calcularon mapas omitiendo en la estructura los sitios de unión a Ca^{2+} , con el fin corroborar la presencia de los 4 iones Ca^{2+} (Figura 2.3.15) y modelar correctamente las cadenas laterales de los aminoácidos que interaccionaban con estos iones. Finalmente se modelaron las moléculas de agua.

Las estadísticas finales del refinamiento y las características del modelo se muestran en la Tabla 2.3.10. La estereoquímica del modelo se verificó con el programa PROCHECK (Laskowski *et al.*, 1993). En la Figura 2.3.14 se representa el diagrama de Ramachandran, donde se representan los ángulos de torsión ψ y ϕ de los residuos de la cadena polipeptídica. El 83.6% de los residuos se encuentran en las zonas más favorables, el 16.4% en zonas permitidas y ninguno en zonas prohibidas.

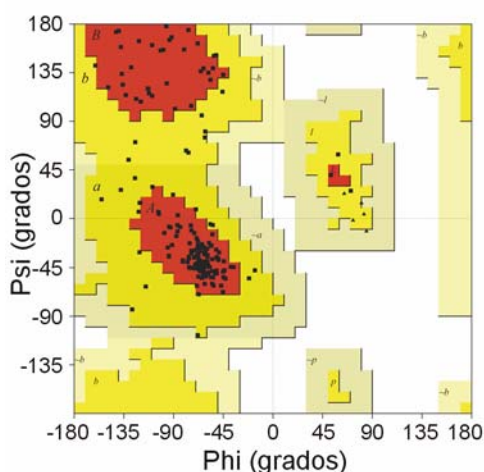


Figura 2.3.14: Diagrama de Ramachandran del modelo estructural del complejo SOS3 Ca^{2+} (Laskowski *et al.*, 1993).

La estructura del complejo SOS3 Mn^{2+} Ca^{2+} se determinó mediante “síntesis de Fourier diferencia”, a partir del modelo de SOS3 Ca^{2+} refinado, ya que los cristales eran isomorfos. El proceso de refinamiento fue análogo al seguido para el complejo SOS3 Ca^{2+} . Se calculó un mapa de diferencias anómalas con los datos del complejo SOS3 Mn^{2+} Ca^{2+} recogidos a la longitud de onda de 1.75 Å para identificar la posición del ión Mn^{2+} en la mano EF-4 (Figura 2.3.15). Los cálculos se realizaron utilizando programas del paquete de CCP4 (Bailey, 1994). Para evitar el posible sesgo del modelo de SOS3 Ca^{2+} sobre el del complejo SOS3 Mn^{2+} Ca^{2+} en el sitio de unión a Mn^{2+} , EF-4, se calcularon mapas omitiendo del modelo inicial los aminoácidos que interaccionaban con el ión Ca^{2+} . Las estadísticas del refinamiento y las características del modelo final se detallan en la Tabla 2.3.10.

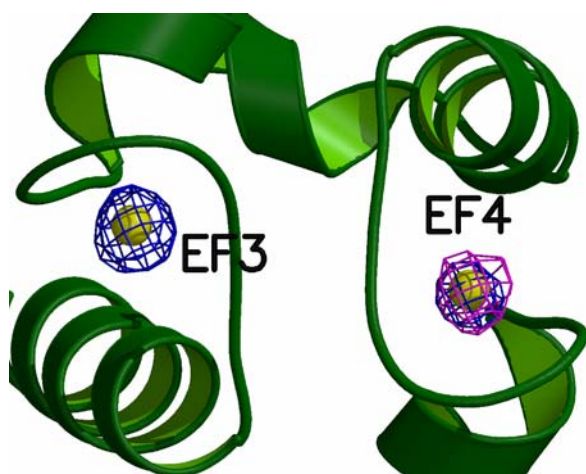


Figura 2.3.15: Secciones del mapa de densidad electrónica ($2F_o-F_c$) (en azul) del complejo SOS3 Ca^{2+} , en donde se han omitido los iones Ca^{2+} y del mapa de diferencias anómalas (F_o-F_c) (en magenta) de los sitios de unión a Ca^{2+} y Mn^{2+} EF3 y EF4 del complejo SOS3 Ca^{2+} Mn^{2+} . Ambos mapas están contorneados a 4σ . Los iones Ca^{2+} se representan con bolas amarillas.

Refinamiento	SOS3 Ca^{2+}	SOS3 Mn^{2+} Ca^{2+}
Resolución (Å)	20.0-2.75	20.0-3.0
R (%)	23.5(41.6)	23.7(32.9)
R-free (%)	26.4(37.1)	26.2(39.8)
Número de átomos	1614	1614
Número de reflexiones	9213	6609
Modelo		
Moléculas por unidad asimétrica	1	1
Aminoácidos	188	188
Iones calcio(II)	4	3
Iones manganeso(II)	0	1
Iones yoduro	7	7
Moléculas de MPD	2	2
Moléculas de agua	40	40
RMSD		
Longitud de enlace (Å)	0.014	0.016
Ángulo de enlace (grados)	1.735	1.926
Grupos planares (Å)	0.005	0.006
Observaciones	Los residuos A1-A14 y A203-A222 no están incluidos en el modelo. El mapa de de densidad electrónica en el bucle A67-A75 no está bien definido.	

Los valores entre paréntesis corresponden a la concha de mayor resolución: 2.82-2.75 Å para el complejo SOS3 Ca^{2+} y 3.08-3.0 Å para el complejo SOS3 Mn^{2+} Ca^{2+} .

Tabla 2.3.10: Estadísticas del refinamiento de los complejos SOS3 Ca^{2+} y SOS3 Mn^{2+} Ca^{2+} .

Las figuras de los modelos que se muestran a lo largo de esta memoria se realizaron con MOLSCRIPT (Kraylis, 1991) y RASTER (Merrit & Murphy, 1994). Las coordenadas y las amplitudes de los factores de estructura de los complejos SOS3 Ca^{2+} y SOS3 Mn^{2+} Ca^{2+} se depositaron en el PDB con los códigos 1v1g y 1v1f, respectivamente.

2.3.3. Análisis de la estructura de los complejos de SOS3

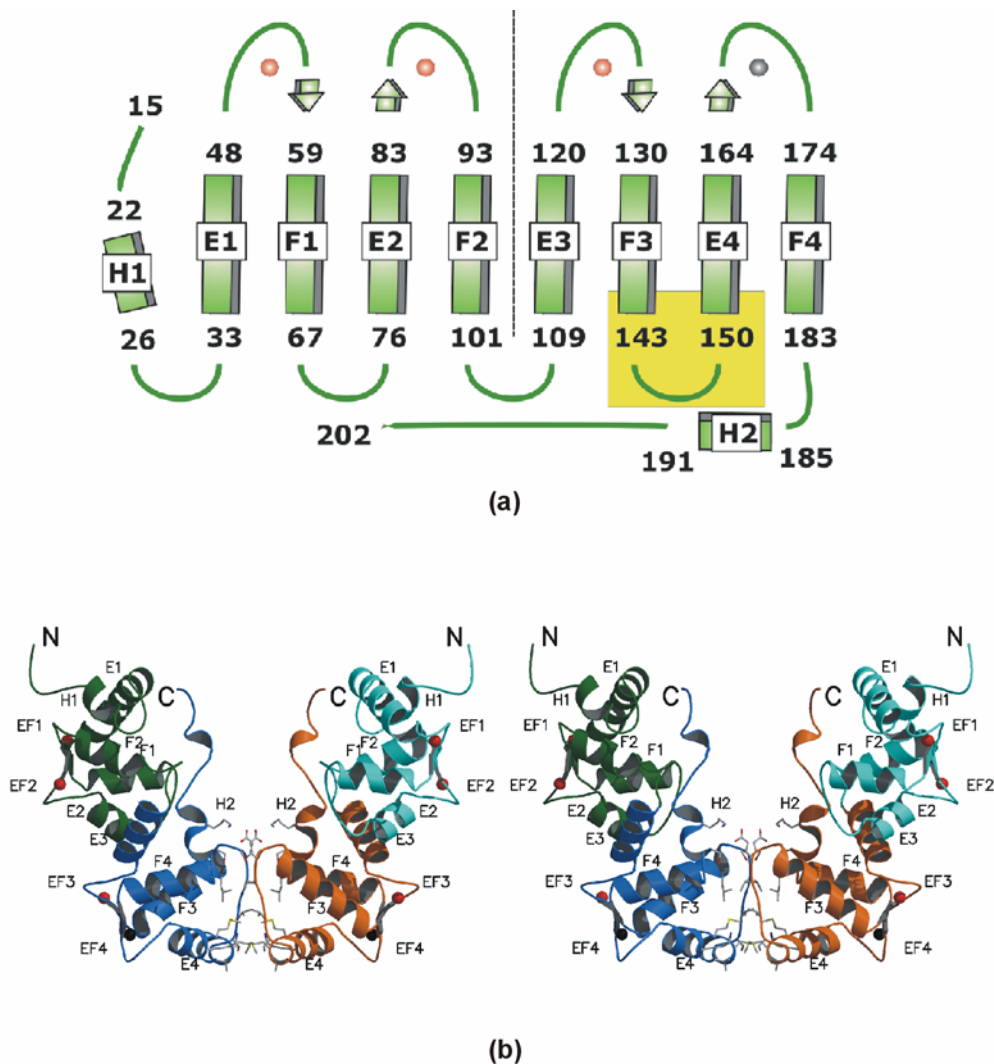


Figura 2.3.16: (a) Representación esquemática de la topología de la estructura de SOS3. Los elementos de estructura secundaria están definidos por RasMol (Sayle & Milner-White, 1995). Los elementos involucrados en la dimerización están destacados mediante un rectángulo amarillo. La línea discontinua separa el dominio N-terminal del C-terminal. Las bolas rojas y negra representan a los iones Ca^{2+} y Mn^{2+} , respectivamente. (b) Representación estereográfica en forma de lazo de la estructura en dominios del dímero de SOS3. Las cadenas laterales que participan en la dimerización se representan en modo de bastones.

Las estructuras de SOS3 en complejo con Ca^{2+} y con Ca^{2+} y Mn^{2+} son casi idénticas, la desviación cuadrática media de los átomos C α del esqueleto polipeptídico (r.m.s.d.) es de 0.2 Å (Sutcliffe *et al.*, 1987), con ligeros cambios confinados al sitio de unión a Ca^{2+} y Mn^{2+} EF-4, que no impiden que la coordinación sea igual en ambos casos. Por tanto, a partir de ahora, a menos que se indique, todos los resultados se referirán al complejo SOS3 Ca^{2+} . El plegamiento general de la proteína es muy parecido al de SCaBP1/AtCBL2 (Nagae *et al.*, 2003) y al de los homólogos estructurales CnB (Cyert & Thorner, 1992) y NCS1 (Schaad *et al.*, 1996).

La estructura de SOS3 está compuesta por dos dominios unidos por un pequeño conector. Cada dominio está formado por dos pares adyacentes de manos EF. El dominio N-terminal contiene una hélice α adicional y el C-terminal una hélice α , más una cola desestructurada que apunta hacia el dominio N-terminal (Figura 2.3.16(a)).

El análisis de la estructura cristalina de SOS3 muestra que la proteína es dimérica (Figura 2.3.16(c)). En el dímero, los protómeros están relacionados por un eje de simetría binario que coincide con un eje cristalográfico de orden 2. Por tanto, en la unidad asimétrica solo se encuentra una subunidad peptídica. El dímero cristalográfico tiene forma de V y expone al solvente 8 sitios de unión a metales, el sitio de miristoilación N-terminal y el extremo C-terminal. Las moléculas que conforman el dímero interactúan a través del extremo C-terminal de la hélice F3, la hélice E4 y el bucle que conecta ambas hélices. En la formación del dímero, cada protómero pierde un 4% de su superficie accesible al solvente; el 80% del área ocluida corresponde a residuos hidrofóbicos (Wallace *et al.*, 1995). La dimerización implica la formación de un único puente de hidrógeno entre el oxígeno hidroxílico de la cadena lateral del residuo Ser 144 y su pareja relacionada por simetría, los demás contactos son hidrofóbicos.

El análisis de la superficie molecular del dímero de SOS3 indica que éste es altamente hidrofóbico, el 61% de la superficie molecular accesible es hidrofóbica (Wallace *et al.*, 1995). Así, si se estudia el empaquetamiento de moléculas de SOS3 en el cristal, se puede observar cómo los contactos cristalográficos entre moléculas tienden a proteger las regiones hidrofóbicas del solvente. En el dímero se exponen fundamentalmente dos zonas hidrofóbicas, una localizada en la interfase formada por los dominios, y la otra en la interfase entre las subunidades que constituyen el dímero.

2.4. CARACTERIZACIÓN ESTRUCTURAL DE SOS3 EN SOLUCIÓN

2.4.1. Introducción

Existe todo un conjunto de técnicas de caracterización de proteínas en solución que permiten obtener información estructural a baja resolución. Los sistemas macromoleculares son complejos, por lo que la utilización conjunta de técnicas de alta resolución, como la cristalografía, y técnicas en solución resulta muy interesante para caracterizar el sistema en estudio. Las técnicas en solución poseen características que las hacen versátiles e interesantes de combinar con los estudios cristalográficos ya que:

- Aportan información estructural que puede ser de gran utilidad para encontrar las condiciones óptimas de cristalización de la proteína.
- Complementan los resultados estructurales a alta resolución cuando la estructura de la proteína es conocida. Asimismo facilitan la interpretación de los resultados estructurales obtenidos a partir del modelo atómico.
- Son técnicas de bajo coste y el gasto de muestra es moderado.

Con objeto de caracterizar a SOS3 en solución, se han empleado las siguientes aproximaciones experimentales que a continuación se detallan:

- **Ultracentrifugación analítica:** en este caso, el objetivo concreto es caracterizar el comportamiento hidrodinámico de SOS3 en diferentes condiciones químico-físicas.
- **Espectroscopía de dicroísmo circular:** proporcionar información sobre la estructura secundaria (UV-lejano) y terciaria (UV-cercano) de SOS3.
- **Espectroscopía de fluorescencia:** va a permitir la obtención de información de carácter local (estructura terciaria) del único residuo de triptófano de SOS3 en presencia de diferentes cationes relevantes fisiológicamente.

2.4.2. Comportamiento hidrodinámico de SOS3: análisis mediante ultracentrifugación analítica

Las técnicas de centrifugación analítica permiten determinar la masa molecular y los parámetros hidrodinámicos de las macromoléculas biológicas en solución. Además, posibilitan la detección y cuantificación de las interacciones macromoleculares reversibles proteína-proteína o proteína-ligando. Estas interacciones son generalmente débiles y cuando las fuerzas del campo cristalino son equivalentes o de mayor magnitud que ellas, la ultracentrifugación analítica es una de las herramientas más potentes para su caracterización cuantitativa.

Hay dos tipos de experimentos básicos en ultracentrifugación analítica: la velocidad de sedimentación y el equilibrio de sedimentación. Mediante los primeros es posible estimar los coeficientes de sedimentación y difusión de las macromoléculas estudiadas, los cuales están relacionados con su tamaño y forma. Por otro lado, los experimentos de equilibrio de sedimentación permiten determinar la masa molecular promedio de las macromoléculas estudiadas, así como poner de manifiesto la existencia de equilibrios de asociación y llevar a cabo su caracterización cuantitativa mediante la determinación de las constantes de equilibrio y la estequiometría de los complejos.

2.4.2.1. Experimentos de equilibrio de sedimentación

Con el fin de analizar el comportamiento asociativo de SOS3, así como los potenciales efectos que sobre él pudieran tener el Ca^{2+} y otros iones metálicos fisiológicamente relevantes, se llevaron a cabo experimentos de equilibrio de sedimentación.

Los experimentos se realizaron en una ultracentrífuga Beckman Optima XL-A de seis canales (paso óptico de 12 mm) equipada con rotor AnTi50. Las muestras de SOS3 se equilibraron en tampones que contenían 20 mM Tris-HCl pH 7.5, 50 mM NaCl y CaCl_2 , MnCl_2 , o MgCl_2 según correspondía, de manera que se registraron los perfiles de sedimentación para mezclas de proteína:ión metálico (Mn^{2+} y Mg^{2+}), en relaciones molares 1:1, 1:10 y 1:100. Los experimentos que se realizaron con la proteína nativa y en presencia de Ca^{2+} aparecen resumidos en la [Tabla 2.4.1](#). La concentración de referencia de SOS3 fue 27.2 μM . Las diferentes muestras se centrifugaron a 17000, 20000 y 23000 r.p.m. a una temperatura de 25°C. Para cada velocidad, se realizaron medidas radiales de absorbancia a 280 nm a las 12, 14 y 16 horas. Las tres resultaron idénticas, indicando que el equilibrio se había alcanzado. La masa molecular promedio (M_w) de SOS3 se determinó con el programa EQASSOC (Minton, 1994), empleando para ello un valor para el volumen específico parcial de SOS3 a 25°C de 0.730 ml/g, calculado a partir de su composición aminoacídica (Laue *et al.*, 1992). Los cálculos se realizaron con los datos obtenidos a 17000 r.p.m.

[SOS3]:[Ca ²⁺]	1:0	0.5:0	1:1	1:3	1:10	1:30	1:100
M_w (KDa)	42.1	37.1	42.1	44.4	45.7	47.5	50.6
Fracción de dímero (α)	0.45	0.27	0.45	0.55	0.61	0.71	0.90

[SOS3]:[Ca²⁺] indica la relación molar de SOS3 y Ca²⁺. Los datos están referidos a [SOS3] = 27.2 μM . La masa molecular teórica para el monómero de SOS3 es 25974.7 Da (<http://us.expasy.org>).

Tabla 2.4.1: Comportamiento asociativo de SOS3 en función de la concentración de proteína y de Ca²⁺.

Los resultados muestran, por un lado, que la masa molecular promedio de SOS3 es siempre superior a la del monómero, y por otro, que ésta varía con la relación molar [SOS3]:[Ca²⁺], indicando la existencia de procesos asociativos. En concreto, los datos experimentales se ajustan a un modelo de asociación monómero-dímero (Figura 2.4.1(a)). De esta manera, si consideramos la existencia de un equilibrio de asociación monómero-dímero:



la constante de dimerización (K) viene definida por la Ecuación 2.4.1, donde c es la concentración inicial de proteína y α la fracción de proteína en forma dimérica (Rivas *et al.*, 1992).

$$K = \frac{\alpha}{c(1-\alpha)^2}$$

Ecuación 2.4.1: Constante de dimerización

La constante de dimerización calculada para SOS3, a una concentración de 13.6 μM , es de $3.7 \cdot 10^4 \text{ M}^{-1}$. Por otro lado, el análisis del equilibrio de dimerización de SOS3 en función de la concentración de Ca²⁺ muestra que éste se desplaza a la forma dimérica (Figura 2.4.1(a)) a medida que aumenta con la concentración de Ca²⁺ (Tabla 2.4.1). Así, a concentraciones de Ca²⁺ próximas a 10^{-3} M , la práctica totalidad de SOS3 se encuentra en forma dimérica. Por el contrario, no se observa ningún efecto significativo de los iones Mg²⁺ o Mn²⁺ en el equilibrio de dimerización de SOS3 (datos no mostrados).

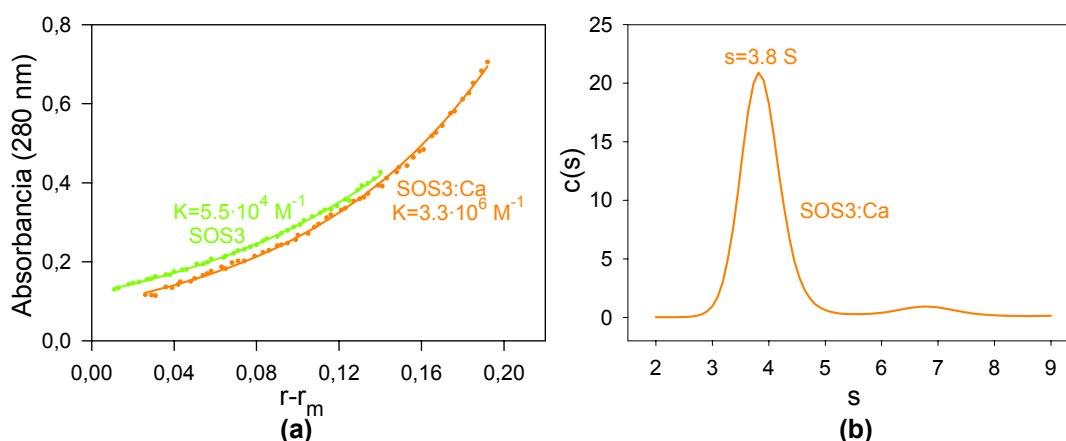


Figura 2.4.1: (a) Efecto del Ca²⁺ sobre el equilibrio de sedimentación de SOS3. Variación de la absorbancia con la distancia al menisco ($r-r_m$) (cm) en el equilibrio para SOS3 (27.2 μM) y SOS3:Ca (relación molar 1:100). En los dos casos aparecen los puntos experimentales y las curvas que representan el mejor ajuste realizado con el programa EQASSOC (Minton, 1994). (b) Distribución de probabilidad de los coeficientes de sedimentación (s) para una solución de SOS3:Ca 1:100 (relación molar referida a una concentración de SOS3 de 36.2 μM). $c(s)$ indica la probabilidad en %.

2.4.2.2. Experimentos de velocidad de sedimentación

Se llevaron a cabo experimentos de velocidad de sedimentación para analizar las propiedades hidrodinámicas del dímero de SOS3. Para ello fue necesario asegurar la existencia de una única especie en solución, por lo que los experimentos se realizaron con muestras de SOS3 que contenían CaCl_2 a una relación molar 1:100. En estas condiciones, la relación dímero:monómero es 9:1, de acuerdo con los análisis previamente descritos (Apartado 2.4.2.1).

Los experimentos se realizaron en una ultracentrífuga Beckman Optima XL-A de seis canales (paso óptico de 12 mm) con rotor AnTi50, empleando una velocidad angular de 45000 r.p.m. La muestra de SOS3 (36.2 μM) se equilibró en tampón 20 mM Tris-HCl pH 7.5, conteniendo 50 mM NaCl y 3.6 mM CaCl_2 . El coeficiente de sedimentación (s) se calculó con el programa SEDFIT (Schuck, 2000) (Tabla 2.4.2). Por otro lado, el coeficiente de fricción translacional (f) de la especie en solución se determinó a partir de la masa molecular teórica del dímero de SOS3 (51948 Da, <http://us.expasy.org>) y el coeficiente de sedimentación. El coeficiente de fricción translacional de la esfera hidratada equivalente, f_o , se estimó haciendo uso de un coeficiente de hidratación (δ) de 0.30 g de agua/g de proteína, estimado a partir de la secuencia de aminoácidos (Kuntz, 1971). De esta manera se estimó la relación f/f_o , que proporciona información de la forma de la partícula en solución (Tabla 2.4.2). El volumen específico parcial a 25°C era de 0.730 ml/g (Laue *et al.*, 1992). La densidad de la disolución se fijó en 1 g/cm³ y la viscosidad en 0.01 P.

SOS3 Ca^{2+}	
s_{exp} (S)	3.8
f/f_o	1.29

Tabla 2.4.2: Parámetros hidrodinámicos de SOS3 en presencia de Ca^{2+} .

La distribución experimental de coeficientes de sedimentación (Figura 2.4.1(b)) es consistente con la existencia de una especie mayoritaria en solución, el dímero. Además, aparece un pico minoritario que correspondería a agregados de alto peso molecular. El valor de f/f_o es superior a 1, indicando que el comportamiento del dímero en solución se aleja del esperado para una partícula esférica.

Conociendo el valor de esta relación friccional y asumiendo modelos simples de elipsoides de revolución para el dímero de SOS3, se pudieron estimar las dimensiones de los mismos mediante el programa SEDFIT (Schuck, 2000). Los resultados así obtenidos no se aproximaron a las

dimensiones del dímero cristalográfico. Mientras que las relaciones axiales (a/b) estimadas fueron 3.55 y 3.73 para un elipsoide oblongo y prolato, respectivamente (Schuck, 2000), las dimensiones de la caja mínima, que contiene al dímero cristalográfico, son $75 \times 57 \times 40 \text{ \AA}^3$. Este resultado indica que un modelo simple, como el de los elipsoides de revolución, no representa adecuadamente al comportamiento hidrodinámico del dímero de SOS3 en solución.

2.4.2.3. Predicción de propiedades hidrodinámicas a partir de un modelo atómico

La comparación del coeficiente de sedimentación experimental con el calculado a partir del modelo tridimensional indica si la forma del modelo es compatible con el comportamiento en solución observado. Por este motivo, se abordó el cálculo del coeficiente de sedimentación teórico a partir del modelo atómico del dímero de SOS3 obtenido por técnicas cristalográficas. Estos cálculos se realizaron con el programa HYDROPRO (García de la Torre *et al.*, 2000; García de la Torre, 2001).

HYDROPRO genera un modelo hidrodinámico primario a partir de las coordenadas de la estructura, de manera que reemplaza todos los átomos que no son hidrógeno por esferas de un radio fijo (AER, de *atomic element radius*) que incluye la hidratación (se ha estimado que 3.2 \AA es un buen valor). Posteriormente, este modelo de esferas solapadas se modela como una envoltura de miniesferas, que tienen que poseer un radio tal que el número de miniesferas esté entre 1800 y 2000. Sobre este último modelo hidrodinámico generado se calculan las propiedades hidrodinámicas. Para los coeficientes de sedimentación, se ha estimado que el cálculo lleva asociado un error del 2% (García de la Torre *et al.*, 2000; García de la Torre, 2001). Los parámetros que necesita el programa para hacer el cálculo son: el AER, el número de miniesferas y el radio (σ , en \AA) de las mismas, la temperatura (en Kelvin), la viscosidad del solvente (en Poises), la masa molecular de la partícula (en Daltons), el volumen específico parcial (en cm^3/g) y la densidad del solvente (en g/cm^3).

En el caso de SOS3, se introdujeron las coordenadas del dímero cristalográfico y se fijó el AER en 3.2 \AA . El cálculo del radio y número de miniesferas lo estimó previamente el programa, de manera que consideró como valores óptimos 1.608 \AA y 1920, respectivamente. Los demás parámetros necesarios para el cálculo se detallan en el [Apartado 2.4.2.2](#). Con todos estos datos, HYDROPRO calculó un valor teórico para el coeficiente de sedimentación, $s_T = 3.89 \text{ S}$.

El valor teórico calculado para el coeficiente de sedimentación, considerando el error asociado a estos cálculos, coincide con el experimental obtenido mediante experimentos de velocidad de

sedimentación, indicando que el comportamiento hidrodinámico observado en solución (Tabla 2.4.2) es compatible con la forma del dímero de SOS3 observada mediante técnicas cristalográficas.

2.4.3. Análisis estructural de SOS3 mediante espectroscopía de dicroísmo circular

El dicroísmo circular (CD) es una técnica espectroscópica sensible a la estructura secundaria (UV-lejano) y terciaria (UV-cercano) de las proteínas. La técnica más utilizada es el CD en el UV-lejano, donde los enlaces peptídicos proporcionan señales distintas según estén dispuestos en hélices α , láminas β o en segmentos no periódicos de la estructura. En esta región, el espectro de una proteína se puede representar como una combinación lineal de los espectros de los elementos de estructura secundaria “puros” más el ruido, así como la contribución de los cromóforos aromáticos (Johnson, 1990; Greenfield, 1996). Actualmente existen métodos de análisis de espectros de CD en el UV-lejano que consideran bases de datos de proteínas de estructura conocida y que permiten deducir los porcentajes de estructura secundaria de la proteína en estudio (Perczel *et al.*, 1991). El CD en el UV-cercano permite fácil y rápidamente establecer, con independencia del contenido en estructura secundaria, si una proteína está plegada. En el UV-cercano los cromóforos son esencialmente los residuos aromáticos localizados en entornos anisotrópicos, pudiendo también contribuir a la actividad óptica los puentes disulfuros. El CD es una técnica espectroscópica muy útil para detectar cambios conformacionales inducidos por las condiciones químico-físicas de la solución (pH, temperatura, presencia o ausencia de un ligando, etc).

2.4.3.1. Experimentos de dicroísmo circular en el UV-lejano y el UV-cercano

Se realizaron medidas de CD en el UV-lejano y cercano con SOS3, en presencia de Ca^{2+} y otros iones metálicos (Tabla 2.4.3), para estudiar si la unión de los mismos producía cambios conformacionales en la proteína nativa.

Los espectros de CD se registraron en un espectropolarímetro JASCO J-810 equipado con un sistema Peltier de control de temperatura. En la región del UV-lejano, las medidas se realizaron en cubetas de cuarzo de 0.1 cm de paso óptico y a una concentración proteica de 0.2 mg/ml (7.7 μM). Los espectros de CD en la región del UV-cercano se registraron en cubetas con un paso óptico de 1 cm y a una concentración de proteína de 0.8 mg/ml (30.8 μM). Para el registro de los espectros se realizaron seis acumulaciones. Todos los espectros se registraron a una temperatura de 20°C. Una vez sustraída la contribución del tampón de medida, las elipticidades obtenidas se

convirtieron en elipticidades molares promedio por residuo $[\Theta]$, utilizando una masa molecular promedio por residuo de 115.4 Da.

Muestra	Relación molar [SOS3]:[M ⁿ⁺]
SOS3:Ni ²⁺ / Mg ²⁺ / Na ⁺ / K ⁺ / Rb ⁺ / Cs ⁺	1:10 y 1:100
SOS3:Ca ²⁺	1:3, 1:6, 1:10, 1:20, 1:30, 1:60, 1:100 y 1:1000
SOS3:Mn ²⁺	1:1, 1:2, 1:3, 1:10 y 1:100
SOS3:Mn ²⁺ :Ca ²⁺	1:10:10
SOS3:EDTA	1:1250

Tabla 2.4.3: Muestras empleadas para los experimentos de CD

Las muestras de SOS3 se prepararon del mismo modo que para los experimentos de equilibrio de sedimentación (Apartado 2.4.2.1). Los espectros se registraron para las mezclas proteína:ión metálico (Mⁿ⁺) y proteína:EDTA que se indican en la Tabla 2.4.3. La proteína nativa, obtenida de la purificación, fue utilizada como control.

En la Figura 2.4.2(a) se muestra el espectro en la región del UV-cercano de la proteína nativa. La señal dicróica en esta región es poco intensa. Los espectros de las demás muestras analizadas no presentaban cambios significativos si se considera la relación señal-ruido que presentan los espectros en esta zona (datos no mostrados). Si se comparan los espectros en la región del UV-lejano, éstos tampoco presentan cambios significativos con la adición de ninguno de los cationes o EDTA. En la Figura 2.4.2(b) se muestran como ejemplo los espectros de SOS3 en presencia de Ca²⁺ y Mn²⁺ a una relación molar proteína:cation metálico de 1:10.

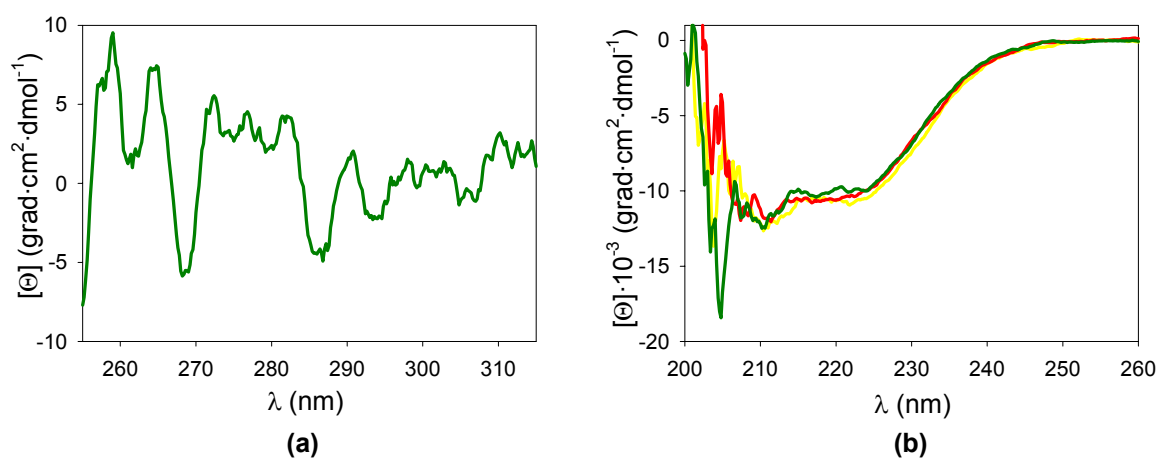


Figura 2.4.2: (a) Espectro de CD en la región del UV-cercano de la proteína SOS3 nativa. (b) Espectro de CD en la región del UV-lejano de la proteína nativa (en verde), de SOS3:Ca 1:10 (en rojo) y de SOS3:Mn 1:10 (en amarillo).

Se realizó un análisis del espectro de CD en el UV-lejano de la proteína nativa para estimar el contenido de estructura secundaria de SOS3. Para ello se hizo uso de distintos programas: CONTIN (Provencher & Glöckner, 1981) y CDNN (Böhm *et al.*, 1992). El contenido de hélice α que predecían estos programas variaba entre un 36 y un 43%, resultado que coincide con el calculado a partir de la estructura cristalográfica (42.7%).

2.4.3.2. Desnaturalización térmica registrada por CD

La desnaturalización térmica de una proteína a concentraciones crecientes de un ligando se puede emplear para investigar la unión proteína-ligando (Varea *et al.*, 2000). Por tanto, se registró la estabilidad térmica por CD de SOS3, así como en presencia de EDTA y a concentraciones crecientes de distintos iones metálicos (Tabla 2.4.3). Además, se registró la estabilidad térmica de una muestra de SOS3 que previamente se había dializado exhaustivamente frente a un tampón Tris-HCl 20 mM pH 7.5, 50 mM NaCl, 10 mM EDTA, 0.05% NaN₃.

