

Alicia Forcada Nadal

Hacia el entendimiento del control de la expresión génica del regulón cianobacteriano del nitrógeno mediado por el factor de transcripción NtcA y la proteína reguladora PipX

Dirigida por:

Dr. Vicente Rubio Zamora

Dr. José Luis Llácer Guerri



VNIVERSITAT
ID VALÈNCIA

Doctorado en Biomedicina y Biotecnología

**Hacia el entendimiento del control de la expresión génica
del regulón cianobacteriano del nitrógeno mediado por el
factor de transcripción NtcA y la proteína reguladora PipX**

Memoria presentada por

Alicia Forcada Nadal

Para optar al grado de Doctor por la Universitat de València

Directores:

Dr. José Luis Llácer Guerri

Dr. Vicente Rubio Zamora

Laboratorio de Enzimopatología Estructural

Departamento de Genómica y Proteómica

Instituto de Biomedicina de Valencia

Consejo Superior de Investigaciones Científicas

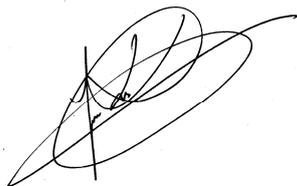
Valencia, a 31 de mayo de 2018

José Luis Llácer Guerri, Científico Titular del Consejo Superior de Investigaciones Científicas con expectativa de toma de posesión en el Instituto de Biomedicina de Valencia, y Vicente Rubio Zamora, Profesor de Investigación de esta misma institución y en el mismo Instituto y Profesor Titular de Universidad de Bioquímica en excedencia con seis sexenios de actividad investigadora reconocida (fecha del último concedido, 2007-2012)

CERTIFICAN

1. Que D^a. Alicia Forcada Nadal, Licenciada tanto en Biología como en Bioquímica por la Universidad de Valencia, ha realizado bajo la dirección de ambos su trabajo de Tesis Doctoral que lleva por título *“Hacia el entendimiento del control de la expresión génica del regulón cianobacteriano del nitrógeno mediado por el factor de transcripción NtcA y la proteína reguladora PipX”*.
2. Que a juicio de ambos este trabajo de tesis reúne los requisitos de volumen de datos, originalidad y avance en el conocimiento adecuados para una Tesis Doctoral, y que el ámbito temático de la tesis se inscribe en el área de conocimiento de la Bioquímica y Biología Molecular, razón por la que se somete a este Departamento.
3. Que la memoria de Tesis aportada por la Dra. Forcada Nadal cumple suficientemente con los requisitos expositivos, de detalle técnico y formales para ser sometida al Tribunal correspondiente con la finalidad de que la doctoranda opte al grado de Doctor.
4. Que además avala el proceso formativo de la doctoranda necesario para hacerla acreedora al grado doctoral el publicismo de la misma (ver CV de la doctoranda), siendo autora en siete publicaciones científicas internacionales de amplia difusión y razonable impacto, de las que es primera o co-primera autora en tres de las seis publicaciones realizadas en el seno de este grupo, estando las temáticas y/o las tecnologías de estas seis publicaciones del grupo directamente o indirectamente relacionadas con los contenidos esta Tesis.

Valencia, a 31 de Mayo de 2018



Fdo: José Luis Llácer Guerri



Fdo: Vicente Rubio Zamora

Este trabajo de Tesis Doctoral se ha llevado a cabo en el grupo de Enzimopatología Estructural del Instituto de Biomedicina de Valencia (IBV-CSIC) que es también grupo 739 del CIBER de Enfermedades Raras (CIBERER-Instituto de Salud Carlos III) bajo los siguientes proyectos financiados de investigación:

- BFU2008-05021 del Ministerio de Economía y Competitividad del Gobierno de España, titulado "Complejos macromoleculares, pluriempleo e implicaciones en enfermedades raras de los miembros de la familia aminoácido quinasa" (IP, V. Rubio)

- BFU2011-30407 del Ministerio de Economía y Competitividad del gobierno de España, titulado "Luz estructural sobre señalización y regulación por nitrógeno y sobre biosíntesis de arginina/urea, sus errores congénitos, y su conexión con biología del envejecimiento" (IP, V. Rubio)

- BFU2014-58229-P del Ministerio de Economía y Competitividad del gobierno de España, titulado "Una mirada molecular al control de la detoxificación de amonio y a sus patologías y errores congénitos y a la señalización por nitrógeno. En busca del papel de la proteína CutA" (IP, V. Rubio)

- PrometeoII/2014/029 de la Conselleria d'Educació de la Generalitat Valenciana, titulado "BioMeder. Genes, proteínas y rutas de señalización en enfermedades raras" (Investigadores, P. Sanz, A. Marina y V. Rubio)

- Beca de Formación de Personal Investigador (FPI) (BES-2009-028001) del Ministerio de Educación y Ciencia.

AGRADECIMIENTOS

Durante el desarrollo de esta tesis doctoral he tenido la suerte de conocer a gente maravillosa y la que no olvidaré, algunos de ellos han sido muy importante en mi vida, otros se han cruzado en ella, pero todos me han aportado mucho.

Para empezar, me gustaría agradecer a mis directores de tesis, a Vicente Rubio por darme la oportunidad de realizar una tesis doctoral en su grupo, por el tiempo que me ha dedicado y por todo lo que he podido aprender de él, y a José Luis Llácer por haber sido mi “maestro” y haberme enseñado tanto en mis primeros años en el laboratorio.

Un agradecimiento muy importante es para todos los compañeros de laboratorio con los que he tenido la suerte de coincidir, tanto los que aún siguen como los que pasaron por aquí. A Lorena, Clara, Nadine y Belén por ser tan buenas compañeras y amigas y estar siempre más que dispuestas a ayudar en lo que haga falta, gracias de verdad. A los que ya se fueron, en especial a Carles que fue como mi “hermano” en el laboratorio, pero también por supuesto a Fer (el gran dragón), Juanma, Sergio, Enea, Mariano, Carmen, Leo y Laurita. Todos me habéis ayudado y enseñado muchísimo.

Al profesor Karl Forchhammer por tratarme tan bien durante mi estancia y motivarme tanto en mi trabajo, y a toda la gente de su laboratorio por hacerme sentir tan a gusto con ellos (y también por esa cerveza que hacían ellos mismos y nos bebíamos los viernes!). También a la Dra. Asunción Contreras y a su laboratorio por acogerme durante un tiempo y estar tan dispuestos a ayudar.

A mis queridos amigos de la CdF, Carol, Carles, Carla, Lorena, Raquel, Jordi y Edu, por ser mis amigos dentro y fuera del laboratorio, por los muchísimos buenos momentos que hemos pasado juntos, por las juntas, las paellas en moncloa, el IBV newsletter, HoL, las fiestas en XL, las batallas de dj's, el arenal y los anillos del poder, todas las risas y tantísimos buenos recuerdos. También

quiero agradecer a las nuevas generaciones de CdF que he ido formando por escuchar mis consejos como experta en el tema y mantener el espíritu festivo y las ganas de organizar cosas. Estoy orgullosa de vosotros.

Un agradecimiento especial a toda la gente del IBV, a todos lo que he visto pasar por aquí y a todos los que han ido viniendo porque me han ayudado en todo lo que he podido necesitar y, sobre todo, porque que han hecho que aquí me sintiera siempre como en casa.

A mis amigos, a todos ellos, porque son muy importantes para mí. A los de toda la vida, en especial a Ximo, Glor, Carol, Evita, Cris, Miguel, Vicen, Borjita, Jorge, Cano y todos los demás, por apoyarme aunque os rierais de mis “partículas” y “bacterias”. A los de la universidad, Maxon, Mire, Angels, Indra, David, Ángela, Mariano y Rubén, por animarme y confiar en mí. A mis amigos de Erasmus, que son como una familia para mí, en especial a Palo, Marta, Yuse, Merce y Ana. Y por supuesto a mis amigos del swing, Anuki, Keko, Raymon, Carlos, Javi, Quique, Anna y tantos más, porque hacéis que mi día a día sea muy feliz y no hay nada más importante que eso.

Por ultimo quiero agradecer a mi familia, a mis padres y a mi hermano Javi por todo lo que me han dado y apoyado durante toda mi vida y por creer siempre en mí.

Esta tesis se la dedico a toda esa gente que se ha cruzado en mi camino durante estos años y me ha ayudado ser una mejor versión de mi misma.

ABREVIATURAS

2OG: 2-oxoglutarato, α -cetoglutarato

α -CTD: Dominio C-terminal de la subunidad alfa de la RNA polimerasa

α -NTD: Dominio N-terminal de la subunidad alfa de la RNA polimerasa

ADP: Adenosín difosfato

ATP: Adenosín trifosfato

BSA: Albúmina sérica bovina

cAMP: Adenosín monofosfato cíclico

CRP: Proteína receptora de cAMP (factor de transcripción bacteriano)

DBD: Dominio de unión a DNA (*DNA Binding Domain*) de los factores de transcripción de la familia CRP

DMSO: dimetilsulfóxido

DTT: Ditiotreitól

EBD: Dominio de unión del efector (*Effector Binding Domain*) de los factores de transcripción de la familia CRP

EDTA: Ácido etilendiaminotetraacético

EMSA: Ensayo de retraso de movilidad en gel

FC: Cámara de flujo (*Flow Chamber*)

FPLC: Una modalidad de cromatografía líquida para proteínas, de alta velocidad

GOGAT: Glutamato sintasa, glutamina:2-Oxoglutarato aminotransferasa

GS: Glutamina sintetasa

HTH: Motivo hélice-giro-hélice (*Helix-Turn-Helix*) de unión a DNA

IPTG: Isopropil- β -D-1-tiogalactopiranosido

MPD: 2-metil-2,4-pentanodiol

NAGK, acetilglutamato quinasa: quinasa de N-acetil-L-glutamato

NCS: Simetría no cristalográfica

NOE: Efecto nuclear Overhauser

PCR: Reacción en cadena de la polimerasa (*Polymerase Chain Reaction*)

PDB: Banco de datos de proteínas (*Protein Data Bank*). También utilizado en referencia al código de acceso en la misma base de datos

PEG: Polietilén glicol

PMSF: Fluoruro de fenilmetilosulfonilo

PLP: Piridoxal fosfato

RMSD: Raíz cuadrada de la media de los cuadrados de las desviaciones (*Root Mean Square Deviation*)

RMN: Resonancia magnética nuclear

RNAP: RNA polimerasa

SDS: Dodecilsulfato sódico

SDS-PAGE: Electroforesis en gel de poliacrilamida en condiciones desnaturalizantes

SPR: Resonancia en superficie de plasmón (*Surface Plasmon Resonance*)

SUMO: *Small Ubiquitin-like Modifier*

TBE: Tampón Tris-borato-EDTA

TBS: *Tris Buffered Saline*

TLD: Dominio de tipo Tudor (*Tudor-Like Domain*)

TLS: *Translation, Libration, Screw*

ÍNDICE

INTRODUCCIÓN	1
Las cianobacterias.....	3
Asimilación de nitrógeno en cianobacterias	5
Proteínas PII y NtcA, reguladas por 2OG en cianobacterias	10
El regulador transcripcional NtcA y su regulón.....	14
NtcA y la familia CRP de reguladores transcripcionales.....	17
A) CRP.....	18
B) Promotores dependientes de CRP.....	22
C) Estructuras con DNA de miembros de la familia CRP	23
D) Estructuras sin ligando de miembros de la familia CRP	25
La proteína PipX y el sistema PII-PipX-NtcA.....	27
OBJETIVOS	37
MATERIALES Y MÉTODOS	41
Producción y purificación de NtcA y PipX de <i>S. elongatus</i> para estudios estructurales..	43
Preparación de cristales de proteína.....	45
A) El procedimiento cristalográfico.....	45
B) Preparación de los diferentes cristales.....	46
C) Detección de PipX y DNA en los cristales.....	49
Recogida de datos y determinación de estructuras.....	50
A) Metodología general.....	50
B) Aplicación a las estructuras determinadas aquí	51
Mutagénesis dirigida de NtcA.....	56
Ensayos de retardo en gel (EMSA).....	57
Producción de NtcA, PipX y CRP de <i>Synechocystis</i>	58

Ensayos de resonancia de plasmón (SPR).....	62
Inmunodetección de proteínas.....	65
RESULTADOS	69
<u>Capítulo 1:</u>	
<u>El regulador transcripcional NtcA y sus estructuras cristalinas</u>	69
El complejo NtcA-2OG-DNA.....	71
A) Estructura del c complejo	71
B) Reconocimiento de la caja NtcA.....	79
Estructuras del complejo PipX-NtcA-DNA.....	83
Nuevas estructuras de NtcA sin 2OG unido.....	94
<u>Capítulo 2:</u>	
<u>Análisis de SPR de la unión a promotores de NtcA y CRP</u>	99
El contexto: promotores de <i>Synechocystis</i> para NtcA y CRP.....	101
Unión de NtcA y/o CRP a los promotores de <i>glnA</i> , <i>glnN</i> y <i>cccS</i>	104
Influencia del promotor en la unión de 2OG a NtcA.....	109
Influencia de PipX sobre el sistema NtcA/2OG.....	111
A modo de sumario.....	114
<u>Capítulo 3:</u>	
<u>Estructura de PipX en solución, e inmunodeterminación en</u>	
<u><i>S. elongatus</i> de proteínas de la red reguladora de PipX</u>	115
Estructura de PipX en solución	117
A) El contexto.	117
B) Mi contribución.....	118
C) Los resultados del trabajo de RMN.....	120
Inmuno-determinación en <i>S. elongatus</i> de proteínas de la red	
reguladora de PipX.....	125

A) PII, PipX, PlmA.....	125
B) El operón <i>pipX-pipY</i>	128
DISCUSIÓN	131
NtcA y la familia de CRP de reguladores transcripcionales.....	133
A) Especificidad de reconocimiento.....	133
B) Comparación estructural de las formas "activa" e "inactiva" de NtcA. Activación por 2OG.....	141
C) Transmisión de la señal.....	143
D) Activación de la transcripción por otros miembros de la familia de CRP.....	146
El complejo PipX-NtcA-DNA.....	150
Análisis funcional de las interacciones en el complejo de activación: estudios de SPR con NtcA y CRP de <i>Synechocystis</i>	152
La red de regulación de PII-PipX, y las relaciones con nuevos elementos de dicha red.....	158
CONCLUSIONES	167
BIBLIOGRAFÍA	173

Introducción

Las cianobacterias

Las cianobacterias, las cuales se encuentran entre los organismos más antiguos de la tierra (se estima que aparecieron hace 3000 millones de años) han tenido un papel crucial en la evolución de la vida en nuestro planeta al generar tanto carbono orgánico como una atmósfera oxidante (Beck *et al.*, 2012; Steinhauser *et al.*, 2012; Hamilton *et al.*, 2016). Estos organismos forman un grupo monofilético de bacterias gram negativas capaces de realizar la fotosíntesis oxigénica, generando electrones que son utilizados para fijar dióxido de carbono en materia orgánica mediante el ciclo de Calvin (Tandeau de Marsac & Houmard, 1993). Además, algunas cianobacterias son capaces también de crecer en condiciones heterotróficas en presencia de otras fuentes de carbono. Estos organismos presentan una enorme diversidad en morfología, fisiología, habilidades metabólicas y propiedades moleculares, lo que son factores clave para su éxito ecológico (Beck *et al.*, 2012). Como resultado, pueden encontrarse en prácticamente en todos los ambientes en que esté presente el agua.

Hoy en día es generalmente aceptado que los eucariotas primitivos incorporaron bacterias ancestrales dando origen a mitocondrias y cloroplastos, en un proceso llamado endosimbiosis, adquiriendo así nuevas potencialidades y funciones (Gray *et al.*, 2001; Hohmann-Marriott & Blankenship, 2011). El endosimbionte de los eucariotas fotoautótrofos que les dio la capacidad de la fotosíntesis oxigénica fue sin lugar a dudas una cianobacteria, ancestro del plástido (Hohmann-Marriott & Blankenship, 2011). Así, muchos genes de plantas se originaron a partir del endosimbionte cianobacteriano, incluyendo aquellos que codifican para proteínas implicadas en fotosíntesis, respiración y otras funciones

metabólicas y reguladoras (Martin *et al.*, 2002). De hecho, las cianobacterias han sido usadas durante décadas como modelo biológico sencillo para estudiar eucariotas fotoautótrofos. En los últimos años también han ganado mucha importancia como fuente de compuestos bioactivos y han sido consideradas uno de los grupos más prometedores para producirlos (Abed *et al.*, 2009). Estos metabolitos producidos por cianobacterias incluyen antibacterianos, antifúngicos, antivirales, antitumorales, antipalúdicos, alguicidas y agentes inmunosupresores (Abed *et al.*, 2009), por lo que el estudio de estos organismos presenta cada vez más interés.

Los resultados presentados en esta tesis han sido obtenidos trabajando con materiales derivados o relacionados principalmente con la cianobacteria: *Synechococcus elongatus* PCC 7942, también conocida como *Anacystis nidulans* R2 (*S. elongatus* a partir de ahora), aunque en una parte de esta tesis también hemos trabajado con secuencias génicas y proteínas de *Synechocystis* sp. PCC 6803 (de aquí en adelante *Synechocystis*). Ambas son cianobacterias unicelulares, de agua dulce y no fijan N₂. *S. elongatus* es un fotoautótrofo obligado mientras que *Synechocystis* puede crecer en condiciones heterotróficas en ausencia de luz si se le proporciona una fuente orgánica de carbono (Stanier *et al.*, 1979). Además, ambas cianobacterias son usadas como modelo de estos organismos al estar sus genomas totalmente secuenciados y ser fácilmente transformables con DNA exógeno.

Asimilación de nitrógeno en cianobacterias

Para hacer frente a las fluctuaciones en la disponibilidad de nutrientes los seres vivos han desarrollado una gran variedad de sistemas de transducción de señales con los que perciben el estatus nutricional predominante, lo que les permite generar una respuesta metabólica apropiada. Dos elementos clave de la materia viva son el nitrógeno y el carbono. Las cianobacterias, como la mayoría de organismos, particularmente los microorganismos, coordinan y regulan sus procesos metabólicos dependiendo de la disponibilidad de ambos elementos en formas utilizables. Ello les obliga a disponer de sensores para percibir la riqueza del medio en uno y otro elemento en dichas formas utilizables, transduciendo esta información a la regulación de las rutas metabólicas pertinentes (Commichau *et al.*, 2006). Un metabolito utilizado por las cianobacterias para coordinar la regulación del metabolismo del nitrógeno y del carbono es el 2-oxoglutarato (2OG), un compuesto intermedio del ciclo de los ácidos tricarboxílicos que es usado como esqueleto carbonado en las reacciones de asimilación de nitrógeno (Commichau *et al.*, 2006), siendo posteriormente incorporado en el glutamato, la glutamina, la ornitina y arginina, moléculas precursoras de las proteínas y, por ejemplo, de las poliaminas (Bender, 2012)

Aunque también son capaces de utilizar como fuentes de nitrógeno el nitrito, nitrato, urea, algunos aminoácidos, y en ciertas especies el dinitrógeno N_2 , el amonio es la fuente favorita de nitrógeno para la mayoría de cianobacterias (y de muchos otros microorganismos) bajo condiciones aeróbicas (Herrero *et al.*, 2001). Otras fuentes de nitrógeno inorgánico más oxidadas han de ser convertidas en amonio para su

incorporación a la sustancia orgánica, lo que conlleva el costoso proceso de reducirlas. Por ejemplo, la asimilación de nitrato requiere su reducción a amonio a través de dos reacciones secuenciales catalizadas por la nitrato reductasa y la nitrito reductasa con su correspondiente gasto de poder reductor. Así, en general en condiciones de abundancia de amonio la asimilación de nitrógeno a partir de fuentes alternativas a éste se encuentra reprimida (Leigh & Dodsworth, 2007).

Para la asimilación de amonio, muchos procariotas, incluidas las cianobacterias utilizan el ciclo de la glutamina sintetasa/glutamato sintasa (o, abreviado, GS/GOGAT). La glutamina sintetasa incorpora amonio en el glutamato, dando glutamina, desde la cual se transfiere el grupo amino a una molécula de 2OG por acción de la glutamato sintasa, generando así dos moléculas de glutamato (Fig. 1). De este modo, los niveles de 2OG fluctúan de acuerdo con la disponibilidad de amonio (Muro-Pastor *et al.*, 2001; Laurent *et al.*, 2005), mostrando a su vez una correlación inversa entre dicha disponibilidad de amonio y la acumulación de 2OG. Por esta razón, este metabolito es muy adecuado para actuar como señal metabólica de fluctuaciones en la disponibilidad de nitrógeno, lo que ha permitido la evolución de proteínas sensoras de 2OG que coordinan el metabolismo del nitrógeno y del carbono en muchos procariotas (Huerigo & Dixon, 2015).

Sólo algunas cianobacterias son diazótrofes, es decir, son capaces de fijar nitrógeno atmosférico. Esta fijación la lleva a cabo el enzima nitrogenasa, el cual es extremadamente sensible a la presencia de oxígeno por lo que estas cianobacterias han de separar o bien espacialmente o bien

temporalmente los procesos de fotosíntesis oxigénica y fijación de nitrógeno.

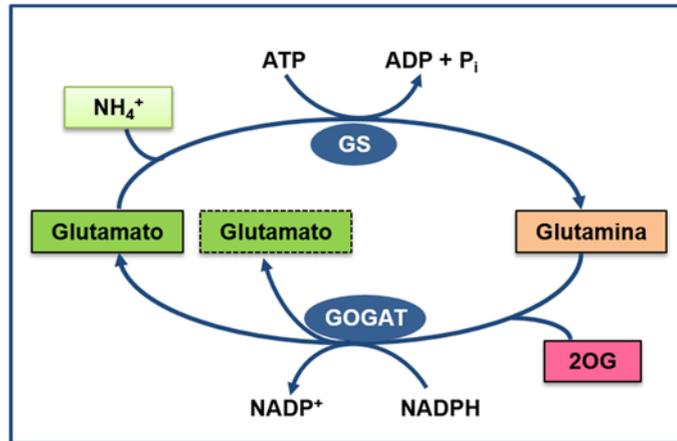


Figura 1. El ciclo GS/GOGAT para la incorporación de amonio. Adaptado de (van Heeswijk et al., 2013)

Cianobacterias como las del género *Anabaena* o *Nostoc*, han desarrollado unas células diferenciadas especializadas en la fijación de nitrógeno molecular llamadas heterocistos (Herrero *et al.*, 2001) donde la nitrogenasa queda confinada en un ambiente libre de oxígeno y alcanza por tanto su máxima actividad. Otras cianobacterias no poseen heterocistos pero son capaces de fijar nitrógeno gracias a una separación temporal de la fotosíntesis y la fijación de nitrógeno de manera que la fotosíntesis se realiza durante las horas de luz y la fijación de dinitrógeno solamente por la noche (Mitsui *et al.*, 1986).

Durante muchos años se ha considerado que las cianobacterias poseían un ciclo de los ácidos tricarbónicos (o ciclo de Krebs) incompleto debido a la falta del enzima 2OG deshidrogenasa (Stanier & Cohen-Bazire, 1977; Steinhäuser *et al.*, 2012), de manera que no podrían convertir el 2OG en

succinil-CoA y todo el 2OG generado sería utilizado como esqueleto carbonado para la asimilación de amonio, convirtiéndose así el 2OG en el indicador del balance N/C en la célula. Sin embargo, recientemente se han descubierto en distintas cianobacterias rutas alternativas que completan el ciclo de Krebs (Fig. 2). Así, este ciclo puede completarse por la acción de dos enzimas alternativas, la 2OG descarboxilasa y la succínico semialdehído deshidrogenasa (Zhang & Bryant, 2011), las cuales convierten el 2OG en succinato y cuyos genes se encuentran presentes en todos los genomas cianobacterianos con la excepción de *Prochlorococcus* y especies marinas de *Synechococcus* (Zhang & Bryant, 2011). Además también puede cerrarse el ciclo por las vías del GABA (Xiong *et al.*, 2014) o del glioxilato (Xiong *et al.*, 2014; Zhang & Bryant, 2015), aunque estas se encuentran en una fracción mucho menor de cianobacterias (que incluyen en el caso de la vía del GABA a *Prochlorococcus* y especies marinas de *Synechococcus*). De esta forma las cianobacterias si presentarían un ciclo de los ácidos tricarbónicos completo aunque no canónico. Sin embargo, en *S.elongatus* (y otras cianobacterias) un estudio muy reciente (Broddrick *et al.*, 2016) apunta a una vía lineal y no cíclica de este ciclo de forma que el 2OG representa el último paso de la vía del ciclo de Krebs.

De esta forma se habrían descrito hasta el momento cuatro variantes del ciclo de los ácidos tricarbónicos en cianobacterias: la clásica y la abierta, la complementación por 2OG descarboxilasa y succínico semialdehído deshidrogenasa (Zhang & Bryant, 2011), y la derivación a las vías de GABA (Xiong *et al.*, 2014) o del glioxilato (Xiong *et al.*, 2014; Zhang & Bryant, 2015) (Fig. 2). De cualquier manera, todos estos trabajos también

indican que las cianobacterias producen energía de forma fotoautófica más que por respiración (excepto durante períodos de oscuridad), por lo que la función primaria del ciclo de Krebs sería la de producir precursores metabólicos para el crecimiento, siendo el 2OG un metabolito clave en la asimilación de amonio, regulación del nitrógeno y metabolismo del carbono en cianobacterias (Zhang & Bryant, 2011).

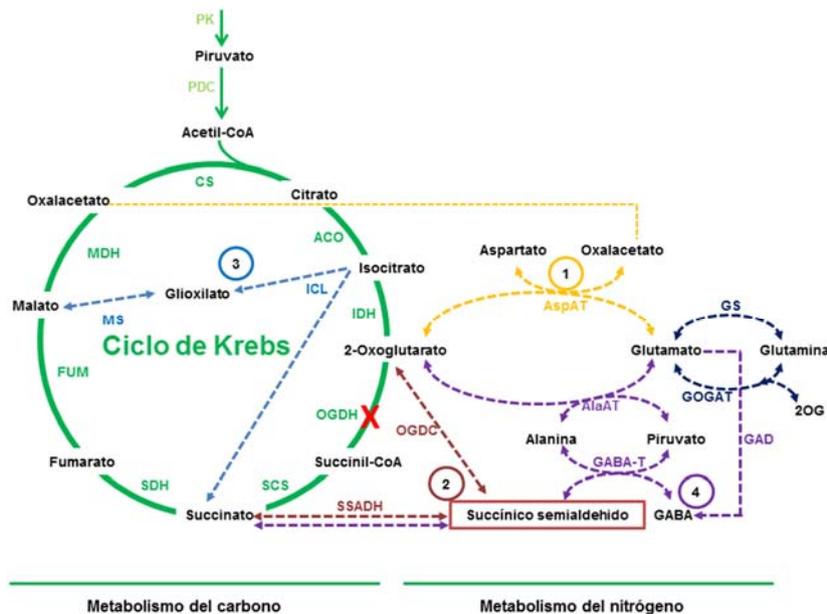


Figura 2. Esquema del ciclo de Krebs y sus variantes encontradas en cianobacterias (adaptado de (Steinhauser *et al.*, 2012). (1) Reacción final del ciclo incompleto clásico. (2) Complementación por 2OG descarboxilasa y succinico semialdehído deshidrogenasa. (3) Vía del glioxilato. (4) Vía del GABA. La cruz roja indica la falta de OGDH en cianobacterias. Abreviaturas: 2OG: 2-oxoglutarato, ACO: aconitasa, AlaAT: alanina aminotransferasa, AspAT: aspartato aminotransferasa, CL: citrato liasa, CS: citrato sintasa, FUM: fumarasa, GABA-T: GABA aminotransferasa, GAD: glutamato descarboxilasa, GOGAT: glutamato sintasa, GS: glutamina sintetasa, ICL: isocitrato liasa, IDH: isocitrato deshidrogenasa, MDH: malato deshidrogenasa, MS: malato sintasa, OGDC: 2OG descarboxilasa, OGDH: 2OG deshidrogenasa, PDC: piruvato deshidrogenasa, PK: piruvato quinasa, SCS: succinil-coA ligasa, SDH: succinato deshidrogenasa, SSADH: succinico semialdehído deshidrogenasa.

Proteínas PII y NtcA, reguladas por 2OG en cianobacterias

Entre las proteínas de cianobacterias reguladas por 2OG, compuesto del que son sensoras y transductoras, se encuentran la proteína señalizadora PII, de amplia distribución en todos los dominios de la vida excepto en animales (Sant'Anna *et al.*, 2009), y el regulador transcripcional NtcA, también conocido como el regulador global del nitrógeno en cianobacterias, organismos en los que se encuentra en forma exclusiva (Körner *et al.*, 2003; Matsui *et al.*, 2013). Dada su difusión, quizá las proteínas PII estén entre los sensores de 2OG más antiguos, habiendo sido PII además el primer sensor de 2OG reconocido, descubierto en estudios bioquímicos sobre la regulación de la glutamina sintetasa de *Escherichia coli* (Adler *et al.*, 1975).