Se realizó una serie de medidas de CD a una longitud de onda de 222 nm, aumentando la temperatura desde 20 a 90°C a una velocidad de 30°C/h y permitiendo que la temperatura se equilibrara durante 1 min, antes de que se registrara la medida. Se registraron datos cada 1°C. La elección de la longitud de onda vino motivada por el análisis previo de los espectros en la región del UV-lejano, ya que a esta longitud de onda la señal era cercana al mínimo y poco ruidosa. Los experimentos se realizaron utilizando las muestras y cubetas que se emplearon para registrar los espectros antes descritos (Apartado 2.4.3.1). Después de cada ciclo de desnaturalización se volvió a registrar el espectro en el lejano, comprobándose que la señal disminuía drásticamente y que la muestra precipitaba, indicando la no reversibilidad del proceso de desnaturalización. Para el análisis de los perfiles de desnaturalización se tomó como referencia el perfil de desnaturalización de la proteína nativa obtenida de la purificación (Apartado 2.1.2.2).

Los resultados muestran que el Ca²⁺ y el Mn²⁺ son los únicos cationes que afectan a la desnaturalización térmica de SOS3, provocando un desplazamiento del perfil de desnaturalización a temperaturas superiores (Figura 2.4.3). Esto demuestra la unión de la proteína a estos ligandos. Por otro lado, el hecho de que en presencia del resto de cationes no se observe ningún cambio significativo en el perfil de desnaturalización, indica que el efecto que se observa con los iones Ca²⁺ y Mn²⁺ no está motivado por el incremento de la fuerza iónica.

El hecho de que los perfiles de desnaturalización de SOS3 en presencia y ausencia de 10 mM de EDTA coincidan indica que la proteína control no intercambia fácilmente cationes metálicos con el medio. Por otro lado, el perfil de desnaturalización térmica de SOS3, previamente dializada frente

a EDTA, cambia drásticamente (Figura 2.4.3(a)). Este efecto se revertía adicionando a la muestra dializada Ca^{2+} en una relación proteína: Ca^{2+} 1:10, sugiriendo que la proteína controla un Ca^{2+} con alta afinidad y lenta velocidad de intercambio y que éste es responsable de la estabilización general de la proteína. En el caso del experimento de $\text{SOS3}:\text{Ca}^{2+}:\text{Mn}^{2+}$ (1:10:10), el perfil de desnaturalización es igual al de $\text{SOS3}:\text{Mn}^{2+}$ (1:10), indicando que a esas relaciones molares el Ca^{2+} no desplaza al Mn^{2+} del sitio de unión EF4 que comparten.

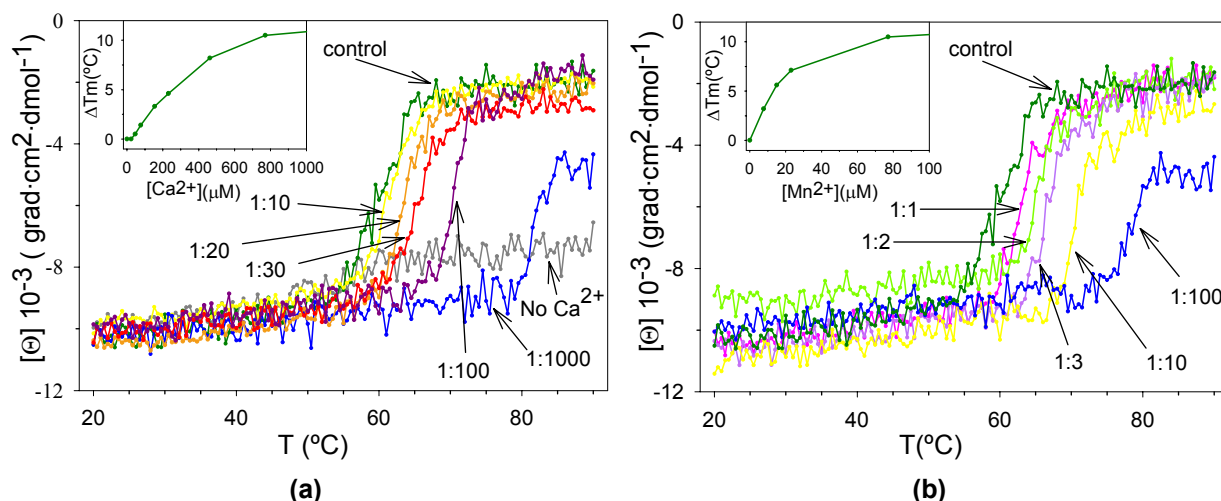


Figura 2.4.3: Perfiles de desnaturalización por dicroísmo circular en función de la concentración de (a) Ca^{2+} y (b) Mn^{2+} . La relación molar SOS3:ión metálico correspondiente a cada curva aparece indicada. Insertada en cada gráfica aparece representada la variación de la T_m con la concentración de ión metálico para cada caso.

A partir de los perfiles de desnaturalización se puede calcular la temperatura de desnaturalización (T_m) de SOS3, que se define como la temperatura a la que el incremento total de la señal dicroica es del 50%. Considerando que las muestras de SOS3 precipitan tras la desnaturalización térmica, los valores de la T_m tienen valor exclusivamente en términos comparativos. En la Tabla 2.4.4 y en la Figura 2.4.3 se muestran los valores de T_m de SOS3 a concentraciones crecientes de Ca^{2+} y Mn^{2+} .

[SOS3]:[M ²⁺]	1:0	1:1	1:2	1:3	1:6	1:10	1:20	1:30	1:60	1:100	1:1000
[M ²⁺] (μM)	0	7.7	15.4	23.1	46.2	77	154	231	462	770	7700
$T_m(\text{Ca}^{2+})$	59.7		59.7	-	60.2	61.1	63.0	64.3	67.9	70.2	81.1
$T_m(\text{Mn}^{2+})$	59.7	62.9	65.3	66.8	-	70.2	-	-	-	77.2	-

Tabla 2.4.4: Variación de la T_m con la concentración de Ca^{2+} y Mn^{2+} . [SOS3]:[M²⁺] indica la relación molar proteína:ión metálico.

2.4.4. Análisis de la estructura terciaria de SOS3 mediante espectroscopía de fluorescencia

La espectroscopia de fluorescencia en el UV-cercano es una técnica muy sensible que ofrece información sobre la estructura terciaria de las proteínas. Empleando una longitud de onda de excitación de 280 nm es posible excitar tanto a tirosinas como a triptófanos, por lo que el espectro de emisión resultante sería la composición de las emisiones de los dos tipos de fluoróforos. Por otro lado, se puede determinar la contribución específica de los residuos de triptófano de la molécula si se excita a una longitud de onda a la cual la tirosina no absorba, o lo haga muy poco, y la elegida suele ser 295 nm. El análisis de los espectros de emisión resultantes proporciona información sobre la naturaleza del entorno en el que se encuentran los triptófanos, ya que sus propiedades espectroscópicas son muy sensibles a la polaridad del medio. Cuando éstas cambian como consecuencia de la unión de un determinado ligando, se pueden llevar a cabo estudios de unión del mismo (García-Segura *et al.*, 1996).

2.4.4.1. Experimentos de emisión de fluorescencia

SOS3 representa un caso especialmente favorable para su análisis por espectroscopía de fluorescencia por cuanto posee un único residuo de triptófano (A177), localizado en las cercanías de los sitios de unión de Ca^{2+} EF3 y EF4 (donde también se une el Mn^{2+}) (Figura 2.4.4(a)). Con objeto de comprobar si la unión de Ca^{2+} o Mn^{2+} a SOS3 se puede analizar empleando el triptófano como sonda espectroscópica, se registraron distintos espectros de emisión.

Los espectros de emisión de fluorescencia del triptófano se obtuvieron en un espectrofluorímetro SLM Aminco 8100, equipado con un arco de xenon de 450 W. La longitud de onda de excitación se fijó en 295 nm, y los espectros de emisión de la proteína se registraron en un intervalo de longitudes de onda comprendido entre 300-450 nm. La anchuras de rendija para los haces de excitación y emisión de luz fueron de 4 nm en los dos casos. Se emplearon cubetas de cuarzo con un paso óptico de 0.5 cm tanto de emisión como de excitación. La temperatura de la muestra se mantuvo a 20°C. Se partió de una concentración de proteína de 0.5 mg/ml (19.2 μM) y se fueron añadiendo concentraciones crecientes de Ca^{2+} y Mn^{2+} hasta alcanzar relaciones molares proteína:cación metálico 1:0, 1:1, 1:5, 1:10, 1:100 y 1:1000. Además, también se partió de una muestra SOS3 Mn^{2+} a unas concentraciones de 19.2 y 192 μM , respectivamente, sobre la que se fueron añadiendo concentraciones crecientes de Ca^{2+} para que las relaciones molares proteína: Mn^{2+} : Ca^{2+} fueran 1:10:0, 1:10:1, 1:10:5, 1:10:10, 1:10:100, 1:10:1000. En todos los casos la dilución de SOS3 se podía considerar despreciable. Todos los espectros fueron corregidos por

el factor de respuesta instrumental del aparato, el factor de dilución y sustraída la línea base correspondiente.

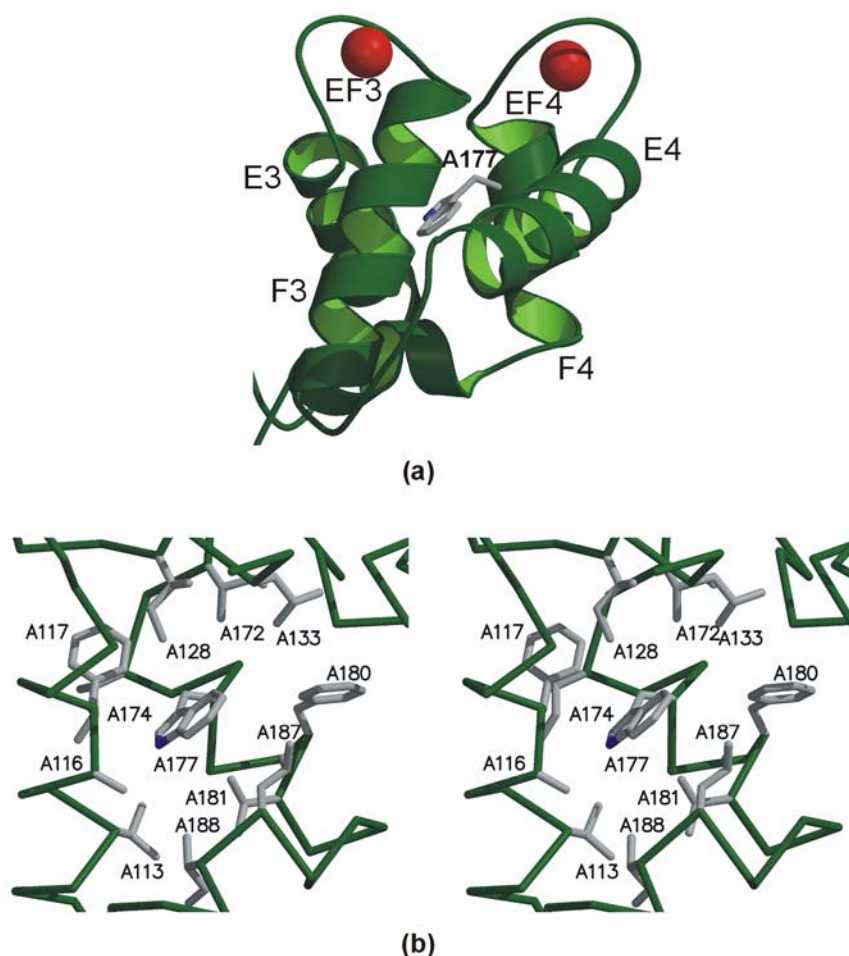


Figura 2.4.4: (a) Imagen de la situación del triptófano A177 con respecto a las manos EF3 y EF4 en el complejo SOS3 Ca^{2+} . El triptófano se representa en modo de bastones y los cationes Ca^{2+} como bolas rojas. (b) Imagen donde se muestra el entorno hidrofóbico que rodea al triptófano A177. Los aminoácidos implicados se representan en modo de bastones.

Los espectros registrados presentan un único máximo de emisión alrededor de 314 nm (Figura 2.4.5). Este valor es característico de un triptófano localizado en un ambiente muy apolar, tal y como por otro lado, se puede observar en su estructura cristalográfica (Figura 2.4.4(b)). La valoración con Mn^{2+} produce una disminución de la intensidad de fluorescencia proporcional a la concentración de Mn^{2+} (Figura 2.4.5). Por el contrario, no se observan cambios significativos en el espectro de emisión de la proteína nativa con la adición de Ca^{2+} . Esto indica que el Trp A177 es sensible a la unión del Mn^{2+} pero no a la del Ca^{2+} . Como los estudios cristalográficos indican que únicamente el Mn^{2+} se une a la EF4 (Apartado 2.3.2.3 y Figura 2.3.15), los estudios mediante fluorescencia permiten analizar la competición del Mn^{2+} y del Ca^{2+} por el sitio EF4. La valoración con Ca^{2+} de la proteína previamente incubada con Mn^{2+} , en relación molar 1:10, no muestra

cambios en el espectro de emisión. Este resultado demuestra que a las concentraciones en que se realizó la valoración, el Ca^{2+} no es capaz de desplazar al Mn^{2+} del sitio que comparten, ya que si compitieran, la intensidad del espectro debería aumentar.

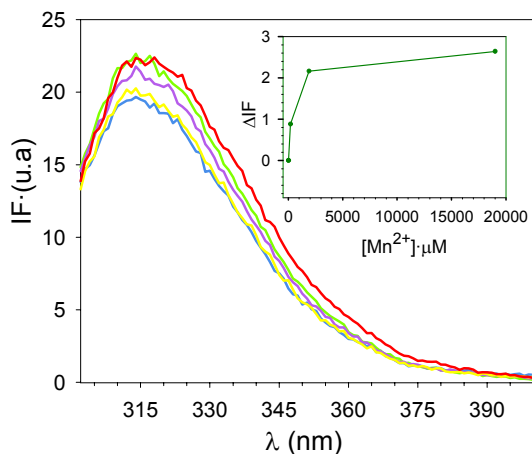


Figura 2.4.5: Espectros de emisión de fluorescencia del Trp en la valoración de SOS3 con Mn^{2+} . Los espectros correspondientes a las soluciones con relaciones molares $\text{SOS3}:\text{Mn}^{2+}$ 1:0, 1:1, 1:10, 1:100 y 1:1000 se representan en rojo, verde, lila, amarillo y azul, respectivamente.

2.4.5. Cromatografía de interacción hidrofóbica

La cromatografía de interacción hidrofóbica (HIC) es una técnica de purificación de proteínas que se basa en la diferente capacidad que poseen las proteínas de interactuar con ligandos hidrofóbicos que se encuentran inmovilizados en una matriz. A concentraciones moderadamente altas de sal (alrededor de 1-2 M) se fomenta la adsorción de las proteínas y mediante un descenso gradual de la concentración salina se consigue la elución de las mismas, reteniéndose más las que presentan mayor hidrofobicidad (Rozanas, 1998).

2.4.5.1. Experimentos de HIC con SOS3

Como ya se ha mencionado en la Introducción ([Capítulo 1](#)), las quinasas de la familia PKS/CIPK interactúan con las proteínas que unen Ca^{2+} de la familia CBL/SCaBP a través de un motivo estructural hidrofóbico conservado de 21 aminoácidos, conocido como motivo NAF/FISL (Guo *et al.*, 2001; Albrecht, *et al.*, 2001). Con objeto de investigar si la unión de Ca^{2+} a SOS3 inducía un aumento de la superficie hidrofóbica accesible al solvente de la macromolécula que favoreciera la unión a SOS2, se realizó una cromatografía de interacción hidrofóbica con SOS3, en presencia y ausencia de Ca^{2+} .

De este modo, se cargaron 1.5 ml de proteína SOS3 nativa ($5.7 \mu\text{M}$) en una columna *Phenyl Sepharose* de 1 ml (Amersham Biosciences Limited, UK) que previamente se había equilibrado con un tampón 20 mM Tris-HCl pH 7.0, 1.7 M $(\text{NH}_4)_2\text{SO}_4$, 0.05% (m/v) (tampón 1). A continuación

se lavó la columna con 5 volúmenes de tampón 1 y se aplicó un gradiente de 1.7 a 0 M de $(\text{NH}_4)_2\text{SO}_4$ en 20 volúmenes. Esta misma cromatografía se realizó empleando tampones que contenían CaCl_2 a una concentración de 10 mM. Posteriormente se lavó la columna con 5 volúmenes de tampón 20 mM Tris-HCl pH 7.0, 10 mM EDTA.

En un medio libre de Ca^{2+} la proteína SOS3 eluía a una concentración de 0.9 M de $(\text{NH}_4)_2\text{SO}_4$ (Figura 2.4.6). Por otro lado, en presencia de Ca^{2+} , SOS3 se quedaba fuertemente adsorbida a la resina, incluso a bajas concentraciones de sal. La elución se consiguió lavando la columna con un tampón que contenía EDTA (Figura 2.4.6), eliminando de esta manera el Ca^{2+} unido a SOS3. Este resultado muestra que la unión de Ca^{2+} a SOS3 lleva asociada un aumento de la superficie hidrofóbica accesible de la macromolécula.

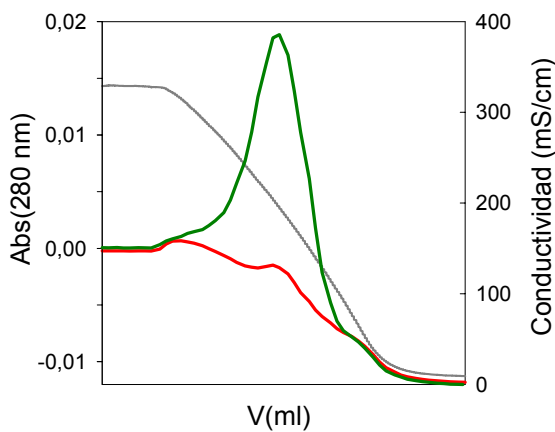


Figura 2.4.6: Cromatografía de interacción hidrofóbica de SOS3. La línea gris representa la conductividad de la solución. Las líneas verde y roja representan la absorbancia registrada a 280 nm en ausencia y presencia de 10 mM CaCl_2 , respectivamente.

3. DISCUSIÓN

3. DISCUSIÓN

La mayoría de las proteínas que unen Ca^{2+} contienen motivos estructurales del tipo mano EF como base estructural para la unión de dicho catión. Las manos EF canónicas constituyen uno de los elementos estructurales más conservados filogenéticamente. Los sitios EF están formados por un bucle de doce aminoácidos flanqueado por dos hélices α . A cada mano EF se une un Ca^{2+} a través de residuos pertenecientes al bucle. Los aminoácidos que participan en la unión se encuentran en las posiciones X, Y, Z, $-Y$, $-X$ y $-Z$ (Figura 3.1.1(a) y (b)). Los situados en las posiciones X, Y, Z y $-Z$, unen el Ca^{2+} a través de oxígenos donadores de la cadena lateral. Los residuos en posiciones $-Y$ y $-X$ participan en la coordinación del ión metálico a través de un oxígeno carbonílico del esqueleto proteico y de una interacción mediada por una molécula de agua, respectivamente (Figura 3.1.1(a) y (b)). La respuesta celular a un elevado número de estímulos está mediada por cambios en la concentración citosólica de Ca^{2+} . El arquetipo de las proteínas que son capaces de sentir estos cambios y de transmitir la señal son las calmodulinas, que están constituidas por un número variable de manos EF funcionales que fluctúa entre uno y cuatro. La unión a Ca^{2+} induce una reorganización estructural que posibilita la interacción y regulación de diferentes proteínas, como las fosfatasas y las quinasas. A pesar del alto grado de conservación, las calmodulinas son capaces de diferenciar multitud de señales mediadas por Ca^{2+} y de desencadenar diferentes procesos celulares.

Las proteínas CBL constituyen una nueva familia de proteínas que contienen manos EF y que participan en sistemas de señalización celular mediados por Ca^{2+} . Sin embargo, presentan características diferenciales con las calmodulinas (CaMs):

- Mientras que estas últimas están muy conservadas entre especies, las CBLs solo se han identificado en plantas superiores, sugiriendo que actúan en procesos de señalización específicos de plantas.
- Las CaMs interaccionan con una gran variedad de proteínas, sin embargo las CBLs interaccionan únicamente con la familia de las PKS/CIPK (Shi *et al.*, 1999).
- En general, las manos EF de las CBLs no son canónicas y presentan una mutación muy conservada en la posición Y, perdiendo de esta manera un ligando implicado directamente en la interacción con el Ca^{2+} (Figura 3.1.1(a) y Figura 3.1.2).

Como se menciona en la Introducción ([Apartado 1.3](#)), la familia de las SCaBPs/CBLs está compuesta por diez miembros, sin embargo poco se sabe de las implicaciones biológicas y del papel que desempeñan, a excepción de SOS3, cada una de ellas (Kolukisaoglu *et al.*, 2004). En relación con el número de sitios canónicos que presentan las CBLs, esta familia se puede dividir en tres grupos, aquellas que presentan un sitio canónico (CBL8 y CBL10), aquellas que presentan dos sitios canónicos (CBL1 y CBL9) y las que no presentan ninguno (CBL2, CBL3, CBL5, CBL6, CBL7 y SOS3) ([Figura 3.1.1\(a\)](#) y [3.1.2](#)). Tal y como ocurre con los sensores neuronales de calcio (NCS) o las recoverinas (Haeseleer *et al.*, 2002), se piensa que la afinidad por el Ca^{2+} , y por tanto la capacidad de reconocer un determinado estímulo celular mediado por dicho catión, está relacionada con el número de sitios canónicos que presenta la proteína y con el grado de conservación de los mismos.

Otra característica importante que presentan algunas CBLs, en concreto SOS3, CBL1, CBL5 y CBL9, es que conservan motivos de miristoilación y palmitoilación (Batistic & Kudla, 2004; Resh, 1999) ([Figura 3.1.3](#)). Se ha comprobado que las cuatro se miristoilan *in vitro*, y además, se ha demostrado que SOS3, CBL1 y CBL9 están localizadas en la membrana (Batistic & Kudla, 2004). Esto sugiere que tanto la afinidad por Ca^{2+} , como la localización celular juegan un papel importante en la transducción de la señal mediada por las CBLs. Además, estas modificaciones permitirán que las proteínas pasen del citosol a la membrana como consecuencia de un determinado estímulo.

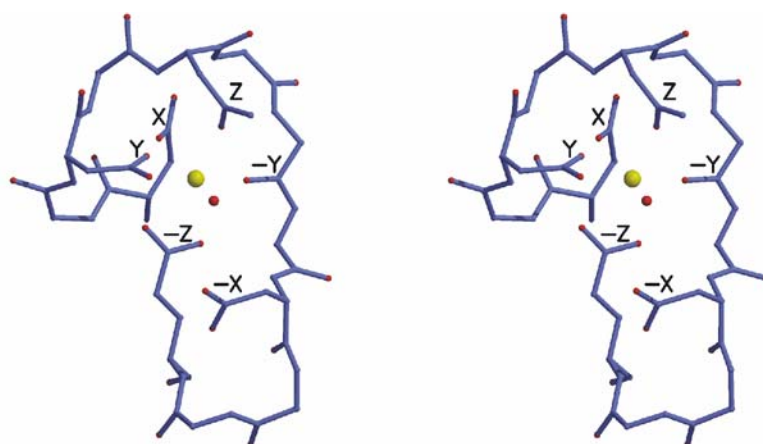
Hasta ahora solo había sido caracterizada desde un punto de vista cristalográfico la proteína CBL2 (Nagae *et al.*, 2003), de cuya su función biológica poco se sabe. Por este motivo, la caracterización estructural de la proteína SOS3 es de gran importancia ya que a partir de los resultados obtenidos y de la comparación con la estructura de CBL2 se van a poder inferir resultados extrapolables al resto de proteínas de la familia.

A continuación se van a discutir los aspectos fundamentales que derivan del análisis de los resultados y que se pueden resumir en los siguientes puntos: los sitios de unión a metales, la dimerización de SOS3, el aumento de la hidrofobicidad de SOS3 por la unión de Ca^{2+} , la comparación estructural de SOS3 con CBL2 y el mecanismo de SOS3 de respuesta al Ca^{2+} e interacción con SOS2.

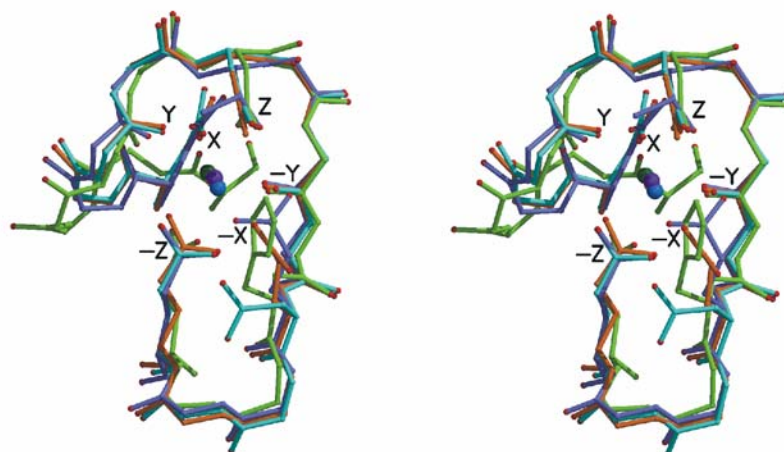
3.1. LOS SITIOS DE UNIÓN A METALES DE SOS3

		X		Y	Z	-Y	-X		-Z						
Consenso		D	-	-	*	D	G	D	G	*	I	*	*	*	E
		N				N	N			S					
EF1	47	S	S	S	I	I	D	D	G	L	I	H	K	E	E
EF2	84	D	-	-	V	K	R	N	G	V	I	E	F	G	E
EF3	112	D	-	-	L	R	Q	T	G	F	I	E	R	E	E
EF4	165	D	-	-	R	K	N	D	G	K	I	D	I	D	E

(a)



(b)



(c)

Figura 3.1.1: (a) Comparación de la secuencia de los sitios EF de unión a Ca^{2+} de SOS3 con la clásica de la superfamilia de las calmodulinas. Los residuos implicados en la unión a Ca^{2+} están indicados con X, Y, Z, -X, -Y, -Z, de acuerdo con la anotación consenso de las manos EF. Los colores rojo y azul corresponden a los oxígenos donadores de cadena lateral y principal, respectivamente. Los residuos conservados aparecen en letra negra (b) Representación estereográfica de la estructura de las manos EF clásicas (las coordenadas corresponden al motivo EF2 de la CnB, código PDB 1aii). (c) Vista estereográfica de la superposición de las estructuras de los sitios de unión a Ca^{2+} de SOS3. Los sitios EF1, EF2, EF3 y EF4 se representan en color verde, lila, naranja y cian.

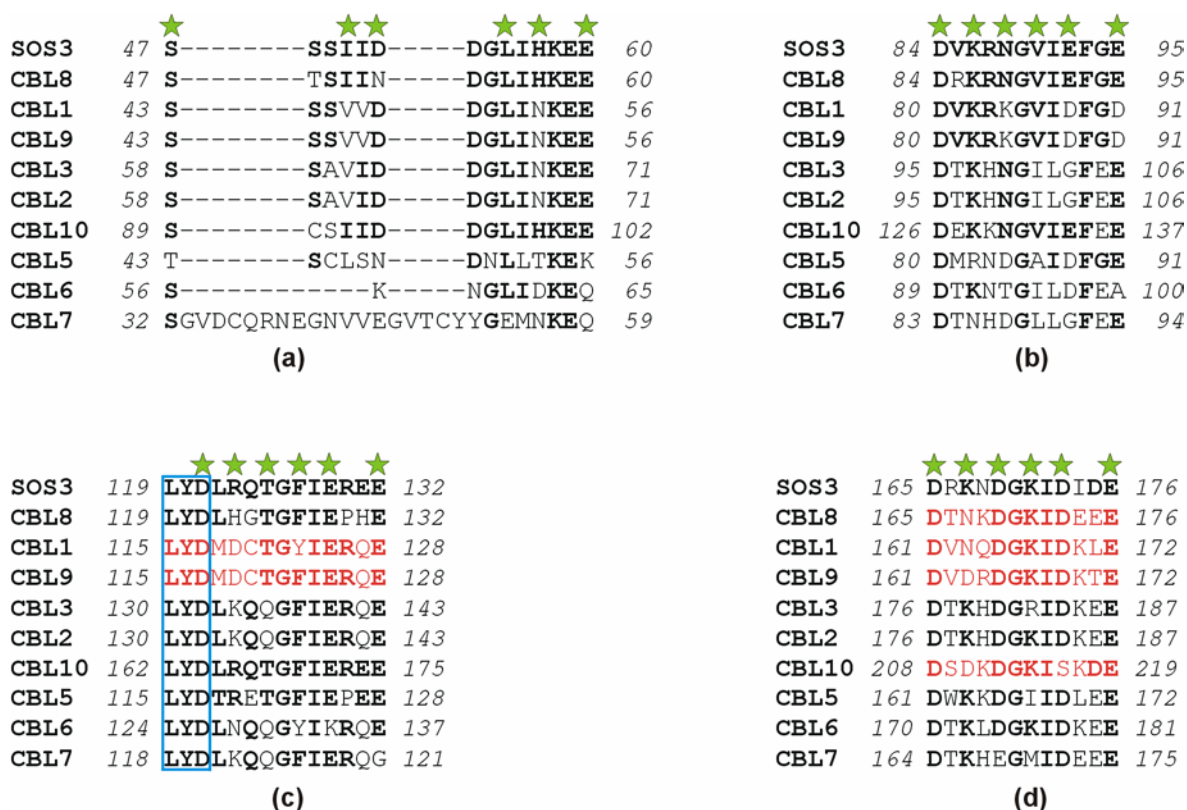


Figura 3.1.2: Alineamiento secuencial de los sitios EF de las CBLs en *A. thaliana* (Pearson & Lipman, 1998). (a) EF1, (b) EF2, (c) EF3 y (d) EF4. Las coincidencias en la secuencia se muestran con letra negra y las manos EF canónicas en rojo. Las estrellas verdes indican los residuos X, Y, Z, -Y, -X y -Z que unen Ca²⁺. El recuadro azul muestra los residuos suprimidos en el mutante inactivo *sos3-1*, que se conservan en toda la familia.

Aunque ninguna de las manos EF de SOS3 es canónica, el mapa de densidad electrónica experimental (Figura 2.3.9(c)) y omitido (Figura 2.3.15) del complejo SOS3 Ca²⁺ ha permitido la identificación de cuatro Ca²⁺ y de sus ligandos peptídicos. Respecto a las manos canónicas, los bucles de SOS3 acomodan las diferencias en secuencia para unir Ca²⁺ (Figura 3.1.1(a), Figura 3.1.1(c) y Figura 3.1.4). Así, el motivo EF1 presenta una inserción de dos aminoácidos entre las posiciones X e Y, y el oxígeno carbonílico de la cadena principal del aminoácido en posición X coordina con el Ca²⁺. Por otro lado, la ausencia de oxígeno donador en la cadena lateral del residuo en la posición Y se suple con un oxígeno carbonílico del esqueleto. A pesar de estos cambios, la geometría de la coordinación del Ca²⁺ es similar a la observada en las manos EF canónicas (Figura 3.1.1(b), Figura 3.1.1(c) y Figura 3.1.4). No obstante, se observan algunas diferencias. En los sitios EF1 y EF2, el Ca²⁺ únicamente interacciona con seis ligandos proteicos, ya que la interacción entre las cadenas laterales de los residuos en las posiciones Z y -X, en el

caso de EF1, e Y y -X, en el caso de EF2, impide la aproximación de otro ligando, dejando al catión inaccesible al solvente (Figura 3.1.4(a) y (b)). En el sitio EF3, el Ca²⁺ presenta una coordinación de bipirámide pentagonal en la que participan siete ligandos proteicos, en lugar de seis y una molécula de agua. Este hecho se debe a que el residuo en la posición -X interacciona directamente con el catión (Figura 3.1.4(c)). Finalmente, el sitio EF4 coordina con siete ligandos, seis proteicos y una molécula de agua (Figura 3.1.4(d)).

SOS3	1	-----MGCSVSKKKKKNAMRPP--G	18
CBL8	1	-----MLAFVKCFSLKRAKHPR--G	18
CBL1	1	-----MGCFHS---KAAKEFR--G	14
CBL9	1	-----MGCFHS---TAAREFP--D	14
CBL3	1	-----MSQCIDGFKHVCSSFFRCFDIDIIYKQS	27
CBL2	1	-----MSQCVDGIKHLCTSVLGCFDLDLYKQS	27
CBL10	1	MTTGRPNNILALKISTRSSSLTVGEQFCAVFIPFFAIIDVLSVSGQCFDCRSTSPRTCQ	60
CBL5	1	-----MGCVCS---KQLEGR--R	13
CBL6	1	-----MMMQCLDGLKHLALILLTCCDADPPKVR	28
CBL7	1	-----MDST--R	5

Figura 3.1.3: Alineamiento secuencial del extremo N-terminal de las CBLs en *A. thaliana* (Pearson & Lipman, 1998). En azul se recuadran los sitios de miristoilación y en amarillo, la secuencia de unión a membrana predicha para SOS3. En letra negra y color rojo se indican los residuos palmitoílable y en negro, los miristoilables.

El análisis conjunto de las estructuras de los sitios de unión a metales de SOS3 (Figura 3.1.1(c)) y CBL2 revela un modo único de unión para las manos EF no canónicas de la familia de las proteínas SCaBP/CBL. Como ocurre con SOS3, el sitio EF1 de CBL2 presenta una inserción de dos aminoácidos entre las posiciones X e Y, y además, el ligando en la posición Y es un carbonilo del esqueleto proteico (Nagae *et al*, 2003). De las diez proteínas que constituyen la familia de las SCaBPs/CBLs, siete de ellas (SOS3, CBL2, CBL1, CBL3, CBL8, CBL9 y CBL10) presentan el sitio EF1 muy conservado, por lo que es de esperar que todas éstas unan Ca²⁺ en este sitio, tal y como lo hacen SOS3 y CBL2. En los restantes sitios (EF2, EF3 y EF4) es de esperar que todas las proteínas que conservan los residuos implicados en la unión del Ca²⁺, coordinen con el oxígeno carbonílico en posición Y, tal y como lo hacen SOS3 y CBL2 (Figura 3.1.2). El análisis comparado de los sitios EF2 y EF3 de SOS3 y CBL2 sugiere por qué estos sitios están desocupados en CBL2. La posición -X del sitio EF2 en CBL2 es una glicina, de forma que se pierde un oxígeno donador para coordinar con el Ca²⁺ (Figura 3.1.2(c)). Por otro lado, en el sitio EF3, CBL2 presenta en la posición Z una glutamina (Figura 3.1.2(d)), cuya cadena lateral es larga e impide el posicionamiento de los otros ligandos peptídicos.

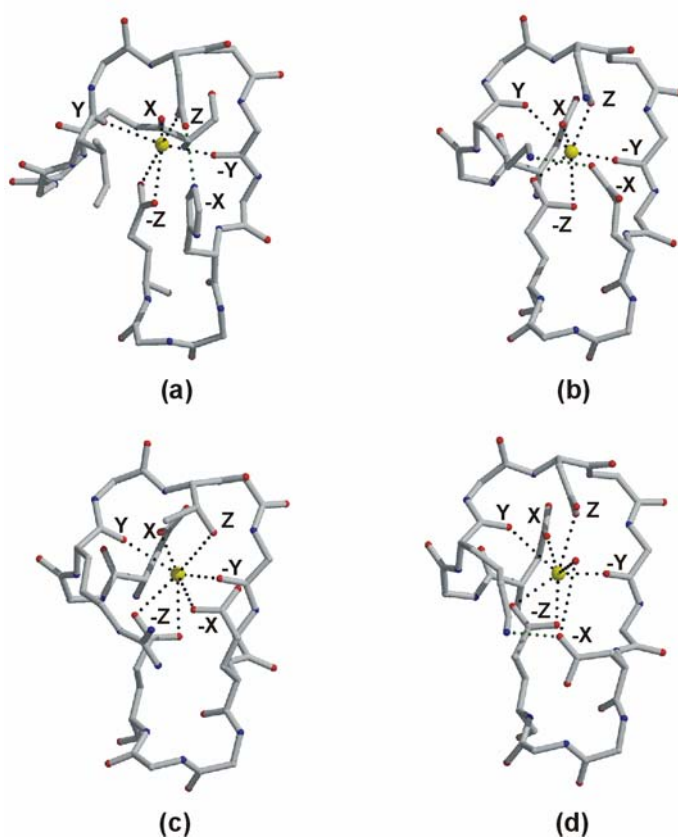


Figura 3.1.4: Estructura de los sitios de unión a Ca^{2+} de SOS3: (a) EF1, (b) EF2, (c) EF3 y (d) EF4. Los ligandos implicados en la unión al catión aparecen etiquetados con la nomenclatura consenso. Las líneas discontinuas negras indican la interacción de los ligandos con el Ca^{2+} , que se representa mediante una bola amarilla. Las líneas discontinuas verdes indican la interacción entre las cadenas laterales de los residuos que impiden la accesibilidad del Ca^{2+} al solvente.

En la estructura de SOS3, tanto el dominio N como C-terminal presentan una pequeña hoja β antiparalela localizada entre sitios adyacentes de unión a Ca^{2+} (Figura 3.1.5 y Figura 2.3.16(a)). Estas hojas β están estabilizadas por la interacción hidrofóbica entre las cadenas laterales de los residuos en posición $-Y$, y el residuo isoleucina conservado en la posición $-Y+1$. Estas interacciones sugieren que SOS3 muestra cooperatividad en la unión a Ca^{2+} entre manos EF adyacentes, tal y como se ha observado para otras proteínas de la superfamilia de las manos EF (Zhang *et al.*, 1995; Kuboniwa *et al.*, 1995; Finn *et al.*, 1995; Verdino *et al.*, 2002). En las calmodulinas, la cooperatividad en la unión del Ca^{2+} es importante para la funcionalidad de las mismas, ya que asegura una activación total de las proteínas en un rango de concentración de Ca^{2+} muy estrecho. Si se realiza una comparación con la estructura de CBL2, es interesante remarcar que este patrón de interacciones no se observa, ya que esta proteína solo presenta un sitio de unión a Ca^{2+} funcional en cada dominio. Esta característica diferencial entre SOS3 y CBL2 debería de ser suficiente como para que dichas proteínas muestren una diferente transducción de la señal del Ca^{2+} .

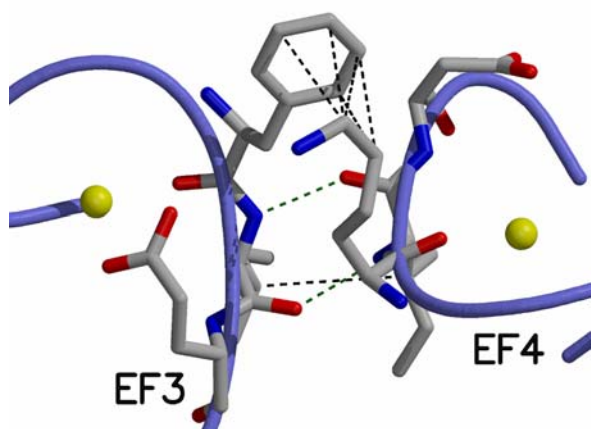


Figura 3.1.5: Interacciones entre los sitios EF3 y EF4 del complejo SOS3 Ca²⁺. Los residuos que forman la hoja β antiparalela y que están implicados en las interacciones se representan en modo de bastones. Las líneas discontinuas verdes corresponden a las interacciones de puente de hidrógeno y las negras a hidrofóbicas. Los iones Ca²⁺ se representan como bolas amarillas.

El análisis de los datos de difracción muestra que SOS3 es capaz de unir Ca²⁺ y Mn²⁺ (Apartados 2.3.2.3 y 2.3.3). Mediante dicroísmo circular, se registró la desnaturalización térmica de SOS3 en presencia de sales de metales, para estudiar qué cationes metálicos es capaz de unir la proteína (Apartado 2.4.3.2). Los resultados confirman que SOS3 une Ca²⁺ y Mn²⁺. Sorprendentemente, nuestro análisis demuestra que, al contrario que otras proteínas con manos EF, SOS3 no une Mg²⁺. Algunas calmodulinas unen este catión en condiciones celulares de reposo, y solo cuando las concentraciones de Ca²⁺ en el citosol son elevadas, el Mg²⁺ es desplazado por el Ca²⁺, de manera que el mecanismo de regulación de las mismas depende de la relación entre las concentraciones de Mg²⁺ y Ca²⁺ (Cates *et al.*, 1999). Nuestros resultados sugieren que la regulación de SOS3, y probablemente la de otras proteínas de la familia, está regida por otro mecanismo.

Es interesante recalcar que la unión de Mn²⁺ a SOS3 podría ser de relevancia biológica, ya que la proteína quinasa SOS2 y otras de la familia de las PKs/CIPKs utilizan el Mn²⁺ como cofactor y/o muestran un aumento de su actividad en presencia del mismo (Shi *et al.*, 1999; Gong *et al.*, 2002b; Luan *et al.*, 2002; Gong *et al.*, 2004). De esta manera, SOS3 y otros miembros de la familia, presentarían un nivel adicional de regulación de sus correspondientes quinasas actuando como transportadores del Mn²⁺ o, alternativamente, como tamponadores de las concentraciones locales de Mn²⁺ libre, para evitar una activación continuada de las quinasas. No obstante, para poder plantear esta hipótesis con solidez, es necesario caracterizar la afinidad de las PKs y las SCaBPs por este catión y medir cuál es la concentración de Mn²⁺ libre en el citosol.

El análisis de la variación de la T_m con el aumento de la concentración de un ligando se puede utilizar para estimar la afinidad de la proteína por ese ligando en particular. Si la desnaturalización

es reversible y la proteína presenta un único sitio de unión, o si presenta varios pero de igual afinidad y donde la unión no esté acoplada, el análisis puede llevarse a cabo con un formulismo similar al descrito en el trabajo de Rivas *et al.* (1992). Este no es el caso de SOS3. Si embargo, de los resultados obtenidos puede deducirse, al menos de forma cualitativa, que la constante de afinidad aparente de SOS3 por el Mn^{2+} debe de ser al menos $10^6 M^{-1}$, ya que se produce un efecto en el desplazamiento de la T_m cuando la concentración de Mn^{2+} libre es del orden de $10^{-6} M$ (Tabla 2.4.2 y Figura 2.4.3). En relación con esta estimación, la afinidad aparente por el Ca^{2+} debería de ser al menos un orden de magnitud inferior, $10^5 M^{-1}$, a la observada para el Mn^{2+} , ya que el efecto que se observa en la T_m a una concentración dada de Mn^{2+} es el mismo que se consigue cuando la concentración de Ca^{2+} es aproximadamente diez veces superior (Tabla 2.4.2 y Figura 2.4.3). Aunque los sensores de Ca^{2+} tradicionales implicados en rutas de transducción de señales en el citosol son capaces de responder a variaciones de Ca^{2+} que se encuentran en el rango del micromolar o submicromolar (Veeraraghavan *et al.*, 2002), se ha especulado que SOS3, al estar miristoilada, podría desempeñar su función en las inmediaciones de la membrana plasmática, cerca de los transportadores de Ca^{2+} , donde la concentración transitoria y local de dicho catión podría ser muy superior, pudiendo de esta manera responder a la señal del Ca^{2+} (Ishitani *et al.*, 2000, White & Broadley, 2003; Evans *et al.*, 2001). Por otro lado, la concentración celular de Mn^{2+} debe ser muy baja, ya que es citotóxico. Por tanto, si la unión de Mn^{2+} por SOS3 tiene sentido fisiológico, sería de esperar que la constante de afinidad fuera alta, tal y como se deduce de los experimentos de desnaturalización.

Como ya se ha comentado en el Apartado 2.4.3.2, la adición de EDTA en la solución de SOS3 no produce ningún efecto en el perfil de desnaturalización de la proteína. Sin embargo, la diálisis exhaustiva de SOS3 frente a EDTA produce una desestabilización general de la proteína, que se refleja en un cambio pronunciado en el perfil de desnaturalización (Figura 2.4.3(a)). Este resultado sugiere que SOS3 presenta sitios estructurales de unión a Ca^{2+} de alta afinidad, de manera que la disociación del catión provoca la desestabilización de la proteína. El análisis de la estructura de SOS3 sugiere que los sitios EF1 y EF2 sean los sitios estructurales, ya que los Ca^{2+} unidos a estos sitios son inaccesibles al solvente (Figura 3.1.4), y que, por tanto, las variaciones en la T_m registradas por CD (Figura 2.4.3(a)) se relacionen con la unión/disociación de los metales en los sitios EF3 y EF4. Algunas calmodulinas también presentan sitios de unión a Ca^{2+} con diferente afinidad, y se ha demostrado que los sitios de baja afinidad tienen una función fisiológica, mientras que los de alta afinidad tienen una función estructural (Burroughs *et al.*, 1994; Stemmer *et al.*, 1994; Kissinger *et al.*, 1995; Otterbein *et al.*, 2002). En este sentido, realizando ensayos de incubación con Ca^{2+} radioactivo, Ishitani *et al.* (2000) mostró que la supresión de tres aminoácidos en la mano EF3 (Figura 3.1.2(c)) producía un mutante no funcional que drásticamente reducía, pero no anulaba, la afinidad por el Ca^{2+} . Según la estructura tridimensional de SOS3, esta

supresión afectaría la unión a Ca^{2+} en los sitios EF3 y EF4, por tanto, este resultado apoya el modelo en el que hay dos tipos de sitios de Ca^{2+} con diferente afinidad. Los sitios EF1 y EF2 darían cuenta de la unión residual con alta afinidad y baja velocidad de intercambio con el Ca^{2+} radioactivo, y EF3 y EF4 de la unión funcional.

3.2. LA ESTRUCTURA DIMÉRICA DE SOS3

En biología, las interacciones moleculares transitorias explican mecanismos para la regulación de la función y los procesos celulares, y dicha transitoriedad en las interacciones refleja la naturaleza dinámica de la célula. Entre otras funciones, los complejos moleculares transitorios están implicados en la transducción de señales intracelulares, ya que su función requiere una rápida asociación y disociación. Una de las características estructurales distintivas de estos complejos, frente a los permanentes es el reducido tamaño de las interfases de contacto, donde solo un número limitado de interacciones tienen que romperse o generarse, para acomodarse más fácilmente a los cambios que se producen en la célula. En estos complejos, el estado de oligomerización depende de factores tales como la concentración de proteína, la unión de un ligando en particular o el entorno físico-químico. Este es el caso de SOS3, que como se ha demostrado, presenta un equilibrio monómero-dímero dependiente de la concentración de Ca^{2+} ([Apartado 2.4.2.1](#))

La estructura cristalina de SOS3 en presencia de Ca^{2+} es un dímero en forma de V ([Apartado 2.3.3](#)) que es compatible con el comportamiento hidrodinámico del dímero en solución ([Apartado 2.4.2](#)) y con su funcionalidad, ya que deja accesible al medio los sitios metálicos y los sitios de miristoilación ([Figura 3.2.1](#)). Además, también deja accesibles los extremos N y C-terminal, que son esenciales para la interacción y activación de SOS2 (Ishitani *et al.*, 2000). La interfase de interacción está constituida por un 4% de la superficie molecular accesible de cada protómero (Wallace *et al.*, 1995). Un 80% de esta área la componen residuos hidrofóbicos (Wallace *et al.*, 1995). Las características de la superficie de contacto no son compatibles con las que se observan para homodímeros estables en solución y, por el contrario, coinciden con los que presentan los homodímeros débiles transitorios (Lo Conte *et al.*, 1999; Nooren & Thornton, 2003a; Nooren & Thornton, 2003b). De hecho, la constante de dimerización calculada para SOS3 ([Apartado 2.4.2.1](#)) está en el rango de las que presentan este tipo de complejos.

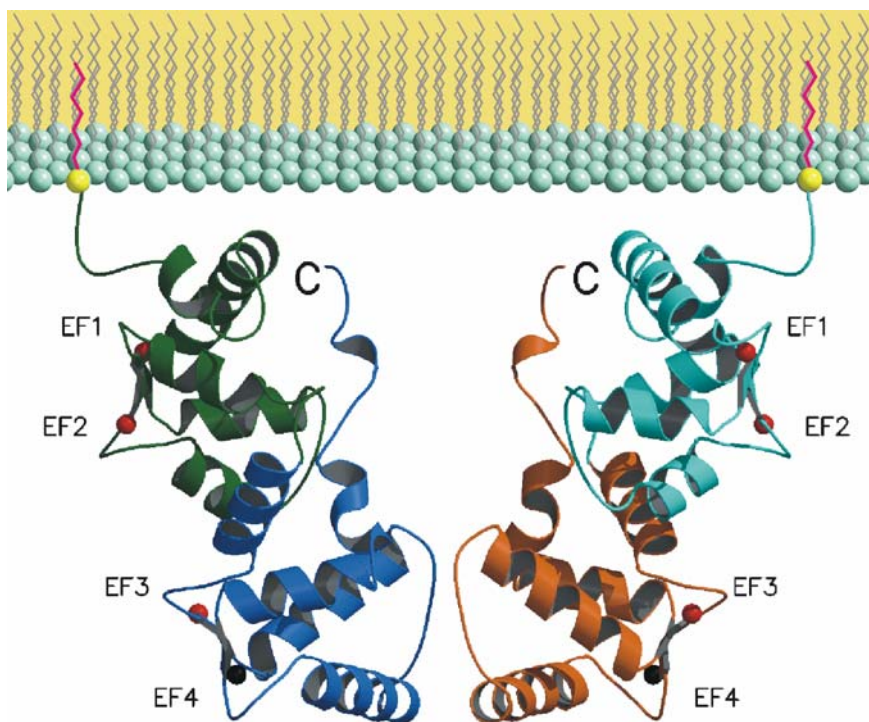


Figura 3.2.1: Imagen de la estructura de SOS3 anclada en una membrana.

Mediante experimentos de equilibrio de sedimentación ([Apartado 2.4.2.1](#)) se ha demostrado que el Ca^{2+} promueve la dimerización de SOS3. Parece razonable que la unión del Ca^{2+} a los sitios EF3 y EF4, o la del Ca^{2+} al sitio EF3 y el Mn^{2+} al sitio EF4 ([Apartado 2.3.2.3](#) y [Figura 2.3.15](#)) sea responsable de la autoasociación de la macromolécula, ya que éstos se encuentran en la interfase de dimerización ([Figura 3.2.1](#), [Apartado 2.3.3](#) y [Figura 2.3.16](#)). Por otro lado, el hecho de que la supresión de tres aminoácidos en la mano EF3 ([Figura 3.1.2\(c\)](#)) (Ishitani *et al.* 2000) produzca un mutante no funcional, y que además este sitio se encuentre implicado en la dimerización de la proteína, sugiere que la unión del Ca^{2+} al sitio EF3 y EF4 y la dimerización son procesos acoplados y fisiológicamente relevantes.

Ishitani *et al.* (2000) demostró que la miristoilación de SOS3 es necesaria para la activación de SOS2, pero no para su asociación a la membrana. De hecho, SOS3 presenta una secuencia de cinco lisinas en el extremo N-terminal ([Figura 3.1.3](#)), cuya interacción con los fosfolípidos tendría que ser suficiente para la unión de SOS3 a la membrana (Monné *et al.*, 1998; Killian & von Heijne, 2000). Esto sugiere que la miristoilación debe estar relacionada con el correcto posicionamiento de SOS3 respecto a la membrana ([Figura 3.2.1](#)). El dímero de SOS3 duplicaría la eficiencia de la unión con respecto al monómero, ya que proporcionaría dos puntos accesibles de unión a la membrana. Además, éstos estarían en la misma orientación, por lo que las restricciones geométricas que implican la formación del dímero, más las restricciones impuestas por la

miristoilación obligarían un posicionamiento del dímero más definido, posiblemente clave para la posterior activación de SOS2.

3.3. AUMENTO DEL CARÁCTER HIDROFÓBICO DE SOS3 POR LA UNIÓN AL Ca^{2+}

Las proteínas de la familia de SOS3 interaccionan con las de la familia de SOS2 a través de un motivo estructural hidrofóbico denominado FISL (Guo *et al.*, 2001) o NAF (Albrecht, *et al.*, 2001) (Apartado 1.2 y 1.3). Los experimentos de cromatografía de hidrofobicidad que se han realizado con SOS3, en presencia y ausencia de Ca^{2+} (Apartado 2.4.5.1 y Figura 2.4.3), han puesto de manifiesto que la unión del catión a SOS3 induce un aumento de la hidrofobicidad de la macromolécula que podría promover la interacción con el motivo FISL de SOS2.

La elevada hidrofobicidad observada para la forma activa de SOS3 por técnicas cromatográficas concuerda con el elevado número de residuos hidrofóbicos expuestos que se observan en la estructura tridimensional del dímero y que significan un 61% de la superficie molecular del mismo (Apartado 2.3.3 y Figura 3.3.1). Estos resultados indican que el elevado carácter hidrofóbico de SOS3 debe ser consecuencia de un cambio conformacional de la macromolécula mediado por la unión al Ca^{2+} que conlleva una exposición elevada de residuos hidrofóbicos en la superficie.

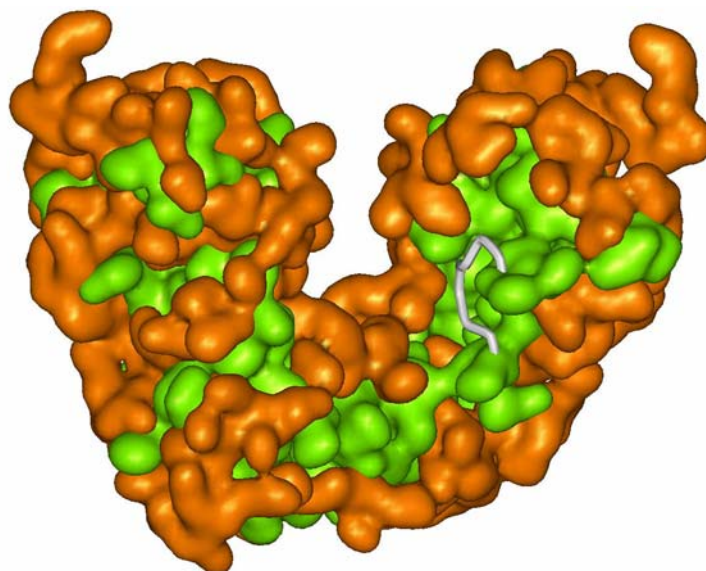


Figura 3.3.1: Representación de la superficie molecular del dímero de SOS3, realizada con Web lab (<http://www.msi.com>). SOS3 está rotada 15° alrededor del eje del dímero respecto a la vista que se muestra en la Figura 3.2.1. Los residuos hidrofóbicos de la superficie se muestran en verde. El bucle móvil, desde el residuo A67 al A73, se ha excluido del cálculo de la superficie molecular y se representa en modo de lazo blanco.

3.4. COMPARACIÓN ESTRUCTURAL DE SOS3 Y CBL2

Las diferencias estructurales y secuenciales entre SOS3 y su homólogo estructural CBL2 se reflejan en un cambio en la forma global y en las propiedades de superficie de las macromoléculas, que dan cuenta de su diferente funcionalidad y de la manera específica que tienen de transmitir la señalización celular mediada por Ca^{2+} (Figura 3.4.1). Por tanto, el análisis comparado de las estructuras de SOS3 y CBL2 con sus homólogos (NCS y CnB) pondrá de manifiesto los determinantes estructurales asociados a sus propiedades.

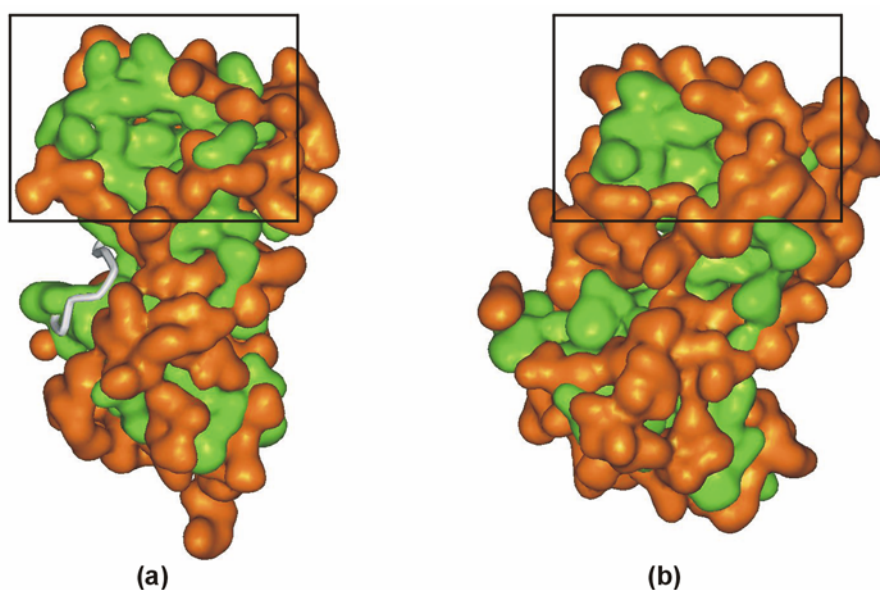


Figura 3.4.1: Representación de la superficie molecular de (a) SOS3 y (b) CBL2, realizada con Web lab (<http://www.msi.com>). La interfase de dimerización aparece recuadrada en negro. Los residuos apolares de la superficie se representan en verde. El bucle móvil de SOS3 (A67-A73) se ha excluido del cálculo de la superficie molecular y se representa en modo de lazo blanco.

La comparación de las estructuras de SOS3 y CBL2 con las de sus homólogos (CnB y NCS1) muestra que el plegamiento general está muy conservado y que las proteínas presentan un elevado número de aminoácidos topológicamente equivalentes (Figura 3.4.2) (Sutcliffe *et al.*, 1987). Las mayores diferencias encontradas entre las proteínas se centran en la posición relativa de los dominios N y C-terminal (Figura 3.4.3(a)). Este efecto puede medirse superponiendo los dominios N-terminales y calculando la distancia entre los centros de masa de los dominios C-terminales y el ángulo que los relaciona (<http://bioinfo1.mbfys.lu.se/~guoguang/fit.html>) (Figura 3.4.4). En la Tabla 3.4.1 se muestran estos parámetros junto con la identidad secuencial y la desviación cuadrática media de los carbonos $\text{C}\alpha$ de SOS3 respecto a CBL2, CnB y NCS1. El análisis de estos datos muestra que no existe ninguna relación entre las posiciones relativas de los dominios, la r.m.s.d. local y la identidad secuencial.

	CBL2	CnB	NCS1
Identidad (%)	50.4	19.8	26.6
Giro entre dominios (grados)	33	26	24
SOS3 Distancia entre C _M (Å)	4.3	3.2	1.4
r.m.s.d. (dominio N-terminal) (Å)	1.67(77)	1.04(70)	1.27(75)
r.m.s.d. (dominio C-terminal) (Å)	1.25(92)	1.40(70)	1.51 (68)

Los valores entre paréntesis corresponden al número de átomos Ca equivalentes (distancia de corte 3.5 Å) en el alineamiento estructural (Sutcliffe *et al.*, 1987). C_M indica centro de masas. Los códigos PDB para SOS3, CBL2, CnB y NCS1 son 1v1g, 1uhn, 1aui y 1g8i, respectivamente.

Tabla 3.4.1: Comparación entre las estructuras de SOS3 y sus proteínas homólogas CBL2, CnB y NCS1.



Figura 3.4.2: Alineamiento estructural de SOS3, CBL2, CnB y NCS1 (códigos PDB 1v1g, 1uhn, 1aui y 1g8i, respectivamente). Los elementos de estructura secundaria de SOS3 (Sayle & Milner-White, 1995) se indican en la parte superior, las hélices α (H), las hebras β (B), los bucles y las regiones desordenadas se representan en color morado, verde, amarillo y azul, respectivamente. Las hélices α de SOS3 y de las proteínas homólogas se indican mediante recuadros morados.

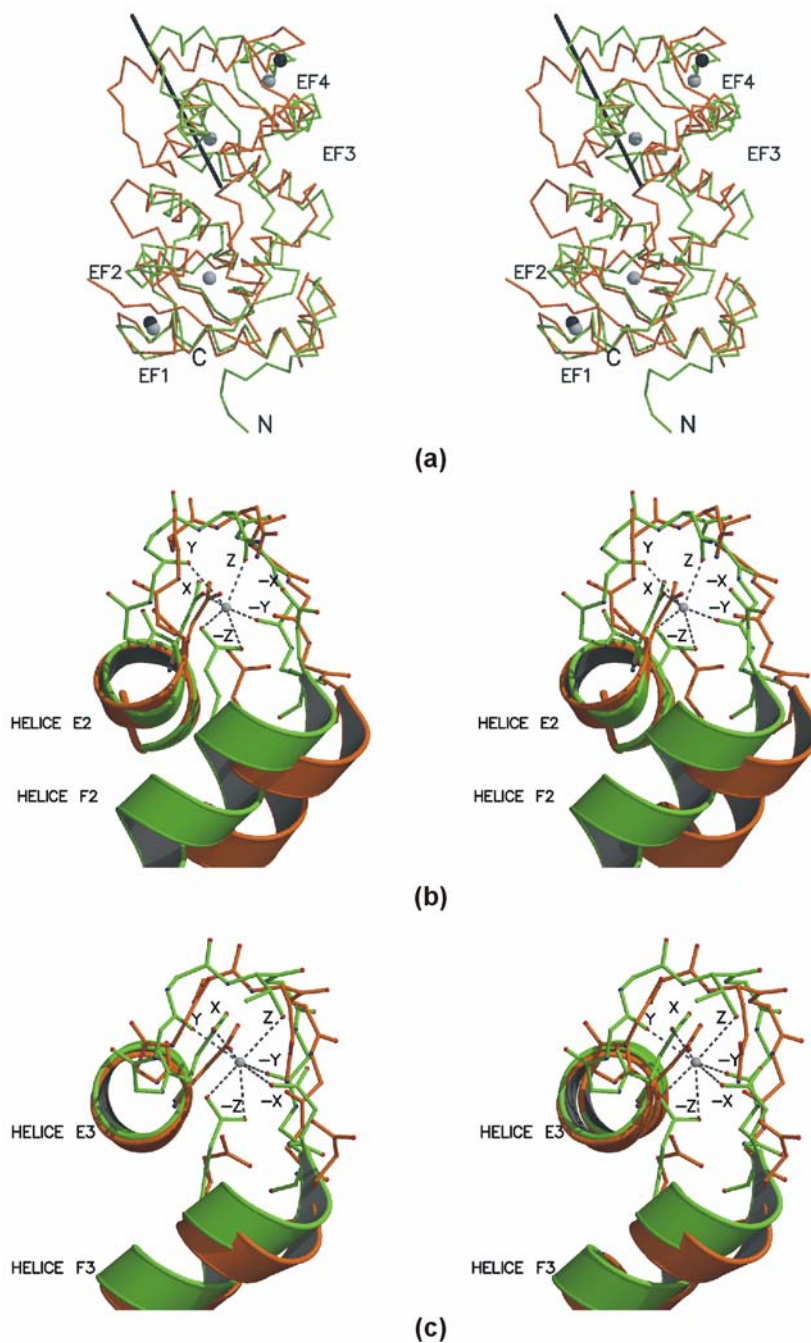


Figura 3.4.3: Vistas estereográficas de la superposición de las estructuras de SOS3 (en verde) y CBL2 (en naranja). (a) Diferencias en el esqueleto proteico cuando se superpone el dominio N-terminal. Los cationes Ca^{2+} se representan con bolas grises y negras en las estructuras de SOS3 y CBL2, respectivamente. El bastón negro representa el eje de rotación que relaciona los dominios C-terminales. Diferencias estructurales en los sitios EF2 (b) y EF3 (c) cuando las hélices E2 y E3 se superponen.

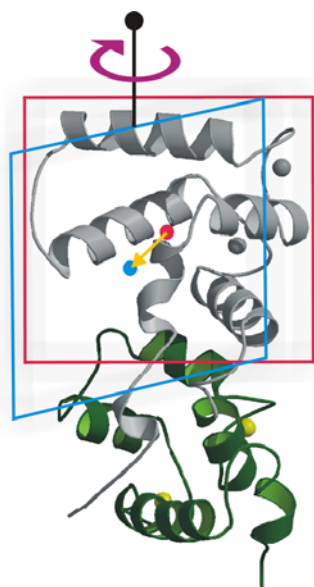


Figura 3.4.4: Esquema del cálculo del giro del dominio C-terminal (en gris) respecto del N-terminal (en verde) entre dos proteínas, cuando el dominio N-terminal se superpone. Para una mejor comprensión únicamente aparece la estructura de una proteína, SOS3. El recuadro rojo representa el dominio C-terminal de SOS3 y el punto del mismo color, el centro de masas de dicho dominio. El recuadro azul representa el dominio C-terminal de una proteína homóloga y el punto, su centro de masas. La flecha amarilla indica la distancia entre los centros de masa. El bastón negro muestra el eje de rotación que relaciona los dominios C-terminales de ambas proteínas. La flecha morada representa el giro que hay que realizar para llevar un dominio al otro.

El análisis del alineamiento estructural de las proteínas (Figura 3.4.2) sugiere que las diferencias estructurales deben estar relacionadas con la distinta longitud de los bucles que conectan las manos EF, así como con pequeños cambios locales en la secuencia de aminoácidos. Sin embargo, SOS3 y CBL2 muestran las mayores diferencias entre la posición relativa de los dominios, a pesar de que secuencialmente son las más parecidas, presentan inserciones de la misma longitud en todos sus bucles y pertenecen a la misma familia de proteínas (Figura 3.4.3(a)). Por tanto, la diferente ocupación de los sitios metálicos podría dar cuenta de las diferencias encontradas en las estructuras tridimensionales. Como ya se ha mencionado, CBL2 une Ca^{2+} en los sitios EF1 y EF4, mientras que los sitios EF2 y EF3 se encuentran en conformación abierta estabilizados por interacciones electrostáticas con moléculas de agua. La comparación estructural de SOS3 y CBL2 en estos sitios revela que existen diferencias apreciables en las posiciones relativas de las hélices E2 y F2 y E3 y F3. En la Figura 3.4.3 se muestra la superposición de los sitios EF2 y EF3 de SOS3 y CBL2. Como se puede observar, la interacción del ligando bidentado glutamato, en posición $-Z$, con el Ca^{2+} provoca un desplazamiento general de las hélices F2 y F3 con respecto a las hélices E2 y E3. La reorganización en los sitios EF2 y EF3 por la unión al Ca^{2+} podría ser la responsable del cambio en la posición relativa de los dominios N y C-terminal, ya que estos sitios se encuentran en la interfase de los dominios.

Teniendo en cuenta la relación entre la unión a Ca^{2+} y la dimerización, los datos presentados en esta memoria sugieren que la unión de Ca^{2+} , el giro entre los dominios y la dimerización son procesos acoplados. Estos cambios globales en la estructura, motivados por la unión del Ca^{2+} , han sido demostrados para otras proteínas con manos EF (Lewit-Bentley & Rety (2000); Otterbein *et al.*, 2002; Ames *et al.*, 2002). No se han realizado experimentos biofísicos para la caracterización del estado de asociación de otras proteínas de la familia, sin embargo, la estructura tridimensional de CBL2 muestra que la proteína es monomérica. El análisis comparado de las superficies moleculares de SOS3 y CBL2 muestra que CBL2 presenta en su superficie un 25% menos de residuos hidrofóbicos expuestos (Wallace *et al.*, 1995) (Figura 3.4.1) en la zona por donde SOS3 dimeriza. Estas diferencias vienen fundamentalmente determinadas por cambios locales en la secuencia aminoacídica de las hélices F3 y E4 y el bucle que las une, que son los elementos estructurales que constituyen la interfase de dimerización (Figura 3.4.2). Por el contrario, la elevada identidad secuencial que presenta SOS3 con algunas de las proteínas de la familia (como CBL8, CBL1 y CBL9, con identidades secuenciales del 59.9, 59.0 y 58.5%, respectivamente) sugiere que otros miembros de las proteínas CBL puedan realizar el mismo mecanismo de regulación por Ca^{2+} basado en la dimerización (Figura 3.4.5).



Figura 3.4.5: Detalle del alineamiento secuencial (Pearson & Lipman, 1998) de SOS3, CBL8, CBL1, CBL9 y CBL2 en la región de dimerización de SOS3. Los elementos de estructura secundaria de SOS3 y CBL2 se muestran en la parte superior. El recuadro amarillo indica la región de dimerización y los residuos de SOS3 implicados directamente en la misma se muestran en verde. Las coincidencias en la secuencia se indican con letras en negrita. Las variaciones en la secuencia más graves, que probablemente afectarían a la dimerización de las proteínas comparadas, se muestran en rojo. CBL2 presenta las mayores diferencias, de ahí que no dimerice tal y como lo hace SOS3.