Las proteínas PII típicas son homotrómeros de 12-13 kDa por subunidad, las cuales poseen un plegamiento tipo ferredoxina ($\beta\alpha\beta\beta\alpha\beta$) y tres lazos denominados B, T, y C (Carr *et al.*, 1996; Cheah *et al.*; Xu *et al.*, 2003). De forma general, el trímero de PII presenta un cuerpo hemisférico del que emergen los lazos T (Fig. 3A). Estos lazos son muy largos (unos 20 residuos), expuestos al solvente y vitales para la función de estas proteínas, ya que generalmente median los contactos con sus proteínas diana (Conroy *et al.*, 2007; Gruswitz *et al.*, 2007; Llácer *et al.*, 2007, 2010; Zhao *et al.*, 2010a; Chellamuthu *et al.*, 2014). Cada trímero es capaz de unir hasta tres moléculas de 2OG y de ATP (Fig. 3B), ADP (Jiang & Ninfa, 2009) y al menos en algunas especies, también AMP (Palanca *et al.*, 2014), compitiendo estos nucleótidos por el mismo sitio de unión. La unión de estos ligandos favorece distintas conformaciones de los lazos T, modulando así su habilidad para unirse a sus proteínas

diana (Conroy *et al.*, 2007; Gruswitz *et al.*, 2007; Llácer *et al.*, 2007, 2010; Fokina *et al.*, 2010; Truan *et al.*, 2010; Zhao *et al.*, 2010a; Chellamuthu *et al.*, 2014). Para que el 2OG se una a PII, es necesaria la unión de ATP, colaborativa con la de 2OG (Fokina *et al.*, 2010; Truan *et al.*, 2010), por tanto, en ausencia de ATP o cuando están unidos ADP o AMP, no hay unión de 2OG a PII. El hecho de que ATP y ADP/AMP compitan por el mismo sitio y que la unión del 2OG requiera la presencia de ATP le da a PII un gran potencial para integrar la señal de 2OG con la percepción del estado energético de la célula (Zeth *et al.*, 2014). Además, la unión del ATP y del 2OG a sus tres sitios presenta una cooperatividad negativa por lo que PII puede detectar un amplio rango de concentraciones de estos dos metabolitos (Forchhammer, 2008).

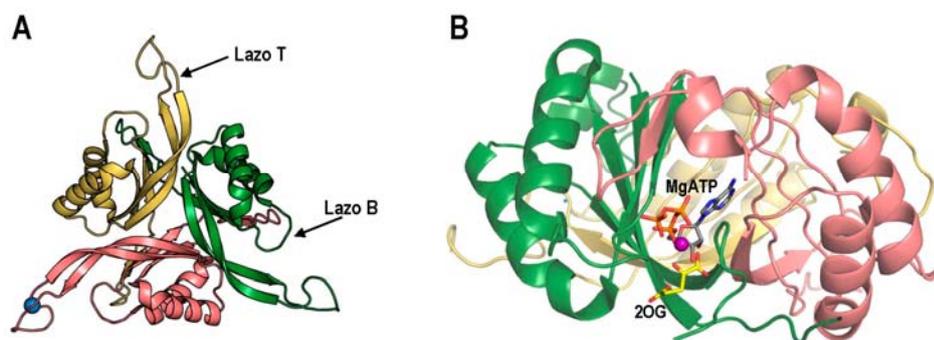


Figura 3. Representación en cintas de la estructura de la proteína PII de *S. elongatus* (A) sin ligandos (PDB 1QY7) y (B) con MgATP unido (PDB 2XUL). Cada subunidad está coloreada distintamente. En (A) se señalan los lazos T y B en una de las subunidades. La esfera azul en uno de los lazos corresponde a la serina que es fosforilada en cianobacterias. En (B) se señalan el 2OG (átomos de carbono en amarillo), el ATP (átomos de carbono en gris) y el magnesio (esfera de color morado).

Además, la estructura de la proteína PII puede estar afectada por modificaciones covalentes en residuos del lazo T en condiciones de escasez de nitrógeno (Forchhammer, 2008). Dichas modificaciones son

muy relevantes para la función de PII ya que pueden afectar a su capacidad de formar complejos con sus proteínas diana, debido a que, como se ha dicho, su interacción con éstas es principalmente a través de estos lazos T. En algunas cianobacterias como *S. elongatus* y *Synechocystis*, PII es fosforilada en la serina 49 (numeración de *S. elongatus*) (Forchhammer, 2004) (Fig. 3A) en condiciones de escasez de nitrógeno. Esta serina se modifica en condiciones de bajo nitrógeno por una quinasa aún desconocida (Forchhammer & Tandeau de Marsac, 1994) y se desfosforila en abundancia de nitrógeno por una fosfatasa cuya actividad está controlada por 2OG y MgATP (Irmmler & Forchhammer, 2001).

Las proteínas PII presentan una gran cantidad de proteínas diana que se podrían agrupar en tres grupos: enzimas con propiedades reguladoras como glutamina sintetasa o, en cianobacterias y plantas, acetilglutamato quinasa; transportadores de amonio, nitrato, nitrito y bicarbonato; y proteínas involucradas en la regulación de la expresión génica. En cianobacterias y plantas, una de las dianas clave de PII es la enzima N-acetilglutamato quinasa (NAGK), paso inicial y regulador de la síntesis de arginina, aminoácido que al contener cuatro átomos de nitrógeno en su molécula, es un buen almacén de nitrógeno como tal en proteínas ricas en arginina (el caso de las semillas) o en los péptidos ricos en arginina fabricados por las cianobacterias (denominados cianoficinas) (Llácer *et al.*, 2008). El complejo formado entre PII y NAGK activa el enzima y lo hace menos inhibible por arginina, el producto final de la ruta, posibilitando así la acumulación de nitrógeno como arginina ya que su síntesis no se reprime por inhibición *feed-back* (Llácer *et al.*, 2007, 2008).

Niveles intracelulares altos de 2OG o la fosforilación de la serina 49 en PII favorecen la disociación del complejo, de forma que la formación de éste se da solamente en condiciones de riqueza energética y abundancia de nitrógeno (Heinrich *et al.*, 2004)(Fig. 4).

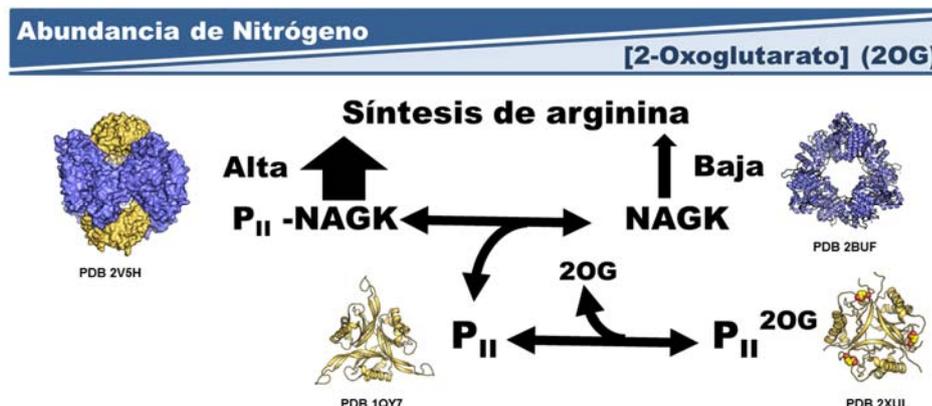


Figura 4. Esquema del sistema PII-NAGK en cianobacterias. Se muestran las estructuras resueltas de los miembros del sistema con su correspondiente código PDB.

Otra diana importante de PII en cianobacterias es la proteína PipX, a la cual se une cuando las concentraciones de 2OG son bajas (reflejando altos niveles de amonio). Tal y como se verá más adelante, esta proteína también se une y activa al regulador transcripcional NtcA cuando los niveles de amonio son bajos. Así, a través de PipX, PII controla la transcripción del regulón de NtcA (ver más adelante). Sin embargo, no es esta la única consecuencia de la formación del complejo PII-PipX. En experimentos recientes en colaboración con el grupo de la Dra. Asunción Contreras (Labella *et al.*, 2016) se ha establecido claramente la formación de un complejo ternario PII-PipX-PlmA, en el que PipX interacciona con PlmA. Como esta última proteína es, a juzgar por su secuencia, un regulador transcripcional de tipo GntR, hay que suponer que la formación

de este complejo debe tener consecuencias reguladoras sobre el aún inexplorado regulón de PlmA. Además, la observación reciente de interacciones funcionales entre PipX y el producto de su gen adyacente, PipY (*Labella et al.*, 2017) (una proteína de unión de piridoxal fosfato de función no completamente aclarada), añade complejidad a las interacciones reguladoras entre PII y PipX.

El regulador transcripcional NtcA y su regulón

NtcA pertenece a la familia CRP/ FNR de factores de transcripción, y se encuentra presente únicamente y de forma altamente conservada en cianobacterias (*Herrero et al.*, 2001), donde es esencial (*Rubin et al.*, 2015). NtcA es considerado el regulador global del nitrógeno en estos organismos, ya que controla un regulón muy amplio que incluye mayoritariamente genes implicados en el metabolismo de nitrógeno, aunque también genes relacionados con el metabolismo del carbono, fotosíntesis y respuestas al estrés (*Herrero et al.*, 2001).

La transcripción dependiente de NtcA está regulada por el nivel de 2OG, ya que este metabolito activa alostéricamente a NtcA, favoreciendo en la mayoría de los casos que se inicie la transcripción de los genes que regula (*Tanigawa et al.*, 2002; *Vázquez-Bermúdez et al.*, 2002a). Por lo tanto, es a niveles bajos de amonio (la fuente de nitrógeno preferida) cuando se activa el regulón de NtcA.

En cuanto a su estructura (*Llácer et al.*, 2010; *Zhao et al.*, 2010b), NtcA es una proteína homodimérica de 222 aminoácidos por subunidad. Cada subunidad se compone de un dominio N-terminal denominado regulador ya que es el dominio que al que se une el 2OG, y un dominio C-terminal

de unión al DNA el cual contiene un motivo hélice-giro-hélice (Fig.5). En los promotores de los genes que regula, NtcA se une a la secuencia consenso **GTAN₈TAC**, o, en forma más general, **GTN₁₀AC** situada a -41.5 nucleótidos respecto al sitio de inicio de la transcripción (Picossi *et al.*, 2014), siendo muy similar la forma en la que NtcA activa la transcripción a la regulación mediada por CRP en los promotores de clase II (discutidos más adelante). Además, NtcA también puede actuar como represor transcripcional de algunos genes al unirse a secuencias que solapan con el sitio de unión de la RNA polimerasa (Mitschke *et al.*, 2011).

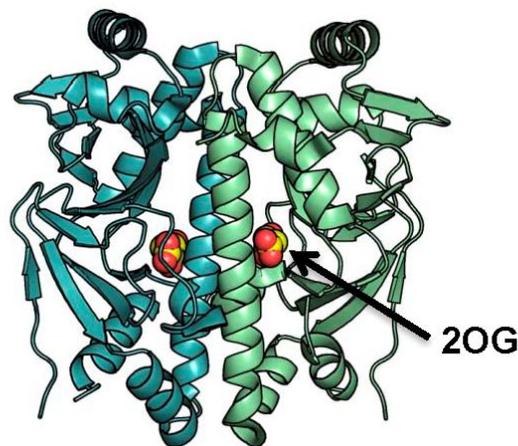


Figura 5. Estructura del factor de transcripción NtcA de *S. elongatus* (PDB 2XHK). NtcA se representa en cintas con cada subunidad coloreada distintamente. El 2OG unido se representa en esferas. (Llácer *et al.*, 2010)

En respuesta a la falta de nitrógeno en la célula, NtcA activa la expresión de los genes implicados en la utilización de trazas de nitrógeno, activando la expresión de traslocadores de amonio (genes *amt*), de nitrato y nitrito (*nrtABCD*) o urea (*urtABCD*), así como genes implicados en la

asimilación de amonio como *glnA* y *glnN*, que codifican dos isoformas de la glutamina sintetasa (Luque *et al.*, 1994; Reyes *et al.*, 1997; Vázquez-Bermúdez *et al.*, 2002b); y genes implicados en la asimilación de nitrato, como el gen *nirA* (Su *et al.*, 2005), que codifica para la nitrito reductasa. En cianobacterias que forman heterocistos, NtcA es necesario también para la diferenciación y función de estas células especializadas, en respuesta a una carencia prolongada de nitrógeno (Picossi *et al.*, 2014). NtcA también regula genes que codifican proteínas reguladoras, como la proteína PII (*glnB*) (García-Domínguez & Florencio, 1997), el factor regulador PipX (*pipX*) (Valladares *et al.*, 2011), e incluso *ntcA* ya que este factor de transcripción activa también su propia expresión, expresándose a niveles basales en presencia de amonio y aumentando su transcripción en ausencia de éste (Luque *et al.*, 1994).

Sin embargo, no todos los genes regulados por NtcA están implicados en el metabolismo del nitrógeno. Se han encontrado genes que codifican proteínas implicadas en transcripción y traducción, incluyendo varios genes que codifican proteínas del ribosoma (Picossi *et al.*, 2014), así como genes para proteínas implicadas en la asimilación de carbono, metabolismo de ácidos grasos y fosfolípidos, fotosíntesis y también genes implicados en la biosíntesis de cofactores y grupos prostéticos (Picossi *et al.*, 2014).

Por otra parte, NtcA es capaz de actuar como represor de ciertos genes, como los que codifican para los factores inactivantes de la glutamina sintetasa (*gifA* y *gifB*) (García-Domínguez *et al.*, 2000) o del gen *rbcL*, (Fadi Aldehni *et al.*, 2003) que codifica para la ribulosa-1,5-bifosfato carboxilasa/oxigenasa (RuBisCo).

NtcA y la familia CRP de reguladores transcripcionales

La regulación de la expresión génica en bacterias es un proceso crucial para su adaptación a ambientes cambiantes, colonización de nichos o para su patogenicidad. Los factores de transcripción son clave para mediar esta regulación. Entre todos ellos, la familia CRP-FNR (*cAMP receptor protein-fumarate/nitrate reductase regulator*), a la que pertenece NtcA, tiene una importancia crucial (Körner *et al.*, 2003) ya que una proporción muy elevada de genes está controlada sólo por miembros de esta familia (Browning & Busby, 2004). Esta familia de factores de transcripción se encuentra ampliamente distribuida en bacterias y sus cientos de miembros se encuentran agrupados en diferentes subfamilias como son: CRP, FNR, ArcR, CooA, CprK, Dnr, FixK, FnrN, Flp, DnR, MalR, NnrR, NtcA, PrfA y YeiL, las cuales controlan un número muy amplio de genes implicados en funciones vitales como son la respuesta al hambre, estrés oxidativo, fijación de nitrógeno, virulencia, fotosíntesis, halorespiración o catabolismo de arginina (Körner *et al.*, 2003). La mayoría de estos factores actúan como activadores de la transcripción, uniéndose a -61.5 (clase I) o a -41.5 (clase II) nucleótidos del lugar de inicio de la transcripción (Browning & Busby, 2004; Lawson *et al.*, 2004); aunque algunos pueden actuar también como represores y en estos casos su sitio de unión se sitúa muy cerca del inicio de la transcripción impidiendo la unión de la RNA polimerasa al promotor. Su actividad reguladora está controlada normalmente por la unión de moléculas efectoras, las cuales responden a un amplio espectro de señales ambientales como pueden ser el cAMP, 2OG, anoxia, estado redox, monóxido de carbono, temperatura, óxido nítrico, etc. (Körner *et al.*, 2003).

Los miembros de la familia CRP se caracterizan por ser diméricos y poseer un dominio N-terminal de unión a ligando, un dominio C-terminal con un motivo hélice-giro-hélice de unión a DNA y una larga hélice α que conecta los dos dominios y permite la dimerización. Entre todos los miembros de esta familia, CRP (acrónimo para *cAMP receptor protein*), también llamado CAP (*catabolite activator protein*) de *E.coli* es considerado el paradigma y ha sido estudiado en profundidad (Lawson *et al.*, 2004).

A) CRP

El factor de transcripción CRP en *E. coli*, es conocido por regular la expresión de cientos de genes, muchos de ellos requeridos para catabolizar fuentes de carbono alternativas a la glucosa e implicados en el catabolismo de carbohidratos (Kolb *et al.*, 1993). Una reducción en los niveles de glucosa resulta en un aumento de la concentración intracelular de cAMP, que al unirse a CRP provoca en éste un cambio conformacional responsable de su unión al DNA (Popovych *et al.*, 2009).

En algunas cianobacterias se pueden encontrar genes que codifican para la proteína CRP (Xu & Su, 2009), habiéndose descrito la implicación de esta proteína en un amplio espectro de funciones que engloban entre otros, fotosíntesis y fijación de carbono, asimilación de nitrógeno, transportadores y sistemas de dos componentes (Ohmori *et al.*, 1988; Suzuki *et al.*, 2007; Xu & Su, 2009). En esta tesis hemos trabajado con la proteína CRP de *Synechocystis* (llamada SYCRP1 en este organismo (Yoshimura *et al.*, 2000), la cual se ha descrito que se expresa desde su gen homónimo y regula de forma dependiente de cAMP la expresión de genes como *cccS* y *cccP*, los cuales están implicados en la construcción

de componentes de la superficie celular (Yoshimura *et al.*, 2010), habiéndose descrito también la implicación de SYCRP1 en motilidad celular (Yoshimura *et al.*, 2002a).

La primera estructura de un activador transcripcional resuelta fue la de CRP de *E. coli* en el año 1981 (McKay & Steitz, 1981). Desde entonces se han descrito diversas estructuras de CRP de *E. coli* tanto en su forma apo (Fig. 6 A y B) (Popovych *et al.*, 2009; Sharma *et al.*, 2009) como unido a su activador alostérico, el cAMP (Fig. 6C)(McKay & Steitz, 1981; Weber & Steitz, 1987; Passner *et al.*, 2000). También se han resuelto las estructuras de CRP unido a su DNA diana en presencia de cAMP (Fig. 6D) (Parkinson *et al.*, 1996; Schultz *et al.*, 1991) y en complejo con el dominio CTD de la subunidad α de la RNA polimerasa (Benoff *et al.*, 2002) (Fig. 6E). Además de las estructuras de CRP de *E.coli*, también se han resuelto las estructura de CRP de *Mycobacterium tuberculosis* tanto en su forma apo (Gallagher *et al.*, 2009) como unido a cAMP (Reddy *et al.*, 2009) y DNA (PDB 3MZH, sin publicación asociada).

En todas las estructuras, CRP es una proteína homodimérica formada por dos subunidades de 23kDa cada una. Cada monómero consiste en dos dominios: un dominio N-terminal de unión a cAMP y un dominio C-terminal de unión a DNA el cual contiene el típico motivo hélice-giro-hélice (HTH) formado por las hélices α F y α E (Fig. 6D). Tras la unión del cAMP al dominio sensor de cada subunidad, CRP pasa de un estado “inactivo” en el que presenta una baja afinidad para unir DNA, a un estado “activo” en el cual se une al DNA con alta afinidad.

La determinación de la estructura de CRP en complejo con su DNA diana aclaró cómo se une CRP específicamente a las secuencias que reconoce (Schultz *et al.*, 1991). Esta unión está mediada principalmente por la hélice F perteneciente al motivo HTH que se encuentra en el dominio C-terminal (llamado el dominio de unión a DNA, abreviado DBD) de cada subunidad

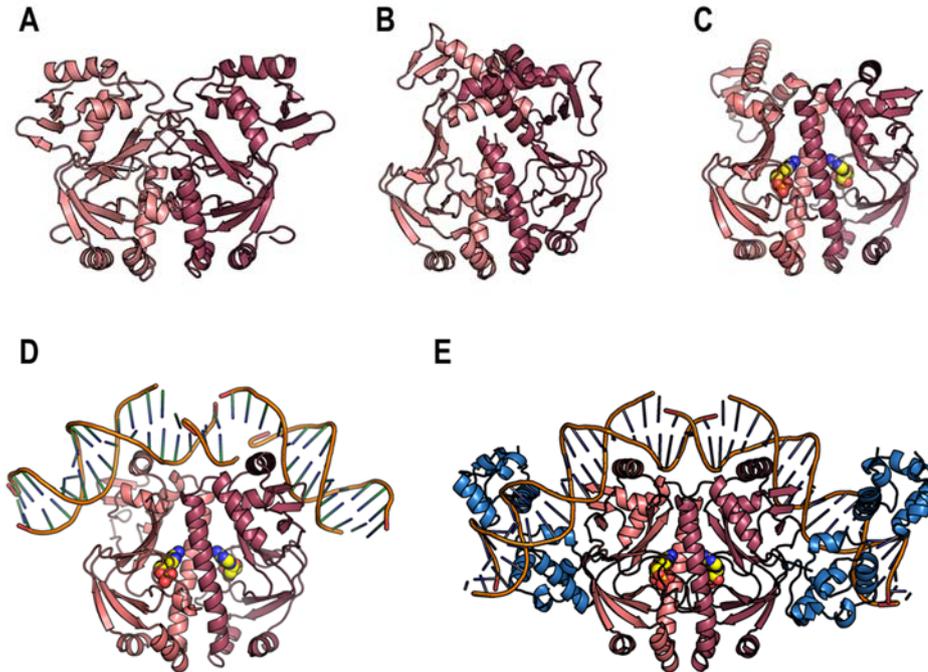


Figura 6. Estructuras de CRP de *E.coli*. En (A) (PDB 2WC2) y (B) (PDB 3HIF) se muestran formas inactivas del factor de transcripción. (C) Forma activa con cAMP unido (PDB 1G6N), (D) forma activa unida al DNA (PDB 1J59) y (E) forma activa unida al DNA y al dominio C-terminal de la subunidad α de la RNA polimerasa (α -CTD) (en color azul) (PDB 1LB2). NtcA se muestra en cintas con cada subunidad coloreada de un tono de rosa diferente y el cAMP se muestra como esferas.

del homodímero. Las hélices F de ambas subunidades de CRP interaccionan con el surco mayor en dos vueltas consecutivas de la doble hélice de DNA, donde se encuentran las secuencias génicas específicas que reconoce (la caja CRP) (Schultz *et al.*, 1991; Parkinson *et al.*, 1996;

Chen *et al.*, 2001a; Benoff *et al.*, 2002; Napoli *et al.*, 2006), (Fig. 6D). Estas cajas CRP, presentes en los promotores de los genes que regula, poseen la secuencia consenso **TGTGAN₆TCACA** (Ebright *et al.*, 1989), bastante parecida a la secuencia reconocida por NtcA, **GTAN₈TAC**. El reconocimiento de la caja CRP lo realiza este factor de transcripción a través de un mecanismo de “lectura directa”, que es el reconocimiento de la secuencia de DNA mediado por puentes de hidrógeno directos o interacciones de van der Waals con pares de bases del DNA; combinado con otro de “lectura indirecta”, que se refiere al reconocimiento de la secuencia del DNA por la detección de los efectos dependientes de secuencia en la posición de los fosfatos del DNA o la susceptibilidad a la deformación del DNA (Lawson *et al.*, 2004). La unión de CRP además se asocia con una curvatura del DNA de aproximadamente 90°, lo que parece necesario para la activación de la transcripción.

Por otra parte, la determinación de la estructura del complejo de CRP unido simultáneamente al DNA y a un dominio de la subunidad α de la RNA polimerasa (RNAP) mostró interacciones directas entre CRP y el dominio C-terminal (α CTD), y de este último con el DNA cerca de la caja CRP (Benoff *et al.*, 2002). De esta forma, se propuso que estos contactos proteína-proteína son los que reclutan a la RNAP a los promotores dependientes de CRP, contribuyendo así a la activación de la transcripción (Browning & Busby, 2004; Lawson *et al.*, 2004; Ghosh *et al.*, 2010).

B) Promotores dependientes de CRP

Los promotores dependientes de CRP más simples requieren solo la presencia de este factor para la activación de su transcripción y se agrupan en dos clases según la posición del DNA a la que se une CRP y el correspondiente mecanismo de activación de la transcripción (Lawson *et al.*, 2004).

En los promotores de clase I (Fig. 7A), CRP se une a una secuencia centrada a -61.5 nucleótidos del lugar de inicio de la transcripción e interacciona con el dominio C-terminal de la subunidad α de la RNA polimerasa (α -CTD; siendo la composición del holoenzima de una RNA polimerasa bacteriana: $\beta\beta'\alpha2\omega\sigma$). Esta interacción facilita la unión del α -CTD, y por consiguiente, de la RNA polimerasa (RNAP) al promotor, estimulando de este modo el inicio de la transcripción por un mecanismo de reclutamiento directo (Lawson *et al.*, 2004; Lee *et al.*, 2012).

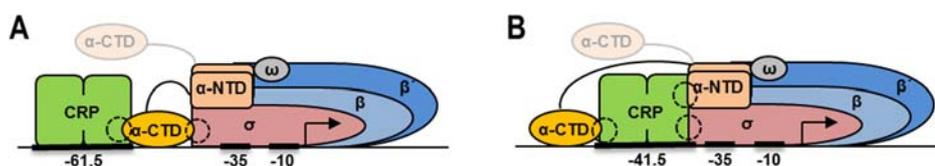


Figura 7. Esquema de la activación de la transcripción en promotores bacterianos de clase I (A) y II (B). Los contactos entre CRP y la RNA polimerasa (concretamente con las subunidades α -CTD, α -NTD y σ) se representan como círculos con borde discontinuo. El DNA se representa de forma lineal señalándose las secuencias de reconocimiento de CRP a -61.5 (A) y -41.5 (B) y los elementos -35 y -10 del promotor. El origen de transcripción se señala con una flecha. Adaptado de (Lawson *et al.*, 2004).

En los promotores de clase II (Fig. 7B), la secuencia a la que se une CRP está centrada a -41.5 nucleótidos del lugar de inicio de la transcripción, solapando parcial o totalmente el elemento -35 del promotor (Decker & Hinton, 2013). En esta posición, CRP interacciona con las subunidades α

(tanto con el dominio N-terminal (α -NTD) como con el C-terminal (α -CTD) y σ de la RNA polimerasa. De esta forma CRP activa la transcripción en este tipo de promotores por reclutamiento directo de la RNAP (interacción con α -CTD, como pasaba en los promotores de clase I) y sobre todo por un mecanismo post-reclutamiento mediado por las interacciones con α -NTD y σ , las cuales facilitan la tasa de isomerización de complejo cerrado a abierto de la RNAP (etapa en la cual se separan ambas hebras del DNA para formar la burbuja de transcripción) (Lawson *et al.*, 2004; Lee *et al.*, 2012).

C) Estructuras de miembros de la familia CRP unidos a DNA

En este laboratorio se han determinado estructuras de NtcA con 2OG unido induciendo una estructura con características apropiadas para unirse al DNA (Llácer *et al.*, 2010), pero no se había generado el complejo de NtcA-DNA. En esta tesis presentamos dicho complejo. Previamente, y pese a que la forma en la que se une CRP al DNA es considerada la forma en la cual factores de transcripción de esta familia se unen al DNA, además de la estructura de CRP de *E. coli* (ver Fig. 6) y de *M. tuberculosis* (Fig. 8A) solo se había determinado la estructura en complejo con DNA en otros cuatro miembros de la familia.

Dichos miembros son CprK de *Desulfitobacterium* spp, (Fig. 8B) (Levy *et al.*, 2008), FixK₂, de *Bradyrhizobium japonicum* (Fig. 8C), (Bonnet *et al.*, 2013), PrfA de *Listeria monocytogenes*, (Hall *et al.*, 2016) unido al DNA en presencia (Fig 8D) y en ausencia (Fig. 8E) de su ligando glutathion (GSH), y TAP de *Thermus thermophilus* (Fig. 8F), unido al holoenzima RNAP σ^A y a un promotor dependiente de TAP de clase II (Feng *et al.*, 2016). A diferencia de otros miembros de la familia, FixK₂

es capaz de activar la transcripción en ausencia de ligando (Mesa *et al.*, 2005) y de hecho su estructura unido a DNA en ausencia de ligando es muy similar a las de CRP y CrpK unidas a DNA también (Bonnet *et al.*, 2013). Además, las estructuras de PrfA unido a DNA con y sin ligando son prácticamente iguales ya que este factor de transcripción también es capaz de unirse al DNA en ausencia de ligando aunque con menor afinidad.