3.5. MECANISMO DE ACTIVACIÓN DE SOS3

La interacción del dominio FISL de SOS2 con SOS3 debe implicar necesariamente contactos hidrofóbicos entre ambas proteínas, por lo que la forma activa de SOS3 debe presentar residuos hidrofóbicos en su superficie para llevar a cabo dicha interacción.

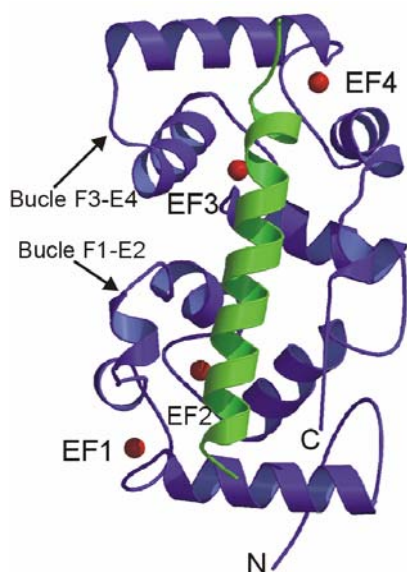


Figura 3.5.1: Detalle de la estructura de la Calcineurina. La subunidad reguladora, CnB, se muestra en morado. La hélice α de color verde corresponde a la región de la CnA que interacciona con la CnB. Los cationes Ca^{2+} se representan con bolas rojas.

Nagae *et al.* (2003) propusieron un mecanismo de reconocimiento entre CBL2 y su quinasa diana similar al observado para la calcineurina. En esta proteína, la subunidad fosfatasa (CnA) expande una hélice hacia una hendidura hidrofóbica definida entre el extremo C-terminal y los bucles F1-E2 y F3-E4 de la subunidad reguladora CnB (Figura 3.5.1). La estructura monomérica de CBL2 muestra una conformación cerrada de esta hendidura y se sugiere que el Ca^{2+} podría inducir un giro entre los dominios que llevara a un cambio conformacional, de forma que se expusieran los residuos hidrofóbicos ocultos en esta cavidad (Nagae *et al.*, 2003). Esto permitiría una interacción de CBL2 con el motivo FISL de su quinasa diana. Sin embargo, el análisis estructural de SOS3 sugiere que la activación de SOS2 debería estar regida por otro mecanismo, fundamentalmente por dos motivos, primero porque la superposición estructural de SOS3, que tiene todos los sitios de Ca^{2+} ocupados, y CBL2 (Figura 3.4.3(a)), muestra que el giro entre dominios tiende a cerrar esta hendidura y segundo, porque la estructura dimerica de SOS3 (Figura 2.3.16(a)) impediría la interacción con SOS2, tal y como se observa en el complejo CnB-CnA.

El análisis de la superficie molecular del dímero de SOS3 indica la presencia de una zona hidrofóbica con forma alargada que incluye residuos de la interfase entre los dominios, residuos de la interfase de dimerización y residuos EF que se encuentran bajo un bucle móvil localizado entre las hélices F1 y E2 (Figura 3.3.1). El bucle está altamente cargado, siendo la secuencia del mismo Arg-Asn-Arg-Asn-Arg-Arg-Asn. Esta zona hidrofóbica podría ser la responsable de la interacción con el dominio FISL de SOS2. Es interesante comprobar que en CBL2, este bucle está conservado y se encuentra estabilizado mediante contactos con el dominio C-terminal, de forma que la zona hidrofóbica queda parcialmente cubierta. En particular, se produce un puente salino entre el residuo Lys A78 del bucle (Arg A67 en SOS3) y el Glu A154 del extremo C-terminal de la

hélice F3 (Glu A143 en SOS3) (Figura 3.5.2). Aunque secuencialmente SOS3 conserva estos residuos, la disposición relativa de los dominios N y C-terminal impide la formación de estos contactos.

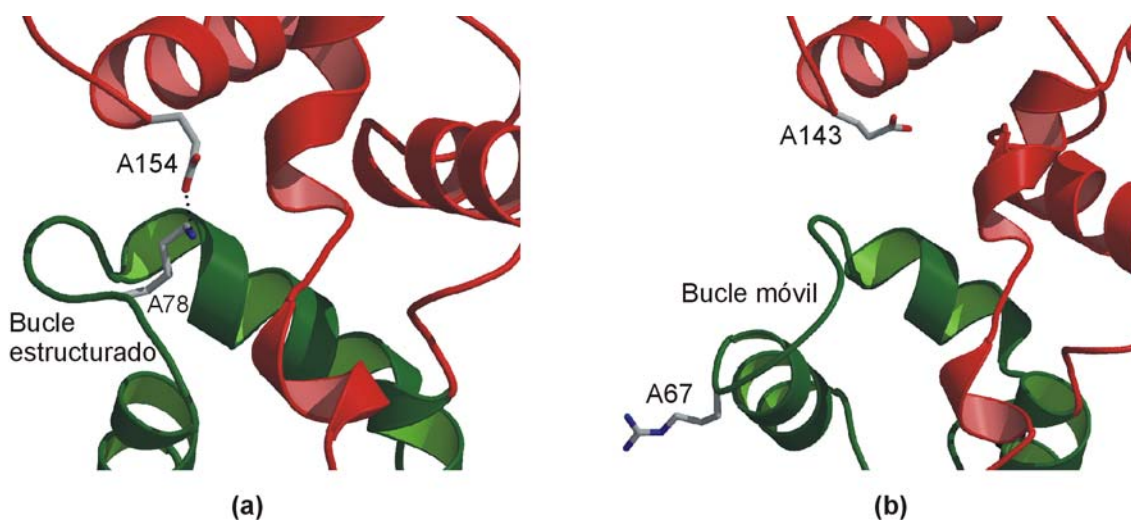


Figura 3.5.2: Detalle de la estructura CBL2 (a) y SOS3 (b). El dominio N-terminal se representa en modo de lazo verde y el C-terminal en rojo. Los aminoácidos implicados en la estabilización del bucle F1-E2 de CBL2 (y los equivalentes de SOS3) se representan en modo de bastones. Se han superpuesto los dominios C-terminal.

El análisis comparado de la unión del Ca^{2+} a SOS3 y CBL2 sugiere que, como ocurre con otras proteínas, las manos EF actuarán como bisagras moleculares que se abrirán o cerrarán con la unión a dicho catión (Figura 3.4.2(b) y (c)). La elevada identidad secuencial de SOS3 y CBL2 permite construir un modelo tridimensional de SOS3 en el que los sitios EF2 y EF3 están desocupados (Figura 3.5.3(a)). La superposición de los dominios N y C-terminal de SOS3 en CBL2 resultaría en un modelo de SOS3 (híbrido-1) con una arquitectura molecular que impediría la formación del dímero en forma de V, tal y como ocurre con CBL2, y que estabilizaría el bucle F1-E2 sobre la zona hidrofóbica que se ha propuesto que interacciona con el FISL de SOS2. No obstante, los datos experimentales presentados en esta memoria sugieren que el sitio EF2 no intercambia Ca^{2+} con el medio. Por tanto, se ha construido otro modelo (híbrido-2) en el que se superpone la región comprendida entre la hélice F3 y el extremo C-terminal de SOS3 en CBL2 (Figura 3.5.3(b)). En este modelo, la reorganización estructural entre dominios es menor pero interesantemente, el extremo C-terminal de la hélice F3 se aproxima al bucle F1-E2, permitiendo la estabilización del bucle sobre la región hidrofóbica.

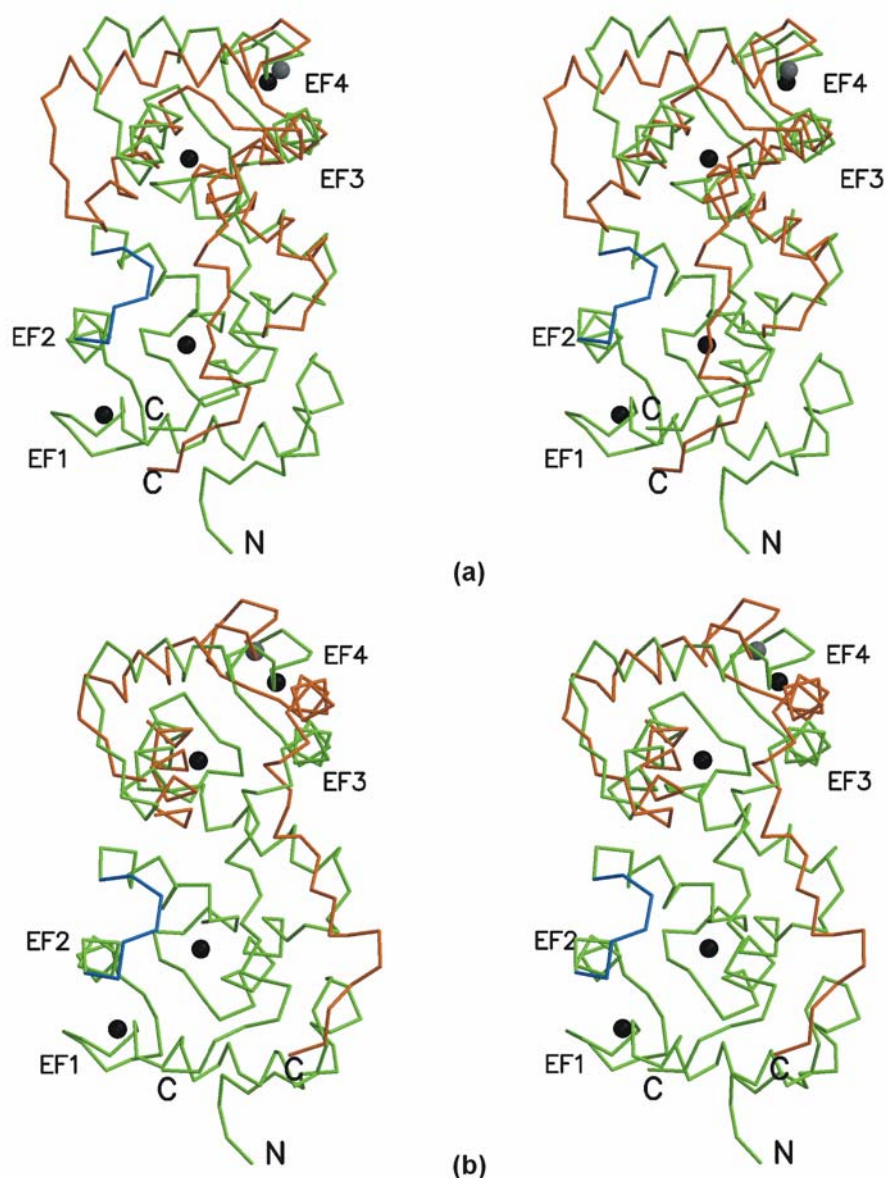


Figura 3.5.3: Vistas estereográficas de la superposición de la estructura del complejo SOS3 Ca²⁺ (en verde) y los modelos de SOS3 libres de Ca²⁺ propuestos (en naranja), (a) híbrido-1 y (b) híbrido-2. Únicamente se muestra las regiones de los modelos que difieren respecto a SOS3. Los cationes Ca²⁺ del complejo y de los modelos híbridos se representan mediante bolas negras y grises, respectivamente. El bucle móvil de SOS3 se indica en color azul.

Por tanto, sobre la base de los datos experimentales obtenidos (estructura cristalográfica y técnicas estructurales a baja resolución) y a la comparación de CBL2 y SOS3, es posible especular acerca de un modelo de activación de SOS3 según el cual, la unión del Ca²⁺ provocará cierto giro entre los dominios N y C-terminal que romperá la interacción entre el bucle F1-E2 y el dominio C-terminal. Así, este bucle se desestabilizará y se generará el área hidrofóbica alargada que podría ser adecuada para la interacción con el motivo FISL de SOS2.

3.6. MECANISMO MOLECULAR DE SOS3 EN LA RUTA SALT OVERLY SENSITIVE

Con los resultados obtenidos, es posible proponer un mecanismo de acción de SOS3 en el contexto de la ruta SOS y de su medio celular.

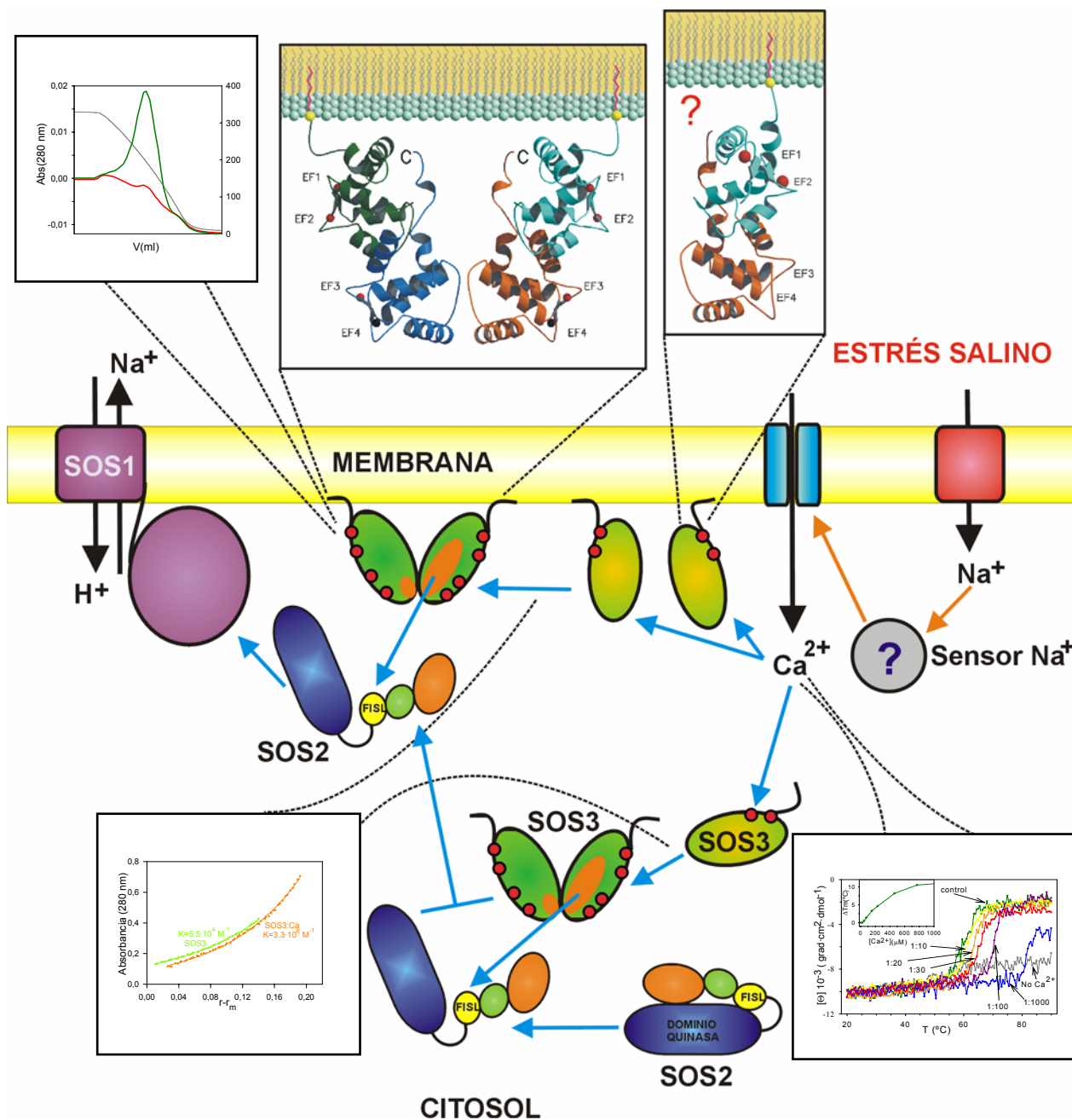


Figura 3.6.1: Esquema de la ruta SOS. Funcionalidad del dímero de SOS3 en la ruta.

En condiciones de estrés salino, se produce la entrada de Na^+ a la planta a través de las raíces, Na^+ que ha de ser eliminado. El aumento de la concentración de Na^+ en el citosol lo percibirá un sensor de Na^+ , que activará los mecanismos de entrada de Ca^{2+} en el citosol (Figura 3.6.1). Hasta el momento, se desconoce el sensor de Na^+ capaz de activar la respuesta celular a este estímulo externo. El aumento de la concentración citosólica de Ca^{2+} activa a SOS3, que en condiciones de reposo se encuentra monomérica en el citosol o anclada en la membrana (Figura 3.6.1). La unión del Ca^{2+} a SOS3 trae consigo un cambio conformacional en la proteína, que provoca la dimerización de la misma y el aumento de su hidrofobicidad (Figura 3.6.1). SOS3 activa interaccionará con el dominio FISL de SOS2 a través de residuos hidrofóbicos expuestos, activando a la quinasa. La miristoilación de SOS3 determinará el posicionamiento correcto del complejo SOS3-SOS2 y permitirá que SOS2 interaccione y active al antiporteador SOS1, que se encargará de la evacuación del Na^+ al medio extracelular. Además, la especie activa de SOS3 podrá rescatar a SOS2 del citosol y conducirla hacia la membrana, permitiendo la colocalización de SOS2 y SOS1 (Figura 3.6.1).

Como se ha mencionado en la Introducción (Apartado 1.3), las CBLs son sensores de Ca^{2+} que responderán a estímulos ambientales de un modo análogo a como lo hace SOS3 en la ruta SOS. Cada CBL tendrá que descifrar de un modo específico la señal del Ca^{2+} . Para entender dichos mecanismos de respuesta sería necesario realizar estudios experimentales para corroborar que las diferencias en afinidad esperadas para esta familia de proteínas contribuyen a descifrar las distintas señales del Ca^{2+} , correspondientes a distintos estímulos ambientales. Para ello queda un largo camino por recorrer, ya que no solo habría que investigar acerca de las concentraciones de Ca^{2+} libre en el citosol como respuesta a un estímulo concreto, sino que también habría que estimar las afinidades que presentan las CBLs por el Ca^{2+} , conocer las parejas calmodulina-quinasa que intervienen en la respuesta a un estímulo concreto, caracterizar la interacción entre ellas, la dependencia del Ca^{2+} en dicha interacción, la localización subcelular de los complejos, el sustrato de la quinasa y los cofactores necesarios para su actividad y por último, la localización del sistema en la planta.

4. CONCLUSIONES

4. CONCLUSIONES

1. Se han puesto de manifiesto las capacidades del anión yoduro como faseador de estructuras cristalográficas y su versatilidad en la aplicación de distintas metodologías de resolución estructural (SAD, SIRAS, MAD, MIR, etc) mediante el análisis de dos casos prácticos, el faseado de las estructuras de las proteínas SOS3 y AS-48.
2. La estructura de SOS3 presenta el plegamiento típico de las calmodulinas, que consiste en dos dominios unidos por un conector. Cada dominio está formado por un par de manos EF. La proteína une cuatro Ca^{2+} , pudiendo unir también Mn^{2+} al cuarto sitio EF. La estructura cristalográfica de SOS3 en complejo con Ca^{2+} y Ca^{2+} y Mn^{2+} es un dímero en forma de V. La estructura cuaternaria deja accesible al solvente los ocho sitios de unión a Ca^{2+} , el sitio de miristoilación y los extremos N y C-terminal. Las características de la superficie de contacto sugieren que la dimerización es transitoria.
3. La superficie molecular del dímero de SOS3 es altamente hidrofóbica, exponiendo un 61% de residuos hidrofóbicos al solvente.
4. El análisis de los sitios de unión a Ca^{2+} y el modo de unión de SOS3 a dicho catión ha revelado una nueva manera de unión a Ca^{2+} en proteínas con manos EF no canónicas. Las interacciones que se han encontrado entre los pares de manos EF sugiere que existe cooperatividad en la unión al Ca^{2+} . La comparación estructural de los sitios EF de SOS3 y CBL2 ha permitido demostrar que las proteínas de la familia presentan un modo único de unión a Ca^{2+} para las manos EF no canónicas.
5. Los experimentos de desnaturalización monitorizada por dicroísmo circular y de emisión de fluorescencia del triptófano han demostrado que SOS3 sólo une Ca^{2+} y Mn^{2+} . La afinidad por el Ca^{2+} es similar a la observada para otros sensores. El papel de la unión al Mn^{2+} puede ser fisiológicamente relevante dada la alta afinidad que presenta SOS3 por el mismo.
6. Los experimentos de cromatografía de interacciones hidrofóbicas han revelado que la unión del Ca^{2+} a SOS3 provoca un aumento del carácter hidrofóbico de la macromolécula.
7. Los experimentos de equilibrio de sedimentación de SOS3 en presencia de diversos cationes metálicos han puesto de manifiesto que únicamente el Ca^{2+} promueve la dimerización de la macromolécula. La comparación de los resultados de los experimentos

de velocidad de sedimentación con la predicción teórica de propiedades hidrodinámicas de SOS3, a partir de la estructura cristalina, ha demostrado que el comportamiento del dímero en solución es compatible con la estructura cristalográfica.

8. La integración de los resultados obtenidos, los estudios bioquímicos y genéticos existentes y la comparación de la estructura de SOS3 con sus homólogos estructurales han permitido proponer un mecanismo de respuesta al Ca^{2+} e interacción con SOS2 según el cual la unión del Ca^{2+} a SOS3 provoca un cambio conformacional que induce la dimerización de la proteína y un aumento del carácter hidrofóbico de la misma. Esto permite la interacción de SOS3 con el motivo hidrofóbico FISL de SOS2 y la activación de la función quinasa de la misma.

5. BIBLIOGRAFÍA

5. BIBLIOGRAFÍA

- Abrahams, J.P. & Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F₁ ATPase. *Acta Cryst. D*, **52**, 30-42.
- Albrecht, V., Ritz, O., Linder, S., Harter, K. & Kudla, J. (2001) The NAF domain defines a novel protein-protein interaction module conserved in Ca²⁺-regulated kinases. *EMBO J.*, **20**(5), 1051-1063.
- Albrecht, V., Weinl, S., Balzevic, D., D'Algelio, D., Batistic, O., Kolukisaoglu, U., Bock, R., Schulz, B., Harter, K. & Kudla, J. (2003) The calcium sensor CBL1 integrates plant responses to abiotic stresses. *Plant J.*, **36**, 457-470.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.
- Ames, J.B., Hamasaki, N. & Molchanova, T. (2002) Structure and calcium-binding studies of a recoverin mutant (E85Q) in an allosteric intermediate state. *Biochemistry*, **41**(18), 5776-5787.
- Amtmann, A. & Sanders, D. (1999) Mechanism of Na⁺ uptake by plant cells. *Adv. Bot. Res.*, **29**, 75-112.
- Anderson, K.M. & Hovmöller, S. (2000) The protein content in crystals and packing coefficients in different space groups. *Acta Cryst. D*, **56**, 789-790.
- Apse, M.P. & Blumwald, E. (2002) Engineering salt tolerance in plants. *Curr. Op. Biotechnol.*, **13**, 146-150.
- Bailey, S. (1994) The CCP4 suite, programs for protein crystallography. *Acta Cryst. D*, **50**, 760-763.
- Batistic, O & Kudla, J. (2004) Integration and channeling of calcium signaling through the CBL calcium sensor/CIPK protein kinase network. *Planta*, **219**(6), 915-924.
- Blessing, R.H. & Smith, G.D. (1999) Difference structure-factor normalization for heavy-atom or anomalous-scattering substructure determinations. *J. Appl. Cryst.* **32**(4), 664-670.
- Blumwald, E. (2000) Sodium transport and salt tolerance in plants. *Curr. Opin. Cell. Biol.*, **12**, 431-434.
- Blundell, T.L. & Johnson, L.N. (1976) *Protein Crystallography*. Academic Press, Nueva York.
- Böhm, G., Muhr, R. & Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neuronal networks. *Protein Engineering*, **5**, 191-195.
- Bourne, Y., Dannenberg, J., Pollmann, V., Marchot, P. & Pongs, O. (2001) Immunocytochemical localization and crystal structure of human frequenin (neuronal calcium sensor 1) *J. Biol. Chem.* **276**(15), 11949-11955.
- Branden, C. & Tooze, J. (1991) *Introduction to protein structure*. Garland Pub., Nueva York.

- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L. (1998) Crystallography and NMR System. *Acta Cryst. D*, **54**, 905-921.
- Burgoyne, R.D. & Weiss, J.L. (2001) The neuronal calcium sensor family of Ca²⁺-binding proteins. *Biochem. J.*, **353**, 1-12.
- Burroughs, S.E., Horrocks, W.D. Jr., Ren, H. & Klee, C.B. (1994) Characterization of the lanthanide ion-binding properties of calcineurin-B using laser-induced luminescence spectroscopy. *Biochemistry*, **33**(34), 10428-10436.
- Cates, M.S., Berry, M.B., Ho, E.L., Li, Q., Potter, J.D. & Phillips, G.N. (1999) Metal-ion affinity and specificity in EF-hand proteins: coordination geometry and domain plasticity in parvalbumin. *Structure*, **7**, 1269-1278.
- Cheong, Y.H., Kim, K.N., Pandey, G.K., Gupta, R., Grant, J.J. & Luan, S. (2003) CBL1, a calcium sensor that differentially regulates salt, drought, and cold responses in *Arabidopsis*. *Plant Cell*, **15**, 1833-1845.
- Cowtan K. & Main, P. (1998) Miscellaneous algorithms for density modification. *Acta Cryst. D*, **54**, 487-493.
- Cyert, M.S. & Thorner, J. (1992) Regulatory subunit (CNB1 gene product) of yeast Ca²⁺/calmodulin-dependent phosphoprotein phosphatases is required for adaptation to pheromone. *Mol. Cell Biol.*, **12**(8), 3460-3469.
- Dauter, Z., Dauter, M & Rajashankar, R. (2000) Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Cryst. D*, **56**, 232-237.
- Dauter, Z. & Dauter, M. (2001) Entering a new phase: using solvent halide ions in protein structure determination. *Structure* **8**, R21-R26.
- Dauter, Z., Li, M. & Wlodawer, A. (2001) Practical experience with the use of halides for phasing macromolecular structures: a powerful tool for structural genomics. *Acta Cryst. D*, **57**, 239-249.
- Debreczeni, J.E., Bunkoczi, G., Ma, Q., Blaser, H. & Sheldrick, G.M. (2003) In-house measurement of the sulfur anomalous signal and its use for phasing. *Acta Cryst. D*, **59**, 668-696.
- Ducruix, A. & Giegé, R. (1992) Crystallization of protein and nucleic acids: a practical approach, IRL Press at Oxford University Press, Oxford, Nueva York.
- Dunn, M.F., Pattison, S.E., Storm, M.C. & Quiel, E. (1980) Removal of metals from enzyme solutions. *Biochem.* **19**, 718.
- Evans, G & Bricogne, G. (2003) Triiodine derivatization in protein crystallography. *Acta Cryst. D*, **59**, 1923-1929.
- Evans, N.H., McAinsh, M.R., and Hetherington, A.M. (2001) Calcium oscillations in higher plants. *Curr. Opin. Plant Biol.*, **4**(5), 415-420.

- Evans, P.R. (1993) SCALA: Data reduction. *Proceedings of CCP4 Study Weekend on Data Collection & Processing*, 114-122.
- Finn, B.E., Evenas, J., Drakenberg, T., Waltho, J.P., Thulin, E. & Forsen, S. (1995) Calcium-induced structural changes and domain autonomy in calmodulin. *Nat. Struct. Biol.*, **2(9)**, 777-83.
- de la Fortelle, E. & Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous methods. *Methods Enzimol.* **276**, 472-494.
- García de la Torre, J. (2001) Hydration from hydrodynamics. General considerations and applications of bead modelling to globular proteins. *Biophysical Chemistry* **93**, 159-170.
- García de la Torre, J., Huertas, L. & Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophysical Chemistry* **78**, 719-730.
- García-Segura, J.M., Gavilanes, J.G., Martínez del Pozo, A. Montero, F., Oñaderra, M. & Vivanco, F. (1996) *Técnicas instrumentales de análisis de bioquímica*. Editorial Síntesis. Madrid.
- Garman, E.F. & Schneider, T.R. (1997) Macromolecular crystallography. *J. Appl. Cryst.* **30**, 211-237.
- Gavira, J.A., Toh, D., López-Jaramillo, J., García-Ruiz, J.M. & Ng, J.D. (2002) *Ab initio* crystallographic structure determination of insulin from protein to electron density without crystal handling. *Acta Cryst. D*, **58**, 1147-1154.
- Gómez-Moreno C. & Sancho J. coords (2003) *Estructura de proteínas*. Editorial Ariel. Barcelona.
- Gong, D., Zhang, C., Chen, X., Gong, Z. & Zhu, J.K. (2002a) Constitutive activation and transgenic evaluation of the function of an Arabidopsis PKS protein kinase. *J. Biol. Chem.*, **277**, 42088-42096.
- Gong, D., Guo, Y., Jagendorf, A.T., & Zhu, J.K. (2002b) Biochemical characterization of the Arabidopsis protein kinase SOS2 that functions in salt tolerance. *Plant Physiol.*, **130(1)**, 256-264.
- Gong, D., Guo, Y., Schumaker, K.S. & Zhu, J.K. (2004) The SOS3 Family of Calcium Sensors and SOS2 Family of Protein Kinases in Arabidopsis. *Plant Physiol.*, **134(3)**, 919-926.
- Guo, Y., Halfter, U., Ishitani, M. & Zhu, J.K. (2001) Molecular characterization of functional domains in the protein kinase SOS2 that is required for salt tolerance. *Plant Cell*, **13**, 1383-1399.
- Guo, Y, Qiu, Q, Quintero, F.J., Pardo, J.M., Otha, M., Zhang, C., Schumaker, K.S. & Zhu, J.K. (2004) Transgenic evaluation of activated mutant alleles of SOS2 reveals a critical requirement for its kinase activity and C-terminal regulatory domain for salt tolerance in Arabidopsis thaliana. *Plant Cell*, **16**, 435-449.
- Greenfield, .J. (1996) Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry* **235**, 1-10.

- Haeseleer, F., Imanishi, Y., Sokal, I., Filipek, S., Palczewski, K. (2002) Calcium-binding proteins: intracellular sensors from the calmodulin superfamily. *Biochem. Biophys. Res. Commun.*, **290**(2), 615-623.
- Halfter, U., Ishitani, M. & Zhu, J.K. (2000) The *Arabidopsis* SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3. *Proc. Natl. Acad. Sci. USA*, **97**, 3735-3740.
- Hardie D.G. (1999) Plant protein serine/threonine kinases: Classification and functions. *Annu. Rev. Plant. Physiol.*, **50**, 97-131.
- Harmon A.C., Gribskov M. & Harper J.F. (2000) CDPKs, a kinase for every Ca²⁺ signal? *Trends Plant Sci.*, **5**, 154-159.
- Hasegawa, P. M., Bressan, R. A., Zhu, J.K. & Bohnert, H. J. (2000) Plant cellular and molecular responses to high salinity. *Annu. Rev. Plant Mol. Plant Physiol.*, **51**, 463-499.
- van Holde, K.E. (1985) *Sedimentation. Physical Biochemistry*. Englewood Cliffs. NJ: Prentice Hall.
- Howell, L. & Smith, D. (1992) SCALEIT: Normal probability analysis. *J. Appl. Cryst.*, **25** 81-86.
- Huang, Q.Q., Teng, M.K. & Niu, L.W. (1999) Protein crystallization with a combination of hard and soft precipitants. *Acta Cryst. D*, **55**, 1444-1448.
- Ishitani, M., Liu, J., Halfter, U., Kim, C.S., Shi, W. & Zhu, J.K. (2000) SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell*, **12**(9), 1667-1678.
- Jankarin, J. & Kim, S.H. (1991) Sparse matrix sampling: A screening method for crystallization of proteins. *J. Appl. Cryst.*, **24**, 409-411.
- Johnson, W.C. (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins. Structure, Function and Genetics*. **7**, 205-214.
- Jones, T. A., Zon, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991) Improved methods for model building in electron density maps and the location of errors in these models. *Acta Cryst. A*, **47**, 110-119.
- Kantardjieff, K.A. & Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science* **12**, 1865-1871.
- Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H. & Phillips, D.C. (1958) A three dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature (London)*, **181**, 662-666.
- Killian, J.A. & von Heijne, G. (2000) How proteins adapt to a membrane-water interface. *TIBS*, **25**, 429-434.
- Kim, K.N., Cheong, Y.H., Gupta, R. & Luan, S. (2000) Interaction specificity of *Arabidopsis* calcineurin B-like calcium sensors and their target kinases. *Plant Physiol.*, **124**, 1844-1853.

- Kissinger, C.R., Parge, H.E., Knighton, D.R., Lewis, C.T., Pelletier, L.A., Tempczyk, A., Kalish, V.J., Tucker, K.D., Showalter, R.E., Moomaw, E.W. et al. (1995) Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Nature* **378(6557)**, 641-644.
- Kleywegt, G.J. & Brünger, A.T. (1996) Checking your imagination: applications of de free R value. *Structure* **4**, 897-904.
- Kolukisaoglu, U., Wein, I. S., Blazevic, D., Batistic, O., Kudla, J. (2004) Calcium sensors and their interacting protein kinases: genomics of the Arabidopsis and rice CBL-CIPK signaling networks. *Plant Physiol.*, **134(1)**, 43-58.
- Kraut, J., Sieker, L.C., High, D.F. & Freer, S.T. (1962) FHSCALC. *Proc. Natl. Acad. Sci. USA*, **48**, 1417-1424.
- Kraylis, P.J. (1991) MOLSCRIPT, a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946-950.
- Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B. & Bax, A. (1995) Solution structure of calcium-free calmodulin. *Nat. Struct. Biol.*, **2(9)**, 768-776.
- Kudla, J. (1999) Novel protein kinases associated with calcineurin B-like calcium sensors in *Arabidopsis*. *Plant Cell*, **11(12)**, 2393-2405.
- Kuntz, I.D. (1971) Hydration of macromolecules IV. Polypeptide conformation in frozen solutions. *J. Am. Chem. Soc.* **93**, 516-518.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993) PROCHECK, a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. D*, **26**, 283-291.
- Laue, T.M., Shah, B.D., Ridgeway, T.M. & Pelletier, S.L. (1992) Analytical Ultracentrifugation. Biochemistry & Polymer Science (Harding, S.E., Horton, H.C. Rowe, A.J., eds), Royal Society of Chemistry, London.
- Leslie, A.G.W. (1987a) A reciprocal-space method for calculating a molecular envelope using the algorithm of B.C. Wang. *Acta Cryst. A*, **43**, 134-136.
- Leslie, A.G.W. (1987b) MOSFLM. In *Proceedings of the CCP4 Study Weekend*. Machin, J.R. y Papiz, M.Z. editores, SERC Daresbury Laboratory, Warrington, UK, 39-50.
- Lewit-Bentley, A. & Rety, S. (2000) EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.*, **10(6)**, 637-643.
- Liu, J., Ishitani, M., Halfter, U., Kim, C.S. & Zhu, J.K. (2000) The Arabidopsis thaliana SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc. Natl. Acad. Sci. USA*, **97(7)**, 3730-3734.
- Liu, J. & Zhu, J.K. (1998) A calcium sensor homolog required for plant salt tolerance. *Science*, **280**, 1943-1945.
- Lo Conte, L., Chothia, C. & Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.*, **285(5)**, 2177-2198.