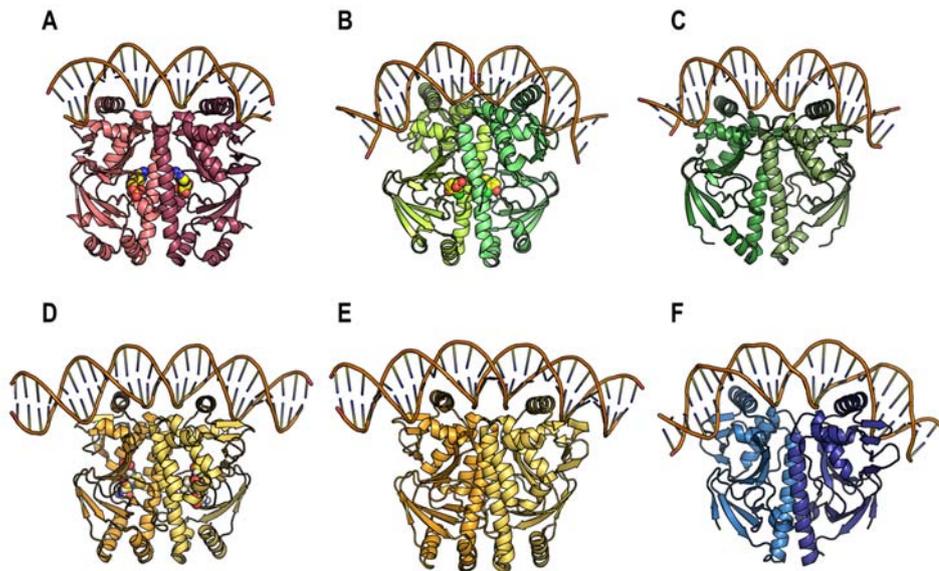


Figura 8. Miembros de la familia CRP cuyas estructuras unidas a DNA han sido resueltas (además de la estructura CRP de *E.coli* unida al DNA que se muestra en la Fig. 6). (A) CRP de *M.tuberculosis* (PDB 3MZH), (B) CprK de *Desulfitobacterium* sp (PDB 3E6C), (C) FixK2 de *B. japonicum* (PDB 4I2O), (D) PrfA de *L. monocytogenes* en presencia de GSH (PDB 5LRS) o en ausencia de (E) (PDB 5LEJ), (F) TAP de *T.thermophilus* (PDB 5I2D).

En todas estas estructuras de miembros de la familia CRP unidos a DNA, la unión de la proteína al DNA ocurre fundamentalmente de la misma forma que en CRP.

D) Estructuras sin ligando de miembros de la familia CRP

Una característica importante de los reguladores transcripcionales de esta familia es el control alostérico ejercido por moléculas efectoras. Durante muchos años no ha estado claro como estos efectores desencadenan la activación del factor de transcripción. Las estructuras de CRP de *E.coli* no unido a ligando han sido resueltas mediante RMN (Popovych *et al.*, 2009) (Fig. 6A) y difracción de rayos X (Sharma *et al.*, 2009) (Fig. 6B), siendo estas estructuras bastante diferentes entre sí. La estructura de RMN revela que las hélices F de las dos subunidades de CRP se sitúan en una posición expuesta pero inapropiada para unirse simultáneamente a dos surcos mayores consecutivos del DNA, donde se encuentran las secuencias específicas de la caja CRP, mientras que en la estructura cristalina apoCRP presenta una conformación distinta en la cual las hélices F se encuentran ocultas impidiendo su unión al DNA.

Para otros miembros de la familia CRP se han determinado también tanto las formas apo como las formas unidas a ligando. Estos miembros incluyen los ya mencionados PrfA (Eiting *et al.*, 2005; Hall *et al.*, 2016), CrpK (Joyce *et al.*, 2006; Levy *et al.*, 2008), TAP (Agari *et al.*, 2012), FixK₂ (Bonnet *et al.*, 2013) y CRP de *M.tuberculosis* (Gallagher *et al.*, 2009; Kumar *et al.*, 2010) así como SdrP (Agari *et al.*, 2008), GlxR (Townsend *et al.*, 2014), CooA (Lanzilotta *et al.*, 2000; Borjigin *et al.*, 2007), CLP (Chin *et al.*, 2010), DNR (Giardina *et al.*, 2009) y FNR (Volbeda *et al.*, 2015). En todos los casos, las formas unidas a ligando son muy parecidas entre sí y, sin embargo, las formas sin efector unido difieren bastante unas de otras.

En el caso del factor de transcripción NtcA, estructuras tanto en forma activa con 2OG unido correctamente, como sin 2OG o con unión aberrante de éste han sido publicadas para las cianobacterias *Anabaena* sp PCC 7120 (*Anabaena* sp a partir de ahora) (Zhao *et al.*, 2010b) (Fig. 9A y B) y *S. elongatus* (Llácer *et al.*, 2010)(Fig. 9C y D).

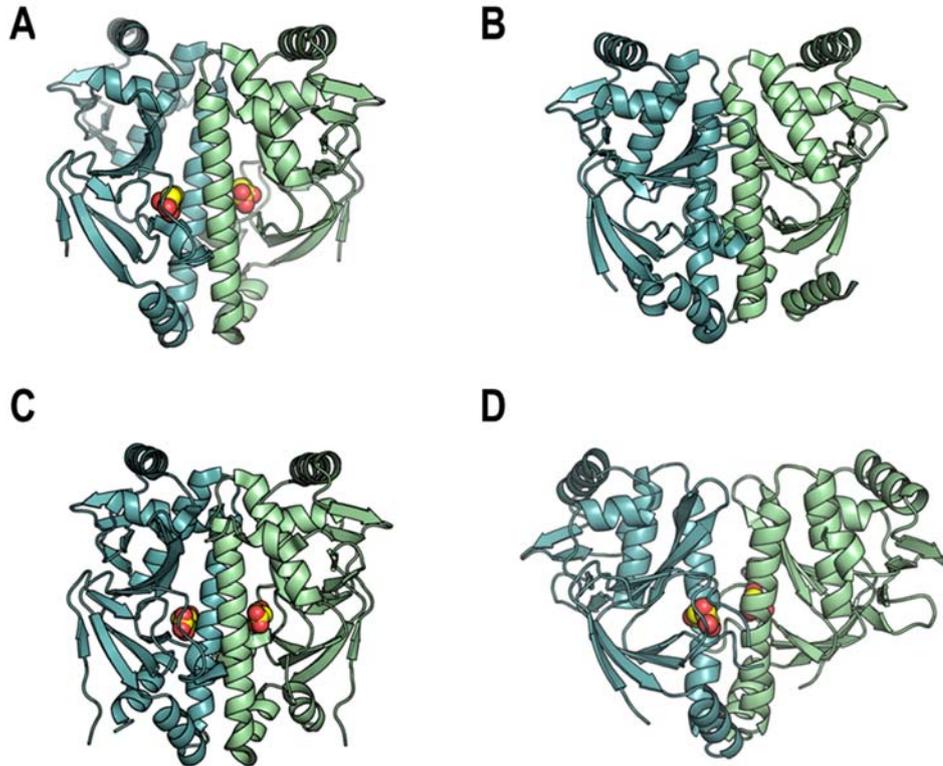


Figura 9. Estructuras publicadas del factor de transcripción NtcA en las cianobacterias *Anabaena* sp (A) en su forma activa (PDB 3LA2) o (B) inactiva (PDB 3LA7) y *S. elongatus* en su forma activa (C) (PDB 2XHK) o (D) inactiva (PDB 2XGX). NtcA se representa en cintas con cada subunidad coloreada distintamente. El 2OG unido de forma aberrante en (D) se representa como esferas.

Las estructuras de las formas con 2OG correctamente unido son prácticamente idénticas entre sí y con una alta similitud con las estructuras de otros miembros de la familia en su forma activa. Sin

embargo, la forma sin 2OG de *Anabaena* sp y la unida aberrantemente a 2OG de *S. elongatus* tienen conformaciones muy diferentes. Puesto que ambas parecen representar formas "inactivas" de NtcA, se podría plantear la duda de si la conformación inactiva es específica de especie, o si, por el contrario, existen varias conformaciones inactivas para cada especie.

Para responder a esta pregunta, en esta tesis presentamos también las estructuras de tres formas inactivas de NtcA en *S. elongatus*.

La proteína PipX y el sistema PII-PipX-NtcA

PipX es una proteína pequeña, de 89 aminoácidos, monomérica, cuya expresión está controlada por NtcA, al menos en *Anabaena* sp (Mitschke *et al.*, 2011; Valladares *et al.*, 2011). Se encuentra exclusivamente en cianobacterias, estando presente en la gran mayoría de ellas. Fue descubierta por el grupo de la Dra. Contreras en la Universidad de Alicante (Burillo *et al.*, 2004), siendo identificada como una diana de la proteína PII en *S. elongatus* mediante sistema de doble híbrido, y así, su nombre viene de "PII interacting protein X".

PipX oscila entre la proteína PII, a la que se une en ausencia de 2OG, y el regulador de la transcripción NtcA, al que se une en presencia de dicho ligando (Espinosa *et al.*, 2006) (Fig. 10). La interacción entre NtcA y PipX sólo se produce en presencia de 2OG ya que, en condiciones de escasez de amonio, PipX actúa como un co-activador transcripcional de NtcA (Espinosa *et al.*, 2006, 2007), aumentando la expresión de los genes que regula. Posiblemente la eficacia co-activadora de PipX varíe dependiendo de la calidad del promotor del gen como lugar de unión de NtcA. Las estructuras de NtcA unido a 2OG solo (Llácer *et al.*, 2010;

Zhao *et al.*, 2010b) o junto a PipX (Llácer *et al.*, 2010) han llevado a la visión de que las dos moléculas de PipX que se unen al dímero de NtcA estabilizan la conformación activa de este factor de transcripción, la cual es inducida por el 2OG (Llácer *et al.*, 2010).

Sin embargo, la presencia de 2OG disocia el complejo formado entre PII y PipX (Espinosa *et al.*, 2006). Así, bajo condiciones de bajo amonio, donde el 2OG debe de ser abundante, la unión del 2OG a PII propicia la liberación de PipX del complejo PII-PipX, dejando a esta última disponible para interactuar con NtcA y aumentar la actividad de este factor de transcripción.

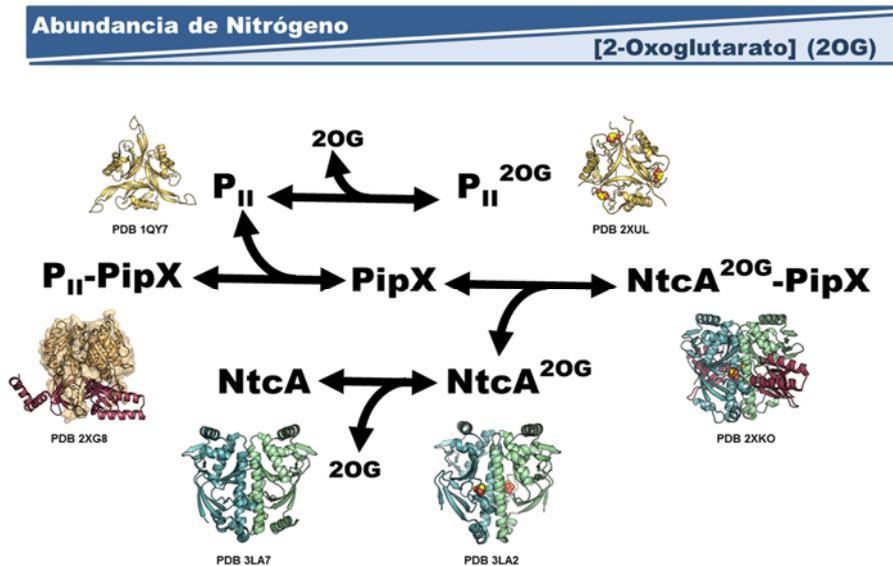


Figura 10. Esquema de la red de regulación PII-PipX-NtcA. PipX interactúa con PII o con NtcA dependiendo de las concentraciones de 2OG. Se indican las estructuras resueltas de cada uno de los componentes del sistema y su correspondiente código de PDB.

Por el contrario, cuando las concentraciones de 2OG son bajas (reflejando altos niveles de amonio) PII secuestra entre sus lazos T extendidos y su cuerpo hemisférico, tres moléculas de PipX de forma que las porciones de la misma que interaccionan con NtcA quedan ocultas, impidiendo así la interacción de PipX con este regulador transcripcional (Llácer *et al.*, 2010). De esta forma, PipX sitúa a NtcA y a su regulón bajo control parcial por PII. Esta regulación de PipX es muy importante para las cianobacterias, ya que no mantener la acción de PipX controlada parece ser tóxico y letal para para las células, tendiendo PipX a sufrir mutaciones que la inactivan en mutantes nulos de PII (Espinosa *et al.*, 2009; Laichoubi *et al.*, 2012).

Nuestro laboratorio ha determinado las estructuras de PipX de *S.elongatus* tanto en complejo con PII como en complejo con NtcA-2OG (Llácer *et al.*, 2010). Otros han determinado también la estructura de PII-PipX de *Anabaena* (Zhao *et al.*, 2010a). Sin embargo, a pesar de múltiples esfuerzos incluidos los míos propios, no hemos conseguido cristales de PipX sola, es decir, sin PII o NtcA unidos. La determinación de la estructura de PipX sin ligandos tiene interés. En los complejos de PipX con PII o con NtcA, el monómero de PipX está compuesto por un dominio compacto N-terminal (residuos 1-52) con un plegamiento de tipo Tudor, (una hoja β antiparalela altamente curvada formando un β -sándwich con 5 elementos β), seguido de dos α -hélices C-terminales ($\alpha 1$ y $\alpha 2$) (Fig. 11A). Este dominio Tudor media las interacciones de PipX tanto con PII como con NtcA, lo cual descarta un posible complejo ternario PII-PipX-NtcA. Por otra parte, las hélices C-terminales de PipX se encuentran expuestas al solvente sin mediar casi contactos con PII o NtcA

(Fig. 11B,C), presentando además la hélice C-terminal ($\alpha 2$) diferentes posiciones en ambos complejos. En el complejo NtcA-PipX (Llácer et al., 2010), formado por un dímero de NtcA unido a dos moléculas de 2OG y dos de PipX, la hélice C-terminal está plegada sobre la hélice previa (Llácer *et al.*, 2010) (Fig. 11B). Sin embargo, en el complejo PII-PipX de *S. elongatus* (Llácer *et al.*, 2010), formado por un trímero de PII y tres moléculas de PipX, PipX se puede observar en un monómero en una conformación extendida en la que la hélice se proyecta lejos del complejo mientras que en otro monómero se encuentra en la misma conformación que en el complejo NtcA-PipX (llamada conformación flexionada) (Fig. 11C).

El hecho de que en el complejo PII-PipX la hélice $\alpha 2$ sea móvil y pueda encontrarse en un conformación extendida proyectándose hacia el exterior sugiere que podría proporcionar un elemento para interacción con otras proteínas (Llácer *et al.*, 2010), indicando una posible implicación en funciones regulatorias adicionales. La existencia de funciones de PipX no ejercidas a través de NtcA fue evidenciada en un estudio de transcriptómica del grupo de la Dra. Contreras (Espinosa *et al.*, 2014) en *S. elongatus*, al identificar procesos que implican a *pipX* y exceden a lo explicable por coactivación/correpresión vía NtcA. PipX podría así controlar otros procesos independientemente de NtcA, tal como sería el caso si lo hiciera mediante la formación del complejo con PII.

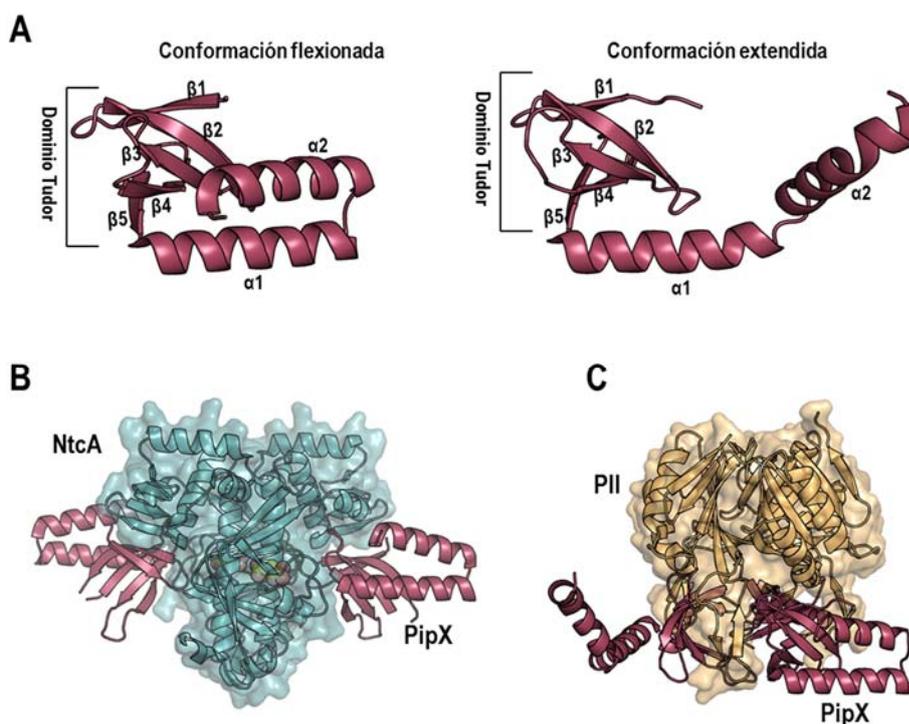


Figura 11. Estructuras de PipX y complejos con NtcA y PII. (A) Representación en cintas de PipX en las conformaciones encontradas en los complejos con NtcA y PII: flexionada y extendida. Los elementos de estructura secundaria han sido indicados. (B) Complejo entre PipX y NtcA (PDB 2XKO), en el que ambas moléculas de PipX presentan la conformación flexionada (C) Complejo formado entre PipX y PII (PDB 2XG8) donde se puede ver como las hélices de PipX adoptan conformaciones distintas. NtcA y PII se muestran en superficie semitransparente con estructura secundaria visible de color azul y amarillo respectivamente. PipX está representada por cintas de color rosa oscuro. El 2OG unido a NtcA se representa como esferas.

Ya se ha comentado anteriormente la detección mediante técnicas de triple híbrido de la formación de un complejo ternario PII-PipX-PlmA (Fig. 12A). Para indagar la especificidad de ese nuevo complejo, se introdujeron en el mismo mutaciones que pudieran afectar a la formación del complejo ternario sin perjudicar la formación del complejo PII-PipX (Labella *et al.*, 2016). Para ello se eligieron residuos de PipX localizados

en las hélices C-terminales así como una delección completa de la hélice 2, y residuos ubicados en los lazos T y B de PII (Fig. 12B).

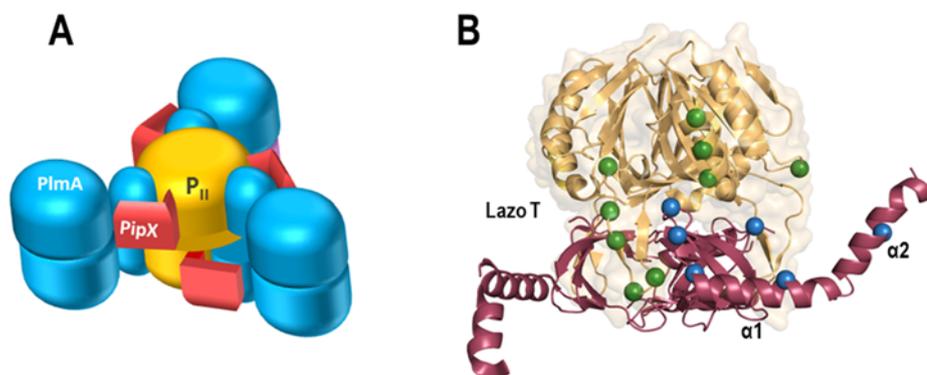


Figura 12. (A) Modelo de la posible estructura del complejo PII-PipX-PlmA. Los trímeros de PII se representan en amarillo, las hélices C-terminales de PipX se representan en rojo y PlmA en color azul. (B) Mutaciones realizadas en (Labella et al., 2016) sobre el complejo PII-PipX de *S. elongatus* para comprobar su efecto sobre la formación del complejo ternario con PlmA. PII se muestra como superficie semitransparente amarilla con la estructura secundaria visible. Las moléculas de PipX se muestran como cintas de color rosa fuerte. Los residuos mutados en PipX se muestran como esferas azules en una subunidad de PipX. Los residuos mutados en PII se muestran como esferas verdes. (PDB 2XG8).

El resultado obtenido fue que ninguna de las mutaciones localizadas en residuos de PII no implicados en la interacción con PipX afectaba a la formación del triple híbrido y sin embargo las mutaciones en los residuos de PipX situados en las hélices C-terminales, en especial la delección de la segunda hélice, afectaban negativamente o suprimían totalmente la formación del complejo ternario. La conclusión extraída de estos resultados fue que PlmA interaccionaría con el complejo PII-PipX a través de las hélices de PipX, siendo la $\alpha 2$ extendida esencial para la formación de las interacciones detectadas (Labella *et al.*, 2016). Además, se descartaría una interacción directa entre PII y PlmA, sugiriendo que la presencia de PII sería necesaria sólo para promover la conformación extendida de PipX necesaria para su interacción con PlmA (Labella *et al.*,

2016). Por ello esclarecer la estructura de PipX libre ha sido importante para responder a la pregunta de si es la unión a PII la que pone en marcha el desplegamiento de la hélice C-terminal de PipX, o si dicho desplegamiento sucede espontáneamente también en PipX libre.

Hay que señalar que la identificación de PlmA como un ligando de PipX en su complejo con PII sitúa a esta proteína en la red de regulación de PipX. A juzgar por su secuencia, PlmA es un miembro de la familia GntR de reguladores transcripcionales (Hoskisson & Rigali, 2009). Los miembros de esta familia se encuentran ampliamente distribuidos en bacterias y regulan procesos biológicos diversos tales como la regulación del metabolismo de ácidos grasos, catabolismo de aminoácidos y del carbono, degradación de complejos orgánicos y regulación de factores de virulencia (Hoskisson & Rigali, 2009). Estos factores son también homodiméricos y tienen un dominio N-terminal de unión a DNA con un plegamiento conservado hélice vuelta hélice alada (tres hélices α y dos hojas β) y un dominio C-terminal de dimerización y unión a ligando que es muy distinto de uno a otro miembro de esta familia (Jain, 2015). PlmA, al igual que ocurre con NtcA, se restringe únicamente a cianobacterias y a juzgar por las secuencias conocidas, se encuentra en todas ellas (Labella *et al.*, 2016), donde además es esencial para la supervivencia de (al menos) *S. elongatus* (Rubin *et al.*, 2015). Para PlmA se han propuesto en *Anabaena* (Lee *et al.*, 2003) funciones relacionadas con el mantenimiento de plásmidos, aunque no tiene lógica que sea esta la función única de PlmA (Suvorova *et al.*, 2015). También se han propuesto funciones en la modulación de la estequiometría de fotosistemas en *Synechocystis* sp.

PCC 6803 (Fujimori *et al.*, 2005). Sin embargo, la estructura y función de PlmA son desconocidas por el momento.

El último miembro reconocido como perteneciente a la red de regulación de PipX es la proteína cianobacteriana PipY, pues su gen (*Synpcc7942_2060* en *S.elongatus*, o, más brevemente, *pipY*) presenta sintenia con *pipX*, encontrándose inmediatamente aguas abajo del mismo (Labella *et al.*, 2017). Esta asociación génica se encuentra presente en la mayor parte de los genomas cianobacterianos estudiados, no existiendo prácticamente distancia entre ambos genes o siendo ésta muy pequeña (Labella *et al.*, 2017). Sin embargo, hasta hace muy poco (Labella *et al.*, 2017; Tremiño *et al.*, 2017), el producto de este gen no había sido estudiado en cianobacterias. *pipY* de *S.elongatus* (Labella *et al.*, 2017) codifica una proteína de 221 aminoácidos que pertenece a la familia COG0325, un grupo muy conservado de proteínas que unen piridoxal fosfato (PLP; la forma activa de la vitamina B₆) las cuales se encuentran universalmente distribuidas (Prunetti *et al.*, 2016). Como hemos podido comprobar al determinar la estructura cristalina de PipY (Tremiño *et al.*, 2017), las proteínas de la familia COG0325 son monoméricas y presentan un plegamiento de tipo III de proteínas de unión a PLP (Percudani & Peracchi, 2009). Este plegamiento consiste en un barril TIM modificado, el cual posee una hélice α N-terminal extra (Tremiño *et al.*, 2017) (Fig. 13), pero al que le falta el segundo dominio que caracteriza a las proteínas de este tipo que tienen actividad enzimática (por ejemplo la ornitina decarboxilasa o la alanina racemasa). Artículos recientes resaltan la implicación de proteínas de la familia COG0325 en homeostasis de vitamina B₆ (Ito *et al.*, 2013; Percudani & Peracchi, 2009), e incluso sus

mutaciones nulas (y otras de cambio de sentido) se han asociado a una epilepsia hereditaria humana dependiente de vitamina B₆ (Darin *et al.*, 2016; Plecko *et al.*, 2017).

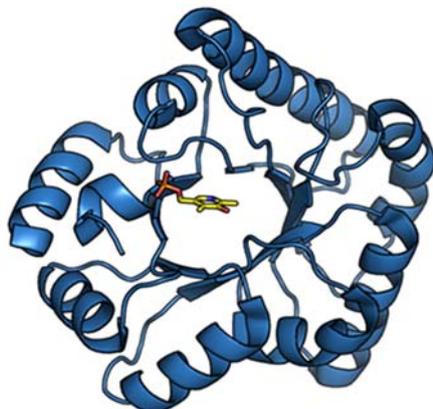


Figura 13. Estructura de la proteína PipY de *S.elongatus* con PLP unido (PDB 2NM8). PipY está representado en cintas de color azul y el PLP se muestra con los carbonos en color amarillo.

La sintenia *pipX-pipY* en cianobacterias y la organización de ambos genes sugieren fuertemente que cuando van juntos en el genoma, forman un operón bicistrónico (solo un 38% de los genes de *S. elongatus* conforman operones (Vijayan *et al.*, 2011)). Este hecho, y evidencias fenotípicas en mutantes de *pipX* y/o *pipY* (Labella *et al.*, 2017; Tremiño *et al.*, 2017) han llevado a incorporar a PipY en la red de regulación por nitrógeno y PipX. Aunque no objeto de esta tesis, hemos determinado la estructura de PipY (Tremiño *et al.*, 2017) e incluso hemos establecido en la proteína homóloga humana, PLPBP (antes PROSC) la patogénesis de la epilepsia vitamina B₆-dependiente por las mutaciones clínicas de cambio de sentido encontradas en pacientes con esa enfermedad (Tremiño *et al.*, 2018).

Objetivos

Sobre la base de lo resumido en la Introducción, esta Tesis aborda los siguientes objetivos, cuya consecución es objeto central de esta memoria:

1. Establecer los determinantes de la especificidad de la unión al DNA de NtcA, determinando cristalográficamente la estructura de NtcA unido a su DNA diana.
2. Investigar estructuralmente el papel de PipX en el complejo PipX-NtcA-DNA, determinando mediante difracción de rayos X la estructura de dicho complejo.
3. Determinar, mediante una aproximación cristalográfica, si hay una o varias conformaciones de NtcA "inactiva" y establecer las bases de su falta de actividad.
4. Estudiar funcionalmente (mediante resonancia de plasmón, SPR) qué hace a algunos genes del regulón de NtcA más o menos sensibles a NtcA, caracterizando la influencia de 2OG y PipX.
5. Someter a prueba mediante SPR la posible existencia de regulación cruzada entre los reguladores transcripcionales NtcA y CRP, dadas sus similitudes y las de sus DNAs dianas.
6. En el contexto de la extensión de la red reguladora de PipX a nuevos elementos: a) propiciar la determinación de la estructura de PipX cuando no está unida a otras proteínas; y b) examinar inmunológicamente las concentraciones relativas de nuevos componentes de esta red.