- Luan, S., Kudla, J., Rodriguez-Concepcion, M., Yalovsky, S. & Gruissem, W. (2002) Calmodulins and calcineurin B-like proteins, calcium sensors for specific signal response coupling in plants. *Plant Cell*, **14**, S389-400.
- Maser, P., Eckelman, B., Vaidyanathan, R., Horie, T., Farbairn, D.J., Kubo, M., Yamagami, M., Yamaguchi, K., Nishimura, M., Uozumi, N., *et. al.* (2002) Altered shoot/root Na⁺ distribution and bifurcation salt sensitivity in *Arabidopsis* by genetic disruption of de Na⁺ transporter AtHKT1. *FEBS Lett.*, **531**, 157-161.
- Mathews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
- McFerrin, M.B. & Snell, E.H. (2002) The development and application of a method to quantify the quality of cryoprotectant solutions using standard area-detector X-ray images. *J. Appl. Cryst.* **35**, 538-545.
- Merritt, E.A. & Murphy, M.E.P. (1994) RASTER3D Versión 2.0, a program for photorealistic molecular graphics. *Acta Crystallogr. D*, **50**, 869-873.
- Miller, R., Gallo, S.M., Khalak, H.G. & Weeks, C.M. (1994) SnB: crystal structure determination via Shake-and-Bake *J. Appl. Cryst.*, **27**, 613-621.
- Minton, A.P. (1994) *Modern Analytical ultracentrifugation*. Schester, T.H. & Lave, T.H. eds., Birkhäuser, Boston. USA.
- Monné, M., Nilsson, I., Johansson, M., Elmhed, N. & von Heijne, G. (1998) Positively and negatively charged residues have different effect on the position in the membrane of a model transmembrane helix. *J. Mol. Biol.*, **284**, 1177-1183.
- Murshudov, G.N., Vagin, A.A. & Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst. D*, **53**, 240-255.
- Nagae, M., Nozawa, A., Koizumi, N., Sano, H., Hashimoto, H., Sato, M., and Shimizu, T. (2003) The crystal structure of the novel calcium-binding protein AtCBL2 from *Arabidopsis thaliana*. *J Biol Chem.*, **278(43)**, 42240-42246.
- Navaza, J. (1994) AMoRe, an automated package for molecular replacement. *Acta Cryst. A*, **50**, 157-163.
- Nolen, B., Taylor, S. & Gourisankar, G. (2004) Regulation of protein kinases: controlling activity through activation segment conformation. *Molecular Cell*, **15**, 661-675.
- Nooren, I.M. & Thornton, J.M. (2003a) Diversity of protein-protein interactions. *EMBO J.*, **22(14)**, 3486-3492.
- Nooren, I.M. & Thornton, J.M. (2003b) Structural characterization and functional significance of transient protein-protein interactions. *J. Mol. Biol.*, **325(5)**, 991-1018.
- Ohta, M., Guo, Y., Halfter, U. & Zhu, J.K. (2003) A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. *Proc. Natl. Acad. Sci. USA*, **100**, 11771-11776.

- Otterbein, L.R., Kordowska, J., Witte-Hoffmann, C., Wang, C.L. & Dominguez, R. (2002) Crystal structures of S100A6 in the Ca(2+)-free and Ca²⁺-bound states, the calcium sensor mechanism of S100 proteins revealed at atomic resolution. *Structure*, **10(4)**, 557-567.
- Otwinowski, Z. & Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, **276**, Macromolecular Crystallography, part A. 307-326.
- Pearson, W.R. & Lipman, D.J. (1998) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA*, **85**, 2444-2448.
- Perczel, A., Hollosi, M., Tusnady, G., Fasman, G.D. (1991) Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins. *Protein Eng.*, **4(6)**, 669-79.
- Perrakis, A., Morris, R.J. & Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.*, **6**, 458-463.
- Perutz, M.F. (1956) Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Cryst.*, **9**, 867-873.
- Provencher, S.W. & Glöckner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, **20**, 33-37.
- Qiu, Q., Barkla, B.J., Vera-Estrella, R., Zhu, J.K. & Schumaker, K.S. (2003) Na⁺/H⁺ exchange activity in the plasma membrane of *Arabidopsis thaliana*. *Plant Physiol.*, **132**, 1041-1050.
- Qiu, Q., Guo, Y., Dietrich, M.A., Schumaker, K.S. & Zhu J.K. (2002) Regulation of SOS1, a plasma membrane Na⁺/H⁺ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc. Natl. Acad. Sci. USA*, **99**, 8436-8441.
- Qiu, Q.S., Guo, Y., , Quintero, F.J., Pardo, J.M., Schumaker, K.S. & Zhu, J.K. (2004) Regulation of vacuolar Na⁺/H⁺ exchange in *Arabidopsis thaliana* by the salt-overly-sensitive (SOS) pathway. *J. Biol. Chem.*, **279**, 207-215.
- Quintero, F.J., Ohta, M., Shi, H., Zhu, J.K. & Pardo, J.M. (2002) Reconstitution of the SOS signaling pathway for Na⁺ homeostasis in plants. *Proc. Natl. Acad. Sci. USA*, **99**, 9061-9066.
- Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V.J. (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95-99.
- Read, R.J. (2003) New ways of looking at experimental phasing. *Acta Cryst. D*, **59**, 1811-1822.
- Read, R.J. & Kleywegt, G.J. (2001) Density modification: theory and practice. *Methods in Macromolecular Crystallography*. IOS Press. Amsterdam.
- Resh, M.D. (1999) Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta*, **1451**, 1-16.
- Rivas, G., Ingham K.C. & Minton A.P. (1992) Ca(2+)-linked association of human complement C1s and C1r. *Biochemistry*, **33(8)**, 2341-2348.
- Rossmann, M.G. & Blow, D.M. (1962) The detection of subunits within the asymmetric unit. *Acta Cryst.*, **15**, 24-31.

- Rozanas, C. (1998) Purification of calcium-binding proteins using hydrophobic interaction chromatography. *Life Science News I*, Amersham Pharmacia Biotech.
- Rus, A., Yokoi, S., Sharkhuu, A., Reddy, M., Lee, B.H., Matsumoto, T.K., Koiwa, H., Zhu, J.K., Versan, R.A. & Hasegawa, P.M. (2001) AtHKT1 is a salt tolerance determinant that controls Na⁺ entry into plant roots. *Proc. Natl. Acad. Sci. USA*, **98**, 14150-14155.
- Sánchez-Barrena M.J., Martínez-Ripoll, M., Galvez, A., Valdivia, E., Maqueda, M., Cruz, V. & Albert, A. (2003) Structure of bacteriocin AS-48, from soluble state to membrane bound state. *J. Mol. Biol.*, **334(3)**, 541-549.
- Serrano, R. & Rodríguez-Navarro, A. (2001) Ion homeostasis during salt stress in plants. *Curr. Op. Cell Biol.*, **13**, 399-404.
- Sanders, D., Pelloux, J., Brownlee, C. & Harper, J.F. (2002) Calcium at the crossroads of signalling. *Plant Cell*, S401-S417.
- Sayle, R.A. & Milner-White E. J. (1995) RasMol, Biomolecular graphics for all. *Trends in Biochemical Sciences*, **20**, 374-376.
- Schaad, N.C., De Castro, E., Nef, S., Hegi, S., Hinrichsen, R., Martone, M.E., Ellisman, M.H., Sikkink, R., Rusnak, F., Sygush, J., & Nef, P. (1996) Direct modulation of calmodulin targets by the neuronal calcium sensor NCS1. *Proc. Natl. Acad. Sci. USA*, **93(17)**, 9253-9258.
- Schneider, T.R. & Sheldrick, G.M. (2002) Substructure solution with SHELXD. *Acta Cryst.* **D58**, 1772-1779.
- Schuck, P. (2000) Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophysical Journal* **78**, 1606-1619.
- Scrase-Field S.A. & Knight M.R. (2003) Calcium, just a chemical switch? *Curr Opin Plant Biol.*, **6**, 500-506.
- Sheldrick, G.M. (1991) The SHELXS system. In *Crystallographic Computing 5* (Podjarny, P.D. & Thiery, J.D., eds), 145-157. IUCR, Oxford University Press, Oxford.
- Sheldrick, G.M., Hauptman, H.A., Weeks, C.M., Miller, R. & Usón, I. (2001) *International Tables of Macromolecular Crystallography vol. F*. Editado por M.G. Rossmann y E. Arnold. Kluwer Academic Publishers.
- Shi, J., Kim, K.N., Ritz, O., Albrecht, V., Gupta, R., Harter, K., Luan, S. & Kudla, J. (1999) Novel protein kinases associated with calcineurin B-like calcium sensors in Arabidopsis. *Plant Cell*, **11(12)**, 2393-2405.
- Shi, H., Ishitani, M., Kim, C. & Zhu, J.K. (2000) The *Arabidopsis thaliana* salt tolerance gene *SOS1* encodes a putative Na⁺/H⁺ antiporter. *Proc. Natl. Acad. Sci. USA*, **97**, 6896-6901.
- Shi, H., Quintero, F.J., Pardo, J.M. & Zhu, J.K. (2002) The putative plasma membrane Na⁺/H⁺ antiporter *SOS1* controls long-distance Na⁺ transport in plants. *Plant Cell*, **14**, 465-477.
- Shi, H., Wu, S.J. & Zhu, J.K. (2003) Overexpression of the plasma membrane Na⁺/H⁺ improves salt tolerance in *Arabidopsis*. *Nat. Biotechnol.*, **21**, 81-85.

- Snedden, W.A. & Fromm, H. (1998) Calmodulin, calmodulin-related proteins and plant response to the environment. *Trends Plant Sci.*, **3**, 299-304.
- Stemmer, P.M. & Klee, C.B. (1994) Dual calcium ion regulation of calcineurin by calmodulin and calcineurin B. *Biochemistry*, **33**(22), 6859-6866.
- Sutcliffe, M.J., Haneef, I., Carney, D. & Blundell, T.L. (1987) Knowledge based modeling of proteins. 1. 3-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng.*, **1**, 377-384.
- Terwillinger, T.C. (2000) Maximum likelihood density modification. *Acta Cryst. D*, **56**, 965-972.
- Terwillinger, T.C. & Berendzen, J. (1999) Automated structure solution for MIR & MAD. *Acta Cryst. D*, **58**, 1937-1940.
- Tester, M. & Davenport, R. (2003) Na⁺ tolerance and Na⁺ transport in higher plants. *Ann. Bot. (Lond)*, **91**, 503-527.
- Trewavas A.J. & Malho, R (1998) Ca²⁺ signalling in plant cells: the big network! *Curr. Op. Plant Biol.*, **1**, 428-433.
- Usón, I. & Sheldrick, G.M. (1999) Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.*, **9**(5), 643-648.
- Vagin, A. & Teplyakov, A. (1997) MOLREP, an automated program for molecular replacement. *J. Appl. Cryst.*, **30**, 1022-1025.
- Verdino, P., Westritschnig, K., Valenta, R. & Keller, W. (2002) The cross-reactive calcium-binding pollen allergen, Phl p 7, reveals a novel dimer assembly. *EMBO J.*, **21**(19), 5007-5016.
- Varea, J., Saiz, J.L., Lopez-Zumel, C., Monterroso, B., Medrano, F.J., Arrondo, J.L., Iloro, I., Laynez, J., Garcia, J.L. & Menendez, M. (2000) Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *J. Biol. Chem.*, **275**(35), 26842-26855.
- Veeraraghavan, S., Fagan, P.A., Hu, H., Lee, V., Harper, J.F., Huang, B., and Chazin, W.J. (2002) Structural independence of the two EF-hand domains of caltractin. *J. Biol. Chem.*, **277**(32), 28564-28571.
- Wallace A.C., Laskowski, R.A. & Thornton J.M. (1995) LIGPLOT, A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, **8**, 127-134.
- Weeks, C.M., Adams, P.D., Berendzen, J., Brünger, A.T., Dodson, E.J., Grosse-Kunstleve, R.W., Schneider, T.R., Sheldrick, G.M., Terwillinger, T.C., Turkenburg, M.G.W. & Usón, I. (2003) Automatic solution of heavy-atom substructures. *Methods in Enzymology*, **374**, 37-83.
- White, P.J. & Broadley, M.R. (2003) Calcium in plants. *Annals of Botany*, **92**(4), 487-511.
- Yao, J.-X. (1981) On the application of phase relationships to complex structures. XVIII. RANTAN-random MULTAN. *Acta Cryst. A*, **37**, 642-644.

-
- Yokoi, S., Quintero, F.J. Cubero, B., Ruiz, M.T., Bressan, R.A., Hasegawa, P.M. & Pardo, J.M. (2002) Differential expression and function of *Arabidopsis thaliana* NHX Na⁺/H⁺ antiporters in the salt tolerance response. *Plant J.*, **30**, 529-539.
- Zhang, M., Tanaka, T. & Ikura, M. (1995) Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nat. Struct. Biol.*, **2(9)**, 758-767.
- Zhu, J.K. (2000) Genetic analysis of plant salt tolerance using *Arabidopsis thaliana*. *Plant Physiol.*, **124**, 941-948.
- Zhu, J.K. (2001) Over expression of a delta-pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water and salt stress in transgenic rice. *Trends Plant Sci.*, **6**, 66-72.
- Zhu, J.K. (2003) Regulation of ion homeostasis under salt stress. *Curr. Opin. Plant Biol.*, **6**, 441-445.
- Zielinski, R.E. (1998) Calmodulin and calmodulin-binding proteins in plants. *Ann. Rev. Plant Physiol.*, **49**, 697-752.

6. ANEXO

**SOS3 (salt overly sensitive 3) from *Arabidopsis thaliana*:
expression, purification, crystallization and preliminary X-ray
analysis.**

Sánchez-Barrena, M.J., Martínez-Ripoll, M., Zhu, J.K. & Albert, A.
(2004) *Acta Crystallogr. D*, **60(7)**, 1272-1274.

SOS3 (salt overly sensitive 3) from *Arabidopsis thaliana*: expression, purification, crystallization and preliminary X-ray analysis

M. J. Sánchez-Barrena,^a
M. Martínez-Ripoll,^a J. K. Zhu^b
and A. Albert^{a*}

^aGrupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química Física 'Rocasolano', Consejo Superior de Investigaciones Científicas, Serrano 119, E-28006 Madrid, Spain, and ^bDepartment of Botany and Plant Sciences and the Institute for Integrative Genome Biology, 2150 Batchelor Hall, University of California, Riverside, CA 92521, USA

Correspondence e-mail: xalbert@iqfr.csic.es

The salt-tolerance gene *SOS3* (salt overly sensitive 3) of *Arabidopsis thaliana* encodes a calcium-binding protein that is able to sense the cytosolic calcium signal elicited by salt stress. *SOS3* activates the *SOS2* protein kinase, which activates various ion transporters. *SOS3* was cloned into a plasmid and expressed in *Escherichia coli*, allowing purification of the protein to homogeneity. Two crystals with different additive contents were grown. Both diffract to 3.2 Å resolution and belong to space group *I*₄, with unit-cell parameters *a* = 93.65, *c* = 80.08 Å and *a* = 91.79, *c* = 85.78 Å, respectively. A promising molecular-replacement solution has been found using neuronal calcium-sensor 1 as the search model. Interestingly, no solution was found using AtCBL2 (*A. thaliana* calcineurin B-like protein) structure as a search model, although this protein belongs to the same family and displays 50% sequence identity.

Received 12 March 2004
Accepted 13 April 2004

1. Introduction

The role of calcium as a second messenger controlling many cellular processes has been reviewed (Klee & Means, 2002; Lewit-Bentley & Rety, 2000; Skelton *et al.*, 1994). In plants, many extracellular stimuli, including hormones and environmental signals such as gravity, light, salinity, drought, cold, oxidative stress, anoxia and mechanical perturbation, cause changes in the cytosolic free Ca²⁺ concentration (Poovaiah & Reddy, 1987; Trewavas & Malho, 1998; Bush, 1993).

Cytosolic calcium changes can vary in magnitude, duration, frequency and spatial arrangement within the cell, depending on the nature of the stimulus.

Despite this variety of stimuli, the most common architecture for calcium binding is the EF-hand motif (Lewit-Bentley & Rety, 2000). This motif is highly conserved and involves two helices and a loop of 12 residues that participate in calcium coordination. However, the structures of EF-hand proteins in the calcium-bound or free states or in complex with target peptides or proteins show a variety of conformations, each unique to the particular activation mechanism in which calcium is involved (Klee & Means, 2002).

In *Arabidopsis thaliana*, *SOS3* is a myristoylated calcium-binding protein that is capable of sensing the cytosolic calcium signal elicited by salt stress (Liu & Zhu, 1998; Ishitani *et al.*, 2000). *SOS3* is able to complex with and to activate the serine/threonine protein kinase *SOS2*, which then activates various ion transporters, such as the Na⁺/H⁺ antiporter *SOS1*.

SOS3 is predicted to share the same fold as the homologous AtCBL2 protein, which

consists of two domains of pairs of adjacent EF-hand motifs. However, AtCBL2 interacts with a different target, has no *N*-myristoylation motif and its expression profile suggests a role in light-signal transduction rather than in salt stress (Ishitani *et al.*, 2000; Nagae *et al.*, 2003). In this work, we describe the expression, purification, crystallization and preliminary diffraction data of *SOS3*. Unexpectedly, it is shown that molecular replacement using the structure of AtCBL2 as a search model fails, indicating significant structural differences between *SOS3* and AtCBL2. Comparison of their molecular structures would be of biological significance for the understanding of the different methods of sensing calcium by these proteins.

2. Experimental

2.1. Protein expression and purification

The *SOS3* gene from *A. thaliana* was cloned into a pET14b (Novagen, Madison, WI, USA) expression plasmid to yield pET-His-*SOS3*. The plasmid was transformed into *Escherichia coli* strain BL21(DE3) for protein expression (Ishitani *et al.*, 2000).

5 ml aliquots of an overnight culture were subcultured into 500 ml fresh LB (Luria-Bertani) medium (10 g Bacto tryptone, 5 g yeast extract, 10 g NaCl per litre of solution) plus ampicillin (50 µg ml⁻¹) and allowed to grow to A₆₀₀ = 0.7 at 310 K. Protein expression was then induced for 3 h with 0.3 mM isopropyl-β-D-thiogalactoside and cells were harvested by centrifugation (10 min, 9000 rev min⁻¹).

The cell pellet was then resuspended in buffer 1 [20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 5 mM imidazole, 0.05% (w/v) NaN₃] and cells were disrupted by sonication. After centrifugation (30 min, 17 000 rev min⁻¹) at 277 K, the clear supernatant was filtered (pore diameter 0.45 µm; Millipore Corporation, Bedford, MA, USA) and applied onto a nickel-Sepharose column (5 ml; Amersham Biosciences Limited, UK) equilibrated with buffer 1. A washing step with five column volumes of buffer 1 followed by five volumes of buffer

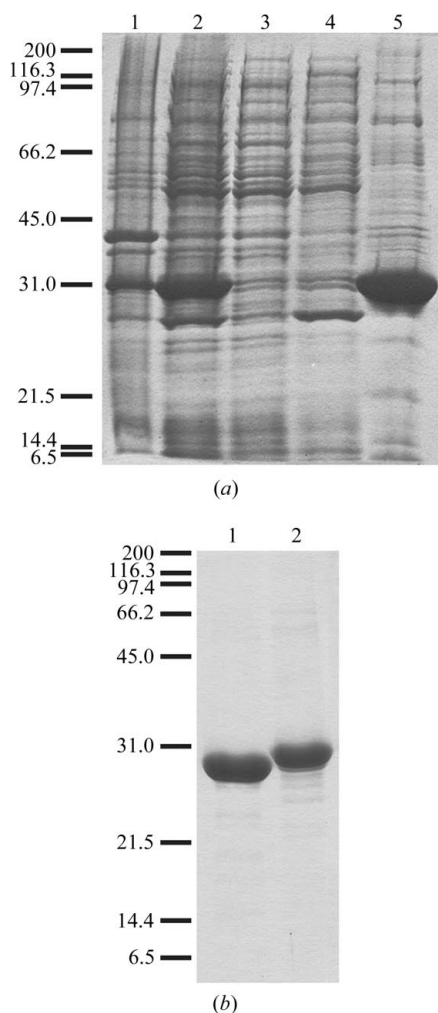


Figure 1

(a) SOS3 expression and purification by nickel-affinity chromatography. Proteins were analyzed by 12% SDS-PAGE gels. Lanes 1 and 2, precipitated and soluble fractions after cell disruption. Lane 3, flowthrough from the affinity column. Lane 4, protein fractions eluted after washing with 30 mM imidazole buffer. Lane 5, protein fractions eluted with 500 mM imidazole. (b) On-column cleavage of the tagged protein. 15% SDS-PAGE gel analysis. Lane 1, eluted cleaved protein after digestion with thrombin protease. Lane 2, the remaining uncleaved protein eluted with 500 mM imidazole. Samples were loaded under reducing conditions. The position of the molecular-weight markers (MW in kDa; BioRad Laboratories Inc., USA) is indicated. Gels were stained using Coomassie Brilliant Blue.

containing 20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 30 mM imidazole, 0.05% (w/v) NaN₃ was performed.

The column was re-equilibrated with buffer 1 and thrombin protease (Amersham Biosciences Limited, UK) was loaded onto the column for an overnight thrombin protease digestion at 293 K (7.5 thrombin units per milligram of SOS3). Cleaved SOS3 was eluted with buffer 1. The time courses of the chromatography and enzymatic digestion were monitored by SDS-PAGE (see Fig. 1). The sample buffer was changed to 20 mM Tris-HCl pH 7.5, 0.05 M NaCl, 0.05% (w/v) NaN₃ (buffer 2) by running through a PD-10 column (Amersham Biosciences Limited, UK). A final polishing step was performed using an anion-exchange Resource Q column (6 ml; Amersham Biosciences Limited, UK) with a salt gradient from 0.05 to 0.5 M NaCl. The protein eluted at 0.1 M NaCl (see Fig. 2).

The protein buffer was changed to buffer 2 using a PD-10 column (Amersham Biosciences Limited, UK). SOS3 was concentrated to a final concentration of 20 mg ml⁻¹ with a 10 kDa cutoff Amicon protein concentrator (YM-10; Millipore Corporation, Bedford, MA, USA). The final protein concentration was determined spectrophotometrically using a molar absorption coefficient of 10 930 M⁻¹ cm⁻¹ at 280 nm. The sample was aliquoted and immediately frozen at 253 K. The final sample purity was examined by SDS-PAGE and mass spectrometry; the measured mass [(M + H) = 25 975] was coincident with the expected value (25 974.7).

2.2. Crystallization and preliminary data collection

Preliminary crystallization conditions were established using the sparse-matrix sampling technique (Jancarik & Kim, 1991) with the hanging-drop vapour-diffusion method at 293 K using Crystal Screens I, II, Lite and Index (Hampton Research, CA,

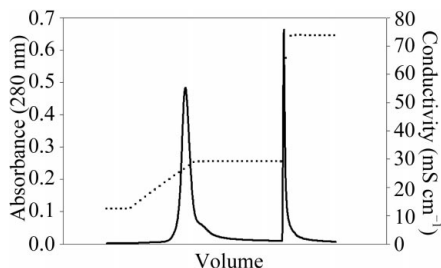


Figure 2

Anion-exchange chromatography of SOS3. The dotted line represents the conductivity of the solution. The black line represents the absorbance recorded at 280 nm.

USA). Drops containing equal volumes (1 µl) of protein (10 mg ml⁻¹) and reservoir solution were equilibrated against 500 µl reservoir solution. Condition No. 40 [(0.1 M trisodium citrate dihydrate pH 5.6, 10% (v/v) 2-propanol] from Crystal Screen Lite (Hampton Research, CA, USA) produced needle-like microcrystals.

Subsequent optimization yielded thin clustered plate-like and needle-like crystals that were not usable for diffraction experiments (see Fig. 3a). Drops containing 2 µl protein (10 mg ml⁻¹, 3.85 mM MnCl₂) and 1 µl reservoir solution were equilibrated against 500 µl reservoir solution [20% (w/v)

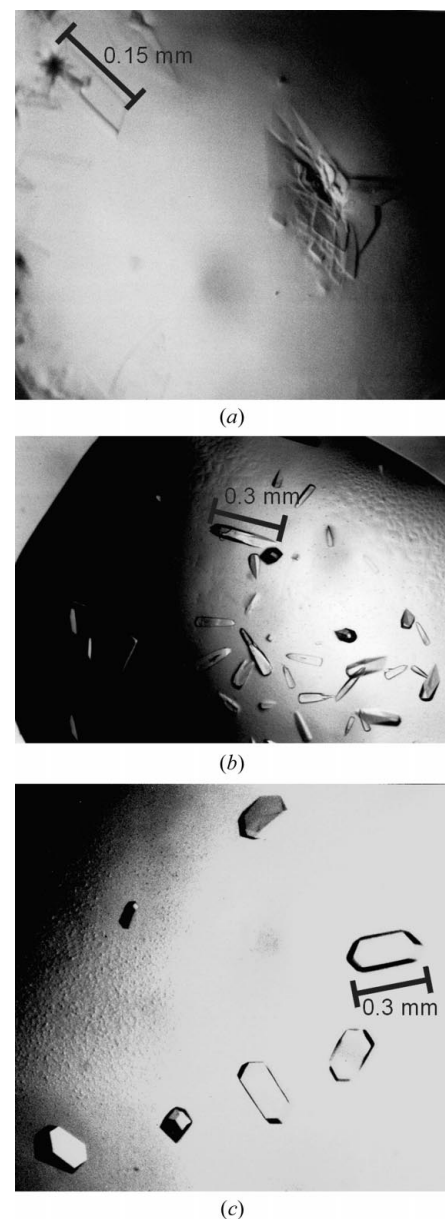


Figure 3

Effect of additives on SOS3 crystals. (a) Crystals grown without additives. (b) Crystals grown in the presence of sucrose and ethanol. (c) Crystals grown with MnCl₂ and NaI.

Table 1

Data-collection and processing statistics of tetragonal crystals of SOS3.

Values in parentheses are for the highest resolution shell (3.20–3.31 Å).

	SOS3 (additives: sucrose and EtOH)	SOS3 (additives: MnCl ₂ and NaI)
Space group	<i>I</i> ₄	<i>I</i> ₄
Unit-cell parameters (Å)	<i>a</i> = <i>b</i> = 93.65, <i>c</i> = 80.08	<i>a</i> = <i>b</i> = 91.79, <i>c</i> = 85.78
Resolution limit (Å)	3.2	3.2
Observations	18957	19633
Unique reflections	5736	5888
Completeness (%)	98.9 (99.5)	99.2 (100)
Multiplicity	3.3 (3.3)	3.3 (3.3)
<i>R</i> _{merge} † (%)	0.09 (0.42)	0.09 (0.43)
<i>I</i> /σ(<i>I</i>)	15.3 (5.2)	14.6 (2.8)

$$\dagger R_{\text{merge}} = \sum_{hkl} |I_{hkl} - \langle I_{hkl} \rangle| / \sum_{hkl} I_{hkl}$$

PEG 4000, 30%(v/v) MPD, 0.1 M sodium citrate buffer pH 4.9]. Different additives were tested around these conditions.

A new rod-shaped crystal form (see Fig. 3*b*) was grown from drops containing 10.7 mg ml⁻¹ SOS3, reservoir solution [20%(w/v) PEG 4000, 30%(v/v) MPD, sodium citrate pH 5.2] and 30%(w/v) D-(+)-sucrose in the ratio 2:0.75:1. The quality of these crystals was improved by adding 0.75 μl 30%(v/v) ethanol to these conditions. Preliminary diffraction data were collected on an in-house Kappa 2000 CCD detector with Cu Kα X-rays generated by a Nonius FR-591 rotating-anode generator with Montel mirrors operated at 45 kV and 100 mA. It was not necessary to use a cryoprotectant solution before flash-freezing the crystal to 120 K as the crystals were grown under cryoprotectant conditions. The crystal-to-detector distance was set to 70 mm with Δφ = 0.5° and 10 min exposure per image. The maximum resolution reached was 3.2 Å.

The quality of the crystals was improved by treating the protein sample with a chelating resin (Hampton Research, CA, USA) that preferentially complexes trace divalent metal ions (Dunn *et al.*, 1980) prior to crystallization experiments and by using NaI as additive. These crystals (see Fig. 3*c*) grew from drops containing SOS3 (10 mg ml⁻¹, 3.85 mM MnCl₂), 1 M NaI and reservoir solution [20%(w/v) PEG 4000, 26%(v/v) MPD, sodium citrate buffer pH 4.8] in the ratio 2:0.75:1. A complete 3.2 Å data set was collected using an in-house

X-ray source at 120 K. The crystal-to-detector distance was set to 70 mm with Δφ = 0.5° and 10 min exposure per image. All data sets were processed with the *HKL2000* suite of programs (Otwinowski & Minor, 1997).

3. Results and discussion

Recombinant SOS3 protein was purified to homogeneity using a two-step procedure. The nickel-affinity column facilitated the purification by providing relatively high purity in one step. After enzymatic cleavage of the His tag, the purity of the protein was further improved by anion-exchange chromatography. The sample purity at each purification step was monitored by SDS-PAGE (see Fig. 1). The purity of the protein sample used for crystallization was examined by mass spectrometry and the measured mass was the same as the expected value.

A 3.2 Å resolution data set for SOS3 crystals grown in the presence of sucrose and ethanol as additives was collected. The space group was tetragonal *I*₄, with unit-cell parameters *a* = *b* = 93.65, *c* = 80.08 Å. Specific volume calculations (Matthews, 1968; Kantardjieff & Rupp, 2003) yielded eight molecules of SOS3 in the unit cell, with a solvent content of 63.6% (*V*_M = 3.38 Å³ Da⁻¹), corresponding to one monomer per asymmetric unit. Data-collection and processing statistics are summarized in Table 1.

Crystals grown in the presence of NaI and MnCl₂ diffracted to 3.2 Å. The space group was also tetragonal *I*₄, with unit-cell parameters *a* = *b* = 91.79, *c* = 85.78 Å. The asymmetric unit contains one SOS3 molecule and has 64.6% solvent content (*V*_M = 3.48 Å³ Da⁻¹; Matthews, 1968; Kantardjieff & Rupp, 2003). The statistics for this data set are shown in Table 1.

Molecular replacement with the program *AMoRe* (Navaza, 1994), using coordinates from the neuronal calcium sensor 1 (NCS1; PDB code 1g8i; Bourne *et al.*, 2001) as the search model (sequence identity of 26.6%), has yielded a promising but not conclusive solution that allows unambiguous determination of the space group (correlation coefficient = 32.1%, *R* factor = 50.3%, resolution range 15–3.5 Å). However, the

electron-density map calculated with this model was extremely noisy and was clearly biased by the NCS1 structure. Interestingly, no molecular-replacement solution has been found using the AtCBL2 structure as a search model, although AtCBL2 belongs to the SOS3 family and has a high percentage of sequence identity (50.4%) with SOS3. This suggests large structural dissimilarities between them that are probably motivated by different calcium substitution, as their sequence differences are located at the calcium-binding EF-hand motifs (Nagae *et al.*, 2003). Synchrotron diffraction measurements at longer wavelength (from 1.5 to 1.7 Å) are being prepared to phase the structure using the anomalous signal from the I atoms. Further heavy-atom derivative and selenomethionine-substituted protein preparations are also projected.

MJS-B was supported by a FPU student-ship from Ministerio de Educación, Cultura y Deporte. This work was funded by grant BMC2002-04011-C05-03 from the Spanish 'Plan Nacional' (MCYT).

References

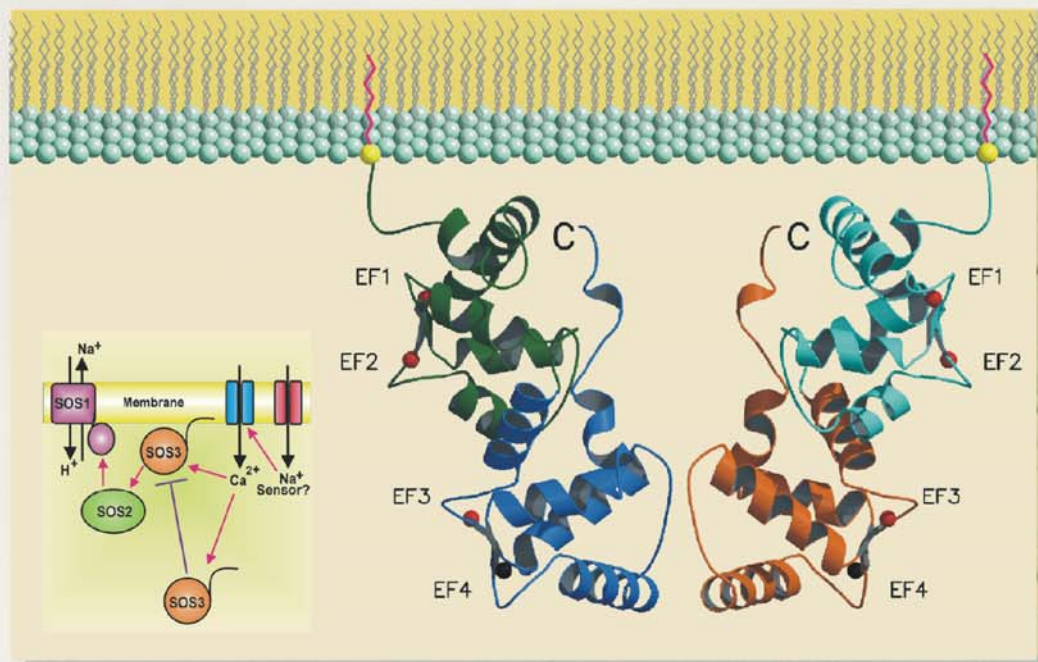
- Bourne, Y., Dannenberg, J., Pollmann, V., Marchot, P. & Pongs, O. (2001). *J. Biol. Chem.* **276**, 11949–11955.
- Bush, D. S. (1993). *Plant Physiol.* **103**, 7–13.
- Dunn, M. F., Pattison, S. E., Storm, M. C. & Quiel, E. (1980). *Biochemistry*, **19**, 718–725.
- Ishitani, M., Liu, J., Halfter, U., Kim, C. S., Shi, W. & Zhu, J. K. (2000). *Plant Cell*, **12**, 1667–1678.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
- Klee, C. B. & Means, A. R. (2002). *EMBO Rep.* **3**, 823–827.
- Lewit-Bentley, A. & Rety, S. (2000). *Curr. Opin. Struct. Biol.* **10**, 637–643.
- Liu, J. & Zhu, J. K. (1998). *Science*, **280**, 1943–1945.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nagae, M., Nozawa, A., Koizumi, N., Sano, H., Hashimoto, H., Sato, M. & Shimizu, T. (2003). *J. Biol. Chem.* **278**, 42240–42246.
- Navaza, J. (1994). *Acta Cryst. A* **50**, 157–163.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Poovalah, B. W. & Reddy, A. S. (1987). *CRC Crit. Rev. Plant Sci.* **6**, 47–103.
- Skelton, N. J., Kordel, J., Akke, M., Forsen, S. & Chazin, W. J. (1994). *Nature Struct. Biol.* **1**, 239–245.
- Trewavas, A. J. & Malho, R. (1998). *Curr. Opin. Plant Biol.* **1**, 428–433.

The structure of the *Arabidopsis thaliana* SOS3: molecular mechanism of sensing calcium for salt stress response.

Sánchez-Barrena, M.J., Martínez-Ripoll, M., Zhu, J.K. & Albert, A. (2005)
J. Mol. Biol., **345(5)**, 1253-1264.

JMIB

JOURNAL OF MOLECULAR BIOLOGY



The Structure of the *Arabidopsis thaliana* SOS3: Molecular Mechanism of Sensing Calcium for Salt Stress Response

María José Sánchez-Barrena¹, Martín Martínez-Ripoll¹, Jian-Kang Zhu² and Armando Albert^{1*}

¹Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química Física "Rocasolano" Consejo Superior de Investigaciones Científicas Serrano 119, E-28006 Madrid Spain

²Department of Botany and Plant Sciences, and the Institute for Integrative Genome Biology 2150 Batchelor Hall, University of California, Riverside, CA 92521, USA

The *Arabidopsis thaliana* SOS3 gene encodes a calcium sensor that is required for plant salt tolerance. The SOS3 protein binds to and activates the self-inhibited SOS2 protein kinase, which mediates the expression and activities of various transporters important for ion homeostasis under salt stress. SOS3 belongs to a unique family of calcium-binding proteins that contain two pairs of EF hand motifs with four putative metal-binding sites. We report the crystal structure of a dimeric SOS3 protein in complex with calcium, and with calcium and manganese. Analytical ultracentrifugation experiments and circular dichroism measurements show that calcium binding is responsible for the dimerization of SOS3. This leads to a change in the global shape and surface properties of the protein that may be sufficient to transmit the Ca²⁺ signal elicited during salt stress.

© 2004 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: calcium signaling; protein structure; salt tolerance; crystallography; EF-hand

Introduction

Due to their sessile nature, plants have to endure adverse environmental conditions such as drought and soil salinity. Excess salts in the soil inhibit plant growth and cause great losses in agricultural productivity worldwide.^{1,2} The excessive sodium ion (Na⁺) present in saline soils is particularly harmful to plants because it is toxic to cellular enzymes if it accumulates to a high level in the cytoplasm. Therefore, maintaining a low concentration of Na⁺ in the cytoplasm is key to plant salt tolerance.^{1–3}

Genetic and biochemical studies in *Arabidopsis* have identified the salt overly sensitive (SOS) pathway for Na⁺ homeostasis and salt tolerance.^{3,4} In this pathway, SOS3 is a calcium-binding protein capable of sensing the cytosolic calcium ion (Ca²⁺) signal elicited by salt stress.^{5,6} SOS3 interacts physically with SOS2, a serine/threonine protein kinase.^{7,8} In the presence of Ca²⁺, SOS3 activates the substrate phosphorylation activity of SOS2. The binding of SOS3 to SOS2 is mediated by the 21

amino acid residue FISL motif in SOS2, which is autoinhibitory to the kinase activity.⁹ SOS3 is myristoylated at its N terminus, and the myristoylation is important for recruiting SOS2 to the plasma membrane and for salt tolerance in plants.^{6,10} Both SOS3 and SOS2 are required for the salt stress activation of SOS1, a Na⁺/H⁺ antiporter on the plasma membrane.^{10,11} The SOS3–SOS2 kinase complex has been shown to phosphorylate SOS1.¹⁰ Loss-of-function mutations in any of the SOS genes disrupts cellular sodium and potassium homeostasis, and render *Arabidopsis* plants hypersensitive to inhibition by salt stress.⁴

SOS3 is the founding member of a unique family of ten EF-hand type calcium-binding proteins in *Arabidopsis*. This family of proteins shares significant sequence similarities to the B subunit of calcineurin (i.e. phosphatase type 2B) and animal neuronal calcium sensors.^{5,9} However, *Arabidopsis* clearly does not have any 2B-type protein phosphatase. Instead, members of the SOS3 family, known as SCaBPs⁹ or CBLs,¹² interact with specific members of the SOS2-like family of protein kinases (PKS or CIPK¹²) to form distinct protein kinase complexes.^{9,13} In plants, diverse environmental stimuli such as salinity, drought, cold, gravity, light, anoxia and mechanical perturbation all are

Abbreviations used: SOS, salt overly sensitive.
E-mail address of the corresponding author:
xalbert@iqfr.csic.es

capable of eliciting Ca^{2+} signals.^{14,15} The SCaBPs together with other EF-hand type calcium-binding proteins may be able to distinguish different Ca^{2+} signals, and transduce the signals to regulate various effector proteins, generating appropriate cellular responses.^{13,14}

In order to understand how SOS3 senses Ca^{2+} and regulates SOS2, and ultimately how it functions specifically to mediate signaling by salt stress but not other stresses that also cause changes in the cytosolic free Ca^{2+} , it is important to solve the three-dimensional structure of SOS3. Here, we report the crystal structure of an SOS3 dimer in complex with Ca^{2+} , and with Ca^{2+} and manganese ion (Mn^{2+}), at 2.75 Å and at 3.0 Å resolution, respectively. The structure suggests a mechanism for Ca^{2+} -sensing by SOS3 and for the activation of its interacting partner, SOS2.

Results and Discussion

SOS3 displays a calmodulin-like fold

The X-ray structure of SOS3 in complex with Ca^{2+} , and with Ca^{2+} and Mn^{2+} was solved by single anomalous scattering phasing provided by seven iodine atoms at 2.75 Å and 3.00 Å, resolution respectively (Figure 1; and see Materials and Methods). The structures of both complexes were refined independently to avoid refinement bias. However, these structures are nearly identical (C^α backbone root-mean-square deviation, RMSD of 0.2 Å),¹⁶ hence, unless stated otherwise, all the results presented here are referred to the SOS3 Ca^{2+} complex.

The overall fold of an SOS3 protomer is almost identical with that found for SCaBP1/AtCBL2¹⁷ and the homologous structures of calcineurin B (CnB)¹⁸ and neuronal calcium-sensor (NCS1).¹⁹ To summarize, SOS3 can be described as a two-domain structure connected by a short linker. Each domain is formed by a pair of adjacent EF-hand motifs. The N-terminal domain contains an additional α -helix and the C-terminal domain a small α -helix plus an unstructured tail that protrudes towards the N-terminal domain.

The dimeric structure of SOS3

Sedimentation equilibrium experiments were performed to investigate the association state of SOS3 in solution and the effect of Ca^{2+} and other physiologically relevant metal ions on SOS3 oligomerization (Table 1; and see Materials and Methods). Our data show that SOS3 displays a monomer/dimer equilibrium with a dimerization constant of around $5 \times 10^4 \text{ M}^{-1}$. However, the analysis of the auto-association equilibrium of SOS3 as a function of the concentration of Ca^{2+} shows that the fraction of dimer increases with the concentration of free Ca^{2+} (Table 1). We do not observe any effect of Mn^{2+} or Mg^{2+} on the

oligomeric state of SOS3 in solution (data not shown).

The analysis of the crystal structure of SOS3 reveals that the protein is a dimer as observed in solution with a 2-fold symmetry axis coincident with the crystallographic 2-fold axis (Figure 1). Hence, only a peptidic subunit is found in the asymmetric unit. The crystallographic dimer displays a V shape. Molecules within the dimer interact through the C-terminal end of helix F3, the N-terminal end of helix E4 and the loop connecting both helices. There is a 4% decrease in the solvent-accessible surface area per protomer in the formation of the dimer; and 80% of this occluded area corresponds to hydrophobic residues.²⁰ The dimerization involves the formation of one hydrogen bond between Ser144 OG and its symmetry-related residue. These values are lower than those expected for stable homodimers in solution, but are similar to those corresponding to other signal transduction complexes.^{21–23} These complexes and SOS3 show a transient oligomerization state that is dependent on factors such as an increase in the local concentration of the protein, the binding of a particular ligand or a change in the local physicochemical environment. SOS3 dimer displays a structure that exposes to the solvent eight metal-binding sites, the N-terminal myristoylation sites, and the C-terminal end of the protein, which has been shown to be essential for SOS2 interaction and subsequent activation.⁶

Ishitani *et al.* showed that SOS3 myristoylation is necessary for SOS2 activation but may not be for membrane association in plant cells.⁶ SOS3 presents a sequence of five lysine residues at the N-terminal end, which may be sufficient for membrane anchorage.^{24,25} This suggests that myristoylation adds a structural determinant for the correct positioning of SOS3 with respect to the membrane. Dimeric SOS3 duplicates the efficiency of the membrane binding with respect to the monomer and would help in positioning the macromolecule, as it provides two accessible anchoring points.

The metal-binding sites of SOS3

The classical EF-hand motif is characterized by a sequence of 12 residues involved in Ca^{2+} binding (Figure 2). The amino acid residues in positions X, Y, Z and $-Z$ ligate Ca^{2+} via side-chain oxygen donors. The residues at $-X$ and $-Y$ participate in the coordination of the metal with a backbone oxygen atom and with a water-mediated interaction, respectively. Sequences of the EF motifs in SCaBPs are different from the classical sequences, since the oxygen donor at position Y is replaced by a hydrophobic or basic amino acid residue for EF1 and for EF2, EF3 and EF4, respectively. In addition, the EF1 motif contains an insertion of two amino acid residues between the X and Y positions.^{9,26}

Both the experimental and the omit electron density maps at SOS3 metal-binding sites allowed

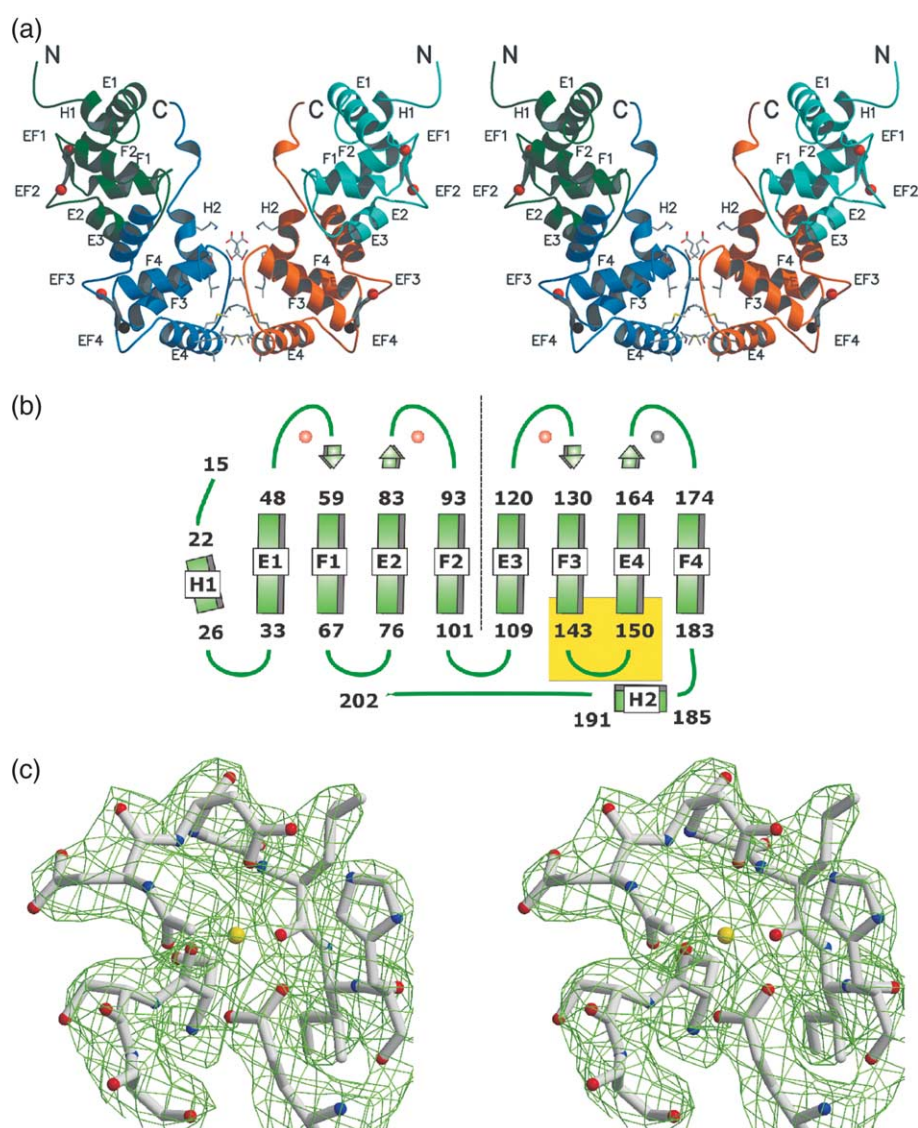


Figure 1. (a) Stereoview of the ribbon structure representation of SOS3 showing the domain structure of the protein. Amino acid side-chains involved in the formation of the dimer are displayed in stick mode. Calcium and manganese ions are displayed as red balls and black balls, respectively. (b) A representation of the topology of SOS3. Secondary structural elements are defined by RasMol.⁵⁶ Elements involved in dimerization are highlighted with a yellow rectangle. A broken line separates the N-terminal domain and the C-terminal domain. The amino acid residues not defined in the crystal structure are displayed in the one-letter code. (c) A section of the experimental electron density map (contoured at 1σ) at the Ca²⁺-binding site EF1 of SOS3.

an unambiguous positioning of four Ca²⁺ and all their peptidic ligands (Figure 1(c)). The joined analysis of the structure of metal-binding sites of SOS3 and SCaBP1/AtCBL2 reveals a unique binding mode for the SCaBP/CBL family of proteins. The SCaBPs Ca²⁺-binding sites substitute the side-

chain oxygen donor at position Y by a backbone carbonyl group and the EF1 loop rearranges to accommodate the insertion of two residues (Figure 2). Interestingly, despite these differences, no significant change is observed in the coordination geometry of Ca²⁺ when the structures of

Table 1. Oligomeric state of SOS3 in solution as a function of protein and Ca²⁺ concentrations

[SOS3]:[Ca ²⁺]	1 : 0	0.5 : 0	1 : 1	1 : 3	1 : 10	1 : 30	1 : 100
Molecular mass (kDa)	42.1	37.1	42.1	44.4	45.7	47.5	50.6
Fraction of dimer, α	0.45	0.27	0.45	0.55	0.61	0.71	0.90

[SOS3]:[Ca²⁺] stands for the ratio between the concentrations of SOS3 and Ca²⁺. Data are referred to [SOS3]= 2.72×10^{-5} M. The calculated molecular mass of SOS3 is 26.0 kDa.

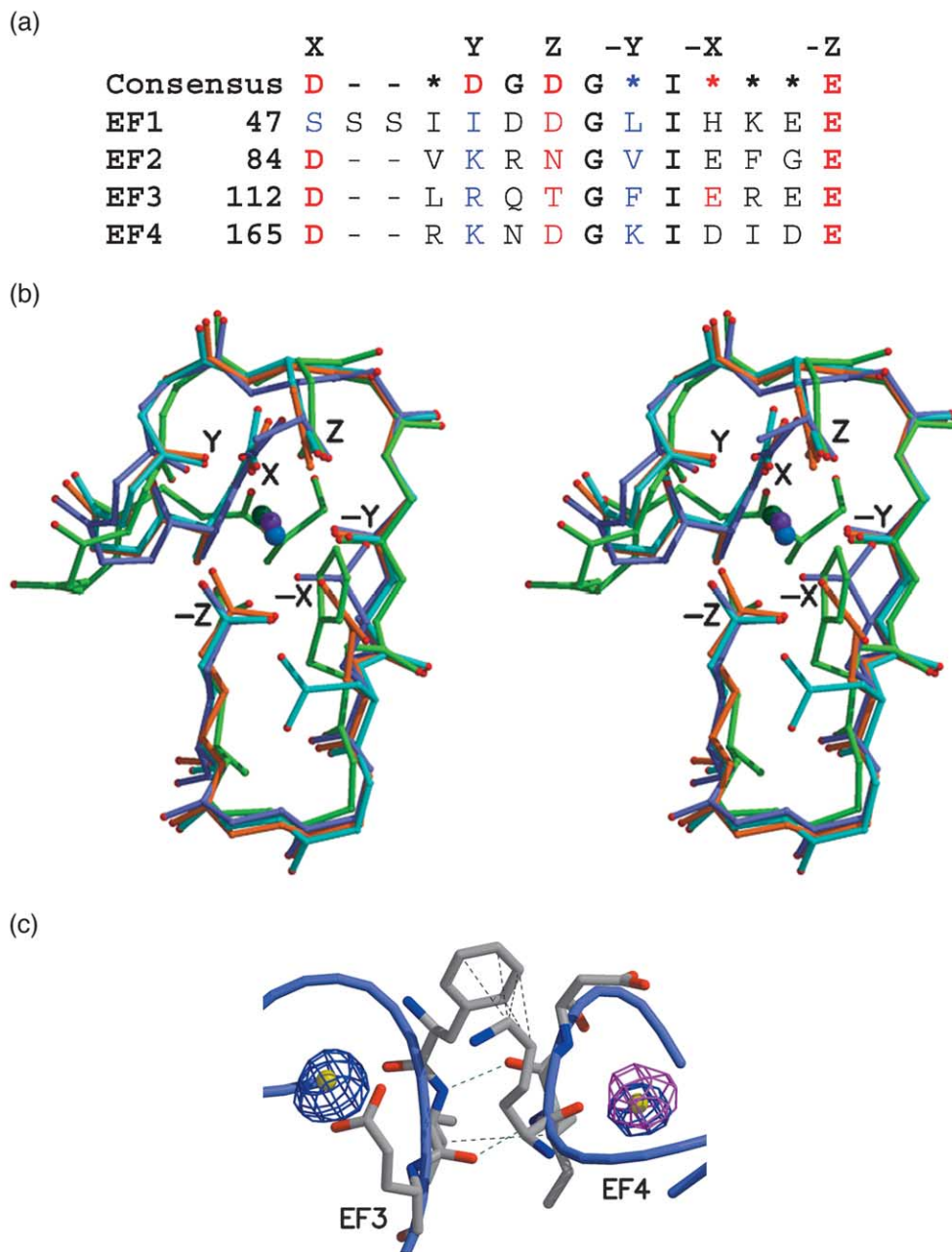


Figure 2. (a) A comparison of the sequence of the EF Ca^{2+} -binding sites of SOS3 and the classical EF-hand superfamily. Residues involved in Ca^{2+} binding are highlighted by X, Y, Z, -X, -Y, -Z according to a classical EF hand. Red and blue colors stand for side-chain or main-chain oxygen donor, respectively. (b) Stereoview of the structural superposition of Ca^{2+} -binding sites of SOS3. EF1, EF2, EF3 and EF4 are depicted in green, cyan, orange and lilac, respectively. Oxygen atoms are displayed in red. (c) The interactions between EF3 and EF4 sites of SOS3. Green and black broken lines stand for hydrogen bonds and for hydrophobic interactions, respectively. Ca^{2+} sites are depicted as yellow balls. Two sections of the electron density omit map (contoured at 4σ) of SOS3 Ca^{2+} complex and the anomalous difference electron density map (contoured at 4σ) of SOS3 Ca^{2+} Mn^{2+} complex are depicted in blue and magenta, respectively.

SCaBPs and a classical EF-hand are compared. Ca^{2+} coordinates seven protein oxygen atoms at site EF3 and six protein oxygen atoms and a water molecule at site EF4. In both cases the metal ion displays the classical bipyramidal pentagonal coordination geometry. On the other hand, Ca^{2+} at sites EF1 and EF2 coordinates only six oxygen

ligands, since the interaction between side-chains at positions Z and -X hinders the approach of another ligand.

Both the N-terminal domain and the C-terminal domain display a short antiparallel β sheet located between adjacent metal-binding sites. These β sheets are stabilized by the hydrophobic interaction

between the side-chains of the residues at position $-Y$ and the conserved isoleucine residue at position $-Y+1$ (Figures 1 and 2). These interactions suggest that SOS3 displays cooperativity in the Ca^{2+} binding between adjacent EF hands, as has been observed for other proteins of the EF hand superfamily.^{27–30} For calmodulins, cooperativity ensures that a full activation occurs in a narrow region of Ca^{2+} concentration during a signaling event. Remarkably, this pattern of interactions is not conserved in the structure of SCaBP1/AtCBL2, where the cooperativity is not observed, since it displays only one functional EF hand in each domain. This may be sufficient for a differential transduction of the Ca^{2+} signal by SOS3 and by SCaBP1/AtCBL2.

The crystal structures of SOS3 in complex with Ca^{2+} and with Ca^{2+} and Mn^{2+} are identical, with changes confined to the EF4 metal-binding site. The anomalous difference map calculated with diffraction data collected at a wavelength of 1.75 Å was used to demonstrate the presence of Mn^{2+} solely at site EF4 (Figure 2(c)). Our data show that Mn^{2+} displays the same coordination as that observed for Ca^{2+} .

The thermal denaturation of a protein at increasing concentrations of a ligand can be used to test if a protein binds a particular ligand.³¹ Consequently, we monitored the thermal stability of SOS3 in the presence of increasing concentrations of different metal ions by circular dichroism spectroscopy (Materials and Methods, and see Figure 3). Native protein purified as described in Materials and Methods was used as a control. Our results show that only Ca^{2+} and Mn^{2+} are able to bind SOS3, since they induce a change in the thermal denaturation temperature (T_m) of the protein. The presence of an excess of EDTA in the medium precludes any variation of the T_m . The analysis of the variation of T_m with the concentration of the ligand indicates that the apparent averaged affinity of SOS3 for Mn^{2+} is at least one order of magnitude higher than that observed for Ca^{2+} . This is not surprising, since the concentration of Ca^{2+} in the cytosol should be much higher than that expected for Mn^{2+} . Mn^{2+} binding to SOS3 may be biologically relevant, since SOS2 kinase and other protein kinases from the PKS/CIPK family display enhanced activity when Mn^{2+} acts as a cofactor.^{12,32,33} Hence, SOS3 or other SCaBP/CBL proteins could act as a carrier for this cofactor or, alternatively, could buffer the availability of free Mn^{2+} to prevent a constitutive activation of the kinase. However, to address this question, it would be necessary to characterize the affinity of this group of proteins for Mn^{2+} and to measure the concentration of free Mn^{2+} in the cytosol.

As mentioned above, Ca^{2+} triggers the dimerization of SOS3. It is likely that the binding at sites EF3 and EF4 is responsible for the self-association of the macromolecule, since they are closer than EF1 and EF2 to the dimerization interface (Figure 1(a)). However, the sedimentation equilibrium

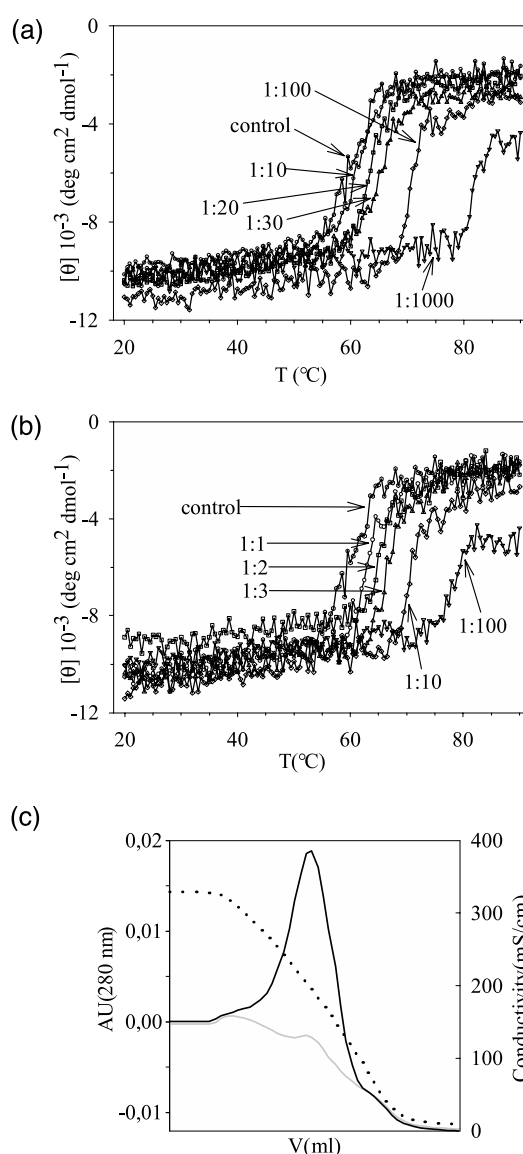


Figure 3. Comparison of the circular dichroism denaturation profiles of SOS3 as a function of (a) Ca^{2+} and (b) Mn^{2+} concentrations. The SOS3:metal ion ratio ($[\text{SOS3}]:[\text{Ca}^{2+}]$) stands for the relationship between the molar concentrations of protein and metal ion in the solution. Data are referred to $[\text{SOS3}] = 7.7 \times 10^{-6}$ M. Profiles corresponding to $[\text{SOS3}]:[\text{Ca}^{2+}]$ ratio 1:3 and 1:6 are omitted for clarity. $[\theta]$ stands for the mean residues ellipticity. (c) Hydrophobic interaction chromatography of SOS3. The dotted line represents the conductivity of the solution. The black and grey lines represent the absorbance recorded at 280 nm in the absence or in the presence of 10 mM CaCl_2 , respectively.

experiments show that Mn^{2+} does not affect the dimerization equilibrium of SOS3. This suggests that Ca^{2+} binding at site EF3 stabilizes the dimer in solution, as Mn^{2+} binds only to EF4. The fact that a deletion of three amino acids at EF3 produces a non-functional mutant protein⁶ suggests that Ca^{2+} binding at EF3 and probably dimerization are physiologically relevant.

SOS3 enhances its hydrophobic character upon calcium binding

As mentioned in Introduction, the SOS2 family of protein kinases interacts with the SCaBPs through a hydrophobic conserved motif of 21 amino acid residues known as the FISL motif.⁹ Hence, it would be expected that calcium activation of SOS3 and subsequent dimerization would enhance the hydrophobic character of the molecular surface of the macromolecule. In order to assess this hypothesis, we performed a hydrophobic interaction chromatography with SOS3. This method is based on the varying strengths of the hydrophobic interactions of proteins with hydrophobic ligands immobilized to a matrix. A moderately high concentration of salt promotes adsorption of the biomolecules and a linear decrease in salt concentration leads to the elution according to their increased hydrophobicity. In a Ca^{2+} -free medium, SOS3 elutes at a medium concentration of salt. On the other hand, in the presence of Ca^{2+} , SOS3 is adsorbed strongly to the resin even at low concentrations of salt. The elution is achieved by adding EDTA to the column. (Figure 3(c); and see Materials and Methods). This demonstrates that Ca^{2+} induces a conformational change in SOS3 leading to an increased exposure of hydrophobic residues.

The structural comparison of SOS3 and SCaBP1/AtCBL2 suggests a model for sensing Ca^{2+}

The main differences between SOS3 and the structural homologs arise from the relative positions of the N and C-terminal domains. This can be measured as the rotation angle that relates the C-terminal domains and the distance between the corresponding centers of mass when the N-terminal domains of the structural homologs are superimposed on that of SOS3. Table 2 shows these parameters together with the sequence identity and the RMSD of the C^α backbone of SOS3 with respect to SCaBP1/AtCBL2, CnB and NCS1. These data indicate that there is no relationship between the differences between domains, the local RMSD and the sequence identity. Rather, as discussed by Nagae *et al.*,¹⁷ local changes motivated by sequence differences at loop regions should account for the structural changes among the

proteins of the family. However, SOS3 and SCaBP1/AtCBL2 display the largest differences in the relative position of the N terminus and the C terminus, despite the fact that they share the highest level of sequence identity, have no insertion at loops and belong to the same family of proteins. This suggests that local differences motivated by Ca^{2+} binding could be responsible for global changes in the structure, as has been reported for other EF hand proteins.^{26,34} SCaBP1/AtCBL2 binds Ca^{2+} at sites EF1 and EF4, whereas EF2 and EF3 display an open conformation stabilized by electrostatic interactions and a water molecule. The structural comparison of SOS3 and SCaBP1/AtCBL2 at these sites reveals noticeable differences between the relative positions of helices E2 and F2, and E3 and F3. Figure 4 shows the superposition of the EF2 and EF3 sites from SOS3 and SCaBP1/AtCBL2. As can be seen, the interaction of a glutamate side-chain at position -Z with Ca^{2+} motivates a general displacement of helix F2 or F3 with respect to helix E2 or E3. These rearrangements upon Ca^{2+} binding could be responsible for the large change in the relative positions of the N and C-terminal domains, since EF2 and EF3 are in the interdomain interface.

The structural differences between SOS3 and SCaBP1/AtCBL2 reflect a change in the global shape and properties of the macromolecule surface that facilitates dimer formation. As mentioned above, SOS3 displays a large accessible hydrophobic path at the dimerization interface, which is around 25% larger than the hydrophobic accessible area observed in SCaBP1/AtCBL2 in this region.²⁰ In addition, the relative position of the N-terminal domain with respect to the C-terminal domain of SCaBP1/AtCBL2 would hinder the formation of the dimer. This fact suggests that calcium binding, domain swiveling and dimerization are coupled processes in SOS3. Unfortunately, there are no biophysical data on the oligomerization state of other SCaBP/CBL proteins; however, the high level of sequence homology shared by the family of proteins suggests that other members of the SCaBP/CBL would display the same mechanism of calcium regulation based on the dimerization.

Nagae *et al.*¹⁷ suggest that the recognition mechanism between SCaBP1/AtCBL2 and its kinase target is similar to that observed for CnB and calcinerurin A (CnA). CnA expands a helix to a

Table 2. Comparison between SOS3 and SCaBP1/AtCBL2, CnB and NCS1

		SCaBP1	CnB	NCS1
SOS3	Identity (%)	50.4	19.8	26.6
	Domain swiveling (deg.)	33	26	24
	Distance between C_M (Å)	4.3	3.2	1.4
	RMSD for the N-term domain (Å)	1.73 (77)	1.05 (70)	1.41 (74)
	RMSD for the C-term domain (Å)	1.22 (84)	1.38 (69)	1.63 (67)

Values within parentheses stand for the number of equivalent (cut-off distance of 3.5 Å) C^α atoms in the structural alignment. C_M stands for the center of mass. The PDB codes are 1UHN, 1AUI and 1G8I for SCaBP1/AtCBL2, CnB and NCS1, respectively.