Materiales y Métodos

Producción y purificación de NtcA y PipX de *S. elongatus* para estudios estructurales

La producción de las proteínas *pipX* y *NtcA* de *S. elongatus* con etiquetas His₆ N-terminales a partir de los plásmidos pTrc99A-*pipX* y pET15b-*NtcA* y su purificación se ha descrito previamente (Espinosa *et al.*, 2006). En resumen, tras electrotransformar con dichos plásmidos células *E. coli* Rosetta (DE3) (de Novagen; expresan abundantemente tRNAs para codones raros de *E. coli*, e incorporan el gen para la polimerasa del fago T7 bajo control de promotor *lac* inducible por IPTG), y de proceder a su crecimiento a 37°C con aireación (180 rpm) hasta OD⁶⁰⁰= 0.6-1 (1-cm camino óptico) en 1 L de medio líquido LB con 0.1 mg/ml ampicilina (resistencia plasmídica) y 35 µg/ml de cloranfenicol (resistencia intrínseca de las células Rosetta), se agregó IPTG y se continuó el cultivo a la temperatura indicada durante 12 horas (temperatura/[IPTG], 25°C/0.25 mM para PipX y 18°C/0.5 mM para NtcA). Se aislaron las células por centrifugación (30 min, 3.200×g), realizando los pasos posteriores a 4 °C. Tras resuspensión de las células en 25 ml de medio de sonicación (para PipX, 50 mM MES pH 6, 10% v/v glicerol, 0.5 M NaCl, 5 mM MgCl₂, 1 mM DTT, 1 mM PMSF, 0.5 mM EDTA; para NtcA, 25 mM citrato sódico pH 6.5, 0.5 M NaCl, 5 mM MgCl₂, 1 mM DTT, 1 mM PMSF y 5 mM imidazol), se lisaron las células por sonicación (Soniprep 150 de MSE-SANYO, sonda mediana, amplitud de 13 µm, 20 ciclos de 15 s separados por periodos de reposo de 45 s), se centrifugó el lisado (30 min, 35.000 x g), el sobrenadante se pasó por un filtro de jeringa y membrana de 0.45 µm, y se sometió a cromatografía de afinidad por Ni²⁺ en columna HisTrap de 1 ml en cromatógrafo ÄKTA FPLC (columnas e instrumento de GE Healthcare) con soluciones tamponadoras frías, monitorizando la

absorbancia a 280 nm del efluente, y recogiendo fracciones sobre hielo fundente. El tampón de equilibrado de la columna y de lavado (hasta absorbancia constante y baja) fue 25 mM HEPES pH 7.5, 0.5 M NaCl, 5 mM MgCl₂, 20 mM imidazol, eluyendo la proteína de interés en gradiente de imidazol (25 ml, 20-500 mM) en el mismo tampón, recogiendo el pico de la proteína y caracterizando su pureza por SDS-PAGE (Laemmli, 1970) en geles de 15% de poliacrilamida (Fig. 14).

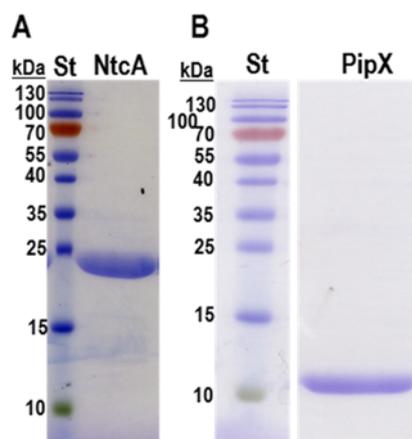


Figura 14. Ejemplos ilustrativos del grado de purificación obtenido para de las proteínas NtcA (A) y PipX (B) de *S. elongatus*. St, proteínas patrón (PageRuler™ Prestained Protein Ladder de Thermo Scientific) con masas polipeptídicas nominales dadas (en kDa) al margen. Las masas de los polipéptidos (incluyendo las etiquetas) en kDa son para NtcA 27 kDa y para PipX 11.5 kDa.

Las proteínas se colocaron en los tampones adecuados y se concentraron a 3.5 mg de PipX/ml o 10 mg de NtcA/ml (determinadas mediante la técnica de Bradford (Bradford, 1976) con el reactivo comercial de Bio-Rad, usando seralbúmina bovina como patrón; mediante ultrafiltración centrífuga (dispositivos Amicon Ultra de Millipore de corte nominal 3 kDa para PipX y 10kDa para NtcA) a 4°C. Las soluciones de conservación fueron, para PipX, 50 mM MES pH 6, 0.5 M NaCl, 1 mM DTT, 50 mM L-arginina hidrocloreuro, 50 mM de ácido L-glutámico (neutralizado con NaOH), y 10 % (v/v) glicerol; y para NtcA, 25 mM

citrato sódico pH 6.5, 0.5 M NaCl, 5 mM MgCl₂, 1 mM DTT, 50 mM L-arginina hidrocloreto y 50 mM de ácido L-glutámico (neutralizado con NaOH). La preparación de NtcA utilizada cuando se obtuvieron complejos con PipX y DNA se conservó en esta misma solución suplementada con 4 mM 2OG.

Cuando se preparó PipX marcada con ¹⁵N solo o junto con ¹³C todo se hizo igual excepto la utilización de medio mínimo M9 (Miller, 1972) en el que las fuentes de nitrógeno y de carbono eran ¹⁵NH₄Cl (1 g/L) y 2 g/L de glucosa, que, cuando fue de interés estaba marcada con ¹³C (ambos reactivos marcados se obtuvieron de Eurisotop, Saint-Aubin Cedex, France), suplementándolo con 0.1 mg/ml de ampicilina y 0.035 mg/ml de cloranfenicol. Los cultivos se realizaron cada vez en cuatro lotes de 0.25L contenidos en matraces Erlenmeyer de 1 L. En general a las 12-16 horas se alcanzaba la densidad apropiada (OD⁶⁰⁰ de 0.6-0.8), induciéndose la expresión con 0.25mM IPTG, durante 24h a 25°C. El resto del procedimiento de purificación de PipX fue como para cristalografía, excepto por el tampón final en que se suspendió (ver Capítulo 3 de Resultados).

Preparación de cristales de proteína

A) El procedimiento cristalográfico. Se empleó la técnica de difusión de vapor en gota asentada basada en el aumento gradual en la concentración de la solución proteica por pérdida de agua de la gota al producirse el equilibrio de vapor con la solución de cristalización presente en el reservorio. Se utilizaron placas MRC2 y MRC3 de 96 pocillos (Corning) y gotas preparadas generalmente mezclando 0.4 µl de solución de cristalización y 0.4 µl de solución de las moléculas a cristalizar,

depositándose las gotas por el servicio de cristalografía robotizado del IBV-CSIC (robot Honeybee X8, de Genomic Solutions) y utilizando cribados mediante el uso de distintos kits comerciales de Hampton, Qiagen, Axygen y Sigma, con sellado mediante plástico flexible adhesivo transparente e incubación a 21 °C, con exámenes visuales de la gota a distintos tiempos usando una lupa binocular de hasta 60 aumentos y luz fría. Cuando aparecieron cristales y estos difractaron, se trató de optimizar las propiedades de difracción de los mismos preparando soluciones a partir de la condición original en las que se variaban gradualmente los agentes previsiblemente importantes para la cristalización (concentraciones de precipitantes, inclusión de pequeñas cantidades de aditivos, cambios de pH o de fuerza iónica).

Para tratar de mejorar la resolución y minimizar el efecto dañino de los rayos X al incidir sobre los cristales, las recogidas de datos de difracción se llevaron a cabo a una temperatura de 100 K, para lo cual, previamente a su difracción, los cristales fueron congelados en nitrógeno líquido. Para evitar la formación de cristales de hielo durante la congelación, los cristales fueron sumergidos en soluciones basadas en la misma condición de cristalización, pero enriquecida en sustancias crioprotectoras (como son los polietilenglicoles de bajo peso molecular o glicerol) antes de su congelación en nitrógeno líquido.

B) Preparación de los diferentes cristales. Para la preparación de cristales de NtcA-DNA, el proceso se inició incubando 10 min a 65°C una mezcla equimolar de los oligonucleótidos ND1 y ND2 en tampón 5 mM Tris-HCl pH 7.5, dejándolos hibridar mediante enfriamiento gradual a 4 °C. Luego se preparó una solución que contenía 0.3 mg/ml del DNA

hibridado y 0.85 mg/ml de NtcA en tampón A (50 mM de sodio citrato pH 6.5, 0.5 M NaCl, 5 mM MgCl₂, 50 mM L-arginina-HCl y 50 mM Na L-glutamato), suplementando esta solución con 10 mM de 2OG. Tras incubar la solución 5 minutos a 22°C, se concentró hasta 7.3 mg/ml de NtcA por ultrafiltración en un dispositivo Amicon Ultracel YM-50. Esta fue la solución que se utilizó en las gotas de cristalización, preparadas mezclando 0.4 µl de solución de proteína-DNA y el mismo volumen de solución de precipitante. El cristal que se ha utilizado aquí para la determinación de la estructura se obtuvo en una solución de cristalización que consistía de 0.1 M Bis-Tris, pH 6.5 y 28% poletilenglicol monometil éter 2K.

El procedimiento para cristalizar el complejo PipX-NtcA-DNA fue similar al usado para el complejo NtcA-DNA. Primero se preparó un complejo NtcA-DNA mezclando 0.6 mg/ml de DNA hibridado y 1.4 mg/ml de NtcA en el tampón tampón A suplementado con 10 mM 2OG, incubando la solución 10 min a 22°C y añadiendo entonces 0.6 mg/ml de PipX, incubando 10 min más. Finalmente se concentró la solución a 6 mg/ml de proteína por ultrafiltración como se ha descrito anteriormente. En este complejo ternario el DNA hibridado fue ligeramente diferente al del complejo con NtcA sin PipX, puesto que contenía la caja NtcA entera en un dúplex simétrico de 30 pares de bases con extremos pegajosos (cebadores NXD-F y NXDR, Tabla 1). A un extremo sobresale una C solitaria y en el opuesto sobresale una G. Las gotas se prepararon como se ha indicado para NtcA-DNA, y, tras varias semanas, se obtuvieron dos tipos diferentes de cristales en soluciones de cristalización compuestas por 0.2M MgSO₄ y 20% PEG 3350 para el tipo I de cristal y 0.1 M Tris-HCl, pH 8.5/2 M fosfato amónico y 0.1M ZnCl₂ para cristales de tipo II.

Los primeros experimentos de difracción con los cristales de tipo II demostraron una resolución muy baja (6 Å) y un alto porcentaje de solvente (62.2%).

Tabla 1. Oligonucleótidos utilizados en este trabajo*

Identificador	Secuencia 5'-3'	Usado para
ND-1	AGCTGATACATAAAAAT	Complejo NtcA-DNA
ND-2	CATTTTTATGTATC	
NXD-F	CATTTTTATGTATCAGCTGATACATAAAAAT	Complejo PipX-NtcA-DNA
NXD-R	GATTTTTATGTATCAGCTGATACATAAAAAT	
V187E-F	TCGGCTCAACGCGGGAGACAGTGACG	Introducir mutación V187E en NtcA
V187E-R	CGTCACTGTCTCCCGCGTTGAGCCGATCG	
EMSA-F	(Flc)CACAACCAGGAACTGAAGAC	Promotor <i>glnA</i> para EMSA
EMSA-R	(Flc)CGCCTGCAAGATTTTCGTTAC	
pLIC-CRP ^{SC} -F	TACTTCCAATCCATG GGCCTAGTCC	Clonación de <i>CRP</i> de <i>Synechocystis</i>
pLIC-CRP ^{SC} -R	TATCAACCTTTACTGTGTCAGGAAATTAGATC	
pLIC-NtcA ^{SC} -F	TACTTCCAATCCATGGATCAGTCCCTAACC	Clonación de <i>ntcA</i> de <i>Synechocystis</i>
pLIC-NtcA ^{SC} -R	TATCAACCTTTACTGTTAGGTAAACTGTTG	
pLIC-PipX ^{SC} -F	TACTTCCAATCCATGAGTAACGAAATTT	Clonación de <i>pipX</i> de <i>Synechocystis</i>
pLIC-PipX ^{SC} -R	TATCAACCTTTACTGTCAAAAAGTGTTT	
p <i>glnA</i> ^{SC} -F	(Biot)AGCCGCAAAAATGGTAGCGAAAAATA CATTTTCTAACTACT	<i>p<i>glnA</i></i> para SPR
p <i>glnA</i> ^{SC} -R	AGTAGTTAGAAAATGTATTTTTTCGCTACCATT TTTTGCGGCT	
p <i>glnN</i> ^{SC} -F	(Biot)GGGTCGTTGTTTTGTATCTATATTGTCT ATTTTAAAAAATC	<i>p<i>glnN</i></i> para SPR
p <i>glnN</i> ^{SC} -R	GATTTTTTAAAAATAGACAATATAGATACAAA ACAACGACCC	
p <i>pccS</i> ^{SC} -F	(Biot)AGATACACAACAGTTGTGATCTGGGTCA CAACCATTGAGTGATT	<i>p<i>pccS</i></i> para SPR
p <i>pccS</i> ^{SC} -R	AATCACTCAATGGTTGTGACCCAGATCACAA CTGTTGTGTATCT	
Kanam-F	(Biot)GCTTTGTTGAATAAATCGAACTTTTGCT GAGTTGAAGGATCA	DNA inespecífico para SPR
Kanam-R	TGATCCTTCAACTCAGCAAAAGTTTCGATTTAT TCAACAAAGC	

*Superíndices SE y SC, *S. elongatus* y *Synechocystis*, respectivamente. (Flc), fluoresceína. (Biot), Biotina.

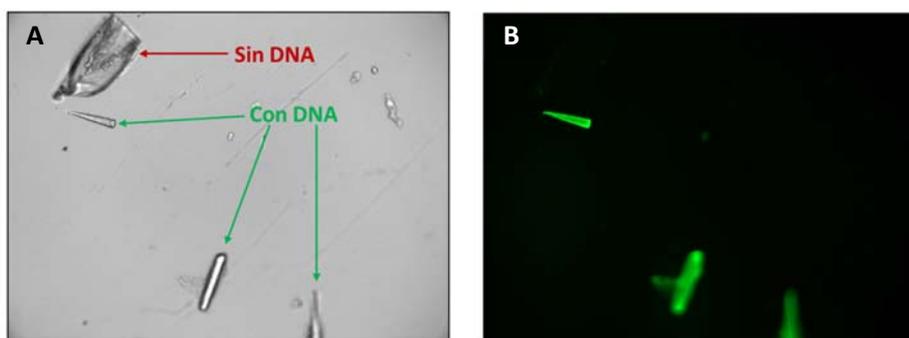


Figura 15. Fluorescencia debida al DNA en los cristales del complejo NtcA-PipX-DNA incubados con SYBR Green. (A) Visión de los cristales en campo claro. (B) Visión mediante microscopía de fluorescencia (epiiluminación).

Para tratar de mejorar la difracción de estos cristales los sometimos a deshidratación parcial con KCl antes de la congelación del cristal. Para ello el cristal se recogió en 5 μ l de solución crioprotectora (solución de cristalización a la que se había suplementado con un 25% de glicerol) y entonces el cristal se equilibró mediante difusión de vapor durante 30 min frente a 50 μ l de una solución 4 M de KCl.

Se obtuvieron cristales de apo NtcA en tres formas distintas a partir de una solución de 6.1 mg/ml de NtcA en tampón A, usando las siguientes soluciones de cristalización: 1) para la forma A1, 0.1 M Na acetato pH 4.5/40% PEG 200; 2) para la forma A2, 0.1 M Na HEPES pH 6.5/20.5% PEG 4 K; 3) para la forma B, 0.05 M Tris-HCl pH 8/2 M NaCl/2% 2-metil-2,4-pentanodiol/10 mM $MgCl_2$ /5% DMSO.

C) Detección de PipX y DNA en los cristales. La presencia de PipX en los cristales se confirmó tras su uso para difracción de rayos X, por SDS-PAGE de cristales lavados en solución de cristalización, seguido de western-blotting e inmunodetección con un antisuero policlonal anti-PipX (donado amablemente por la Dra. A. Contreras, U. de Alicante). La detección de DNA en los cristales la realizó amablemente para nosotros el

Sr. Juan M. Escamilla (técnico de nuestro laboratorio) usando un ensayo de fluorescencia (Kettenberger and Cramer, 2006) basado en la detección de fluorescencia en los cristales con un microscopio de fluorescencia (Leica DM RXA2 con un objetivo HCX PL FLUOTAR 20 X), tras 1 hora de incubación con el colorante fluorescente SYBR Green (de Invitrogen).

Recogida de datos y determinación de estructuras

A) Metodología general. El procedimiento seguido con todos los cristales de forma general fue tomar varias imágenes de difracción para valorar la calidad del cristal, el grupo espacial y la resolución alcanzada, y determinar si era adecuado para su recogida extensa. En caso afirmativo, mediante rotaciones graduales del cristal, se procedía a la recogida de imágenes de la difracción, recogiendo finalmente el número estimado de imágenes necesarias basado en la simetría (grupo espacial) del cristal para obtener un juego de datos completo de difracción. Estas imágenes fueron procesadas utilizando los programas iMOSFLM (Battye *et al.*, 2011) o XDS (Kabsch, 2010) que llevan a cabo el indexado e integrado de las reflexiones recogidas en las imágenes y el programa SCALA (Collaborative Computational Project, Number 4, 1994), con el cual se realiza el escalado. Para obtener los mapas de densidad electrónica a partir de los conjuntos de datos de difracción es necesario conocer tanto la amplitud como la fase de la onda. Para obtener dichas fases cristalográficas hemos utilizado en este trabajo exclusivamente el reemplazo molecular usando los programas PHASER (McCoy *et al.*, 2007) y MOLREP (Vagin & Teplyakov, 2010), ya que existían estructuras resueltas muy similares que podíamos utilizar como modelo inicial. A partir del primer mapa de densidad electrónica obtenido por reemplazo molecular y de un modelo inicial, trazamos el resto de la

proteína en ciclos iterativos usando el programa COOT (Emsley & Cowtan, 2004), donde construimos el modelo de forma manual, seguido de afinado del modelo con REFMAC (Murshudov *et al.*, 1997) o Phenix Refine (Afonine *et al.*, 2012). En este segundo paso, se tienen en cuenta la posición de los átomos construidos, el mapa de densidad electrónica y la estereoquímica de la molécula. Además, durante el afinado se calcula el factor estadístico de correlación R que indica las diferencias entre los factores de estructura experimentales y los calculados a partir del modelo, además del factor R_{free} similar al R, pero calculado a partir de (al menos) un 5 % de las reflexiones descartadas al inicio del afinado y que resulta una herramienta muy útil a la hora de minimizar el error humano por mala interpretación del modelo de densidad electrónica. La estereoquímica de las estructuras se analizó con el programa PROCHECK (Vaguine *et al.*, 1999), con el cual se pudo comprobar que ningún modelo poseía errores importantes, y también reveló buenos valores en la estereoquímica de todas las estructuras. Para la superposición de estructuras se utilizaron los programas SSM (Krissinel & Henrick, 2004) y LSQKAB (Collaborative Computational Project, Number 4, 1994). Las figuras de las estructuras de proteínas han sido realizadas utilizando el programa Pymol (<http://www.pymol.org>).

B) Aplicación a las estructuras determinadas aquí. Los conjuntos de datos para los cristales de NtcA-DNA, de NtcA apo forma B y del cristal tipo I de PipX-NtcA-DNA se recogieron en la solución de cristalización correspondiente, fortificada hasta 38% polietilenglicol monometil éter 2K y 27% 2-metil-2,4-pentanodiol, o hasta 39% polietilenglicol monometil éter 3.35K, respectivamente, o suplementada con 25 % PEG 400 y 25% glicerol en el caso de la forma apo A2 de NtcA

o de los cristales de PipX-NtcA-DNA, respectivamente. Los cristales así crioprotectados, o el cristal de NtcA apo forma A1, que no necesitó adición de crioprotector, fueron rápidamente congelados en N₂ líquido y se sometieron a difracción a 100 K en las fuentes indicadas en las tablas cristalográficas (sección de Resultados) en los sincrotrones ESRF or Diamond. Los conjuntos de datos obtenidos para NtcA apo A1 y apo B a resoluciones respectivas de 2.70 y 3.33 Å, y para los cristales I y II de PipX-NtcA-DNA a resoluciones de 3.8 y 4.30 Å, respectivamente, se procesaron y escalaron con iMOSFLM y SCALA mientras que los conjuntos de datos para NtcA-DNA (3.00 Å resolución) y para NtcA apo A2 (2.85 Å) se procesaron con XDS, y se prepararon con COMBAT (Collaborative Computational Project, Number 4, 1994) para escalarlos con SCALA. Los grupos espaciales para los cristales fueron C2, P12₁1, I4, C222₁, P2₁2₁2₁ and P6₄22 para los cristales de NtcA-DNA, PipX-NtcA-DNA I, PipX-NtcA-DNA II, NtcA apo A1, NtcA apo A2, y NtcA apo B, respectivamente (ver Tablas cristalográficas en la sección de Resultados).

Las fases cristalográficas para el cristal NtcA-DNA (que también contenía una molécula unida de 2OG por subunidad) se determinaron por reemplazo molecular con MOLREP usando como modelo el dímero de NtcA de *S. elongatus* en su conformación activa (número de acceso PDB, 2XHK (Llácer *et al.*, 2010)). El modelo inicial se mejoró mediante afinamiento de cuerpo rígido o restringido de máxima probabilidad usando REFMAC. Las fases obtenidas de esta solución parcialmente afinada se mejoraron por modificación de densidad con el programa PARROT (Collaborative Computational Project, Number 4, 1994). Se construyó manualmente un modelo inicial de DNA usando COOT. A

partir de entonces se realizaron rondas iterativas de construcción de modelo con COOT y afinamiento con REFMAC. Se aplicaron restricciones estrictas de simetría no cristalográfica hasta las rondas finales del afinamiento, cuando se liberaron gradualmente. Se usó TLS (Winn *et al.*, 2001) en los pasos finales del afinamiento, definiendo los grupos de TLS tras el análisis TLSMD (Painter & Merritt, 2006). Se usaron todos los datos de difracción en el proceso de afinamiento, exceptuando el 5% de datos elegidos al azar para calcular R_{free} . Se analizó la calidad estereoquímica del modelo con RAMPAGE (Lovell *et al.*, 2003). El modelo afinado final a 3.00 Å de resolución contenía un dímero de NtcA y dos cadenas de cada oligonucleótido (ver tablas cristalográficas en la sección de Resultados).

Para encontrar las fases para el conjunto de datos para NtcA apo A1 se utilizó también como modelo para el reemplazo molecular con MOLREP una subunidad de la forma “activa” de NtcA de *S. elongatus*. La construcción del modelo con COOT y el afinamiento con REFMAC y TLS fueron como para el complejo NtcA-DNA. Una de las subunidades de este modelo final (que contenía dos subunidades en la unidad asimétrica) se usó para encontrar las fases para el conjunto de datos para NtcA apo A2 usando reemplazo molecular, aunque en este caso con el programa PHASER, obteniéndose una solución satisfactoria para un dímero en la unidad asimétrica. Los demás pasos de construcción de modelo y afinamiento se realizaron como en los casos anteriores, alcanzando una resolución final de 2.85 Å (ver tabla cristalográfica en la sección de Resultados). Para la forma apo B de NtcA el reemplazo molecular con MOLREP tuvo éxito usando como modelo una subunidad de la conformación inactiva encontrada previamente para NtcA de *S.*

elongatus (PDB 2XKP, (Llácer *et al.*, 2010). Esta solución consistió en una única subunidad en la unidad asimétrica, que fue afinada como en los demás casos, y que generó un dímero al aplicar la simetría cristalográfica.

El grupo espacial para el cristal II de PipX-NtcA-DNA se asignó originalmente por XDS al grupo I422. En este grupo espacial, y de acuerdo con los cálculos del coeficiente de Matthews (Collaborative Computational Project, Number 4, 1994) el cristal debería contener probablemente en la unidad asimétrica un dímero de NtcA, una molécula de doble hélice de DNA y dos moléculas de PipX. Sin embargo, los intentos iniciales de encontrar este complejo de tres macromoléculas por reemplazo molecular usando tanto MOLREP como PHASER y utilizando un modelo de búsqueda basado en los complejos NtcA-DNA y NtcA-PipX no tuvieron éxito. Se encontraron soluciones satisfactorias cuando se usaron como modelos de búsqueda el dímero aislado de NtcA, el complejo NtcA-DNA o el complejo ternario con una única molécula de PipX. En todos estos casos la existencia de problemas de empaquetamiento para acomodar razonablemente una segunda molécula de PipX en ese grupo I422 sugirió la posibilidad de “twinning” merohédrico, lo que aumentaría la simetría aparente del cristal, por tanto, obligando a reevaluar la asignación del grupo espacial. Para juzgar sobre dicho twinning del complejo cristalino PipX-NtcA-DNA se realizaron análisis de las estadísticas de las intensidades en *phenix.xtriage* del paquete de software PHENIX (Afonine *et al.*, 2012) para un conjunto de datos I4, determinando la ley de twinning que gobernaba el cristal sujeto a twinning merohédrico como $(-h, k, -l)$, con una fracción de twinning alrededor de 0.45 (0.43 o 0.48 para el test H o para el método de la máxima probabilidad, respectivamente). De hecho, al comparar el valor

de Rsym a partir de los conjuntos de datos para los grupos espaciales I422 e I4 solo se observa una disminución desde 8.5 % a 7.4 %, lo que constituye otra indicación del “twinning” merohédrico casi perfecto ($\alpha=0.5$) del cristal. Así pues, se realizó reemplazo molecular en ese grupo espacial I4 usando PHASER y utilizando dos modelos diferentes, compuestos por un dímero de NtcA, una molécula de DNAY una o dos moléculas de PipX. El reemplazo molecular arrojó una solución satisfactoria de uno de cada uno de los modelos en la unidad asimétrica, con un coeficiente de Matthews de $3.25 \text{ \AA}^3 \text{ D}^{-1}$ y un contenido de solvente del 62.2%. Se afinó este modelo inicial mediante cuerpo rígido usando REFMAC, aplicando el operador “twin” $(-h, k, -l)$ en cada ronda de afinamiento. En este punto, y de acuerdo con el mapa de densidad electrónica se eliminaron del modelo algunos residuos mientras que otras regiones se movieron como dominios rígidos usando COOT. El modelo se afinó finalmente con REFMAC, usando el útil de afinamiento “cuerpo gelatinoso” y aplicando el operador de twin $(-h, k, -l)$ en cada ronda de afinamiento. Se aplicaron de forma automática restricciones locales de simetría no cristalográfica.

Los análisis de las estadísticas para las intensidades realizados en *phenix.xtriage* para el cristal de PipX-NtcA-DNA de tipo I también descubrieron otra patología, la seudosimetría traslacional. Considerando que la simetría resultante para el grupo espacial asignado inicialmente, P21, puede definirse como P12₁1 usando los símbolos universales Herman-Mauguin de grupo espacial, se realizó reemplazo molecular con PHASER usando este último grupo espacial y utilizando como conjunto de búsqueda un dímero de NtcA formando un complejo con dos moléculas de PipX y el DNA del cristal II del complejo PipX-NtcA-

DNA, arrojando una buena solución para dos tales conjuntos en la unidad asimétrica, con coeficiente de Matthews $2.43 \text{ \AA}^3 \text{ D}^{-1}$ y un contenido de solvente del 49.5%. A partir de este punto, tras afinamiento de cuerpo rígido con REFMAC, se alternó la construcción de modelo con COOT y el afinamiento combinado usando Phenix.refine y REFMAC usando de nuevo el instrumento de afinamiento de cuerpo gelatinoso. En REFMAC se aplicaron automáticamente restricciones locales de simetría no cristalográfica. A pesar de la baja resolución, ambas estructuras de PipX-NtcA-DNA (I y II) exhibieron muy buena geometría (Ver tablas cristalográficas en la sección de Resultados).

Mutagénesis dirigida de NtcA

Con el objeto de estudiar la capacidad de unión a DNA del mutante de NtcA V187E en *S. elongatus* con respecto a la proteína silvestre, introdujimos esta mutación en el plásmido pET15-NtcA^{Se} utilizando el método de PCR basado en el sistema Quickchange (Stratagene, Agilent Technologies, La Jolla, CA). Siguiendo las instrucciones de dicho sistema, se realizó una PCR en la cual se utilizaron dos cebadores parcialmente solapantes que introducían la mutación en dirección 5'-3' y 3'-5' y que amplificaban todo el vector parental. La secuencia de dichos cebadores se encuentra recogida en la Tabla 1. El producto de esta PCR, fue digerido con el enzima de restricción DpnI (New England Biolabs) (1U/ μL) durante 1h a 37°C con el objetivo de eliminar el vector parental no mutado. DpnI digiere tan sólo el ADN metilado, el cual se corresponde con el vector parental proveniente de *E. coli*, por lo que el vector mutado amplificado está protegido de la digestión. Seguidamente, 1 μL de producto de digestión fue transformado por electroporación en células DH5 α (Clontech) electrocompetentes. Las diferentes colonias producto de

la transformación se seleccionaron por su resistencia a ampicilina, y para verificar la correcta introducción de la mutación en los clones, el plásmido fue extraído utilizando el kit Wizard Plus SV Minipreps DNA Purification System” (Promega) basado en el método de la lisis alcalina. El rendimiento final se evaluó mediante la medida de absorbancia a 260 nm y las secuencias se determinaron por secuenciación convencional de Sanger automatizada en el servicio de secuenciación del Instituto de Biomedicina de Valencia-CSIC o del Instituto de Biología Celular y Molecular de Plantas (UPV-CSIC).