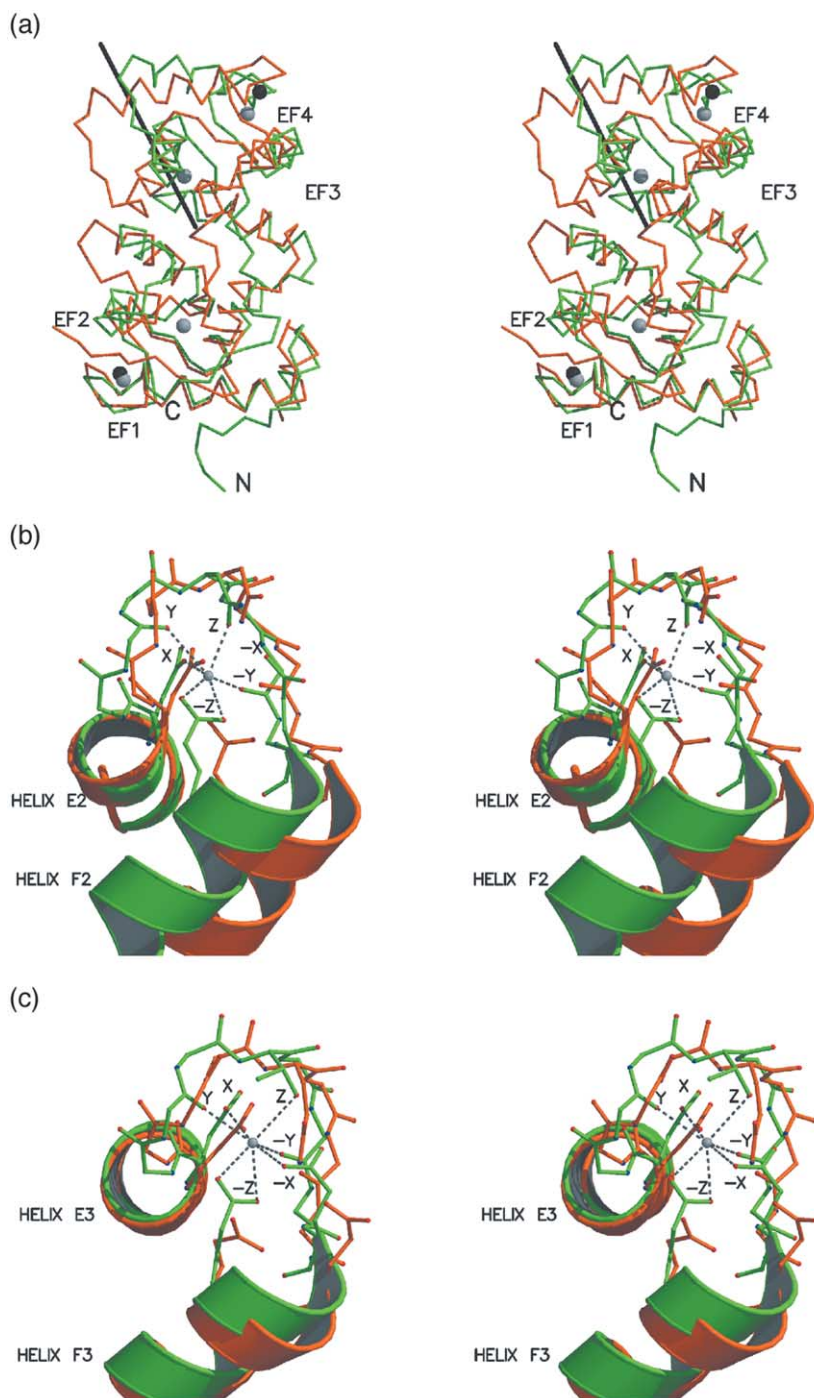


Figure 4. Stereoviews comparing the X-ray structures of SOS3 (green) and AtCBL2 (orange). (a) Backbone differences when the N-terminal domains are superimposed. Ca^{2+} -binding sites are displayed as gray and black balls for SOS3 and AtCBL2, respectively. The black stick represents the rotation axis relating the C-terminal domains. (b) and (c) Structural differences at EF2 and EF3 sites when helices E2 and E3, respectively, are superimposed.

hydrophobic crevice defined between the C-terminal tail and the loops F2-E3 and F3-E4 of CnB. The monomeric SCaBP1/AtCBL2 displays a close conformation of the crevice and it is suggested that Ca^{2+} would induce a domain swiveling, leading to a molecular conformation that would expose the hidden hydrophobic residues of this cavity. This would allow the interaction of SCaBP1/AtCBL2 with the FISL motif of the corresponding kinase. However, the structural analysis of SOS3 suggests that the activation of SOS2 could follow a different mechanism. The structural superposition of the fully Ca^{2+} -occupied SOS3 and SCaBP1/

AtCBL2 shows that the swiveling between domains tends to close the crevice. In addition, the dimeric structure of SOS3 would preclude the interaction between SOS3 and SOS2 in the manner observed in the CnA-CnB complex. Rather, the analysis of the molecular surface properties of SOS3 indicates that there are two hydrophobic patches that are located at the interdomain and the intersubunit interfaces. In addition, the electron density at the F2-E3 loop suggests some degree of mobility for this region of the protein. This loop is highly charged, since it displays sequence Arg-Asn-Arg-Asn-Arg-Arg-Asn and it is shielding a group of hydrophobic residues.

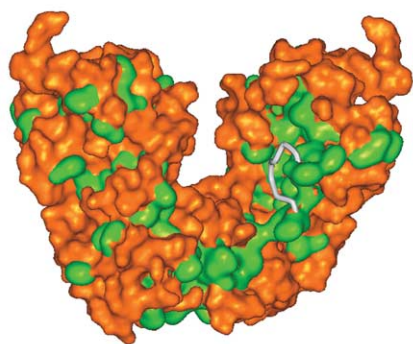


Figure 5. Molecular surface representation of SOS3. SOS3 is rotated 15° along the dimer axis from the view shown in Figure 1. Hydrophobic residues are depicted in green. The mobile loop from A67 to A73 has been excluded from the surface calculation and is represented as a white ribbon.

These occluded residues, together with other hydrophobic residues at the molecular surface of the dimeric SOS3, form an elongated patch involving both subunits within the dimer that may interact with the hydrophobic FISL motif from SOS2 (Figure 5).

Conclusions

The crystal structure of Ca²⁺-bound dimeric SOS3, together with the biophysical characterization of the macromolecule in solution, suggest a mechanism for molecular function of SOS3 and for the activation of its molecular partner, the kinase SOS2. Our data show that SOS3 is able to sense the cytosolic Ca²⁺ signal elicited by salt stress.³ We show that this promotes the dimerization of the macromolecule and a conformational change that reflects an increase of the hydrophobic character of the macromolecule. These changes would be sufficient to transmit the Ca²⁺ signal to SOS2, since they involve a dramatic modification of the global shape and molecular surface properties of the macromolecule.

Materials and Methods

Protein preparation, crystallization and data collection

SOS3 was obtained from cultures of *Escherichia coli* and purified to homogeneity as described by Sánchez-Barrena *et al.*³⁵ To summarize, the overexpressed His-tagged protein was purified in two steps. First, a protocol for nickel-affinity chromatography followed by an in-column cleavage with thrombin protease (Amersham Biosciences Limited, UK) was developed. Second, SOS3 was loaded onto a Resource-Q anionic-exchange column (Amersham Biosciences Limited, UK) and eluted with a NaCl gradient

from 50 mM to 0.5 M. Pure protein was treated with a chelating resin (Hampton Research, CA, USA) that complexes preferentially divalent metal ions.³⁶ Finally, the protein was concentrated to 15 mg/ml in 20 mM Tris-HCl (pH 7.5), 50 mM NaCl.

Prior to crystallization, the protein was incubated with 3.8 mM CaCl₂ (SOS3-Ca) and with 3.8 mM CaCl₂ plus 3.8 mM MnCl₂ (SOS3-Ca-Mn) to reach a final protein concentration of 10 mg/ml and a protein to metal molar concentration ratio of 1 : 10 and 1 : 10 : 10, respectively. Crystallization experiments were carried out at room temperature. Crystals corresponding to the SOS3 Ca²⁺ complex were grown using vapor-diffusion techniques from drops containing SOS3-Ca, reservoir solutions (24% (w/v) methyl-pentanediol (MPD), 18% (w/v) polyethylene glycol 4000 and 0.1 M sodium citrate, pH 4.8) and 1 M NaI in a ratio of 2 : 1 : 0.75 (by vol.). Crystals corresponding to the SOS3 Ca²⁺ Mn²⁺ complex were also grown with the same crystallization technique and solutions except that the concentration of polyethylene glycol 4000 in the reservoir solution was 16% (w/v).

Crystals were mounted in a fiber loop and frozen at 100 K in a nitrogen stream. No cryoprotectant solution was employed, since the reservoir solution had cryoprotectant properties.

X-ray diffraction data were collected in a CCD detector using the ESRF Grenoble synchrotron radiation source at wavelengths 0.92 Å and 1.75 Å at the BM16 beam-line. Diffraction data were processed using MOSFLM (Table 3).³⁷

Structure determination

Attempts to solve the X-ray structure of SOS3 by molecular replacement using the coordinates of the homologous structure AtCBL2/SCaBP1 failed (AMoRe,³⁸ MOLREP³⁹). However, a poor solution was found using the structure of the neuronal calcium-sensor 1 (NCS1) (PDB code 1G8I) as a search model.⁴⁰ Using this solution, it was possible to identify seven iodine atoms in the asymmetric unit (a.u.). Then, the structure of SOS3 was solved by single anomalous scattering (SAD), using the anomalous information corresponding to the iodine atoms from diffraction data of the SOS3 Ca²⁺ complex, collected at wavelength 1.7 Å.^{41,42} Refinement of the iodine substructure and phasing were performed using SHARP.⁴³ Solvent-flattening was performed with DM⁴⁴ and SOLOMON,^{45,46} using the SHARP protocols.

The solvent-flattened electron density map was good enough (Figure 1(c)) to trace an initial model that was first refined using the simulated annealing routine of CNS.⁴⁷ Several cycles of restrained refinement with REFMAC5⁴⁸ and iterative model building with O⁴⁹ were carried out. Water structure and two molecules of MPD from the solvent were also modeled. The myristoylation site (N terminus) and the C terminus of the protein appear disordered in the crystal and it was not possible to model these parts of the macromolecule. The refined structure of SOS3 in the presence of Ca²⁺ at 2.75 Å resolution (data measured at wavelength 0.92 Å) was used to solve the structure of the SOS3 Ca²⁺ Mn²⁺ complex (Table 3). The difference anomalous electron density map calculated from the diffraction data of the SOS3 Ca²⁺ Mn²⁺ complex corresponding to the wavelength 1.75 Å was used to identify a Mn²⁺ at EF4. Calculations were performed using CCP4 programs.⁵⁰

The stereochemistry of the models was verified with PROCHECK.⁵¹ Ribbon Figures were produced using

Table 3. Data collection, structure solution and refinement statistics

	SOS3 Ca ²⁺ complex		SOS3 Ca ²⁺ Mn ²⁺ complex	
A. Data collection				
Wavelength (λ) (Å)	0.92	1.75	0.92	1.75
Space group	I4 ₁		I4 ₁	
Unit cell parameters				
<i>a</i> (Å)	91.6		91.9	
<i>b</i> (Å)	91.6		91.9	
<i>c</i> (Å)	85.4		85.6	
Resolution (Å)	62.0–2.75	45.6–2.8	46.0–3.0	49.9–3.0
<i>I</i> /σ(<i>I</i>)	4.5 (1.8)	2.6 (1.4)	4.5 (1.5)	2.6 (1.9)
<i>R</i> _{merge} ^a (%)	5.7 (41.8)	11.7 (47.5)	8.6 (51.2)	10.6 (39.7)
Completeness (%)	100 (100)	100 (100)	97.7 (97.7)	96.7 (96.7)
Multiplicity	3.7 (3.7)	7.5 (7.4)	4.5 (4.5)	3.9 (3.8)
Temperature (K)	100		100	
B. Structure solution				
Number of sites		7		
<i>R</i> _{Cullis} ^b		0.68		
Phasing power ^b		1.65		
FOM ^b		0.30		
C. Refinement				
Resolution range (Å)	20–2.75		20–3.0	
<i>R</i> -factor (%)	23.5 (41.6)		23.7 (32.9)	
<i>R</i> _{free} (%)	26.4 (37.1)		26.2 (39.8)	
Number of atoms	1614		1614	
Number of reflections	9213		6609	
D. Model				
Molecules per a.u.	1		1	
Solvent content (% v/v)	64.3		64.7	
Amino acid residues	188		188	
Calcium ions	4		3	
Manganese ions	0		1	
Iodine ions	7		7	
MPD molecules	2		2	
Water molecules	40		40	
RMS deviation:				
Bond lengths (Å)	0.014		0.016	
Bond angles (deg)	1.735		1.926	
Planar groups (Å)	0.005		0.006	
Remarks	Residues A1–A14 and A203–A222 are not included in the model. The loop A67–A75 displays poor electron density.			

^a $R_{\text{merge}} = \sum_{hkl} |I_{hkl} - \langle I_{hkl} \rangle| / \sum_{hkl} I_{hkl}$. Values in parentheses are for the highest-resolution shell: 2.90–2.75 Å and 2.95–2.80 Å (for processing statistics of the SOS3 Ca²⁺ complex and λ 0.92 Å and 1.75 Å, respectively), and 3.16–3.00 Å (for processing statistics of the SOS3 Ca²⁺ Mn²⁺ complex). The highest-resolution shell for refinement statistics are 2.82–2.75 Å (for the SOS3 Ca²⁺ complex) and from 3.08–3.00 Å (for the SOS3 Ca²⁺ Mn²⁺ complex).

^b *R*_{Cullis}, Phasing power and FOM stand for the phasing statistics determined by SHARP.⁴² Solvent content calculations were performed following Andersson & Hovmöller.⁵⁷

MOLSCRIPT⁵² and Raster3D.⁵³ The accessible surface area of SOS3 dimer and protomer was calculated with the program NACCESS from the LIGPLOT package.²⁰ The domain swiveling angles of SOS3, NCS1, AtCBL2/SCaBP1 and CnB were calculated with the program FIT†. Sequence alignment data were developed with BLAST.⁵⁴ The molecular surface of SOS3 dimer Figure was produced with Web Lab‡.

Sedimentation equilibrium

Sedimentation equilibrium experiments were performed in a Beckman Optima XL-A ultracentrifuge using a Ti50 rotor and six-channel centerpieces of Epon-charcoal (optical pathlength 12 mm). Samples of SOS3 were equilibrated against 20 mM Tris–HCl (pH 7.5),

50 mM NaCl buffer plus Ca²⁺, Mn²⁺ and Mg²⁺ when appropriate (Table 1), and were centrifuged at 17,000 rpm, 20,000 rpm and 29,000 rpm and 20 °C. Radial scans at 280 nm were taken at 12 hours, 14 hours and 16 hours. The three scans were identical (equilibrium conditions were reached). The weight-average molecular mass (*M*_w) of SOS3 was determined by using the program EQASSOC with the partial specific volume of SOS3 set to 0.730 at 20 °C as calculated from its amino acid composition. Calculations were performed with data at 17,000 rpm. The dimerization constant (*K*) and the fraction of dimer in solution (*α*) were calculated from the observed *M*_w as a function of the concentration of protein and Ca²⁺, following Rivas *et al.*⁵⁵

Circular dichroism spectra

Circular dichroism spectra were recorded in a JASCO J-810 spectropolarimeter. Far-UV spectra were recorded

† <http://bioinfo1.mbfys.lu.se/~guoguang/fit.html>

‡ <http://www.msi.com>

in a 0.1 cm pathlength quartz cell, at a protein concentration of 0.2 mg/ml (7.7×10^{-6} M). Near-UV CD spectra were registered at protein concentration of 0.5 mg/ml (1.9×10^{-5} M) in a 1 cm pathlength quartz cell. The observed ellipticities were converted to mean residue ellipticities $[\theta]$ using an average molecular mass per residue of 115.4 Da. Thermal denaturation was monitored at a wavelength of 222 nm by increasing the temperature from 20 °C to 90 °C at 30 deg. C/hour and allowing the temperature to equilibrate for one minute before recording the spectra.

Samples were prepared as for sedimentation equilibrium experiments. Native protein obtained for purification was considered as a control. Far and near-UV spectra and the thermal denaturation profiles were recorded for: (i) the native protein; (ii) the native protein in the presence of Na^+ , K^+ , Rb^+ , Cs^+ , Ni^{2+} , Mg^{2+} , in SOS3:metal ion molar concentration ratios of 1:10 and 1:100 (data not shown); (iii) the native protein in the presence of Ca^{2+} in molar concentration ratios of 1:3, 1:6, 1:10, 1:20, 1:30, 1:60, 1:100 and 1:1000 (Figure 3(a)); (iv) the native protein in the presence of Mn^{2+} in molar concentration ratios 1:1, 1:2, 1:3, 1:10 and 1:100 (Figure 3(b)).

Hydrophobic interaction chromatography

Native SOS3 (0.22 mg) was loaded onto a phenyl Sepharose column (Amersham Biosciences Limited, UK) previously equilibrated with buffer 1 (1.7 M ammonium sulfate, 20 mM Tris-HCl (pH 7.0), 0.05% (w/v) NaN_3) and subjected to a washing step with five column volumes of buffer 1, followed by a gradient step of 20 column volumes with an ammonium sulfate gradient from 1.7 M to 0 M. In these conditions, SOS3 elutes at 0.9 M ammonium sulfate (Figure 3(c)). The same chromatographic steps were performed in the presence of 10 mM CaCl_2 . In these conditions, SOS3 does not elute from the column. Elution was achieved by removal of bound calcium with 45 mM Tris-HCl (pH 7.0), 10 mM EDTA.

Protein Data Bank accession numbers

The coordinates and structure factors amplitudes of SOS3 Ca^{2+} and SOS3 Ca^{2+} and Mn^{2+} complexes have been deposited in the PDB with accession codes 1V1G and 1V1F, respectively.

Acknowledgements

We thank Dr M. Menéndez for critical comments on the manuscript and Dr J. L. Saiz-Velasco for the analytical ultracentrifugation data analysis. M.J.S.-B. was supported by an F.P.U. studentship from Ministerio de Educación, Cultura y Deporte. A.A. thanks the ESRF (Spanish beam-line BM16) for the access to the synchrotron radiation source. This work was funded by BMC2002-04011-C05-03 grant of the Spanish "Plan Nacional" (MCYT) to A.A. and by US National Institutes of Health grant R01GM59138 to J.K.Z.

References

1. Tester, M. & Davenport, R. (2003). Na^+ tolerance and Na^+ transport in higher plants. *Ann. Bot. (London)*, **91**, 503–527.
2. Hasegawa, P. M., Bressan, R. A., Zhu, J. K. & Bohnert, H. J. (2000). Plant cellular and molecular responses to high salinity. *Annu. Rev. Plant Mol. Plant Physiol.* **51**, 463–499.
3. Zhu, J. K. (2003). Regulation of ion homeostasis under salt stress. *Curr. Opin. Plant Biol.* **6**, 441–445.
4. Zhu, J. (2000). Genetic analysis of plant salt tolerance using *Arabidopsis thaliana*. *Plant Physiol.* **124**, 941–948.
5. Liu, J. & Zhu, J. K. (1998). A calcium sensor homolog required for plant salt tolerance. *Science*, **280**, 1943–1945.
6. Ishitani, M., Liu, J., Halfter, U., Kim, C. S., Shi, W. & Zhu, J. K. (2000). SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell*, **12**, 1667–1678.
7. Liu, J., Ishitani, M., Halfter, U., Kim, C. S. & Zhu, J. K. (2000). The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc. Natl Acad. Sci. USA*, **97**, 3730–3734.
8. Halfter, U., Ishitani, M. & Zhu, J. K. (2000). The *Arabidopsis* SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3. *Proc. Natl Acad. Sci. USA*, **97**, 3735–3740.
9. Guo, Y., Halfter, U., Ishitani, M. & Zhu, J. K. (2001). Molecular characterization of functional domains in the protein kinase SOS2 that is required for salt tolerance. *Plant Cell*, **13**, 1383–1399.
10. Quintero, F. J., Ohta, M., Shi, H., Zhu, J. K. & Pardo, J. M. (2002). Reconstitution of the SOS signaling pathway for Na^+ homeostasis in plants. *Proc. Natl Acad. Sci. USA*, **99**, 9061–9066.
11. Qiu, Q., Guo, Y., Dietrich, M. A., Schumaker, K. S. & Zhu, J. K. (2002). Regulation of SOS1, a plasma membrane Na^+/H^+ exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc. Natl Acad. Sci. USA*, **99**, 8436–8441.
12. Luan, S., Kudla, J., Rodriguez-Concepcion, M., Yalovsky, S. & Gruissem, W. (2002). Calmodulins and calcineurin B-like proteins, calcium sensors for specific signal response coupling in plants. *Plant Cell*, **14**, S389–S400.
13. Gong, D., Guo, Y., Jagendorf, A. T. & Zhu, J. K. (2002). Biochemical characterization of the *Arabidopsis* protein kinase SOS2 that functions in salt tolerance. *Plant Physiol.* **130**, 256–264.
14. Harmon, A. C., Gribskov, M. & Harper, J. F. (2000). CDPKs—a kinase for every Ca^{2+} signal? *Trends Plant Sci.* **5**, 154–159.
15. Scrase-Field, S. A. & Knight, M. R. (2003). Calcium, just a chemical switch? *Curr. Opin. Plant Biol.* **6**, 500–506.
16. Sutcliffe, M. J., Haneef, I., Carney, D. & Blundell, T. L. (1987). Knowledge based modeling of proteins. 1. 3-Dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng.* **1**, 377–384.
17. Nagae, M., Nozawa, A., Koizumi, N., Sano, H., Hashimoto, H., Sato, M. & Shimizu, T. (2003). The crystal structure of the novel calcium-binding protein AtCBL2 from *Arabidopsis thaliana*. *J. Biol. Chem.* **278**, 42240–42246.
18. Cyert, M. S. & Thorner, J. (1992). Regulatory subunit (CNB1 gene product) of yeast Ca^{2+} /calmodulin-

- dependent phosphoprotein phosphatases is required for adaptation to pheromone. *Mol. Cell. Biol.* **12**, 3460–3469.
19. Schaad, N. C., de Castro, E., Nef, S., Hegi, S., Hinrichsen, R., Martone, M. E. *et al.* (1996). Direct modulation of calmodulin targets by the neuronal calcium sensor NCS1. *Proc. Natl Acad. Sci. USA*, **93**, 9253–9258.
 20. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT, a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134.
 21. Nooren, I. M. & Thornton, J. M. (2003). Diversity of protein–protein interactions. *EMBO J.* **22**, 3486–3492.
 22. Nooren, I. M. & Thornton, J. M. (2003). Structural characterization and functional significance of transient protein–protein interactions. *J. Mol. Biol.* **325**, 991–1018.
 23. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198.
 24. Killian, J. A. & von Heijne, G. (2000). How proteins adapt to a membrane–water interface. *Trends Biochem. Sci.* **25**, 429–434.
 25. Monné, M., Nilsson, I., Johansson, M., Elmhed, N. & von Heijne, G. (1998). Positively and negatively charged residues have different effect on the position in the membrane of a model transmembrane helix. *J. Mol. Biol.* **284**, 1177–1183.
 26. Lewit-Bentley, A. & Rety, S. (2000). EF-hand calcium-binding proteins. *Curr. Opin. Struct. Biol.* **10**, 637–643.
 27. Zhang, M., Tanaka, T. & Ikura, M. (1995). Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nature Struct. Biol.* **2**, 758–767.
 28. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C. B. & Bax, A. (1995). Solution structure of calcium-free calmodulin. *Nature Struct. Biol.* **2**, 768–776.
 29. Finn, B. E., Evenas, J., Drakenberg, T., Waltho, J. P., Thulin, E. & Forsen, S. (1995). Calcium-induced structural changes and domain autonomy in calmodulin. *Nature Struct. Biol.* **2**, 777–783.
 30. Verdino, P., Westritschnig, K., Valenta, R. & Keller, W. (2002). The cross-reactive calcium-binding pollen allergen, Phl p 7, reveals a novel dimer assembly. *EMBO J.* **21**, 5007–5016.
 31. Varea, J., Saiz, J. L., Lopez-Zumel, C., Monterroso, B., Medrano, F. J., Arrondo, J. L. *et al.* (2000). Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase? *J. Biol. Chem.* **275**, 26842–26855.
 32. Gong, D., Guo, Y., Schumaker, K. S. & Zhu, J. K. (2004). The SOS3 family of calcium sensors and SOS2 family of protein kinases in *Arabidopsis*. *Plant Physiol.* **134**, 919–926.
 33. Shi, J., Kim, K. N., Ritz, O., Albrecht, V., Gupta, R., Harter, K. *et al.* (1999). Novel protein kinases associated with calcineurin B-like calcium sensors in *Arabidopsis*. *Plant Cell*, **11**, 2393–2405.
 34. Otterbein, L. R., Kordowska, J., Witte-Hoffmann, C., Wang, C. L. & Dominguez, R. (2002). Crystal structures of S100A6 in the Ca(2+)-free and Ca(2+)-bound states, the calcium sensor mechanism of S100 proteins revealed at atomic resolution. *Structure*, **10**, 557–567.
 35. Sánchez-Barrena, M. J., Martínez-Ripoll, M., Zhu, J. K. & Albert, A. (2004). SOS3 from *Arabidopsis thaliana*, expression, purification, crystallization and preliminary X-ray analysis. *Acta Crystallog. sect. D*, **60**, 1272–1274.
 36. Dunn, M. F., Pattison, S. E., Storm, M. C. & Quiel, E. (1980). Removal of metals from enzyme solutions. *Biochemistry*, **19**, 718.
 37. Leslie, A. G. W. (1987). MOSFLM. In *Proceedings of the CCP4 Study Weekend* (Machin, J. R. & Papiz, M. Z., eds), pp. 39–50, SERC Daresbury Laboratory, Warrington, UK.
 38. Navaza, J. (1994). AMoRe, an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
 39. Vagin, A. & Teplyakov, A. (1997). MOLREP, an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.
 40. Bourne, Y., Dannenberg, J., Pollmann, V., Marchot, P. & Pongs, O. (2001). Immunocytochemical localization and crystal structure of human frequenin (neuronal calcium sensor 1). *J. Biol. Chem.* **276**, 11949–11955.
 41. Sánchez-Barrena, M. J., Martínez-Ripoll, M., Galvez, A., Valdivia, E., Maqueda, M., Cruz, V. & Albert, A. (2003). Structure of bacteriocin AS-48, from soluble state to membrane bound state. *J. Mol. Biol.* **334**, 541–549.
 42. Evans, G. & Bricogne, G. (2003). Triiodine derivatization in protein crystallography. *Acta Crystallog. sect. D*, **59**, 1923–1929.
 43. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous methods. *Methods Enzymol.* **276**, 472–494.
 44. Cowtan, K. & Main, P. (1998). Miscellaneous algorithms for density modification. *Acta Crystallog. sect. D*, **54**, 487–493.
 45. Leslie, A. G. W. (1987). A reciprocal-space method for calculating a molecular envelope using the algorithm of B. C. Wang. *Acta Crystallog. sect. A*, **43**, 134–136.
 46. Abrahams, J. P. & Leslie, A. G. W. (1996). Methods used in the structure determination of bovine mitochondrial F₁ ATPase. *Acta Crystallog. sect. D*, **52**, 30–42.
 47. Brunger, A. T. & Adams, P. D. (1998). Crystallography and NMR system. *Acta Crystallog. sect. D*, **54**, 905–921.
 48. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
 49. Jones, T. A., Zon, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for model building in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
 50. Bailey, S. (1994). The CCP4 suite, programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
 51. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK, a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
 52. Kraylis, P. J. (1991). MOLSCRIPT, a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
 53. Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869–873.
 54. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

55. Rivas, G., Ingham, K. C. & Minton, A. P. (1994). Ca(2+)-linked association of human complement C1s and C1r. *Biochemistry*, **33**, 2341–2348.
56. Anderson, K. M. & Hovmöller, S. (2000). The protein content in crystals and packing coefficients in different space groups. *Acta Crystallog. sect. D*, **56**, 789–790.
57. Sayle, R. A. & Milner-White, E. J. (1995). RasMol, biomolecular graphics for all. *Trends Biochem. Sci.* **20**, 374–376.

Edited by R. Huber

(Received 17 September 2004; received in revised form 5 November 2004; accepted 10 November 2004)

Structure of bacteriocin AS-48: from soluble state to membrane bound state.

Sánchez-Barrena, M.J., Martínez-Ripoll, M., Galvez, A., Valdivia, E., Maqueda, M., Cruz, V. & Albert, A. (2003) *J. Mol. Biol.*, **334(3)**, 541-549.

Structure of Bacteriocin AS-48: From Soluble State to Membrane Bound State

M. J. Sánchez-Barrena¹, M. Martínez-Ripoll¹, A. Gálvez², E. Valdivia³
M. Maqueda³, V. Cruz⁴ and A. Albert^{1*}

¹Grupo de Cristalografía
Macromolecular y Biología
Estructural, Instituto de
Química Física "Rocasolano"
Consejo Superior de
Investigaciones Científicas
Serrano 119, E-28006 Madrid
Spain

²Departamento de Ciencias de
la Salud, Área de Microbiología
Facultad de Ciencias
Experimentales, Universidad de
Jaén, 23071 Jaén, Spain

³Departamento de
Microbiología, Facultad de
Ciencias, Universidad de
Granada, 18071 Granada
Spain

⁴Centro Técnico de Informática
Consejo Superior de
Investigaciones Científicas
Pinar 19, 28006 Madrid, Spain

The bacteriocin AS-48 is a membrane-interacting peptide, which displays a broad anti-microbial spectrum against Gram-positive and Gram-negative bacteria. The NMR structure of AS-48 at pH 3 has been solved. The analysis of this structure suggests that the mechanism of AS-48 anti-bacterial activity involves the accumulation of positively charged molecules at the membrane surface leading to a disruption of the membrane potential. Here, we report the high-resolution crystal structure of AS-48 and sedimentation equilibrium experiments showing that this bacteriocin is able to adopt different oligomeric structures according to the physicochemical environment. The analysis of these structures suggests a mechanism for molecular function of AS-48 involving a transition from a water-soluble form to a membrane-bound state upon membrane binding.

© 2003 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: bacteriocin; cyclic peptide; cationic antibacterial peptides; protein crystallography; protein–membrane interaction

Introduction

Small α -helical peptides mediate simple anti-microbial defense mechanisms in a number of organisms.¹ The mechanism by which a soluble protein converts into a membrane bound protein remains a central question in understanding the molecular toxicity of bacterial protein toxins.² Traditionally, small α -helical peptides have been classified either as channel-forming peptides or carpet-like peptides according to their mechanism of membrane interaction. The first type undergoes

a series of remarkable changes in solubility, oligomerization state, structure and dynamics during the processes of membrane binding, assembly, membrane insertion and channel formation that implies the generation of bundles of α -helices.³ The second type forms a carpet of surface-bound helices, which at a sufficiently high peptide concentration, may result in a toroidal pore formation *via* local reorganization of the bilayer structure.¹

AS-48 is a 70 residue α -helical bacteriocin. It is produced by *Enterococcus faecalis* subsp. *liquefaciens* S-48 and it is unique in its natural cyclic structure in which N and C termini are linked by a peptidic bond. Its broad anti-microbial spectrum against Gram-positive and Gram-negative bacteria, together with its remarkable stability and solubility over a wide pH range, suggest that AS-48 could be a good candidate as a natural food preservative.^{4–7}

Abbreviations used: SIRAS, single isomorphous replacement with anomalous scattering; ASA, accessible surface area; DF, dimeric form.

E-mail address of the corresponding author:
xalbert@iqfr.csic.es

These properties are shared with other naturally occurring circular proteins, all of them involved in host defense.⁸ AS-48 appears to be derived from a larger precursor protein; however, nothing is known about the cleavage and cyclation steps to produce a mature protein. Interestingly, AS-48 is encoded in a plasmid together with other genes, which could be involved in maturation and secretion processes.⁹

There is extensive information about antimicrobial activity and physicochemical characteristics of the macromolecule.^{4–7} For instance, the primary target of AS-48 is the bacterial cell membrane, where it forms pores or channels of around 7 Å in diameter. This leads to a membrane permeation not dependent on membrane potential.¹⁰ It is also known that AS-48 activity is pH dependent and that is related to the oligomeric state of the protein.⁷ Furthermore, it has been shown that AS-48 migrates to the hydrophobic phase when a partition water/*n*-octanol experiment is performed.⁷ AS-48 also displays a pH dependent interaction with DPPA monolayers that involves a rearrangement of the structure of the protein.¹¹ These experiments suggest that AS-48 is inserted in the membrane prior to cell disruption. Despite this detailed information the mechanism for molecular function remains obscure.

The NMR structure of AS-48 at pH 3 showed that AS-48 is a monomer consisting of a bundle of five helices arranged with the topology of the saposin fold.¹² González *et al.*¹² proposed a mechanism for its anti-bacterial function which is based on the structural similarities between AS-48 and NK-lysin. It has been suggested that accumulation of positively charged NK-lysin molecules on the membrane surface induces the formation of pores or “molecular electroporation”.^{13,14} This mechanism does not imply that the peptide inserts into the membrane or undergoes a molecular reorganization for its activity. The NMR structure of AS-48 suggests that the positively charged residues of a monomeric AS-48 would bind to the membrane surface. The association of AS-48 molecules at this surface would cause membrane destabilization. However, this model does not agree with experimental studies showing that the permeabilizing activity does not depend on membrane potential.¹⁰ In addition, the distribution of positive charges in NK-lysin and in AS-48 is topologically different. The NK-lysin molecular surface displays a central ring of arginine and lysine residues while AS-48 displays a segregated distribution of positively charged residues and hydrophobic residues. Hence, a deeper analysis of structure–function relationship may help to understand the aspects of the molecular function that remain unclear.

In this work, we provide the crystal structure of the bacteriocin AS-48 together with sedimentation equilibrium data. Our study shows that AS-48 displays at least, two different modes of molecular association according to the physicochemical

environment. The analysis of these structures suggests a mechanism for molecular function of AS-48 that involves a structural transition between them upon membrane binding.

Results and Discussion

Molecular and crystal structure of AS-48

The X-ray structures of AS-48 at pH 4.5 and pH 7.5 (crystal form I) were solved by single isomorphous replacement with anomalous scattering (SIRAS) at 1.46 Å and 1.64 Å resolution, respectively. Despite the different crystallization conditions, both crystals are found to be isomorphous (see Table 1 and Materials and Methods). Overall, the X-ray structure of the AS-48 protomer is a five-helix bundle as observed by NMR techniques¹² (see Figures 1(a) and 2(a)). Four independent molecules of AS-48 are found in the asymmetric unit. The structure of these crystallographic independent molecules is very similar (the highest C^α backbone root-mean-square deviation, RMSD, is 0.4 Å¹⁵ with respect to the various copies in the asymmetric unit). Thus, hereinafter and unless stated, all the results will refer to the chain labeled A of the model at highest resolution (pH 4.5). In addition, a different crystal form of AS-48 was also solved at 2.8 Å resolution (crystal form II), (see Table 1 and Materials and Methods). This work shows that despite the different crystal environment, AS-48 adopts the molecular structure found in crystal form I (the highest C^α backbone RMSD value is 0.6 Å¹⁵ with respect to the various copies in the different crystal forms).