Ensayos de retardo en gel (EMSA)

Para estudiar las diferencias de afinidad entre NtcA silvestre y el mutante V187E llevamos a cabo ensayos de retardo en gel (EMSA, del inglés *Electrophoretic Mobility Shift Assay*). Esta técnica nos permite detectar complejos entre proteínas y DNA ya que los complejos proteína-ácido nucleico migran más lentamente que el DNA al ser sometidos a electroforesis bajo condiciones nativas.

Para estos ensayos utilizamos una sonda de DNA de doble cadena de 159 pares de bases producida mediante amplificación por PCR de una zona de DNA genómico de *S. elongatus* que comprende el promotor del gen *glnA*. Esta sonda fue marcada con fluoresceína incorporada durante la síntesis de los cebadores, cuya secuencia se detalla en la Tabla 1. El fragmento amplificado utilizando estos cebadores y la DNA polimerasa Deep Vent (New England Biolabs) se purificó con GeneClean (MP Biomedicals). 30 ng de esta sonda de DNA marcada se mezclaron con 0 ó 100 ng de NtcA purificado silvestre o mutante (el mutante NtcA V187E fue purificado de como se ha descrito más arriba para la NtcA silvestre) y esta mezcla se

incubó durante 10 min a 32 °C en 20 µl de una solución que contenía 50 mM HEPES pH 8, 3 mM MgCl₂, 20% glicerol, 1 mM DTT, 25µg/ml seralbúmina bovina (BSA) y, como DNA inespecífico, 0.5 µg de poli[d(I-C)] (Roche). Este tampón fue suplementado con 3.2 mM 2OG cuando se indica (ver apartado Resultados). Después de la incubación, se añadió tampón de carga para geles nativos (62,5 mM Tris-HCl pH 7,5, 10% glicerol y 0,02% azul de bromofenol), y la muestra se aplicó a un gel al 6% de poliacrilamida, preparado y corrido en TBE de baja fuerza iónica (Tris-HCl pH 8, 11 mM, ácido bórico 11 mM, EDTA 0.25mM). La electroforesis se realizó durante 90 min a 120 V y 4 °C. Cada experimento fue realizado independientemente y al menos por duplicado. Las posiciones de las bandas de DNA se visualizaron usando el aparato FUJIFILM-FLA-500(Europe, GMBH, Düsseldorf, Alemania).

Producción de NtcA, PipX y CRP de *Synechocystis*

Se clonaron los genes *NtcA*, *pipX*, y *sycrp1* de *Synechocystis* (códigos de identificación de la cianobase <http://genome.microbedb.jp/cyanobase:sll1423>, [ssl0105](http://genome.microbedb.jp/cyanobase:ssl0105) y [sll1371](http://genome.microbedb.jp/cyanobase:sll1371) respectivamente) para que las proteínas expresadas contengan una etiqueta His₆ N-terminal, utilizando el vector pLIC-SGC1 (<https://www.addgene.org/39187/>) donado por el Dr. Opher Gileadi, del Structural Genomics Consortium and Nuffield Department of Medicine, Universidad de Oxford, Inglaterra. Este vector permite la clonación sin ligación, así como la adición de una etiqueta N-terminal de poli-His y de un sitio para la proteasa TEV en el conector entre dicha etiqueta y la proteína producto.

El método de clonación independiente de ligasa se basa en la adición a ambos extremos de la secuencia a clonar extensiones idénticas a las zonas

del vector pLIC-SGC1 que flanquean el sitio para el inserto (Stols *et al.*, 2002) mediante amplificación por PCR con oligonucleótidos adecuados. Este método hace uso combinado de las actividades de síntesis y 3'-exonucleasa de la T4 DNA polimerasa, explotando la primera para una sola base, G o C. Así, las zonas idénticas elegidas del vector y del inserto no tienen las bases G o C hasta unos 15 nucleótidos del extremo más próximo. De este modo el tratamiento con la T4 polimerasa tanto del vector como del inserto en presencia de un único desoxirribonucleótido trifosfato (G o C en este caso) genera a cada lado un extremo libre desapareado de la longitud elegida (12-15 bases, hasta la primera C o G) con secuencia complementaria entre vector e inserto de forma que hibridan entre sí. Esta hibridación es lo suficientemente fuerte como para permitir la transformación en las células huésped sin la previa ligación de los fragmentos, de forma que esta ligación la llevan a cabo los enzimas del huésped una vez el plásmido ha sido introducido.

Como primera fase del proceso de clonación, se amplificaron mediante PCR los genes *NtcA*, *pipX*, y *sycrpl* de *Synechocystis* usando como molde DNA genómico (donado por el doctor Matthias E. Futschik de la Universidad de Algarve, Portugal) utilizando los cebadores indicados en la Tabla 1. La mezcla de reacción (50 μ l) contenía ~100 ng de DNA genómico molde, 1.2 U de la polimerasa de alta fidelidad Deep Vent (New England Biolabs), 0.02 nmoles de cada oligonucleótido cebador (directo y reverso), una mezcla de los cuatro desoxirribonucleótidos (0.6 mM de cada uno) y tampón de reacción compuesto por 20 mM Tris-HCl pH 8,8, 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄ y 0,1% Triton X-100. Se realizaron 30 ciclos de cambio de temperatura con pasos de desnaturalización del DNA (45 segundos a 95°C), hibridación de los

cebadores (45 segundos a 50 °C) y amplificación del DNA (1 min/kb, 72 °C). Tras corroborar la amplificación mediante electroforesis (agarosa 1% y visualización con Gel Red™ Nucleic Acid Gel Stain, Biotium) los productos se purificaron usando los kits comerciales “Illustra GFX PCR DNA and Gel Band Purification Kits” de GE Healthcare o “GeneClean” de MP Biomedicals, siguiendo las instrucciones del fabricante. Los DNAs así obtenidos fueron utilizados en la siguiente fase del proceso de clonación en pLIC-SGC1.

Previamente al tratamiento con la polimerasa T4 el vector pLIC-SGC1 fue digerido con el enzima de restricción BseRI (New England Biolabs, siguiendo indicaciones del fabricante) tanto para linearizarlo como para eliminar el gen *sacB* que porta dicho vector. A continuación, se mezcló el vector linearizado (~2µg) con 15 U de DNA polimerasa del fago T4 (New England Biolabs) y 2,5 mM de dGTP, en una solución que contenía 10 mM Tris-HCl pH 7.9, 10 mM MgCl₂, 5 mM DTT, 50 mM NaCl y 0.1 mg/ml de seralbúmina bovina, incubándose en un termociclador durante 30 min a 22 °C, seguidos de 20 min a 75 °C para desnaturalizar la enzima. El mismo protocolo fue seguido para tratar los genes amplificados (~500ng) pero utilizando en este caso dCTP en vez de dGTP. De esta forma se consiguió obtener extremos cohesivos de una sola hebra entre vector e inserto los cuales deberían tener una longitud de 14 bases. A continuación se mezclaron (para cada gen) 2 µl de la solución que contenía el inserto más 1 µl de la solución del vector, incubando la mezcla 10 min a 25 °C para conseguir la hibridación de las secuencias complementarias del vector y del inserto, tras lo que se usó la mezcla para transformar por choque térmico (Mandel & Higa, 1970) células *E. coli* TOP10 (Invitrogen) termocompetentes. Se realizó una selección doble

para detectar las células que habían incorporado el plásmido. Para ello sembramos las células transformadas en placas con 100 µg/ml de ampicilina (resistencia a dicho antibiótico provista por el plásmido) y también con sacarosa al 10%. La sacarosa provoca la muerte de las células que han incorporado un plásmido que no ha sido digerido y que por lo tanto aún porte el gen *sacB*, el cual debería haber sido sustituido por el inserto. La expresión de este gen en *E. coli* provoca la acumulación de levanos en el periplasma que acaban produciendo la muerte celular (Steinmetz *et al.*, 1983).

Los plásmidos que habían integrado los insertos de interés fueron aislados de las células supervivientes a partir de cultivos de noche de 10-20 ml de células transformadas TOP10 (Invitrogen) usando el kit “Wizard Plus SV Minipreps DNA Purification System” (Promega) basado en el método de la lisis alcalina. El rendimiento final se evaluó mediante la medida de absorbancia a 260 nm. Las secuencias de los genes clonados en los plásmidos se determinaron por secuenciación convencional de Sanger automatizada en el servicio de secuenciación del Instituto de Biomedicina

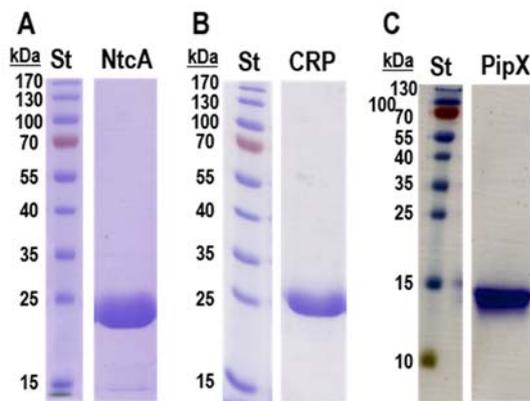


Figura 16. Alto grado de pureza electroforética (SDS-PAGE) de las proteínas de *Synechocystis* producidas. St, proteínas patrón preteñidas de masas nominales dadas al margen (en kDa). Los geles contuvieron 12% (NtcA y CRP) o 15% (PipX) de poliacrilamida. Las masas moleculares de las tres proteínas (incluyendo la etiqueta His₆) son 27.5, 28.6 y 13 kDa para NtcA, CRP y PipX.

de Valencia-CSIC o del Instituto de Biología Celular y Molecular de Plantas (UPV-CSIC).

Los aspectos relacionados con la electrotransformación de células Rosetta (DE3) con los plásmidos recombinantes generados, y el crecimiento de dichas células, así como su inducción y la purificación de las proteínas generadas fueron esencialmente como se han descrito para NtcA y PipX en un apartado anterior de estos Materiales y Métodos, usando para CRP las mismas condiciones de inducción que para NtcA. La figura 16 refleja los altos grados de pureza alcanzados para las tres proteínas de *Synechocystis*.

Ensayos de resonancia de plasmón (SPR)

La técnica SPR consiste en unir a la superficie de un chip el ligando que queremos estudiar y hacer pasar sobre él una muestra que contiene el analito con el que se espera que interaccione (o no) el ligando. La unión de moléculas a la superficie del chip genera una respuesta que es proporcional a la masa de las moléculas unidas (Fig.17).

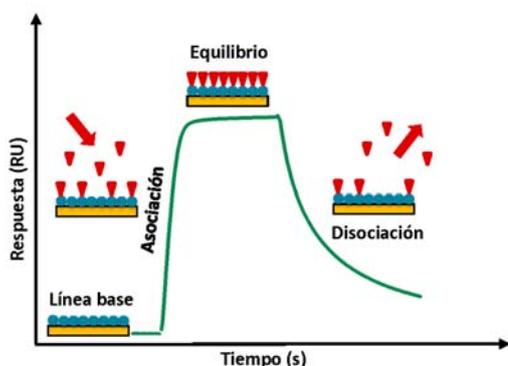


Figura 17. Sensograma SPR. El ligando (representado por círculos azules) es inmovilizado en la superficie del chip (línea base). El analito (representado por triángulos invertidos de color rojo) se hace pasar entonces por la superficie del chip, lo que da lugar a la fase de asociación. Cuando el analito deja de ser proporcionado al sistema, empieza la fase de disociación.

Cada chip utilizado contiene dos cámaras de flujo llamadas FC1 y FC2 (del inglés *flow chamber*) en las cuales se pueden inmovilizar dos DNA

distintos; nosotros usamos FC1 para unir un DNA control (DNA de gen de resistencia a kanamicina) uniendo al FC2 el DNA específico, y determinando la diferencia entre la señal en FC2 y en FC1, (FC2-FC1), determinando así la unión específica tras sustraer la inespecífica entre el DNA y las proteínas estudiadas.

Los experimentos de SPR fueron realizados a 24°C utilizando el sistema biosensor Biacore X (Biacore AB, Uppsala, Suecia) y los chips “*Streptavidin-cover sensorchips*” (GE Healthcare) cuya superficie está recubierta con estreptavidina. Nuestro objetivo era determinar las afinidades de NtcA y CRP por distintos promotores, por lo que quisimos inmovilizar el DNA en el chip y entonces monitorizar la unión de los factores de transcripción dependiendo de la presencia o ausencia de sus respectivos ligandos y de la proteína PipX. La estreptavidina presenta una gran afinidad por la biotina por lo que para inmovilizar el DNA en el chip, utilizamos DNAs de doble cadena con el extremo de una de sus cadenas biotinilada.

Los DNA dúplex que utilizamos para llevar a cabo estos experimentos y que fueron cargados en FC2 de cada chip fueron fragmentos de las regiones promotoras de los genes de *glnA*, *glnN* y *cccS* de *Synechocystis* (códigos de identificación de la cianobase: slr1756, slr0288 y slr1667 respectivamente) los cuales contenían la caja NtcA perfecta (*glnA*) o imperfecta (*glnN*) (Reyes *et al.*, 1997) o caja perfecta para CRP (*cccS*) (Yoshimura *et al.*, 2002b). Como DNA control, se inmovilizó en FC1 de cada chip la misma cantidad de un DNA dúplex inespecífico (derivado del gen *aphA* de resistencia a kanamicina, usado en muchos plásmidos) de la misma longitud y el cual no poseía cajas NtcA o CRP. Los

oligonucleótidos utilizados para preparar estos DNA dúplex, los cuales presentan el extremo 5' de la cadena+ biotinilada, se listan en la Tabla 1.

Para preparar los DNA dúplex (Hart *et al.*, 1999) cada par de oligonucleótidos complementarios biotinilados y no biotinilados fueron mezclados a una concentración de 5 y 6 μM respectivamente, en un tampón que contenía 10 mM Tris-HCl pH 7.4, 100 mM NaCl y 1 mM EDTA, y la hibridación se consiguió al incubar la mezcla primero durante 10 minutos a 95°C, y dejándola enfriar después hasta 25°C durante 30 minutos (Hart *et al.*, 1999). La inmovilización en el chip de los DNA dúplex marcados con biotina a través de unión estreptavidina-biotina, se llevó a cabo inyectando 30 μl de una dilución 1:5.5 del dúplex deseado en tampón (10 mM HEPES pH 7.4, 0.1 M NaCl, 0.2 mM EDTA, 0.005% Nonidet P-40) a un flujo de 10 $\mu\text{l}/\text{min}$ tanto en FC1 (dúplex inespecífico *aphA*) como en FC2 (dúplex del fragmento del promotor deseado) hasta alcanzar una señal de aumento de resonancia de ~1500 unidades relativas (RU), momento en el cual se detenía la inyección. Inmediatamente después se realizaba un lavado inyectando 5 μl de 2 M NaCl y seguidamente se dejaba pasar tampón de flujo (10 mM HEPES pH 7.4, 0.32 M NaCl, 3 mM MgCl_2 , 0.2 mM EDTA y 0.005% Nonidet P-40) a un flujo de 30 $\mu\text{l}/\text{min}$ a través de ambas células de flujo (FC1 y FC2) conectadas en serie.

Para los ensayos de unión, las proteínas NtcA y CRP se diluyeron >10 veces en tampón de flujo y se incubaron 10 min a 25 °C antes de inyectarlas al sistema. Las incubaciones contenían, cuando se indica, las concentraciones especificadas de cAMP, 2OG (sal disódica) y/o proteína PipX. Las mezclas se inyectaron a un flujo de 30 $\mu\text{l}/\text{min}$, pasando a través

de FC1 y FC2 en serie, registrando la diferencia de señales (FC2- FC1). La muestra proteica era entonces reemplazada por tampón de flujo a la misma velocidad de flujo para permitir la disociación del complejo DNA-proteína, regenerando luego el chip con 30 μ l de 2M NaCl y paso subsiguiente de tampón de flujo antes de realizar la siguiente inyección.

Los datos de Biacore se procesaron con el software BIAevaluation 3.0 (Biacore AB). Para el ajuste secundario de datos a hipérbola, sigmoides o exponenciales se utilizó el programa GraphPad Prism (GraphPad Software, San Diego, California).

Inmunodetección de proteínas

Para la determinación de la presencia de PipX de *S. elongatus* en los cristales y también para determinar los niveles celulares de las proteínas PII, PipX, PlmA y PipY en cultivos de *S. elongatus* silvestres y en diferentes cepas mutantes (Labella *et al.*, 2016, 2017) se utilizó la técnica de western blot con anticuerpos policlonales adecuados. En el caso de sedimentos celulares congelados (proporcionados por el grupo de la Dra. Contreras, U. de Alicante) las células se diluyeron en 10 veces su peso de una solución de 10 mM Tris-HCl pH 7.5, 1 mM β -mercaptoetanol, 0.5 mM EDTA y 1 mM PMSF, y tras añadir bolitas de vidrio (*Glass beads, acid-washed*, de 150 μ m, de Sigma) se sometieron a rotura utilizando bien agitación máxima en vórtex durante 5 minutos (5 ciclos de 30 s separados por periodos de 30 s en hielo), bien usando un homogeneizador Precellys Evolution (Bertin Technologies) aplicando 6 pulsos de 25 segundos a 6800 revoluciones por minuto separados por intervalos de 1 minuto en hielo. Una vez realizada la lisis, se separaron las bolitas de vidrio por sedimentación y se determinó la cantidad de proteína total en la

suspensión sin vidrio mediante el método de Lowry usando el “Modified Lowry Protein Assay Kit” de Thermo Scientific, determinando la absorbancia resultante a 750 nm para evitar la interferencia por el color de la clorofila, utilizando un patrón de seralbúmina bovina. Las cantidades indicadas de cada extracto celular (como μg de proteína), en paralelo con cantidades conocidas de las diferentes proteínas puras purificadas se sometieron a a SDS-PAGE en geles al 15% de poliacrilamida para PII, PipX y PipY y al 12% para PlmA. Seguidamente las muestras de los geles se transfirieron a una membrana de nitrocelulosa durante 1 hora a 15V mediante el sistema de transferencia semiseca utilizando para ello el aparato Semi-dry Electrobloetter (Sartorius, Jancos, Dinamarca). Finalizada la transferencia, las membranas se tiñeron durante 1 minuto con rojo Ponceau para comprobar tanto la correcta transferencia de las proteínas, como una carga similar de todos los extractos celulares (control de carga). El rojo Ponceau se destiñó lavando las membranas en una solución de TBS (50 mM Tris-HCl pH 7.5, 200 mM NaCl) con 0.1% de Tween-20, (TBS-T a partir de ahora).

A continuación, para evitar uniones inespecíficas, la membrana fue bloqueada mediante incubación en solución de bloqueo (leche desnatada en polvo al 5% (p/v) en TBS-T) durante 1 hora a temperatura ambiente y en agitación continua pero suave. Una vez bloqueada, la inmunodetección se realizó incubando la membrana durante toda la noche a 4°C en agitación con el anticuerpo específico para cada caso diluido en solución de bloqueo. Estas diluciones fueron de 1:7500, 1:5000, 1:5000 y 1:750 para los anticuerpos contra PII, PipX, PlmA y PipY, respectivamente. Los antisueros de conejo policlonales específicos contra las proteínas PlmA y PipY fueron encargados a Pineda Antikörper Service (Berlin, Alemania)

y Genosphere Biotechnologies (Paris, Francia) respectivamente, usando como antígenos dichas proteínas puras (Labella *et al.*, 2016, 2017). Sin embargo, los antisueros de conejo policlonales específicos contra las proteínas PII y PipX fueron donados por el Dr. Karl Forchhammer (Universidad de Tübingen, Alemania) (Espinosa *et al.*, 2010; Forchhammer & Tandeau de Marsac, 1994).

Tras tres lavados con TBS-T durante 10 minutos a temperatura ambiente, las membranas se incubaron durante 1 hora a temperatura ambiente en agitación con un anticuerpo secundario “*Anti rabbit IgG HRP conjugate*” (Promega) diluido 1:5000 en solución de bloqueo. Tras otros tres lavados con TBS-T, la detección se realizó en un sistema de quimioluminiscencia “*Image Quant LAS4000 Mini*” (GE Healthcare) usando como agente revelador “*ECLTM Western Blotting Detection Reagents*” (Amersham Biosciences, GE Healthcare). La intensidad de las bandas fue cuantificada mediante el programa Multi Gauge V2.1 de (Fujifilm).

Resultados

Capítulo 1

El regulador transcripcional NtcA y sus estructuras cristalinas

El complejo NtcA-2OG-DNA

Determinada la estructura de NtcA de *S. elongatus* y *Anabaena* en ausencia de DNA en formas “inactiva” y “activa” (Llácer *et al.*, 2010; Zhao *et al.*, 2010b), y de *S. elongatus* “activa” con PipX unido (Llácer *et al.*, 2010), en esta tesis hemos abordado la determinación de las estructuras de NtcA unidas a DNA, sin y con PipX unido.

A) Estructura del complejo

Para cristalizar este complejo utilizamos dos hebras de DNA sintético complementarias de distinta longitud (ND-1 y ND-2, Tabla 1) que tienen secuencias basadas en la caja NtcA del promotor del gen *glnA* de *S. elongatus* (a la que NtcA se une fuertemente, (Reyes *et al.*, 1997; Forcada-Nadal *et al.*, 2014). Estas secuencias están elegidas de tal manera que solo está representada la primera mitad de la secuencia consenso, de modo que cuando se hibridan ambos oligonucleótidos se forma una doble hélice con una base desapareada a un extremo (una C, a la que no se dio uso), quedando desapareadas en el extremo contrario cuatro bases con secuencia autocomplementaria, lo que permite el autoensamblado de dos dobles hélices generando una secuencia consenso palindrómica perfecta (Fig. 18).

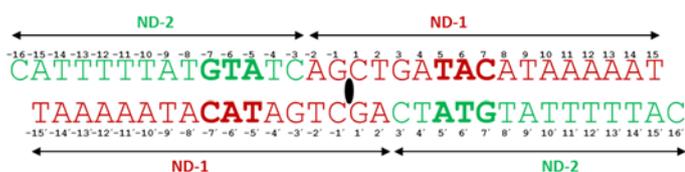


Figura 18. Sitio palindrómico perfecto de unión a NtcA utilizado para cristalización del complejo NtcA-DNA. Las líneas azules señalan las interrupciones en las cadenas de DNA, para facilitar que la doble hélice pueda curvarse. El óvalo negro señala el eje de simetría dos del palíndromo.

Así se reproduce el enfoque utilizado para cristalizar el complejo CRP-DNA (Parkinson *et al.*, 1996), en el que el DNA utilizado contenía puntos de rotura para facilitar la curvatura del DNA.

La obtención de cristales del complejo obligó a preparar dicho complejo en forma diluida y a concentrarlo después por ultrafiltración centrífuga. En resumen, tras incubar toda la noche a 4°C concentraciones equimolares de los oligonucleótidos ND-1 y ND-2 (Tabla 1) totalizando unos 0.3 mg/ml, se agregó una concentración aproximadamente equimolar de NtcA (como dímero; unos 0.85 mg/ml) en presencia de 10 mM 2OG. Se dejó la mezcla a 22°C unos minutos antes de iniciar la ultrafiltración centrífuga a esa misma temperatura hasta alcanzar una concentración en principio apropiada para el crecimiento de cristales (7.3 mg de proteína/ml; para los detalles de la concentración, ver Materiales y Métodos). Usamos esta solución para realizar inmediatamente una búsqueda robotizada de condiciones apropiadas de cristalización por el procedimiento de la difusión de vapor, seguida de mejoras robotizadas y manuales que llevaron finalmente a producir cristales a 21°C conformados como prismas hexagonales (Fig. 19A), de los que uno difractó los rayos X de la línea BM16 del sincrotrón ESRF a 3.0 Å de resolución, obteniéndose un juego de datos completo a esa resolución (Tabla 2).

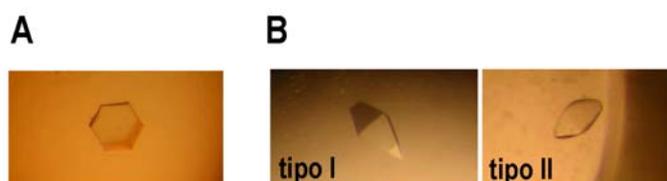


Figura 19. Cristales de A) NtcA con DNA; y B) de PipX-NtcA-DNA.

Las fases se obtuvieron por reemplazo molecular utilizando como modelo la estructura determinada en el laboratorio de NtcA “activa” de *S. elongatus* (PDB 2XHK), construyéndose manualmente la doble hélice de DNA en la densidad no explicada por la proteína. Tras refinamiento como se indica en Materiales y Métodos, se obtuvo un modelo final, a 3.0 Å de resolución en el que la unidad asimétrica contenía un dímero de NtcA y una doble cadena de DNA (Tabla 2, columna NtcA-DNA).

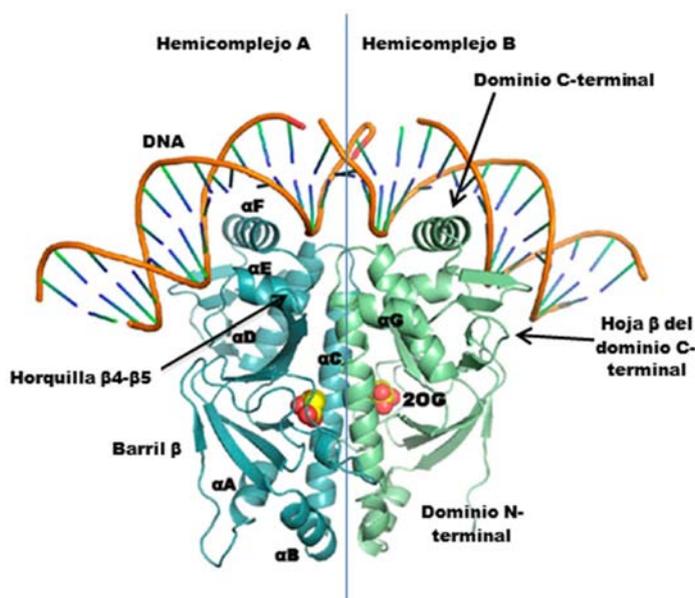


Figura 20. Estructura del complejo NtcA-DNA con 2OG unido (representado en esferas). Una subunidad de NtcA ha sido coloreada en azul y la otra en verde. Se señalan mediante etiquetas en una u otra de las subunidades algunos elementos de estructura secundaria.

Al igual que en la estructura del dímero “activo” de NtcA sin DNA (Llácer *et al.*, 2010; Zhao *et al.*, 2010b) cada subunidad se compone de un dominio N-terminal al que se une el 2OG (abreviado EBD de *effector binding domain*), y un dominio C-terminal de unión al DNA (DBD de *DNA binding domain*), ambos plegados como en la estructura de NtcA “activa” no unida al DNA (Llácer *et al.*, 2010; Zhao *et al.*, 2010b). Así, el dominio N-terminal lo forman un barril β de 8 hebras antiparalelas, con tres hélices adicionales (una N-terminal y dos C-terminales) de las que la

Resultados

Tabla 2. Datos de difracción de rayos X y estadísticas de afinado para el cristal de NtcA-DNA y para el modelo generado, y para los dos cristales de PipX-NtcA-DNA y para los modelos estructurales correspondientes

	NtcA-DNA	PipX-NtcA-DNA I	PipX- NtcA-DNA II
Sincrotrón (línea)	ESRF (BM16)	Diamond (I04.1)	ESRF (ID23-1)
Longitud de onda (Å)	0.9797	0.9173	0.9772
Grupo especial	C2	P12 ₁	I4
Celda unidad: a, b, c (Å)	120.0, 47.0, 139.6	113.3, 68.2, 126.3	165.7, 165.7, 177.6
α, β, γ (°)	90, 98.7, 90	90, 115.1, 90	90, 90, 90
Resolución (Å)	24.2-3.00 (3.16-3.00) ^a	114.4-3.80 (4.01-3.80) ^a	82.9-4.30 (4.53-4.30) ^a
Reflexiones totales/únicas	41,267/15,466	44,255/16,087	59,483 /16,116
Complejidad (%)	98.5 (97.3) ^a	92.2 (86.6) ^a	98.5 (98.5) ^a
Multiplicidad	2.7 (2.4) ^a	2.8 (2.4) ^a	3.8 (3.7) ^a
I/ σ	7.9 (1.8) ^a	3.0 (1.6) ^a	4.8 (1.6) ^a
R _{sym} ^b (%)	8.5 (42.9) ^a	19.3(39.8) ^a	7.4 (47.2) ^a
R _{factor} ^c /R _{free} ^d (%)	22.4/25.4	31.57/36.78	31.3/35.6
rmsd longitud enlace (Å)	0.007	0.009	0.005
rmsd ángulo enlace (°)	1.26	1.05	1.01
Moléculas o átomos			
cadenas de NtcA	2	4	4
cadenas de PipX	0	4	3
átomos de proteína	3343	9548	8833
2-oxoglutarato	2	4	4
dobles hélices de DNA	2 (fusionadas)	2 (fusionadas)	2
átomos de DNA	1230	2460	2460
moléculas de agua	5	0	0
Factores B (medias, Å ²)			
átomos de proteína	62.5	61.39	-
2-oxoglutarato	34.8	54.5	-
átomos de DNA	78.2	73.99	-
moléculas de agua	31.4	-	-
Repres. Ramachandran			
Favorables (%)	95.5	95.3	96.5
Permitidos (%)	3.6	4.3	3.1
No permitidos (%)	0	0.3	0.4

^a Los datos entre paréntesis son los de la celda de mayor resolución

^b $R_{sym} = \sum I - \langle I \rangle / \sum I$, donde I es la intensidad observada e $\langle I \rangle$ la intensidad media.