From the structural point of view, one of the most striking features of the AS-48 protomers is their remarkable amphipathic surface. A plane containing the C^α atoms of Glu4, Glu20, Glu49 and Glu58 segregates a patch of positively charged residues (α -helices H4 and H5 and the loop connecting H1 and H2) from the rest of the hydrophobic or uncharged surface residues (helices H1, H2 and H3) (see Figure 1(a)). According to this, the molecules of AS-48 in the crystal are arranged in chains of pairs of molecules linked either by hydrophobic interactions (dimeric form I, hereinafter abbreviated to DF-I) (see Figure 1(b)) or by hydrophilic interactions (dimeric form II, hereinafter abbreviated to DF-II) (see Figure 1(c)). Interestingly, crystal form II shows only DF-II. The molecules within the DF-I interact through the hydrophobic helices H1 and H2 (see Figure 1(b)). On the other hand, the hydrophilic surfaces of helices H4 and H5 are interacting in DF-II (see Figure 1(c)). Both dimers are built around a non-crystallographic 2-fold axis. The molecules of AS-48 within DF-II are further stabilized by the interaction of TyrB54, GluB58, GluD58 and LysD62 with a sulfate ion (SO₄²⁻) or a phosphate ion (H₂PO₄⁻) (for pH 4.5 and 7.5, respectively). However, at pH 4.5, the high-resolution electron

Table 1. Data collection, structure solution and refinement statistics

	Crystal form I			Crystal form II Native
	Native (pH 4.5)	NaI derivative	Native (pH 7.5)	
A. Data collection				
Space group	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222	<i>P</i> 4 ₁ 2 ₁ 2
Unit cell (Å)	79.47, 83.40, 99.83	79.79, 83.65, 99.43	79.66, 83.88, 99.78	76.49, 76.49, 51.52
Resolution limits (Å)	25.24 (1.46)	33.90 (2.40)	19.43 (1.64)	34.3 (2.70)
<i>I</i> / σ (<i>I</i>)	6.1 (2.4)	4.5 (3.3)	6.1 (1.4)	5.1 (1.3)
<i>R</i> _{merge} (%) ^a	6.4 (28.7)	12.3 (21.5)	8.5 (51.8)	13.8 (59.1)
Completeness (%)	94.6 (94.6)	100.0 (100.0)	90.0 (90.0)	100.0 (100.0)
Multiplicity	4.0 (3.0)	9.9 (9.9)	4.3 (4.1)	9.1 (9.0)
Temp (K)	120	120	120	120
B. Structure solution				
Number of sites		15		
<i>R</i> _{Cullis} (c/a) ^b		0.32/0.29		
Phasing power ^c (c/a)		2.78/4.05		
Mean FOM (c/a) ^d		0.37/0.26		
C. Refinement				
Resolution range (Å)	15.0–1.46		15.0–1.64	10.0–2.80
<i>R</i> factor (%)	22.6 (29.0)		19.4 (38.0)	23.6 (28.6)
<i>R</i> free (%)	22.35 (34.0)		21.48 (40.0)	25.6 (30.9)
Number of atoms	2369		2416	1100
Number of reflections	51,862		35,069	3777
Model				
Molecules per a.u.	4	4	4	2
Solvent content (%)	55	55	55	54
Amino acids	280		280	140
Water molecules	300		384	28
Sulfate ions	5			5
Phosphate ions			2	
Detergent molecule				1
Glycerol	1		1	1
RMS deviation from ideal:				
Bond lengths (Å)	0.009		0.010	0.022
Bond angles (deg.)	1.260		1.314	2.141
Planar groups (Å)	0.005		0.005	0.006
Observations	Loop C47–C49 is not well defined		Loop C47–C49 is not well defined	Loop C47–C49 is not well defined

^a $R_{\text{merge}} = \sum_{hkl} \sum_i |I_{hkl} - I_{hkl}| / \sum_{hkl} I_{hkl}$ for overall resolution range and the resolution shell from 1.54–1.46 Å (for native at pH 4.5), from 1.73–1.64 Å (for NaI derivative), from 1.73–1.64 Å (for native at pH 7.5) and from 2.70–2.85 Å (for crystal form II).

^b $R_{\text{Cullis}} = \sum[|F_H| - (|F_{PH} - |F_P|)] / \sum|F_H|$ for centric (c) and acentric (a) reflections.

^c Phasing power = mean value of the heavy-atom structure-factor amplitudes divided by the residual lack of closure.

^d FOM, figure of merit. Values in parentheses are for the highest resolution shell: 1.54–1.46 Å for native data at pH 4.5, 1.73–1.64 Å for native data at pH 7.5, 1.73–1.64 Å for the derivative and 2.70–2.85 Å for crystal form II. The native structure at pH 4.6 is taken as a reference for the values given for NaI derivative. Solvent content calculations were performed following Andersson & Hovmöller.⁴³

density map clearly shows that Glu58 and Glu58 side-chains display double conformation. Interestingly, only one conformer effectively interacts with the SO_4^{2-} . This should be related with to the protonation state of the carboxylate side-chains. A glycerol molecule from the cryo-protectant has also been found in the interface between protomers. In addition, crystal form II reveals a molecule of detergent (*n*-decyl- β -D-maltoside) in the solvent-exposed hydrophobic interface between the protomers forming DF-II (see Figure 3(a)).

Structural comparison of the X-ray and NMR model

Overall, the X-ray structure of the AS-48 protomers (see Figures 1(a) and 2(a)) is the same as the structure obtained using NMR techniques.¹²

However, the RMSD for C α backbone of the X-ray structure with respect to the average NMR model is 1.6 Å. Following Gronenberg *et al.*,¹⁶ one would expect a value of around 1.0 Å considering the precision reported for the NMR model.¹² The superposition of both models (see Figure 2(a)) shows that the main differences are found in the loop connecting helices H3 and H4 and in the relative position of helices H1, H2 and H4 with respect to helices H3 and H5.¹³ This rearrangement implies significant changes in the composition of the hydrophobic core and in the molecular surface of AS-48¹⁷ (see Figure 2(b)). Remarkably, Val18 and Ile59, which are part of the hydrophobic core in the NMR model, are at the molecular surface in the X-ray model. On the other hand, Leu16 and Ile28 become solvent inaccessible in the X-ray model. The fact that Val18 is hidden in the interface between monomers in the DF-I, suggests that the

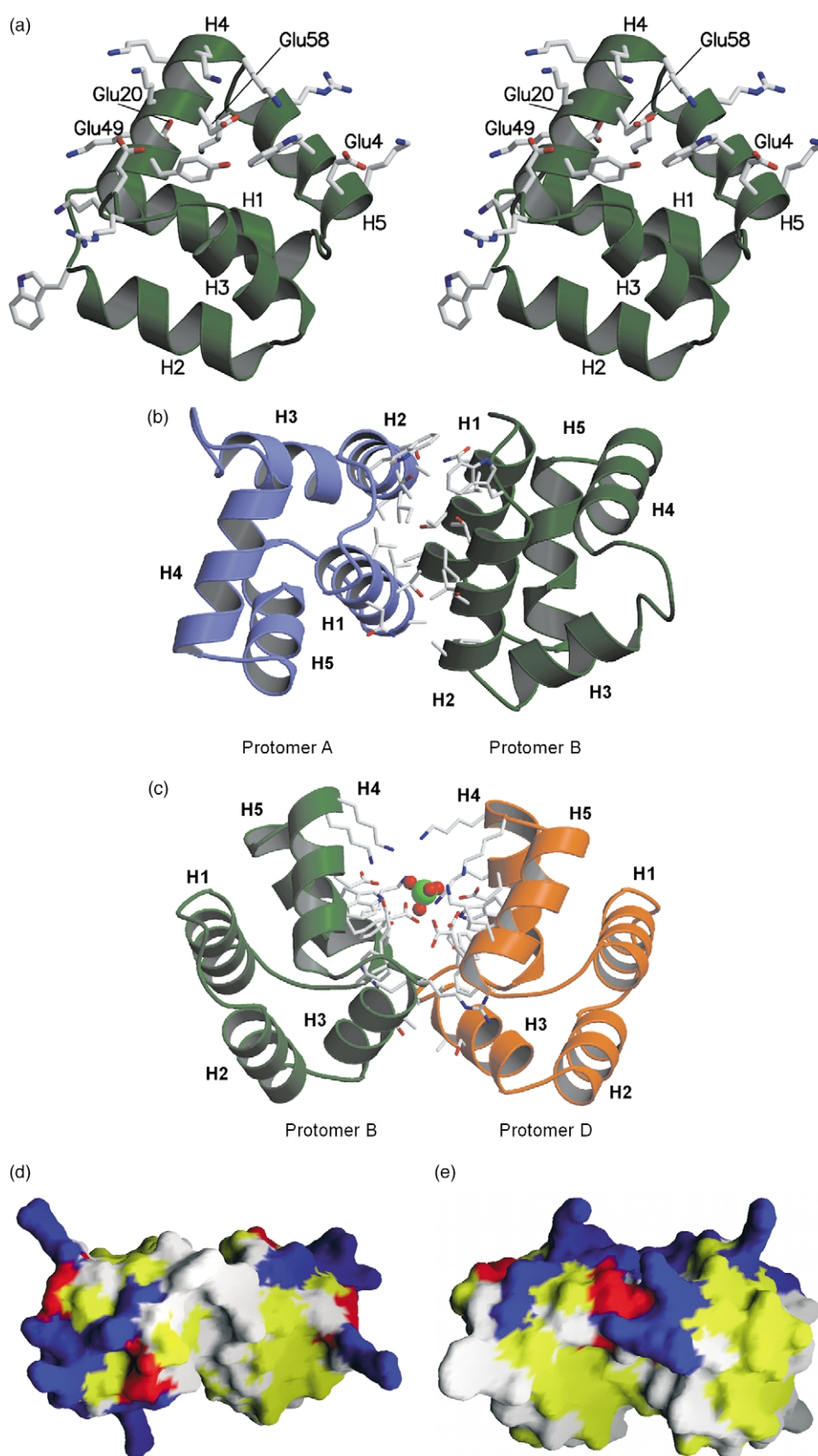


Figure 1. (a) Stereo view of the ribbon structure representation of AS-48 structure. Charged, tryptophan and tyrosine side-chains are displayed in stick mode. α -Helices and Glu side-chains are labeled. Ribbon representations of AS-48 (b) DF-I structure and (c) DF-II structure. Amino acid side-chains involved in the formation of both dimers are displayed in stick mode. A phosphate ion is displayed in ball and stick mode. α -Helices are labeled. Molecular surface representations of (d) DF-I and (e) DF-II. Hydrophobic, negatively charged and positively charged residues are depicted in yellow, in red and in blue, respectively.

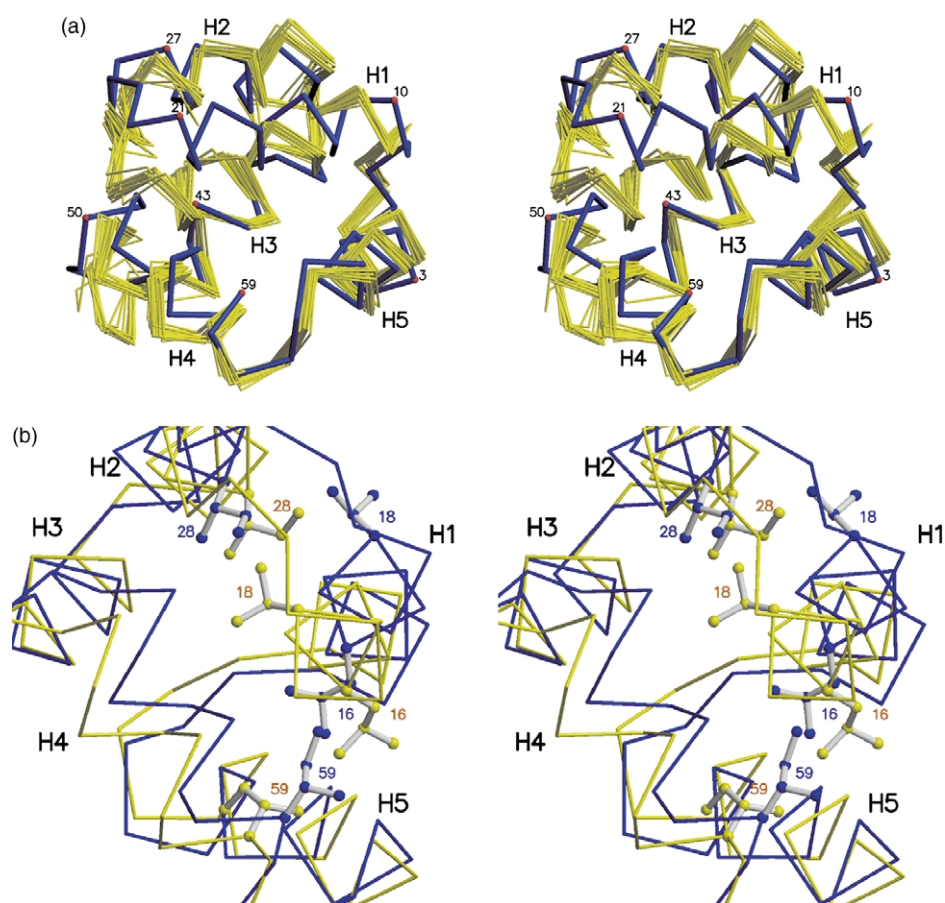


Figure 2. (a) Stereo view of the superposition of the 20 best NMR models and the X-ray structure of bacteriocin AS-48, and (b) a detail showing the main differences between them. Only a representative NMR model is displayed. The key amino acid side-chains are displayed in stick mode and are labeled. The NMR models are displayed in yellow, the X-ray structure is displayed in blue. α -Helices are labeled.

differences between X-ray and NMR model could be a consequence of the crystal packing. However, as mentioned above, it is worth noting that AS-48 displays the same conformation in crystal forms I and II despite it displays completely different crystal environment. Hence, one would expect that AS-48 would be able to adapt its molecular structure to some extent according to the physico-chemical environment. This result supports the unfolding thermodynamic studies of AS-48 showing that AS-48 should display some molecular flexibility, since its extreme stability is mainly due to the entropic constraints induced by the circular organization of the polypeptide chain.¹⁸

Oligomeric state of AS-48 in solution

The association degree of AS-48 in solution at different pH values was tested by sedimentation equilibrium experiments (see Materials and Methods). The results show that AS-48 is dimeric from pH 4.5 to 8.5 (average molecular mass of 15.5 kDa, 14.29 kDa and 15.10 kDa for pH 4.5, 7.5 and 8.5, respectively; the molecular mass of the monomeric AS-48 is 7.2 kDa). Our data show that the dimerization constant for these pH values is at

least 10^{-8} M^{-1} , hence it is likely the dimeric structure is stable at physiological concentrations of the protein. In previous studies, Gonzalez *et al.*¹² showed that AS-48 is monomeric at pH 3. This indicates the existence of a pH dependent dimerization equilibrium of AS-48. Taking into account the amino acid composition of AS-48, the protonation of the four glutamic acid residues at low pH increases the total charge of AS-48 from +6 to +10. Hence, it may well be that this protonation is responsible for the stabilization of the monomeric form of AS-48 observed in solution at pH 3.

Structure of the biological unit of AS-48 in solution

As mentioned above, AS-48 forms dimers in solution at pH in which AS-48 crystals were grown. There is a 10% decrease of the solvent accessible surface area (ASA) per protomer on the formation of DF-I.¹⁷ This value is in agreement with those calculated for molecules of similar molecular mass that form biological homodimers.¹⁹ An 85% of this occluded area corresponds to hydrophobic residues (see Figure 1(b) and (d)). This implies that DF-I displays a higher charged

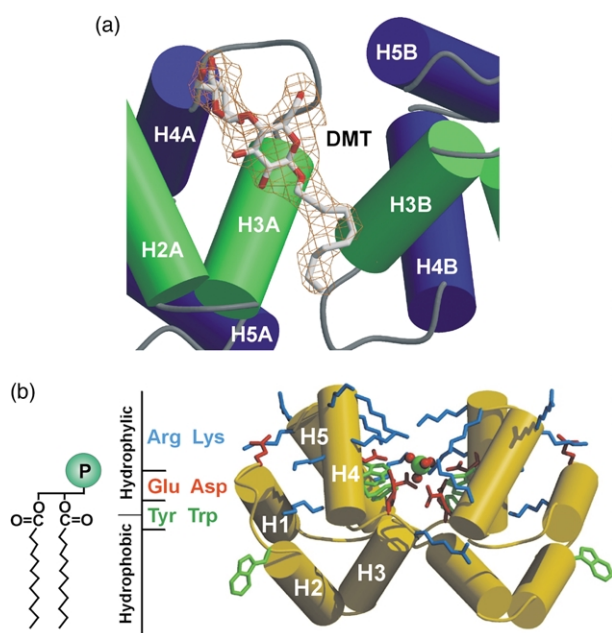


Figure 3. (a) A section of the final electron density map (contoured at 1σ) for the detergent molecule (DMT) bound to DF-II in crystal form II. Hydrophobic and polar helices are depicted in green and blue, respectively. Helices are labeled A or B according to the protomer they belong to. (b) Molecular representation of DF-II showing the layered structure of amino acid at molecular surface. Arg and Lys are depicted in blue, Glu and Asp in red, Tyr and Trp in green and the phosphate ion is displayed in ball and stick mode. A scheme of a phospholipid molecule is also depicted.

to hydrophobic solvent ASA ratio than the monomer. Hence, the solubility of the complex is enhanced with respect to the monomeric structure. These facts suggest that AS-48 adopts the structure of DF-I in solution.

Structure of the membrane bound state of AS-48

The molecular surface of membrane inserted proteins presents a hydrophobic surface exposed to the hydrophobic core of the bilayer and a polar surface in contact with the solvent. In addition, a more complex distribution of amino acid residues is found in the membrane water interface. In this area, protein surface usually displays a striped pattern of lysine and arginine residues, aspartate and glutamate residues, and tryptophan and tyrosine residues. This is a consequence of the more complex chemical nature of this area that includes charged phosphate heads and ester bonds of phospholipids. In general, transmembrane protein segments typically span a half or the total bilayer thickness.^{20–23}

DF-II displays molecular surface properties of membrane-inserted proteins. DF-II segregates the hydrophobic moiety of AS-48, which could be buried into the membrane, from the hydrophilic

moiety that could be exposed to the solvent. In this situation, a belt of positively charged amino acid residues and glutamic side-chains could interact with the phosphate polar heads of phospholipids. Remarkably, the interaction of Glu58 with $\text{H}_2\text{PO}_4^-/\text{SO}_4^{2-}$ in the crystal structure resembles the interaction with lipid head groups. In addition, the aromatic Tyr54 and Trp70 residues, with affinity for the lipid-water interface, could interact with phospholipid ester bonds carbonyl groups as they are located below the plane formed by glutamic side-chains (see Figures 1(c) and (e) and 3(b)). Finally, the crystal structure of AS-48 in the presence of detergent (crystal form II) reveals only DF-II and provides direct evidence of the interaction of the exposed hydrophobic moiety of DF-II with the long aliphatic chain of a molecule of detergent (see Figure 3(a)). Hence, it is likely that AS-48 inserts into the membrane and that it adopts the structure of DF-II. In this situation, AS-48 would reach the center of a fluid bilayer (see Figure 4).

Mechanism of membrane recognition and insertion

The crystal structure of AS-48 shows two different modes of molecular association, DF-I and DF-II. The analysis of these structures suggests a mechanism for molecular function of the toxin that involves a rearrangement of AS-48 from the water-soluble DF-I to the membrane-bound DF-II at the membrane surface (see Figure 4). This transition implies a 90° rotation of each protomer within DF-I, in a way that the partially hidden hydrophobic helices H1 and H2 become solvent accessible. This would permit AS-48 to insert into

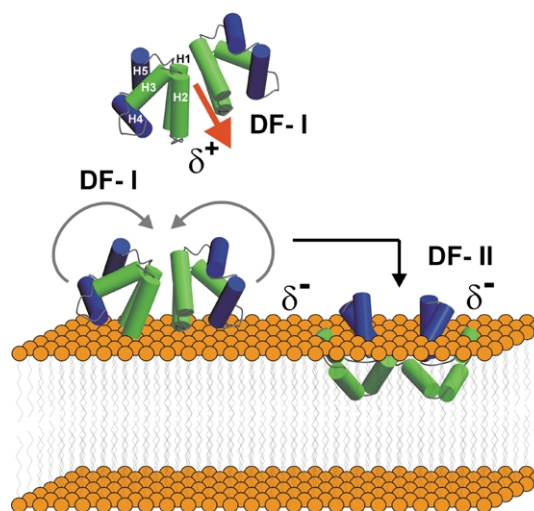


Figure 4. Schematic representation of the mechanism for molecular function of AS-48. The model includes the approach of DF-I to the membrane, the transition from DF-I to DF-II at membrane surface. Hydrophobic and polar helices are depicted in green and blue, respectively. The red arrow represents the direction of the intrinsic dipolar moment of DF-I.

the membrane. The strong dipolar moment of DF-I (84.04 Debyes)²⁴ could drive the approach of AS-48 to the membrane. The sedimentation equilibrium experiments demonstrate the existence of a pH dependent association equilibrium. This suggests that the low pH value provided by the membrane interface would destabilize the DF-I due to the protonation of the glutamic side-chains. As seen in the crystal structures, this situation enables an effective interaction of the carboxylate side-chains with the phospholipid polar heads^{25,26} and the stabilization of the hydrophobic moiety of AS-48 by the interaction with the phospholipid aliphatic chains.

The X-ray structure of AS-48 suggests a mechanism for molecular function, that, as the biochemical experiments suggest,^{4-7,10,11} it is not dependent on the membrane potential, it implies an effective insertion in the membrane and it involves a structural reorganization of a hydrosoluble dimeric form of AS-48. The mechanism resembles the first stages of membrane insertion of some other toxins.³ These toxins show a quaternary structure that shields hydrophobic residues from the solvent. AS-48 DF-I can be described as two layers of hydrophobic helices, H1, H2 and H3, sandwiched between two layers of polar helices H4 and H5 (see Figure 4). These bacterial toxins change their oligomeric state at the membrane interface making accessible the hidden hydrophobic residues.³ The transition from DF-I to DF-II resembles this stage. The insertion of AS-48 into bacterial membranes would produce the accumulation of positive charges at the membrane surface that would destabilize membrane potential and would lead to pore formation and cell leakage.

Materials and Methods

Protein preparation, crystallization and data collection

AS-48 was obtained from cultures of *E. faecalis* and purified to homogeneity as described by Gálvez *et al.*²⁷

The protein was concentrated to 20 mg/ml in buffer comprising 10 mM Tris-HCl (pH 7.5). Experiments were carried out at room temperature. Crystal form I was grown using vapor diffusion techniques from drops containing AS-48 and reservoir solutions (0.2 M ammonium sulfate, 0.1 M sodium acetate trihydrate (pH 4.5), 15% (w/v) polyethylene glycol 4000 and 0.1 M Na-Hepes (pH 7.5), 0.8 M mono-sodium dihydrogen phosphate) in a 1:1 ratio. Crystal form II was grown using vapor diffusion techniques from drops containing AS-48, 18 mM *n*-decyl- β -D-maltoside and reservoir solution (0.2 M ammonium sulfate, 25% (w/v) polyethylene glycol 4000) in a 4:1:5 ratio.

Crystals were mounted in a fiber loop and frozen at 120 K in a nitrogen stream from a solution containing reservoir solution and 20% glycerol. For derivative data collection, a crystal grown at pH 4.5 was soaked for 20 seconds in a mother liquor solution, which contained 1 M NaI and 20% glycerol; afterwards, the crystal was mounted in a fiber loop and frozen at 120 K in a nitrogen stream. X-ray diffraction data were measured using

Cu K α radiation on a Mar-Research imaging plate detector on an Enraf Nonius rotating-anode generator. A high-resolution data set for crystal form I at pH 4.5 was also collected in a CCD detector using the ESRF Grenoble synchrotron radiation source at BM14 beam-line. Diffraction data were processed using MOSFLM.²⁸

Structure determination

Attempts to solve the X-ray structure of AS-48 by molecular replacement using NMR models failed (AMoRe²⁹ and MOLREP³⁰). The AS-48 structure was solved by single isomorphous replacement using isomorphous information corresponding to one wavelength for a NaI derivative (see Table 1).³¹ The coordinates of four sites from the NaI derivative were determined using the Patterson interpretation routine of SHELX.³² All other sites were obtained by difference Fourier techniques. For this proposal and for refinement of the iodine atom sites, all calculations were performed using programs of the CCP4 package³³ and SHARP.³⁴ Solvent-flattening was performed using DM.³⁵

The experimental electron density map was good enough to identify the position of each of the four molecules in the asymmetric unit. Using this information, a molecular mask and the non-crystallographic symmetry operators were calculated. Several cycles of density modification with molecular averaging yield to a high quality electron density map³⁵ and unambiguously allowed the tracing of four polypeptide chains of AS-48 using the program O.³⁶

The initial model was refined first using the simulated annealing routine of CNS.³⁷ Several cycles of restrained refinement REFMAC5³⁸ and iterative model building with O³⁶ were carried out. Water structure was obtained using ARP.³⁹ The refined model at pH 4.5 was used to solve AS-48 structure at pH 7.5. Crystal form II was solved by molecular replacement using AMoRe (correlation coefficient and *R* factor of 67.3/38.3, respectively).²⁹ Several cycles of restrained refinement of the initial model³⁸ followed by iterative model building with O³⁶ were carried out. The stereochemistry of the models was verified with PROCHECK.⁴⁰ Ribbon Figures were produced using MOLSCRIPT⁴¹ and RASTER.⁴² The accessible surface area of the AS-48 dimer and protomer was calculated with the program "naccess" from LIGPLOT package.¹⁷ AS-48 molecular surfaces were calculated with GRASP.²⁴

Sedimentation equilibrium experiments

Sedimentation equilibrium experiments were performed in a Beckman Optima XL-A ultracentrifuge using a Ti50 rotor and six channel centerpieces of Epon-charcoal (optical pathlength 12 mm). Samples of AS-48 were equilibrated in 50 mM acetic acid/sodium acetate (pH 4.5), 50 mM Tris-HCl (pH 7.5) and 50 mM Tris-HCl (pH 8.4), and were centrifuged at 18,000 rpm and 20 °C. Radial scans at 280 nm were taken at 12, 14 and 16 hours. The three scans were identical (equilibrium conditions were reached). The weight-average molecular mass (M_w) of AS-48 was determined by using the program EQASSOC with the partial specific volume of AS-48 set to 0.732 at 20 °C as calculated from its amino acid composition.

Accession numbers

The coordinates and structure factor amplitudes of AS-48 of crystal form I, at pH 4.5 and pH 7.5, and crystal form II have been deposited in the PDB (1O82, 1O83 and 1O84, respectively).

Acknowledgements

We thank Professor M. Rico and Dr J. M. Mancheño for critical comments on the manuscript and Dr J. L. Saiz-Velasco for the analytical ultracentrifugation data analysis. M.J.S.-B. was supported by a F.P.U. studentship from Ministerio de Educación, Cultura y Deporte. A.A. thanks the ESRF (lines BM14 and BM30a) for the access to the synchrotron radiation source. This work was funded by grants PB1998-0565-C04-04, BIO2001-3237 and BMC2002-04011-C05-03 of the Spanish Plan Nacional (MCYT) and 08.5/0055./2000 of the Dirección General de Investigación de la Comunidad Autónoma de Madrid.

References

- La Rocca, P., Biggin, P. C., Tieleman, D. P. & Sansom, M. S. P. (1999). Simulation studies of the interaction of antimicrobial peptides and lipid bilayers. *Biochim. Biophys. Acta*, **1462**, 185–200.
- Lacy, D. B. & Stevens, R. (1998). Unraveling the structures and modes of action of bacterial toxins. *Curr. Opin. Struct. Biol.* **8**, 778–784.
- Gouaux, E. (1997). Channel-forming toxins: tales of transformation. *Curr. Opin. Struct. Biol.* **7**, 566–573.
- Gálvez, A., Maqueda, M., Valdivia, E., Quesada, A. & Montoya, E. (1986). Characterization and partial purification of a broad spectrum antibiotic AS-48 produced by *Streptococcus faecalis*. *Can. J. Microbiol.* **32**, 765–771.
- Martínez-Bueno, M., Maqueda, M., Gálvez, A., Samyn, B., van Beeumen, J., Coyette, J. & Valdivia, E. (1994). Determination of the gene sequence and the molecular structure of the enterococcal peptide antibiotic AS-48. *J. Bacteriol.* **176**, 6334–6339.
- Samyn, B., Martínez-Bueno, M., Bevreese, B., Maqueda, M., Gálvez, A., Valdivia, E. *et al.* (1994). The cyclic structure of the enterococcal peptide antibiotic AS-48. *FEBS Letters*, **352**, 87–90.
- Abriouel, H., Valdivia, E., Gálvez, A. & Maqueda, M. (2001). Influence of physico-chemical factors on the oligomerization and biological activity of bacteriocin AS-48. *Curr. Microbiol.* **42**, 89–95.
- Trabi, M. & Craik, D. J. (2002). Circular proteins—no end in sight. *Trends Biochem. Sci.* **27**, 132–138.
- Martínez-Bueno, M., Valdivia, E., Gálvez, A., Coyette, J. & Maqueda, M. (1998). Analysis of the gene cluster involved in production and immunity of the peptide antibiotic AS-48 in *Enterococcus faecalis*. *Mol. Microbiol.* **27**, 347–358.
- Gálvez, A., Maqueda, M., Martínez-Bueno, M. & Valdivia, E. (1991). Permeation of bacterial cells. Permeation of cytoplasmic and artificial membrane vesicles, and channel formation on lipid bilayers by peptide antibiotic AS-48. *J. Bacteriol.* **173**, 886–892.
- Abriouel, H., Sánchez-González, J., Maqueda, M., Gálvez, A., Valdivia, E. & Gálvez-Ruiz, M. J. (2001). Monolayer characteristics of bacteriocin AS-48, pH effect and interactions with dipalmitoyl phosphatidic acid at the air–water interface. *J. Colloid Interface Sci.* **233**, 306–312.
- González, C., Langdon, G. M., Bruix, M., Gálvez, A., Valdivia, E., Maqueda, M. & Rico, M. (2000). Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proc. Natl Acad. Sci. USA*, **97**, 11221–11226.
- Ruysschaert, J.-M., Goormaghtigh, E., Homblé, F., Andersson, M., Liepinsh, E. & Otting, G. (1998). Lipid membrane binding of NK-lysin. *FEBS Letters*, **425**, 341–344.
- Miteva, M., Andersson, M., Karshikoff, A. & Otting, G. (1999). Molecular electroporation: a unifying concept for description of membrane pore formation by antibacterial peptides, exemplified with NK-lysin. *FEBS Letters*, **462**, 155–158.
- Sutcliffe, M. J., Haneef, I., Carney, D. & Blundell, T. L. (1987). Knowledge based modeling of proteins. 1. 3-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng.* **1**, 377–384.
- Gronenberg, A. M. & Clore, M. (1995). Structures of protein complexes by multidimensional heteronuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.* **30**, 351–385.
- Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* **8**, 127–134.
- Cobos, E. S., Filimonov, V. V., Gálvez, A., Maqueda, M., Valdivia, E., Martínez, J. C. & Mateo, P. L. (2001). AS-48: a circular protein with an extremely stable globular structure. *FEBS Letters*, **505**, 379–382.
- Jones, S. & Thornton, J. M. (1996). Principles of protein–protein interaction. *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
- Killian, J. A. & von Heijne, G. (2000). How proteins adapt to a membrane–water interface. *Trends Biochem. Sci.* **25**, 429–434.
- Wiener, M. C. & White, S. H. (1992). Structure of a fluid dioleoylphosphatidyl choline bilayer determined by joint refinement of X-ray and neutron diffraction data. III. Complete structure. *Biophys. J.* **61**, 437–447.
- Monné, M., Nilsson, I., Johansson, M., Elmhed, N. & von Heijne, G. (1998). Positively and negatively charged residues have different effect on the position in the membrane of a model transmembrane helix. *J. Mol. Biol.* **284**, 1177–1183.
- Ridder, A. N. J. A., Morein, S., Stam, J. G., Kuhn, A., de Kruijff, B. & Killian, J. A. (2000). Analysis of the role of interfacial tryptophan residues in controlling the topology of membrane proteins. *Biochemistry*, **39**, 6521–6528.
- Nicholls, A., Bharradwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166.
- Han, X., Bushweller, J. H., Cafiso, D. S. & Tamm, L. K. (2001). Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nature Struct. Biol.* **8**, 715–720.
- Elder, M., Hitchcock, P., Mason, R. & Shipley, G. G.

- (1977). A refinement analysis of de crystallography of the phospholipid, 1,2-dilauroyl-DL-phosphatidyl-ethanolamine, and some remarks on lipid-lipid and lipid-protein interactions. *Proc. Roy. Soc. ser. A*, **345**, 157-170.
27. Gálvez, A., Giménez-Gallego, G., Maqueda, M. & Valdivia, E. (1989). Purification and amino acid composition of peptide antibiotic AS-48 produced by *Streptococcus (Enterococcus) faecalis* subsp. *liquefaciens* S-48. *Antimicrob. Agents Chemother.* **33**, 437-441.
28. Leslie, A. G. W. (1987). MOSFLM. In *Proceedings of the CCP4 Study Weekend* (Machin, J. R. & Papiz, M. Z., eds), pp. 39-50, SERC Daresbury Laboratory, Warrington, UK.
29. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.
30. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022-1025.
31. Dauter, Z., Dauter, M. & Rajashaukar, K. R. (2000). Novel approach to phasing protein: derivatization by short cryo-soaking with halides. *Acta Crystallog. sect. D*, **56**, 232-237.
32. Sheldrick, G. M. (1991). The SHELXS system. In *Crystallographic Computing 5* (Podjarny, P. D. & Thiery, J. D., eds), pp. 145-157, Oxford University Press, Oxford.
33. Bailey, S. (1994). The CCP4 suite, programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760-763.
34. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous methods. *Methods Enzymol.* **276**, 472-494.
35. Cowtan, K. (1994). Joint CCP4 and ESF-EACBM. *Newsletter Protein Crystallog.* **31**, 34-38.
36. Jones, T. A., Zon, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for model building in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110-119.
37. Brunger, A. T. & Adams, P. D. (1998). Crystallography and NMR System. *Acta Crystallog. sect. D*, **54**, 905-921.
38. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240-255.
39. Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. *Acta Crystallog. sect. D*, **53**, 448-455.
40. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK, a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
41. Kraylis, P. J. (1991). MOLSCRIPT, a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
42. Merritt, E. A. & Murphy, M. E. P. (1994). RASTER3D Versión 2.0, a program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869-873.
43. Anderson, K. M. & Hovmöller, S. (2000). The protein content in crystals and packing coefficients in different space groups. *Acta Crystallog. sect. D*, **56**, 789-790.

Edited by R. Huber

(Received 26 June 2003; received in revised form 22 September 2003; accepted 26 September 2003)