^c $R_{factor} = \sum h ||F_{obs}| - |F_{calc}|| / \sum h |F_{obs}|$, donde $|F_{obs}|$ y $|F_{calc}|$ son los factores de estructura observados y calculados para todas las reflexiones.

^d R_{free} , igual que R_{factor} , pero calculado con el 5% de los datos omitidos del afinamiento.

hélices tienen direcciones en el espacio que, proyectadas sobre un plano, son aproximadamente perpendiculares a la dirección de la cadena de DNA. De esta forma, las cadenas laterales de los aminoácidos de cada hélice F emergen radialmente y pueden interactuar con las bases y fosfatos que les rodean en cada vuelta del surco mayor (Figs. 21 y ver más adelante la Fig. 24).

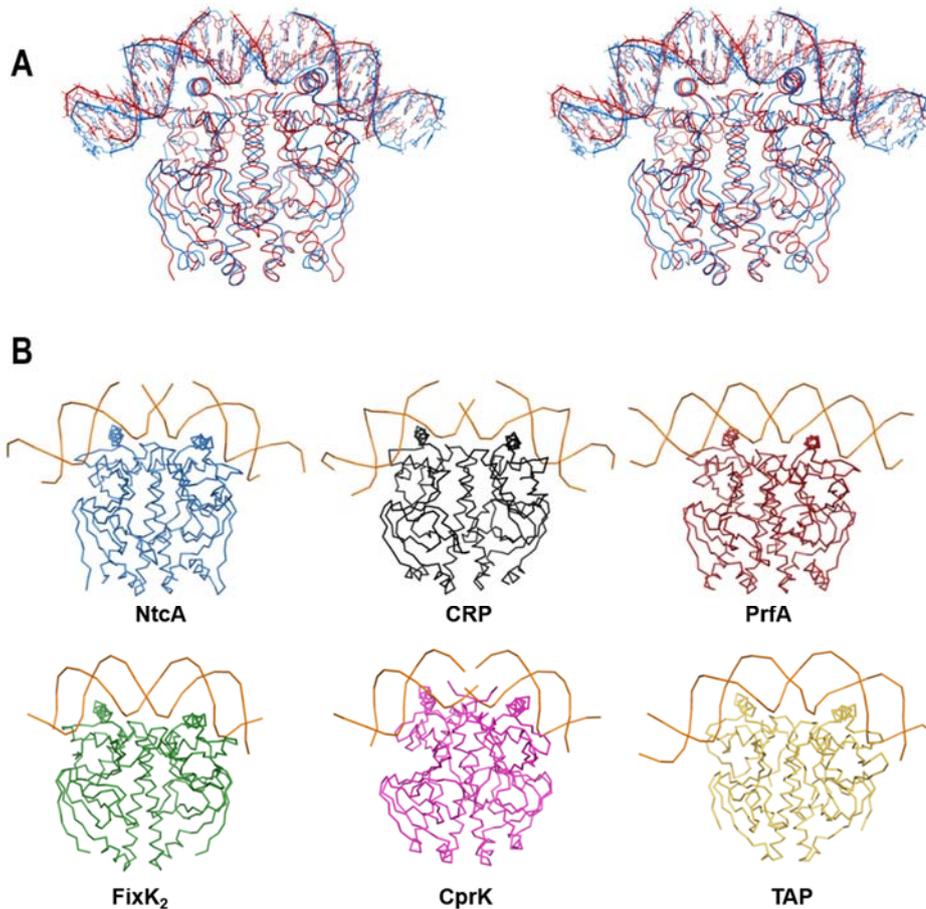


Figura 22.(A) Visión estereoscópica de la superposición de las estructuras NtcA-DNA (rojo) y CRP-DNA (azul) (PDB 1J59). (B) Representación de cadena principal de los miembros de la familia CRP cuya estructura unida a DNA ha sido resuelta. El DNA se muestra en color naranja, NtcA (este trabajo) en azul, CRP (PDB: 1J59) en negro, PrfA (PDB: 5LRS) en rojo, Fixk₂ (PDB: 412O) en verde, CprK (PDB: 3E6C) en rosa y TAP (PDB: 5I2D) en amarillo.

Además, los fosfatos situados en el surco menor del DNA que precede y sigue la secuencia consenso forman también algunos contactos con aminoácidos de la proteína.

La estructura del complejo NtcA-DNA es similar a la del complejo CRP-DNA (Benoff *et al.*, 2002; Chen *et al.*, 2001b; Lawson *et al.*, 2004; Napoli *et al.*, 2006; Parkinson *et al.*, 1996; Schultz *et al.*, 1991) y a las de los complejos con sus respectivos DNAs de otros miembros de la familia CRP: CrpK-DNA (Levy *et al.*, 2008), Fixk₂-DNA (Bonnet *et al.*, 2013), PrfA-DNA (Hall *et al.*, 2016) y TAP-DNA (Feng *et al.*, 2016) (Fig. 22). La alta similitud entre estos complejos sugiere fuertemente que los miembros de la familia CRP activan la transcripción uniéndose al DNA esencialmente de la misma manera

Como ya se ha dicho, el NtcA en este complejo se encuentra en la misma conformación que las formas “activas” de NtcA (unidas a 2OG) de estructuras publicadas previamente (*S. elongatus*, (Llácer *et al.*, 2010); *Anabaena sp.*, (Zhao *et al.*, 2010b) aunque dichas formas no estaban unidas a DNA. Así, los valores de rmsd para la superposición de las subunidades de NtcA-DNA con las formas “activas” de *S. elongatus* y de *Anabaena* son de $0.67 \pm 0.04 \text{ \AA}$ (para $211 \pm 2 \text{ C}^\alpha$) y $0.68 \pm 0.04 \text{ \AA}$ (para $191 \pm 1 \text{ C}^\alpha$), respectivamente, y para la superposición de dímeros de 0.72 \AA (para 421 C^α) o 0.77 \AA (para 388 C^α), respectivamente. También es similar a la forma activa de NtcA cuando forma complejo con 2OG y PipX, pero sin DNA (Llácer *et al.*, 2010) (rmsd para la superposición de los dímeros, 1.01 \AA para 398 C^α), así como cuando forma parte del complejo PipX-NtcA-DNA (ver más adelante; rmsd de la superposición de los dímeros, $0.81 \pm 0.04 \text{ \AA}$ para 416 C^α) (Fig. 23).

Por contraposición, esta conformación es muy diferente a la encontrada en las formas “inactivas” de NtcA (rmsd de la superposición de los dímeros de $2.41 \pm 0.02 \text{ \AA}$ para $375 \pm 11 \text{ C}^\alpha$ y $4.43 \pm 0.03 \text{ \AA}$ para $294 \pm 3 \text{ C}^\alpha$ para las formas inactivas A y B respectivamente, ver apartado formas apo más adelante). La comparación entre todas las formas “activas” revela que las mayores diferencias en la cadena principal se encuentran en los extremos, aunque ninguno de ellos parece participar en la interacción con el DNA. Los cambios atribuibles a la unión del DNA se localizan únicamente en las cadenas laterales de aquellos residuos situados en la hélice de reconocimiento (αF) que interaccionan directamente con éste (Fig. 20).

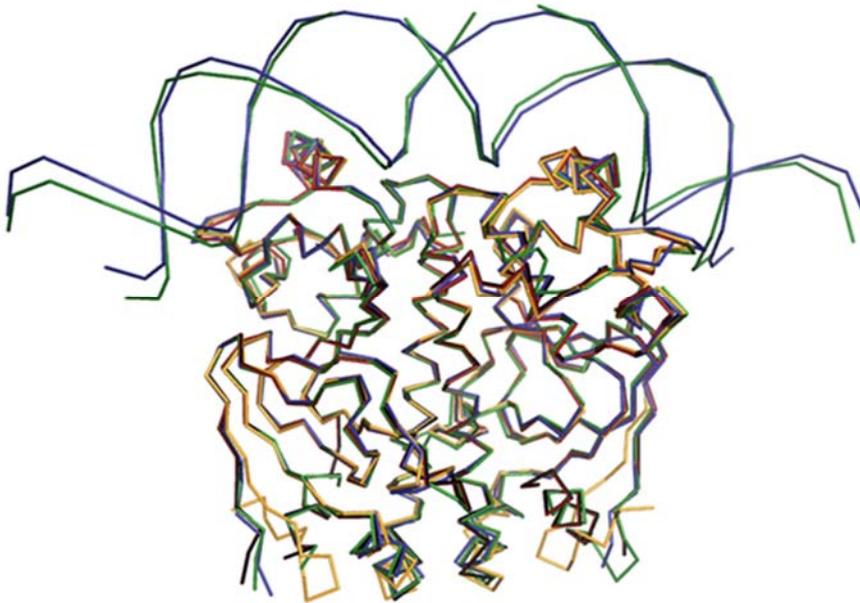


Figura. 23. Comparación de las estructuras de NtcA en sus conformaciones activas: NtcA-DNA (verde), PipX-NtcA-DNA (azul), PipX-NtcA (PDB: 2XKO naranja), y NtcA-2OG (PDB: 2XHK, rojo), todas ellas de *S. elongatus*, y NtcA-2OG de *Anabaena* sp (PDB: 3LA2 negro).

B) Reconocimiento de la caja NtcA

La mayor parte de las interacciones entre NtcA y el DNA están mediadas por residuos localizados en el motivo hélice-giro-hélice de NtcA, especialmente por su hélice de reconocimiento (α F) (Fig. 24). Una parte de estas interacciones se producen con grupos fosfato del DNA, y por lo tanto, aunque dan estabilidad a la unión de la proteína con su DNA diana, no son determinantes de especificidad. Los residuos de NtcA implicados en estas interacciones son R142, M144, H175, S184, T185, T188, R197, K207 (Fig. 24 y Tabla 3), destacando las interacciones mediadas por R142. En las estructuras resueltas de miembros de la familia CRP unidos a DNA, cada subunidad del factor de transcripción interacciona exclusivamente con una mitad de la secuencia consenso de DNA, sin que existan interacciones cruzadas. Sin embargo, en NtcA, el fosfato con el que interacciona el residuo R142, pertenece a la mitad de la caja NtcA que interacciona mayoritariamente con la otra subunidad del factor de transcripción. Además, la posición de R142, que se encuentra situado en la parte más próxima al DNA de la hélice de dimerización (α C), cambia dependiendo de la unión del 2OG a NtcA, de modo que solo en presencia de 2OG, R142 se encuentra situado a una distancia adecuada para interaccionar con el DNA, mientras que en ausencia de 2OG la orientación de la α C es tal que R142 se encuentra lejos de poder establecer contactos con el DNA. Por tanto, las interacciones fosfato-R142 parecen tener particular importancia y pueden ser un determinante clave en el aumento de afinidad de NtcA por su caja cuando se une el 2OG (ver Discusión).

En realidad, el reconocimiento específico de la caja NtcA se puede atribuir a sólo cuatro residuos, R186, V187, T190 y R191, ya que estos

residuos, que son invariables en NtcA, son los únicos que contactan directamente con bases del DNA (Tabla 3). Dos de ellos, R186 y R191 están implicados en puentes de hidrógeno directos con las bases del DNA mientras que V187 y T190 contribuyen con contactos de van der Waals.

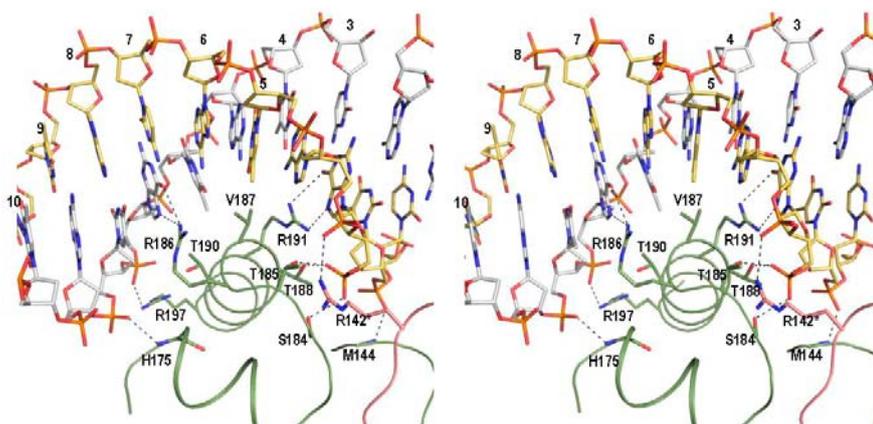


Figura 24. Representación estereoscópica de las interacciones entre NtcA y el DNA. NtcA está coloreada en verde excepto el residuo R142, perteneciente a α C, que al provenir de la otra subunidad se ha coloreado en rosa. Una cadena del DNA se representa con los átomos de carbono en amarillo y la otra con los átomos de carbono en gris.

En el caso de R186 (Fig. 24), el grupo guanidinio de esta arginina forma dos puentes de hidrógeno con los átomos O⁶ y N⁷ de un par G:C invariable de la caja NtcA, estando localizado a 7 pares de bases del eje de simetría en la caja NtcA (Fig. 21). R191 (Fig. 24) también forma dos puentes de hidrógeno con los átomos O⁶ y N⁷ de una guanina correspondiente a la caja de reconocimiento, aunque en este caso, la G reconocida pertenece a los ocho nucleótidos centrales que unen las dos partes del palíndromo, a 3 pares de bases del eje de simetría de la caja NtcA usada aquí, por tanto fuera de las bases constantes de la secuencia

Tabla 3. Contactos NtcA-DNA.

Los hemicomplejos se señalan en la Fig. 20 y las cadenas de DNA en la Fig. 21.

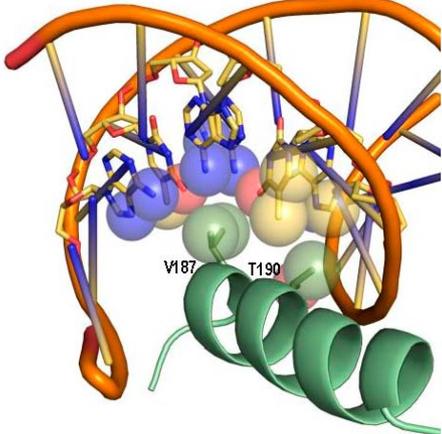
<i>Hemicomplejo A</i>						<i>Hemicomplejo B</i>					
Contactos con bases del DNA						Contactos con bases del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)	Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)
N7	G-7	E	NH2	R186	2.41	N7	G7'	C	NH2	R186	2.17
O6	G-7	E	NH2	R186	3.19	O6	G7'	C	NH1	R186	3.34
O6	G-7	E	NH1	R186	3.41	O6	G7'	C	NH2	R186	3.43
O4	T-6	E	NH2	R186	3.47	N7	G3	F	NH2	R191	3.18
N7	G-3'	D	NH2	R191	3.17	O6	G3	F	NH1	R191	3.39
O6	G-3'	D	NH1	R191	3.19						
<u>Interacciones con grupos metilo de timinas (< 4.2Å)</u>						<u>Interacciones con el grupo metilo de timinas (< 4.2Å)</u>					
C7	T-8	E	NH1	R186	4.01	C7	T8'	C	CZ	R186	4.05
C7	T-8	E	CD	R186	3.84	C7	T8'	C	NE	R186	3.77
C7	T-8	E	NE	R186	3.67	C7	T8'	C	CD	R186	3.81
C7	T-8	E	CZ	R186	3.78	C7	T6'	C	CB	T190	4.12
C7	T-6	E	CG2	T190	3.71	C7	T6'	C	CG2	T190	3.77
C7	T-2'	D	NH2	R191	4.19	C7	T6'	C	CG1	V187	4.13
Contactos con fosfatos del DNA						Contactos con fosfatos del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j	Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j
OP1	T-2'	D	N	M144	2.60	OP1	T2	F	N	M144	2.72
OP1	G-3'	D	OG	S184	2.95	OP1	G3	F	OG	S184	2.65
OP2	G-3'	D	OG1	T185	2.78	OP2	G3	F	OG1	T185	2.89
OP2	G-3'	D	OG1	T188	2.14	OP2	G3	F	OG1	T188	2.06
OP1	T-8	E	NZ	K207	3.00	OP2	A4	F	OG1	T185	3.43
OP2	T-8	E	N	H175	2.75	OP1	T8'	C	N	H175	2.92
OP2	G-7	E	NH2	R197	3.15						
OP1	G-7	E	NH2	R197	2.72						
OP1	A-13'	D	NH1	R29	3.63						
<i>Interacciones cruzadas</i>											
Contactos con fosfatos del DNA											
<u>Puentes de hidrógeno</u>											
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo (hemic)	Distancia i-j						
OP1	A4	F	NH1	R142 (A)	3.32						
OP2	A4	F	NH1	R142 (A)	2.54						
OP2	A-4'	D	NH1	R142 (B)	3.02						
OP1	A4	F	NH1	R142 (A)	3.32						

Figura 25. Detalle de las interacciones hidrofóbicas de V187 y T190 con con el grupo metilo de la timina 6 de la caja NtcA.

consenso para dicha caja (Fig. 21), posiblemente justificando en parte las distintas afinidades que presenta NtcA por diferentes promotores que presentan idéntica secuencia consenso pero variabilidad entre las ocho bases intermedias (Jiang *et al.*, 2000; Vázquez-Bermúdez *et al.*, 2002b). Por otra parte, las cadenas laterales de los residuos V187 y T190 atrapan como una pinza el grupo metilo de la timina perteneciente al par T:A situado a 6 pares de bases del eje de simetría del DNA (Fig. 25). V187 cumple la doble función de interactuar con un grupo metilo que permite diferenciar las cajas para CRP y para NtcA, y de ser el único residuo de entre los que interactúan con las bases del DNA que es invariante en las secuencias de las NtcAs de diferentes cianobacterias, pero que ha sido sustituido por un residuo de glutamato invariante en las CRPs.

Para testar la importancia de la Val187 en la interacción de NtcA con su caja, lo mutamos en NtcA a la secuencia de CRP, glutamato, y estudiamos la unión de NtcA a su caja mediante ensayos electroforéticos de retardo en gel (EMSA). El DNA usado en estos ensayos fue un fragmento de 159 pares de bases del promotor del gen *glnA* de *S. elongatus* conteniendo la caja NtcA de dicho promotor. Este DNA se produjo mediante amplificación por PCR usando DNA genómico de *S. elongatus* y los oligonucleótidos cebadores EMSA-F y EMSA-R (Tabla 1), marcados en su extremo 5' con fluoresceína de forma que las bandas de DNA podían ser detectadas en el gel mediante fluorescencia. La Fig. 26 ilustra la consecuencia de esta sustitución. Mientras que en la muestra con NtcA silvestre (WT) existe un fuerte retraso en la banda fluorescente, siendo la cantidad retrasada considerablemente mayor que sin 2OG, la proteína mutante fue incapaz de inducir dicho retardo ni en presencia ni en ausencia de 2OG. Estos resultados reflejan la importancia de la valina

187 de NtcA para su unión a la caja NtcA. También sugieren fuertemente que la unión de NtcA al DNA en ausencia de 2OG sucede igualmente a través de dicha caja, ya que este cambio de secuencia de la proteína abole igualmente la unión en ausencia de 2OG.

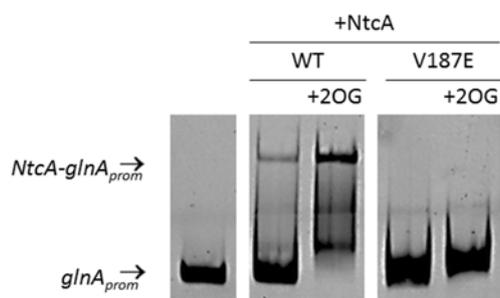


Figura 26. Ensayos de retardo en gel de DNA conteniendo el promotor para *glnA* en ausencia (izquierda) o en presencia (dos paneles de la derecha) de 100 ng de NtcA silvestre o con la mutación V187E, en ausencia o, cuando se indica, en presencia de 3.2 mM 2OG. Para más detalles ver Materiales y Métodos.

Estructuras del complejo PipX-NtcA-DNA

Culminado el objetivo de caracterizar las bases de la afinidad y la especificidad de NtcA por su DNA diana era factible plantearse cómo modificaba PipX dicha unión, y en particular ver si se cumplía la predicción (Llácer *et al.*, 2010) basada en la estructura de NtcA-PipX en ausencia de DNA en la que PipX no debe favorecer la unión de NtcA por un mecanismo de puente, en el que la molécula de PipX interacciona tanto con NtcA como con el DNA diana.

La obtención de cristales que contuvieran simultáneamente todos los componentes del complejo ternario fue difícil de conseguir, obteniéndose en la mayoría de los casos cristales que contenían únicamente NtcA o el complejo NtcA-PipX. Tampoco fue fácil conseguir que los cristales obtenidos del complejo PipX-NtcA-DNA difractaran a resolución casi atómica. Para intentar mejorar la resolución y la proporción de cristales que contuvieran todos los elementos del complejo, probamos tanto diferentes estequiometrias de los componentes como gran número de

dobles hélices de DNA (ensayos no mostrados). Estos DNAs diferían en sus extremos (cohesivos o romos), longitud, presencia o ausencia de puntos de rotura e incluso probamos secuencias de diferentes promotores de genes regulados por NtcA, como el promotor del gen *glnN*. Además, también probamos varios métodos de deshidratación controlada de los cristales para intentar reducir el contenido de solvente de los mismos y así aumentar la resolución. Tras muchos intentos, finalmente se obtuvieron dos tipos de cristales PipX-NtcA-DNA, ambos con hábito de bipirámide tetragonal, con rangos de resolución prometedores desde el punto de vista de la resolución de las estructuras (Fig. 19 y Tabla 2).

El DNA con el que se obtuvieron estos cristales está también basado en la caja NtcA del promotor del gen *glnA* de *S. elongatus*. Como en el caso precedente, se usó la primera mitad del consenso y las ocho bases que lo preceden, pero en este caso esta parte se repitió para lograr un palíndromo idealizado y continuo del sitio consenso de unión a NtcA (Fig. 27). Así, al contrario de lo que ocurría en el DNA utilizado para cristalizar el complejo NtcA-DNA, en este caso el DNA diseñado no presenta puntos de rotura pero sí extremos cohesivos mediante flecos G o C (Fig. 27), que resultaron esenciales para favorecer el empaquetado en los cristales de ambos complejos PipX-NtcA-DNA.

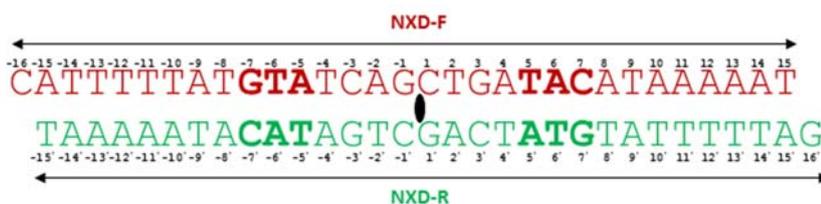


Figura 27. Oligonucleótidos usados para la cristalización del complejo PipX-NtcA-DNA. Obsérvese la C desapareada en un extremo y la G en el otro que han propiciado el encadenamiento de dos dobles hélices como se ha observado en el cristal I.

El primer cristal resuelto (llamado cristal II) difractó solamente hasta 4.3Å y además presentaba una patología importante en forma de “twinning perfecto”, que dificultó la resolución de la estructura del complejo (ver Materiales y Métodos).

Este cristal se resolvió finalmente mediante reemplazo molecular usando modelos basados en los complejos NtcA-DNA y NtcA-PipX, aplicando una simetría menor a la originalmente predicha para el grupo espacial debido a dicho “twinning” del cristal. Tras resolver la ambigüedad de empaquetado debido al “twinning”, se consiguió una solución consistente en un contenido de la unidad asimétrica de dos dobles hélices incurvadas de DNA, cada una unida a un dímero de NtcA que interacciona con el DNA esencialmente como en el complejo NtcA-DNA, con uno de los dímeros unido a dos moléculas de PipX y el otro dímero a una molécula de PipX, estando las tres moléculas de PipX unidas a las moléculas de NtcA como en el complejo NtcA-PipX (Llácer *et al.*, 2010), mediando las interacciones con el dominio Tudor y presentando la hélice C-terminal replegada sobre la hélice previa. (Fig. 28).

Es interesante que las moléculas de PipX que se encuentran entre ambos dímeros de NtcA (designadas en la Fig. 28 como PipX2 y PipX3) establecen algunas interacciones mutuas (aunque escasas) entre sus hélices 1 (la más proximal de las dos hélices) y también interaccionan con el dominio de unión al ligando de una de las dos subunidades del dímero de NtcA vecino, entre el conector de ambas hélices de PipX y el borde del barril beta de dicha subunidad de la molécula de NtcA vecina, así como mediante una glutamina y una lisina de la penúltima vuelta de la hélice 2

de PipX y un lazo conector entre dos hebras beta del barril beta de dicha subunidad del dímero de NtcA vecino.

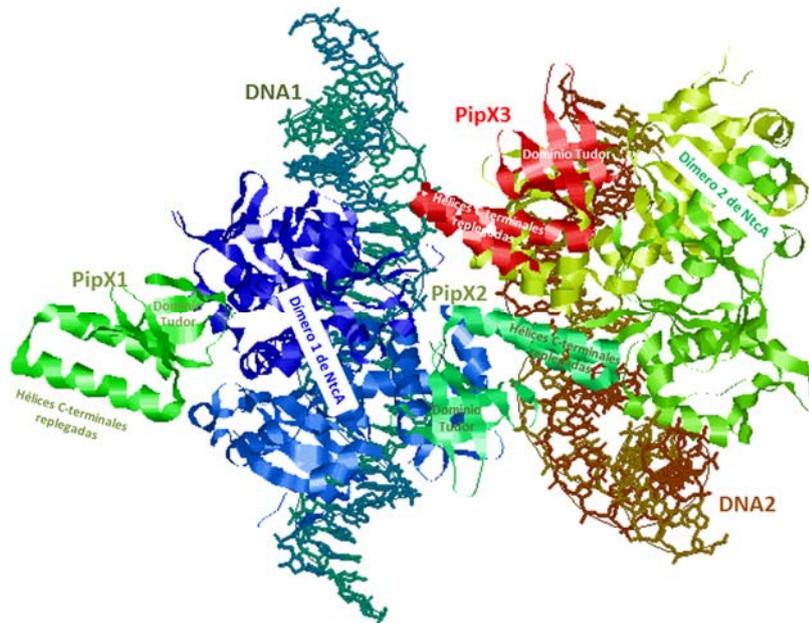


Figura 28. Contenido de la unidad asimétrica en el complejo PipX-NtcA-DNA del cristal II, mostrando los dos dímeros de NtcA (azul y verde) unidos a dos dobles hélices de DNA (verde y azul y rojo y marrón), con el dímero 1 unido a dos moléculas de PipX y el dímero 2 unido a una sola molécula de PipX, todas ellas presentando una forma de unión como en el complejo de NtcA y PipX (Llácer et al., 2010). Es de resaltar que las hélices C-terminales de las moléculas de PipX 2 y 3 se encuentran muy próximas entre sí y al dímero de NtcA vecino, formando incluso algunas interacciones con el EBD de una subunidad.

La estructura de este complejo prueba que PipX no interacciona con el DNA al que se une el dímero de NtcA al que PipX está anclada a través de su dominio Tudor. Sin embargo, cada conector que une ambas hélices C-terminales en las moléculas de PipX2 y PipX3 está muy próximo a uno u otro de los dos extremos de la doble hélice de DNA del otro complejo NtcA-DNA presente en la unidad asimétrica, haciendo plausible suponer que si el dúplex de DNA hubiera sido más largo habría contactado con la molécula de PipX del complejo vecino. Desconocemos el potencial *in vivo* de estas interacciones mediadas por PipX, pero cabría la posibilidad

de que interacciones similares tuvieran lugar *in vivo* entre los complejos unidos a dos cajas NtcA próximas en el espacio.

En cuanto al sitio vacío de PipX (sitio para un PipX4 no unido) en el dímero 2 de NtcA, realmente no está vacío, al ubicarse allí, aparentemente de forma aberrante, el conector de las hélices de PipX1 de NtcA (Figura 29), de nuevo planteando la cuestión de si se trata meramente de una interacción no fisiológica que se da en el cristal, o si dicha interacción podría tener una contrapartida *in vivo*.

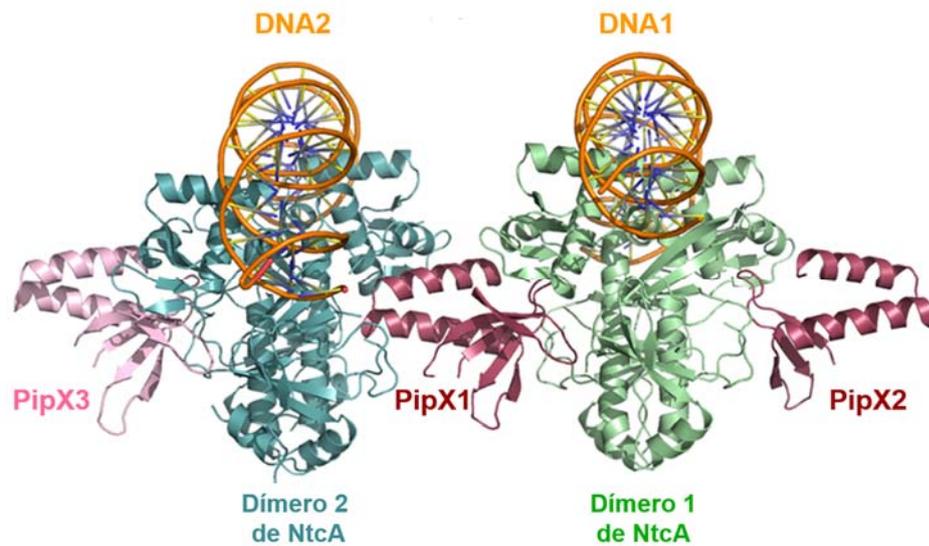


Figura 29. Visión alternativa de la unidad asimétrica del cristal II, para mostrar cómo las dos hélices de PipX1 interactúan de forma aberrante con el sitio para PipX4 del dímero 2 de NtcA.

En todo caso, en este cristal II hay dos poblaciones diferentes de complejos PipX-NtcA-DNA una con dos moléculas de PipX unidas canónicamente y otra con un sitio para PipX vacío ocupado aberrantemente por la parte conectora de las hélices C-terminales de una de las dos moléculas de PipX que se une por su dominio Tudor al otro

tipo de complejo. No queda claro si estas dos poblaciones distintas difieren en su eficiencia a la hora de activar la transcripción, pero la existencia de un complejo con una sola PipX unida tendría sentido si solo una molécula de PipX fuera la que interaccionara con la RNA polimerasa, como sugiere el modelo de Llácer (Llácer *et al.*, 2010) basado en complejo CRP-DNA- α CTD (Benoff *et al.*, 2002).

Hay que atribuir no pequeña parte del éxito en la determinación de la estructura de los complejos presentes en el cristal II, a pesar de la limitada resolución de los datos cristalográficos para ese cristal, a la similitud entre el complejo ternario PipX-NtcA-DNA y las estructuras de los complejos NtcA-DNA y NtcA-PipX, lo que permitió encontrar buenas fases cristalográficas, así como a la muy buena calidad de los mapas de densidad electrónica, con valores de $R_{\text{factor}}/R_{\text{free}}$ razonablemente buenos (Tabla 2).

Cada proteína y el DNA parecen presentar una estructura idéntica a la que presentan en los complejos NtcA-DNA y NtcA-PipX (no mostrado). Solo algunos choques entre lazos expuestos a solvente (acompañada por una falta de densidad electrónica en estos lazos, no indicadas aquí dada la baja resolución) se observan en la solución de reemplazo inicial, atribuibles a pequeñas diferencias en esas regiones.

Hemos utilizado el modelo final construido en el cristal II del complejo de NtcA unido a DNA y a dos moléculas de PipX unidos canónicamente para encontrar mediante reemplazo molecular las fases para los datos de difracción obtenidos con el cristal I a una resolución bastante mejor, de hasta 3.8 Å. Así, encontramos en la unidad asimétrica de este último cristal una doble hélice de DNA comprendiendo dos cajas NtcA unidas

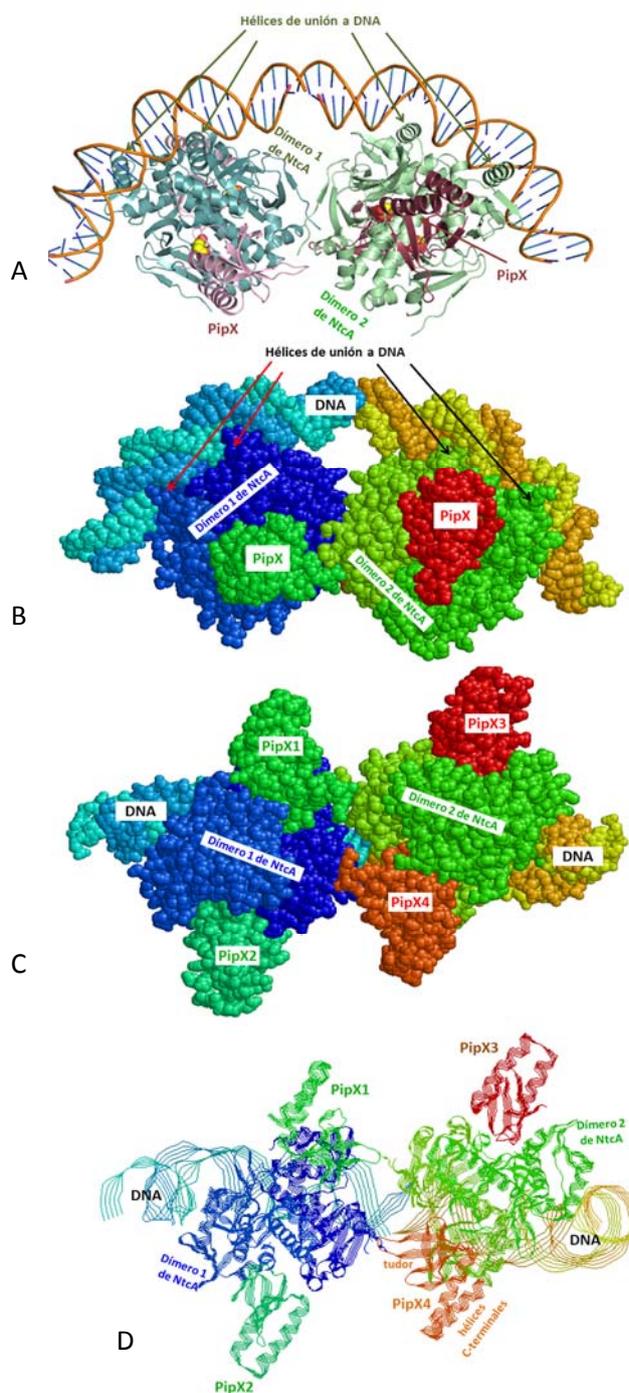


Figura 30. El complejo PipX-NtcA-DNA encontrado en el cristal I. (A-C) Vistas del complejo en dos orientaciones diferentes en formato “cartoon” (A), en visión de esferas de Van der Waals (B y C) y en visión de cintas transparentes (D). Los paneles (A) y (B) presentan la misma orientación, con el DNA hacia arriba. Los paneles (C) y (D) presentan una segunda orientación correspondiente a un giro de 90° alrededor del eje de las X con respecto a la orientación en (A y B), de modo que el DNA está al fondo. Los letreros y las diferencias de color identifican los elementos más relevantes de la arquitectura del complejo. Cada una de las cuatro hebras de DNA también está en un color diferente.

por un par C:G formado por las bases hibridadas del fleco terminal de simple cadena de una base, componiendo un complejo con un DNA curvo que contiene cinco vueltas del surco mayor, de las que las dos más periféricas de cada lado están ocupadas por las hélices de unión al DNA de un homodímero de NtcA en el que cada una de sus subunidades está unida canónicamente a un monómero de PipX (Fig. 30). A pesar de la limitada resolución (3.8 Å, Tabla 2), la calidad del mapa fue tan buena (Fig. 31) que hizo posible la identificación de la mayor parte de las cadenas laterales de las proteínas, especialmente en los puntos de contacto entre los diferentes componentes del complejo.

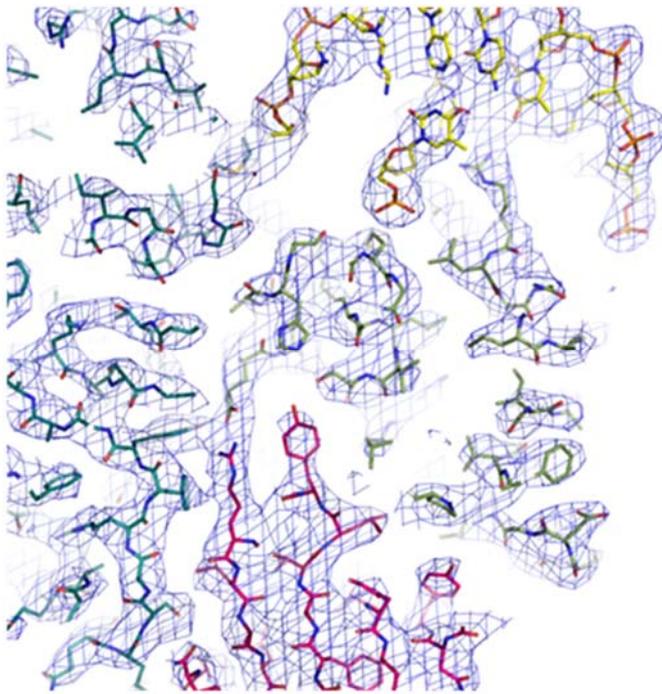


Figura 31. Muestra del mapa de densidad electrónica 2Fo-Fc en el cristal I, a 1σ y 3.8Å de resolución. La región representada incluye elementos de PipX (modelo molecular con los átomos de carbono de color rosa), de NtcA (átomos de C de una subunidad en azul y de la otra en verde) y de DNA (átomos de C en amarillo).

Los dos complejos presentes en la unidad asimétrica son muy similares entre sí (rmsd de 0.97 Å para 594 C^α superpuestos), así como también son

muy similares al complejo PipX-NtcA-DNA completo del cristal II, con una rmsd para la superposición de ambos de $1.00 \pm 0.09 \text{ \AA}$ (para $509 \pm 1 C^\alpha$) y $1.09 \pm 0.08 \text{ \AA}$ (para $594 C^\alpha$). Los componentes individuales son también muy similares, siendo, por ejemplo, el rmsd de la superposición entre todas las moléculas de PipX en los dos complejos PipX-NtcA-DNA y en el complejo NtcA-PipX (Llácer et al., 2010) de $0.62 \pm 0.09 \text{ \AA}$ (para $85 \pm 1 C^\alpha$). De esta forma, la interacción entre NtcA y PipX es prácticamente idéntica a la que se da en el complejo NtcA-PipX (Llácer et al., 2010), y lo mismo ocurre con las interacciones entre NtcA y el DNA (Tabla 4), que son muy similares a las descritas previamente en complejo NtcA-DNA (Tabla 3). Sin embargo, en los complejos ternarios PipX-NtcA-DNA se pueden observar algunas interacciones entre NtcA y el DNA que no se daban en el complejo NtcA--DNA (Fig. 21 y Tabla 4).

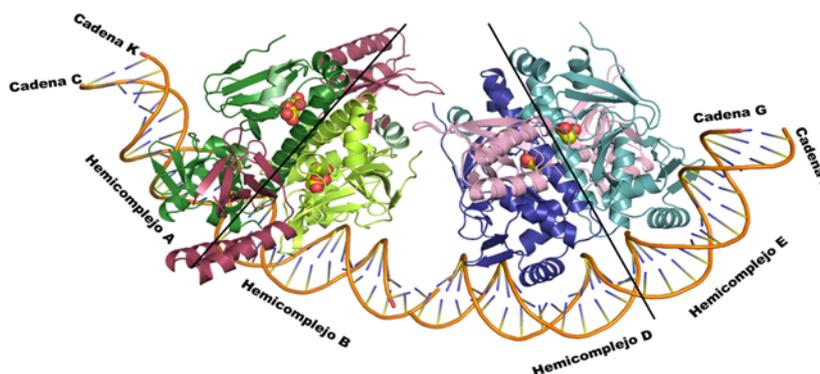


Fig. 32. Elementos de la unidad asimétrica del cristal I, identificados como en la Tabla 4 y en el archivo de coordenadas para el PDB. Los hemicomplejos reciben el identificador alfabético de la cadena de NtcA que se encuentra en cada una

Resultados

Tabla 4. Contactos NtcA-DNA en el complejo PipX-NtcA-DNA (cristal I)

Los hemicomplejos A, B, D y E y las cadenas de DNA se señalan en la Fig 32.

<i>Hemicomplejo A</i>						<i>Hemicomplejo B</i>					
Contactos con bases del DNA						Contactos con bases del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)	Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)
N7	G-3'	C	NH2	R191	3.12	N7	G7'	C	NH2	R186	2.79
O6	G-3'	C	NH1	R191	3.51	O4	T4'	C	NH1	R191	3.36
N7	G-7	K	NH2	R186	2.58	N7	G3	K	NH2	R191	2.92
O6	G-7	K	NH1	R186	3.33						
<u>Interacciones con grupos metilo de timinas (< 4.2Å)</u>						<u>Interacciones con el grupo metilo de timinas (< 4.2Å)</u>					
C7	T-2'	C	CZ	R191	3.56	C7	T8'	C	CZ	R186	3.70
C7	T-2'	C	NH1	R191	3.94	C7	T8'	C	NH2	R186	3.59
C7	T-2'	C	NE	R191	3.73	C7	T8'	C	NE	R186	3.54
C7	T-2'	C	NH2	R191	3.73	C7	T5	K	CG2	V187	3.63
C7	T-5'	C	CG2	V187	3.76						
C7	T-8	K	NE	R186	3.48						
C7	T-8	K	CZ	R186	3.68						
C7	T-8	K	NH2	R186	3.69						
Contactos con fosfatos del DNA						Contactos con fosfatos del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j	Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j
OP1	T-2'	C	N	M144	3.16	OP1	T8'	C	N	H175	3.08
OP2	G-3'	C	OG	S184	2.65	OP1	G7'	C	NH1	R197	3.13
OP2	G-3'	C	N	T185	3.05	OP1	G7'	C	NH2	R197	2.71
OP2	G-3'	C	OG1	T185	2.40	OP1	T2	K	N	M144	3.26
OP2	G-3'	C	OG1	T188	2.53	OP1	G3	K	OG	S184	3.06
OP2	T-8	K	N	H175	2.92	OP2	G3	K	OG1	T188	2.51
OP1	G-7	K	NH2	R197	3.41	OP2	G3	K	OG	S184	2.91
OP2	G-7	K	NH2	R197	3.07	OP2	G3	K	N	T185	3.33
<i>Interacciones cruzadas</i>											
Contactos con fosfatos del DNA											
<u>Puentes de hidrógeno</u>											
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo (hemic)	Distancia i-j						
OP2	A-4'	C	NE	R142	3.36						

Tabla 4 (continuación).

<i>Hemicomplejo C</i>						<i>Hemicomplejo D</i>					
Contactos con bases del DNA						Contactos con bases del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)	Átomo i	Base	Cadena DNA	Átomo j	Residuo	Distancia i-j (Å)
N7	G-3'	G	NH2	R191	2.75	N7	G7'	G	NH2	R186	2.62
O6	G-3'	G	NH1	R191	3.28	O6	G7'	G	NH1	R186	3.39
N7	G-7	I	NH2	R186	3.38	N7	G3	I	NH2	R191	2.97
O4	T-4	I	NH1	R191	3.44	O6	G3	I	NH1	R191	3.39
<u>Interacciones con grupos metilo de timinas (< 4.2Å)</u>						<u>Interacciones con el grupo metilo de timinas (< 4.2Å)</u>					
C7	T-2'	G	CZ	R191	4.20	C7	T8'	G	CZ	R186	3.64
C7	T-5'	G	CG2	V187	3.46	C7	T8'	G	NH2	R186	3.65
C7	T-8	I	NE	R186	3.01	C7	T8'	G	NE	R186	3.48
C7	T-8	I	CZ	R186	3.68	C7	T6'	G	CG2	T190	3.43
C7	T-8	I	NH2	R186	3.98	C7	T5	I	CG2	V187	3.74
C7	T-8	I	CD	R186	3.48	C7	T2	I	NE	R191	4.00
						C7	T2	I	CZ	R191	3.83
						C7	T2	I	NH2	R191	4.06
Contactos con fosfatos del DNA						Contactos con fosfatos del DNA					
<u>Puentes de hidrógeno</u>						<u>Puentes de hidrógeno</u>					
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j	Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo	Distancia i-j
OP1	T-2'	G	N	M144	3.27	OP1	T-2'	G	N	M144	3.27
OP1	G-3'	G	OG	S184	3.11	OP1	G-3'	G	OG	S184	3.11
OP2	G-3'	G	OG	S184	2.92	OP2	G-3'	G	OG	S184	2.92
OP2	G-3'	G	OG1	T185	2.56	OP2	G-3'	G	OG1	T185	2.56
OP2	G-3'	G	OG1	T188	2.44	OP2	G-3'	G	OG1	T188	2.44
OP2	G-3'	G	N	T185	3.39	OP2	G-3'	G	N	T185	3.39
OP2	T-8	I	N	H175	2.80	OP2	T-8	I	N	H175	2.80
OP1	G-7	I	NH2	R197	2.69	OP1	G-7	I	NH2	R197	2.69
<i>Interacciones cruzadas</i>											
Contactos con fosfatos del DNA											
<u>Puentes de hidrógeno</u>											
Átomo i	Fosfato de base	Cadena DNA	Átomo j	Residuo (hemic)	Distancia i-j						
OP2	A-4'	G	NE	R142	3.45						
OP2	A4	I	NH1	R142	3.43						

Hay que hacer constar que en la estructura presente en la unidad asimétrica del cristal I también se observan contactos entre los dos complejos, si bien dichos contactos son pequeños, involucrando de un lado una vuelta del dominio Tudor de PipX y del otro lado una pequeña parte del dominio de unión de efector de una subunidad de NtcA. Por tanto, de nuevo PipX está involucrada en mediar las interacciones inter-

complejos, si bien el elemento utilizado para las interacciones difiere de los usados en los complejos del cristal II.

También hay que hacer notar que las dos hélices C-terminales de los complejos encontrados en ambos cristales I y II se encuentran en todos los casos replegadas (forma flexionada), lo que, como se verá en un capítulo ulterior tiene interés en el contexto de que en el complejo de PII con PipX algunas de las hélices C-terminales se encontraban extendidas, habiendo perdido su relación mutua (Llácer *et al.*, 2010).

Nuevas estructuras de NtcA sin 2OG unido

Este laboratorio publicó previamente (Llácer *et al.*, 2010) una estructura de NtcA de *S. elongatus* en forma “inactiva” (denominada aquí forma Apo B) aunque el sitio para ligando contenía 2OG unido aberrantemente al mismo. Otros (Zhao *et al.*, 2010b) publicaron una estructura de NtcA de *Anabaena sp.* (denominada aquí forma Apo A) sin 2OG, en forma también “inactiva”. Ambas formas diferían importantemente en sus estructuras (rmsd de 3.67 Å para 306 C^α; para la superposición de los dímeros) aunque es cierto que consideraciones de distancia y orientación de sus hélices de unión a DNA corroboraban la incapacidad de unirse a DNA de ambas formas en las conformaciones publicadas de las mismas. Por el contrario, las estructuras de las formas “activas” de NtcA de *S. elongatus* y de *Anabaena* unidas a 2OG fueron virtualmente idénticas (rmsd para la superposición de los homodímeros, 0.56 Å para 390 átomos de C^α), siendo también altamente similares a las formas activas publicadas de otros reguladores transcripcionales de la familia CRP (Agari *et al.*, 2008; Borjigin *et al.*, 2007; Chin *et al.*, 2010; Eiting *et al.*, 2005; Levy *et al.*, 2008; Reddy *et al.*, 2009; Weber & Steitz, 1987). Como la forma B

contenía 2OG, era esencial aclarar si esta forma se correspondía a una forma genuinamente “inactiva” que es posible conseguir en ausencia de 2OG o si era una forma intermedia.

En esta tesis describimos tres estructuras nuevas de NtcA de *S. elongatus* sin 2OG, y por tanto genuinamente “inactivas”. Dos de ellas presentan ligeras variantes de la conformación A y la tercera de la B (Fig. 33). Estas estructuras indican la existencia de diversas formas inactivas de NtcA para la misma especie, y también confirman que la forma B descrita previamente es una conformación genuinamente “inactiva”.

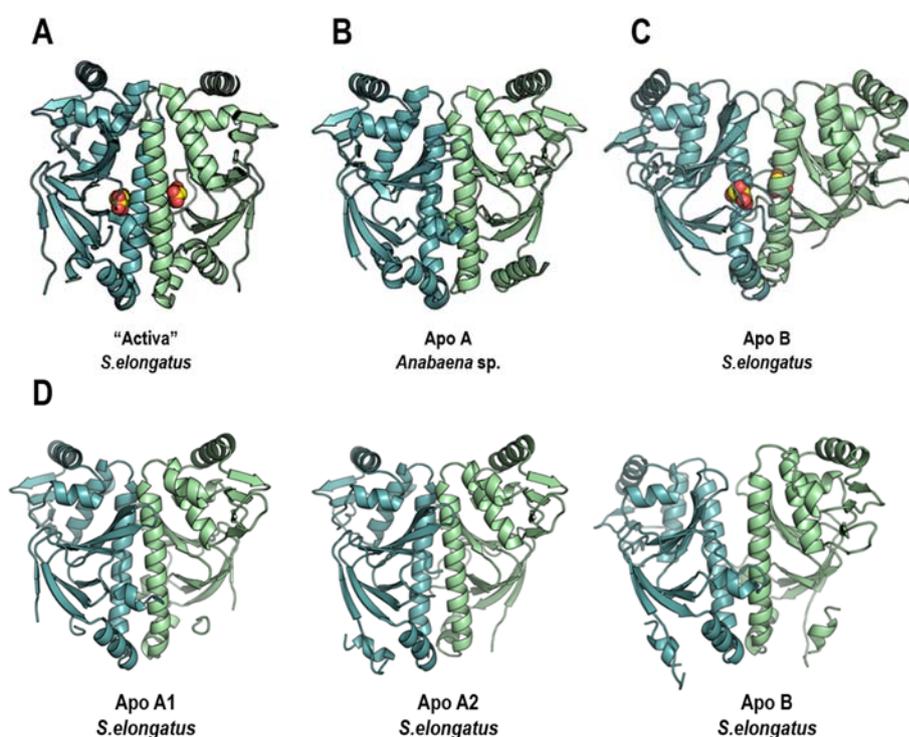


Figura 33. Representación en cintas de las estructuras de NtcA en sus formas “activa” e “inactivas” incluyendo las descritas en esta tesis. Una subunidad de NtcA se muestra en color azul y la otra en color verde. El 2OG se muestra en representación de esferas. (A) Estructura de NtcA en su forma activa en la cianobacteria *S. elongatus* (PDB 2XHK) (Llácer et al., 2010). (B, C) Estructuras de NtcA apo publicadas en la cianobacterias (B) *Anabaena* (PDB 3LA7) (Zhao et al., 2010b) y (C) *S. elongatus*. (PDB 2XGX) (Llácer et al., 2010). (D) Formas apo de NtcA en *S. elongatus* descritas en esta tesis.

Resultados

Tabla 5. Datos de difracción de rayos X y estadísticas de afinado de las estructuras de NtcA en sus formas apo.

	<i>NtcA apoA1</i>	<i>NtcA apoA2</i>	<i>NtcA apoB</i>
Sincrotón (línea)	ESRF (ID23-1)	ESRF (BM16)	ESRF (BM14)
Longitud de onda (Å)	0.9769	0.9797	0.9762
Grupo espacial	C222 ₁	P2 ₁ 2 ₁ 2 ₁	P6 ₄ 22
Celda unidad: a, b, c (Å)	95.2, 95.5, 160.8	66.5, 88.9, 99.7	96.4, 96.4, 126.9
Resolución (Å)	80.4-2.65 (2.79-2.65)	66.3-2.85 (3.00-2.85)	83.5-3.33 (3.51-3.33)
Reflexiones totales/únicas	73,746/21,137	65,283/14,209	36,587 /5,530
Complejidad (%)	98.0 (98.0) ^a	99.0 (95.7) ^a	100 (100) ^a
Multiplicidad	3.5 (3.3) ^a	4.6 (3.6) ^a	6.6 (6.9) ^a
I/σ	5.8 (1.8) ^a	12.1 (2.3) ^a	4.0 (1.8) ^a
R _{sym} ^b (%)	5.7 (40.8) ^a	5.6 (34.1) ^a	13.6 (41.9) ^a
R _{factor} ^c /R _{free} ^d (%)	24.7/25.7	22.5/27.0	24.9/ 28.5
rmsd longitud enlace (Å)	0.009	0.008	0.007
rmsd ángulo enlace (°)	1.25	1.10	1.03
Moléculas o átomos			
Cadenas de NtcA	2	2	1
átomos de proteína	3098	3101	1652
2-oxoglutarato	0	0	0
átomos de DNA	0	0	0
moléculas de agua	12	33	0
Factores B (medias, Å ²)			
átomos de proteína	96.1	73.8	61.5
2-oxoglutarato	0	0	0
átomos de DNA	0	0	0
moléculas de agua	42.4	26.1	0
Repres. Ramachandran			
Favorables (%)	96.2	93.1	91.5
Permitidos (%)	3.8	6.9	8.5
No permitidos (%)	0	0	0

^a Los datos entre paréntesis son los de la celda de mayor resolución

^bR_{sym} = $\sum I - \langle I \rangle / \sum I$, donde I es la intensidad observada e $\langle I \rangle$ la intensidad media. ^c R_{factor} = $\sum h ||F_{obs}| - |F_{calc}|| / \sum h |F_{obs}|$, donde $|F_{obs}|$ y $|F_{calc}|$ son los factores de estructura observados y calculados para todas las reflexiones. ^dR_{free}, igual que R_{factor}, pero calculado con el 5% de los datos omitidos del afinamiento.

Nuestro cristal de la forma apo A1 creció en condiciones similares a las descritas en el cristal de la forma apo A de *Anabaena* (Zhao *et al.*, 2010b) y el grupo espacial fue el mismo también: C222₁ (Tabla 5). Este cristal contiene un dímero de NtcA en la unidad asimétrica, el cual es muy similar a la forma apo A de *Anabaena* (rmsd para la superposición de los dímeros, 1.11 ± 0.01 Å para 387 átomos de C^α). El cristal apo A2 creció en condiciones distintas (ver Materiales y Métodos) a los anteriores cristales y su grupo espacial fue también distinto, P2₁2₁2₁ (Tabla 5).

Este cristal contiene un dímero en la unidad asimétrica el cual, aunque con algunas pequeñas diferencias, se parece bastante a las estructura apo A1 y a la estructura apo A de *Anabaena*, con un rmsd para la superposición de los dímeros de 1.19 ± 0.01 Å (para 379 C^α) y 1.32 Å (para 373 C^α), respectivamente. Por otra parte, el cristal de la forma apo B (Tabla 5) contiene solamente una subunidad en la unidad asimétrica. El correspondiente dímero generado aplicando la simetría cristalográfica se asemeja bastante a la conformación apo B de *S. elongatus*, (rmsd para la superposición de los dímeros, 1.30 Å para 390 C^α). En resumen, en ausencia de 2OG es posible generar dímeros con diferentes conformaciones, todas ellas inadecuadas para su unión a DNA (ver Discusión). Como se verá más adelante, esta situación es bastante general, al menos entre los reguladores transcripcionales de la familia CRP.

Resultados

Capítulo 2

Análisis de SPR de la unión a promotores de NtcA y CRP

El contexto: promotores de *Synechocystis* para NtcA y CRP.

Dada la similitud de CRP y NtcA puestas en evidencia por los estudios estructurales (ver capítulo anterior) y también la gran similitud de las secuencias consenso de DNA reconocidas por NtcA y por CRP (cajas NtcA y CRP, respectivamente) (Berg & von Hippel, 1988; Luque *et al.*, 1994., Herrero *et al.*, 2001; Jiang *et al.*, 2000; Omagari *et al.*, 2008) (Fig. 34, tabla inferior), es concebible la posibilidad de regulación cruzada entre NtcA y CRP, de modo que un regulador transcripcional quizá pudiera interactuar con promotores para el otro factor de transcripción. Esta cuestión es particularmente pertinente para algunas cianobacterias (Xu & Su, 2009) como *Synechocystis* PCC6803 (en adelante *Synechocystis*), en la que existe el gen para una proteína CRP (llamada *Synechocystis* CRP, SYCRP1) y esta proteína se expresa, regulando transcripcionalmente la expresión de un cierto número de genes tales como *cccS* en forma dependiente de cAMP (Yoshimura *et al.*, 2000, 2002b; Omagari *et al.*, 2008). Incluso en cianobacterias sin gen codificante de CRP (el caso de *S. elongatus*) se han detectado varias cajas potenciales CRP en su genoma (Xu & Su, 2009), cajas de función incierta, pero que podrían ser objeto de regulación por NtcA.

Con la finalidad primaria de explorar esta posibilidad de regulación cruzada (Fig. 34) hemos realizado ensayos de unión de NtcA y de CRP a cajas homónimas y cruzadas usando la tecnología de la resonancia de plasmón (SPR), explotando la gran experiencia en este tipo de estudios y la amplia disponibilidad instrumental del laboratorio del Prof. Forchhammer (Tübingen, Alemania), con quien colaboramos para estos estudios, desplazándome a su laboratorio durante 3 meses para realizar

dichos experimentos usando la metodología puesta a punto en el mismo para estudios con DNA (Espinosa *et al.*, 2006, 2007).

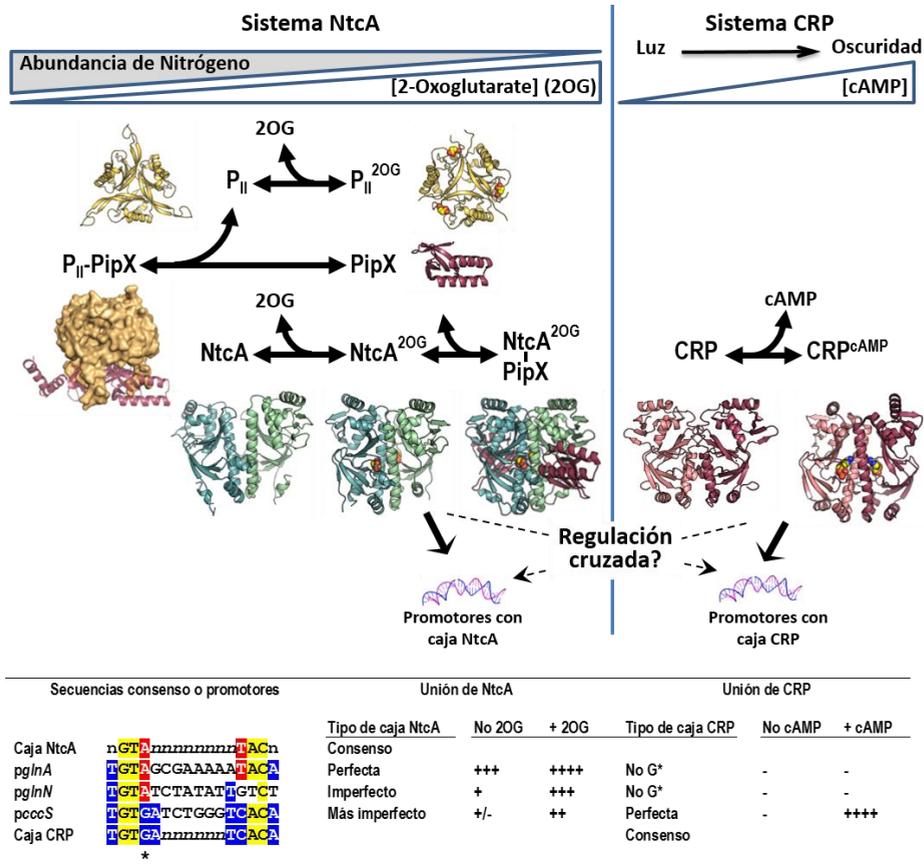


Figura 34. Sistemas controlados por NtcA y CRP. *Izquierda*, complejos formados por NtcA y su maquinaria controladora, que incluye a 2OG, PipX y PII. Las estructuras ilustradas son las determinadas para proteínas de *S. elongatus* (Fokina *et al.*, 2010; Llácer *et al.*, 2010; Xu *et al.*, 2003). *Derecha*, activación por cAMP de CRP. Las estructuras de CRP representadas son las de *E. coli* (Popovych *et al.*, 2009; Schultz *et al.*, 1991). La relación del sistema CRP con la iluminación se basa en la observación de que los niveles de cAMP en una cianobacteria variaron con cambios de iluminación a oscuridad. La tabla inferior da las cajas consenso para las uniones de NtcA y CRP (Berg & von Hippel, 1988; Herrero *et al.*, 2001), así como las versiones reales encontradas en los varios fragmentos de promotor utilizados aquí, sombreando en amarillo, rojo y azul, respectivamente, los nucleótidos conservados en ambas cajas (NtcA y CRP), solo en el consenso para NtcA o solo en el consenso para CRP. Una *n* minúscula denota cualquier nucleótido. El asterisco debajo del consenso para CRP marca una G considerada fundamental para la unión de CRP (Omagari *et al.*, 2004). La columna del tipo de promotor para NtcA indica si la caja es perfecta o menos perfecta, dependiendo del acuerdo con el consenso correspondiente. La tabla también resume la capacidad de NtcA y CRP para unirse a uno u otro promotor según nuestros datos.

Además, estos experimentos ponían también a nuestro alcance la posibilidad de realizar estudios funcionales exploratorios de por qué algunos genes del regulón de NtcA son más o menos sensibles a 2OG o a PipX, lo que quizá podría proporcionarnos claves funcionales que junto con los resultados estructurales permitieran avanzar en el conocimiento molecular del sistema de expresión regulado por NtcA.

En estos experimentos hemos utilizado como DNA inmovilizado en el chip fragmentos de promotor de tres genes de la cianobacteria *Synechocystis sp.* Dos de estos genes, *glnA* y *glnN*, codifican para las glutamina sintetasas I y II (Mérida *et al.*, 1992; Reyes & Florencio, 1994) y poseen cada uno de ellos, en el fragmento de promotor amplificado en estos estudios para su inmovilización en el chip mediante un extremo biotinilado, una caja NtcA, canónica en el caso de *glnA* y no canónica e imperfecta para *glnN* (Reyes *et al.*, 1997) (Fig. 34 tabla inferior). El tercer gen de *Synechocystis* que hemos utilizado, *cccS*, es una diana bien caracterizada de regulación por SYCCRP1 (Yoshimura *et al.*, 2002b, 2010) (Fig. 34 tabla inferior). También hemos producido recombinantemente para testar su unión respectiva a esos promotores los factores de transcripción correspondientes de *Synechocystis*, NtcA y SYCRP1 (abreviado aquí desde ahora CRP), al igual que la proteína coactivadora de NtcA, PipX, del mismo organismo.

Con estos reactivos y siguiendo la metodología perfeccionada en el laboratorio del Prof. Forchhammer y descrita en Materiales y Métodos, hemos investigado en los presentes estudios la unión de NtcA y de CRP a los tres fragmentos de promotores, en ausencia o en presencia de sus

respectivos efectores 2OG y cAMP, o de la proteína co-activadora de NtcA, PipX. Los resultados de estos estudios se describen a continuación.

Unión de NtcA y/o CRP a los promotores de *glnA*, *glnN* y *cccS*

Con la finalidad de proporcionar una percepción directa de los resultados experimentales obtenidos, se ilustran en los paneles A-E de la figura 35 los trazados obtenidos para experimentos con el fragmento del promotor para *glnN* (*pglnN*). En el panel A puede verse como, en presencia de una alta concentración de 2OG (5 mM), la señal aumenta con la concentración de NtcA, el regulador considerado natural para dicho promotor. La señal aumenta con la concentración de NtcA, indicando unión de NtcA al promotor, hasta alcanzar un nivel máximo, que indica la saturación de los sitios existentes pegados al chip para NtcA. El panel B refleja los mismos experimentos en ausencia de 2OG, observándose cómo NtcA es también capaz de unirse a este promotor aún en ausencia de 2OG, si bien y a partir de la comparación de los trazados puede observarse que las concentraciones de NtcA requeridas para alcanzar la saturación son mucho más elevadas que en presencia de 2OG, y que la señal máxima alcanzada es más baja, quizá indicando un modo distinto de unión al de la forma con 2OG. El panel C ilustra cómo CRP no se une a este fragmento de promotor esté o no presente el activador de CRP, cAMP. Este comportamiento es muy diferente al que se observa cuando el fragmento de promotor utilizado es el de *cccS* (*pcccS*, panel F), pues en este caso, en el que también estaba presente a alta concentración cAMP, existe una clara señal de unión de CRP, bastante similar a la obtenida con *pglnN* cuando este se saturó con NtcA en presencia de 2OG (panel A). El panel D muestra cómo es posible utilizar este sistema para titular la afinidad por el efector, al mostrar cómo se incrementa la señal obtenida para *pglnN*

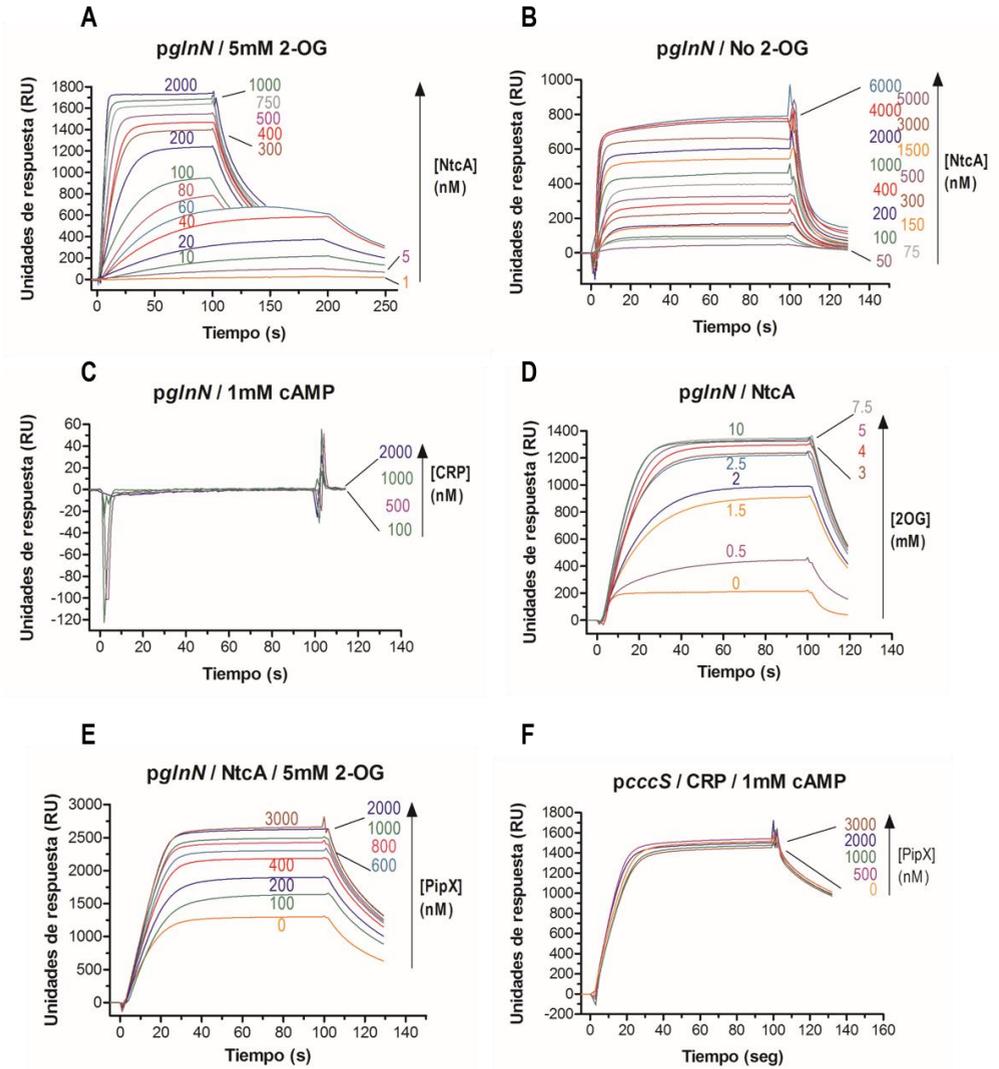


Figura 35. Ejemplos ilustrativos de trazas de SPR. RU, unidades relativas. El título del gráfico sobre cada panel indica el DNA dúplex unido al chip y el ligando presente a concentración fija. La flecha a la parte derecha indica el ligando que se titula y sus concentraciones. (A) y (B) Titulaciones de la unión de NtcA a *glnN* en presencia (A) o en ausencia (B) de 5 mM 2OG. En el panel A, el volumen de las inyecciones se duplicó para las concentraciones más bajas de NtcA (1-60 nM) para permitir alcanzar el equilibrio. (C) CRP no se une a *pglnN* en presencia de 1 mM cAMP. (D) Influencia de las concentraciones de 2OG en la señal producida por 400 nM NtcA sobre el chip para *pglnN*. (E) Influencia de PipX en la señal provocada por 400 nM de NtcA sobre el chip *pglnN* en presencia de 5 mM 2OG. (F) Falta de efecto de PipX sobre la señal provocada por 400 nM de CRP usando el chip *pcccS* en presencia de 1mM de cAMP

con una concentración fija de NtcA dependiendo de la concentración de 2OG, mostrando cómo existe saturación por el efector a concentraciones del mismo en exceso de 3 mM.

Finalmente, el panel E proporciona la oportunidad de observar la unión de PipX sobre la de NtcA a *pglnN* en presencia de 2OG, demostrando que, como era esperable, el proceso se satura a partir de una concentración determinada de PipX. Un control negativo para este efecto lo proporciona el panel F, en el que no se observa aumento alguno de la señal cuando se añade PipX si el promotor es *pcccS* y el factor utilizado es CRP activado por cAMP, como es esperable si PipX no se une a CRP.

El resto de las figuras no muestran resultados directos como los ilustrados en la figura 35 sino análisis cuantitativos sobre los resultados primarios. Así, en la figura 36 se observa que en la presencia respectiva de 2OG y de cAMP, NtcA y CRP produjeron señales de resonancia máxima (señal a saturación del regulador transcripcional correspondiente) que fueron similares cuando se unieron a sus promotores afines, *pglnN* y *pglnA* para NtcA, y *pcccS* para CRP (Fig. 36, paneles A, B, C, curvas superiores). La unión de CRP a *pcccS* exhibió un requerimiento absoluto de cAMP (Fig. 36, panel A). Por el contrario, NtcA exhibió unión sustancial en ausencia de su efector, 2OG, tanto para *pglnN* como para *pglnA* (Fig. 35 B y 36 B y C), si bien en ausencia de 2OG la afinidad aparente de NtcA para ambos promotores se redujo un orden de magnitud y la señal máxima se redujo ~50% y ~30%, respectivamente. De acuerdo con la presencia en *pglnA* de una caja NtcA perfecta mientras que en *pglnN* solo la mitad de la caja NtcA está conservada (Reyes *et al.*, 1997; Aldehni & Forchhammer, 2006) (Fig. 34, tabla inferior), la afinidad de NtcA fue un orden de

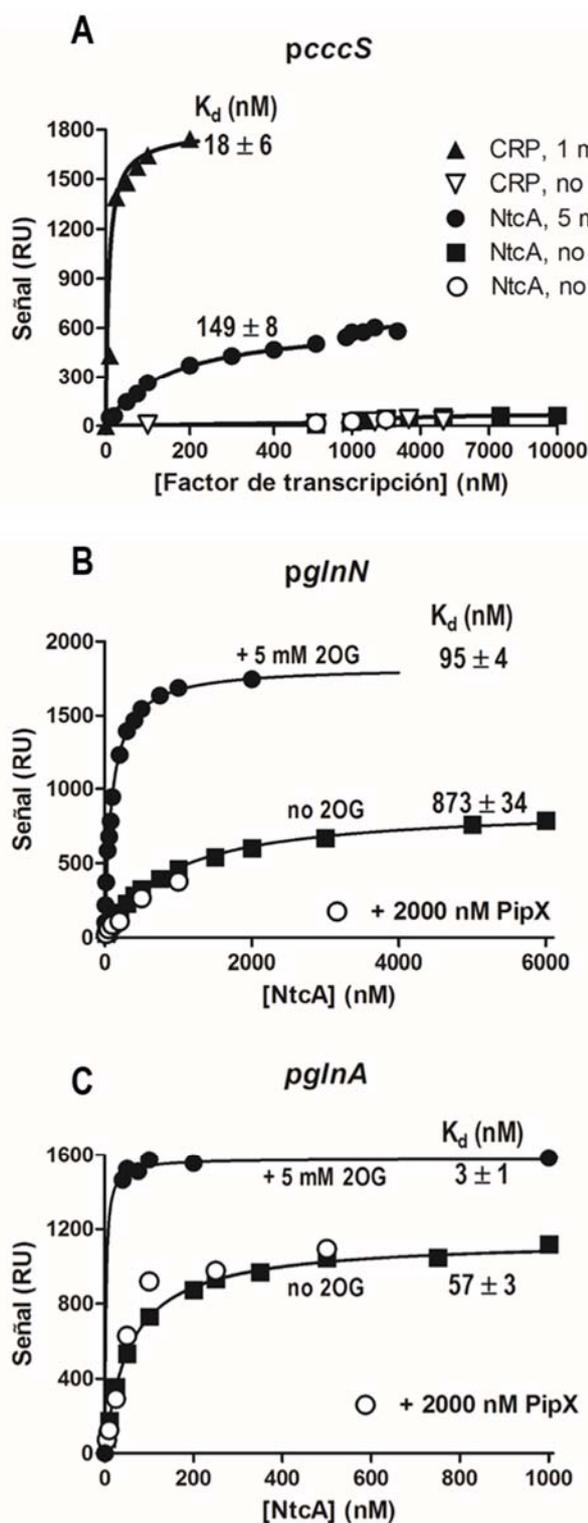


Figura 36. Influencia de la concentración del factor de transcripción en la señal obtenida de los chips portadores de fragmentos de los promotores de los genes (A) *cccS*, (B) *glnN* y (C) *glnA*. Los resultados mostrados corresponden a ensayos realizados en ausencia de efector o en presencia de 1 mM cAMP (CRP) o 5 mM de 2OG (NtcA). Los círculos abiertos indican las señales obtenidas en ausencia de 2OG o en presencia de 2000 nM de PipX. Los datos han sido ajustados a hipérbolas. K_d , concentración del ligando variado para la que la respuesta obtenida es la mitad de la respuesta extrapolada a concentración infinita del ligando.

magnitud superior para *pglnA* que para *pglnN* tanto en presencia como en ausencia de 2OG (Fig. 36, paneles B y C; y Tabla 6).

CRP no se unió a *pglnN* o a *pglnA* (Tabla 6) incluso cuando se usó a una concentración alta, de 2 μ M, o cuando se usaron concentraciones altas de cAMP (hasta 40 mM, datos no mostrados). Por el contrario, NtcA produjo una señal saturable por concentración de la proteína cuando se probó con *pcccS* si estaba presente el 2OG (Fig. 36A y Tabla 6). La afinidad de la unión de NtcA a *pcccS* fue solo un 50% más baja que la que presentó para *pglnN*, lo que sugiere que los promotores con una caja CRP podrían ser dianas razonablemente buenas para NtcA. Sin embargo, esta afinidad fue un orden de magnitud menor que la afinidad de CRP para *pcccS* (Tabla 6), lo que indica que a concentraciones comparables de NtcA y CRP y en presencia de sus correspondientes efectores, CRP debería impedir la unión de NtcA a los promotores característicos para CRP. Curiosamente, a pesar de la similitud de las masas de los dímeros de NtcA y CRP, NtcA provocó una señal máxima 3 veces más baja que CRP al unirse a *pcccS* (Figura 36A), lo que indica que CRP y NtcA no se unen de manera idéntica a *pcccS*.

Tabla 6. Influencia de la presencia de 2OG y cAMP sobre la unión de NtcA y CRP, respectivamente, a los promotores de los genes *glnN*, *glnA* y *cccS*, evaluada por SPR.

Factor	Efactor (mM)	K _a para el promotor indicado (nM)			Señal a saturación del factor de transcripción (RU)		
		<i>glnN</i>	<i>glnA</i>	<i>cccS</i>	<i>glnN</i>	<i>glnA</i>	<i>cccS</i>
NtcA	—	873±34	57±3	1954±267	880 ±12	1141 ±11	78 ±36
	2OG (5)	95 ±4	3±1	149 ±8	1832 ±20	1581 ±14	641 ±9
CRP	—	—*	—*	1166 ±15	—*	—*	45 ±7
	cAMP (1)	>>2000	>>2000	18 ±6	0 (a 2 μ M)	0 (a 2 μ M)	1979±156

* Sin unión detectable

Influencia del promotor en la unión de 2OG a NtcA

Titulamos la señal producida por NtcA (ilustrado en la Fig. 35D para *pglnN*) frente a la concentración de 2OG. Para los tres promotores la señal obtenida dependió sigmoidalmente de la concentración de 2OG (Fig. 37A). Estas sigmoides se ajustan a un número de Hill =2, lo que corresponde al número de sitios para 2OG presentes en el dímero de NtcA (Llácer *et al.*, 2010; Zhao *et al.*, 2010b), indicando también cooperatividad positiva entre los dos sitios para la unión de 2OG. Las concentraciones de 2OG a las que se obtiene la mitad de la señal máxima ($S_{0.5}$) difieren para los tres promotores evaluados (Fig. 36A y Tabla 7), con valores de 0.15, 1.19 y 1.71 mM para *pglnA*, *pglnN* y *pcccS*, respectivamente. Ello indica que la constante de disociación para 2OG difiere para los diferentes complejos NtcA-2OG-promotor. Aunque estos resultados tienen la limitación del pequeño número de promotores analizados, parece existir excelente correlación

Tabla 7. Influencia del promotor y/o de la presencia de PipX sobre la afinidad de NtcA por el 2OG o de CRP por el cAMP

Promotor	Factor de transcripción (nM)	Efector	[PipX] nM	$S_{0.5}^{2OG}$ o K_d^{cAMP} (mM)*
<i>pglnN</i>	NtcA (400)	2OG	---	1.19 ± 0.06
			1000	0.133 ± 0.004
<i>pglnA</i>	NtcA (100)	2OG	---	0.149 ± 0.011
			1000	0.044 ± 0.002
<i>pcccS</i>	NtcA (400)	2OG	---	1.71 ± 0.12
			1000	0.345 ± 0.018
	CRP (400)	cAMP	---	$1.4 \pm 0.2 \times 10^{-3}$
			1000	Igual que sin PipX

*Los valores de $S_{0.5}^{2OG}$ fueron estimados a partir de las curvas sigmoides mostradas en las figuras 36A y 37A (número de Hill=2). La K_d^{cAMP} se estimó a partir de la representación de Scatchard de la figura 37B. No se estimó formalmente la K_d^{cAMP} en presencia de PipX, pero se considera que no cambia en presencia de esta proteína, dada la superposición, dentro del error experimental, de los puntos obtenidos en presencia y ausencia de PipX en la figura 37B.

lineal entre el valor de K_d para la unión de NtcA a cada promotor y el valor de $S_{0.5}$ para 2OG de NtcA en el ensayo realizado con el mismo promotor (Fig. 37A recuadro de la derecha).

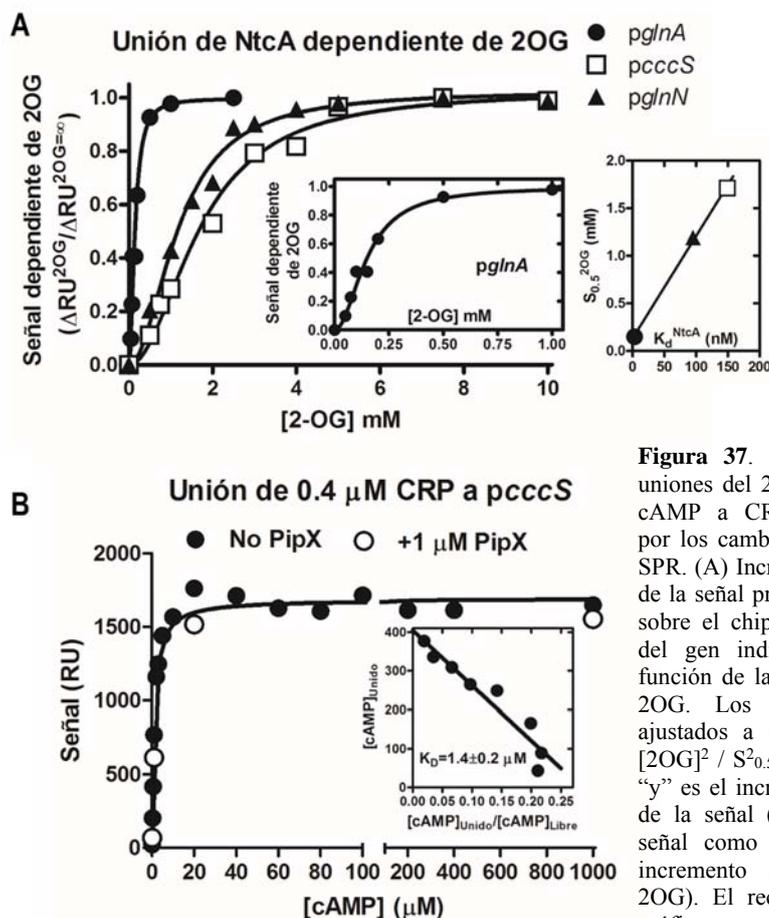


Figura 37. Titulación de las uniones del 2OG a NtcA y del cAMP a CRP, monitorizadas por los cambios en la señal de SPR. (A) Incremento fraccional de la señal producida por NtcA sobre el chip con el promotor del gen indicado, como una función de la concentración de 2OG. Los datos han sido ajustados a la ecuación: $y = [2OG]^2 / S_{0.5}^2 + [2OG]^2$, donde “y” es el incremento fraccional de la señal (incremento en la señal como una fracción del incremento a saturación de 2OG). El recuadro dentro del gráfico muestra en una escala de abscisas expandida la curva

para el chip *pglnA*. El recuadro situado a la derecha representa el $S_{0.5}$ para cada chip como una función de la K_d de NtcA para el mismo chip en presencia de 5 mM 2OG (datos obtenidos de la figura 36). La concentración de NtcA fue de 400 nM en los ensayos realizados con *pglnN* y *pcccS*, y de 100 nM en aquellos realizados con *pglnA*. (B) Influencia de la concentración de cAMP sobre la señal observada con 400 nM de CRP para el chip *pcccS*. Los círculos abiertos ilustran la señal observada al añadir 1000 nM de PipX. La concentración de CRP fue de 400 nM. El recuadro muestra la representación de Scatchard para la unión de cAMP a 400 nM de CRP. Para los detalles sobre las estimaciones de las concentraciones de cAMP libre y unido a partir de los datos en el panel principal, consultar el texto.

También monitorizamos la señal de resonancia producida al unirse CRP a *pcccS* como una función de la concentración de cAMP (Fig. 37B). Los resultados concuerdan con la publicación previa de que SYCRP1 exhibe muy alta afinidad por el cAMP (Yoshimura *et al.*, 2000), requiriéndose con CRP una concentración de cAMP tres órdenes de magnitud más baja que la concentración de 2OG requerida para la unión de NtcA a *pcccS* (Fig. 37 y Tabla 7). Una representación de Scatchard de los datos para el efecto del cAMP (inserto en la Fig. 37 B), preparado calculando las concentraciones de cAMP libre y unido a CRP a partir de la cantidad de complejo cAMP-CRP (estimado por la señal observada a una concentración dada de cAMP respecto a la máxima señal alcanzada a saturación de cAMP) dio un valor de K_d para cAMP de $1.4 \pm 0.2 \mu\text{M}$, similar al valor publicado previamente para CRP de *Synechocystis* utilizando equilibrio de diálisis (Yoshimura *et al.*, 2000).

Influencia de PipX sobre el sistema NtcA/2OG

Como cabe esperar si PipX se une exclusivamente al complejo NtcA-2OG, para los tres promotores estudiados aquí solo en presencia de 2OG aumentó PipX la señal de SPR obtenida con NtcA (Figs. 35E y 38; comparar con la Fig.36, círculos blancos). La especificidad del efecto de PipX sobre NtcA se confirmó por la falta de influencia de PipX en la señal provocada por CRP cuando se une a *pcccS* (Figs. 35F y 37B), como es de esperar si PipX no se une a CRP.

La estructura cristalina del complejo NtcA-PipX sugiere fuertemente que PipX estabiliza la conformación de NtcA que une 2OG (conformación "activa")(Llácer *et al.*, 2010). De acuerdo con esta visión, la presencia de PipX incrementó de 3.5 a 9 veces la afinidad de NtcA por el 2OG cuando

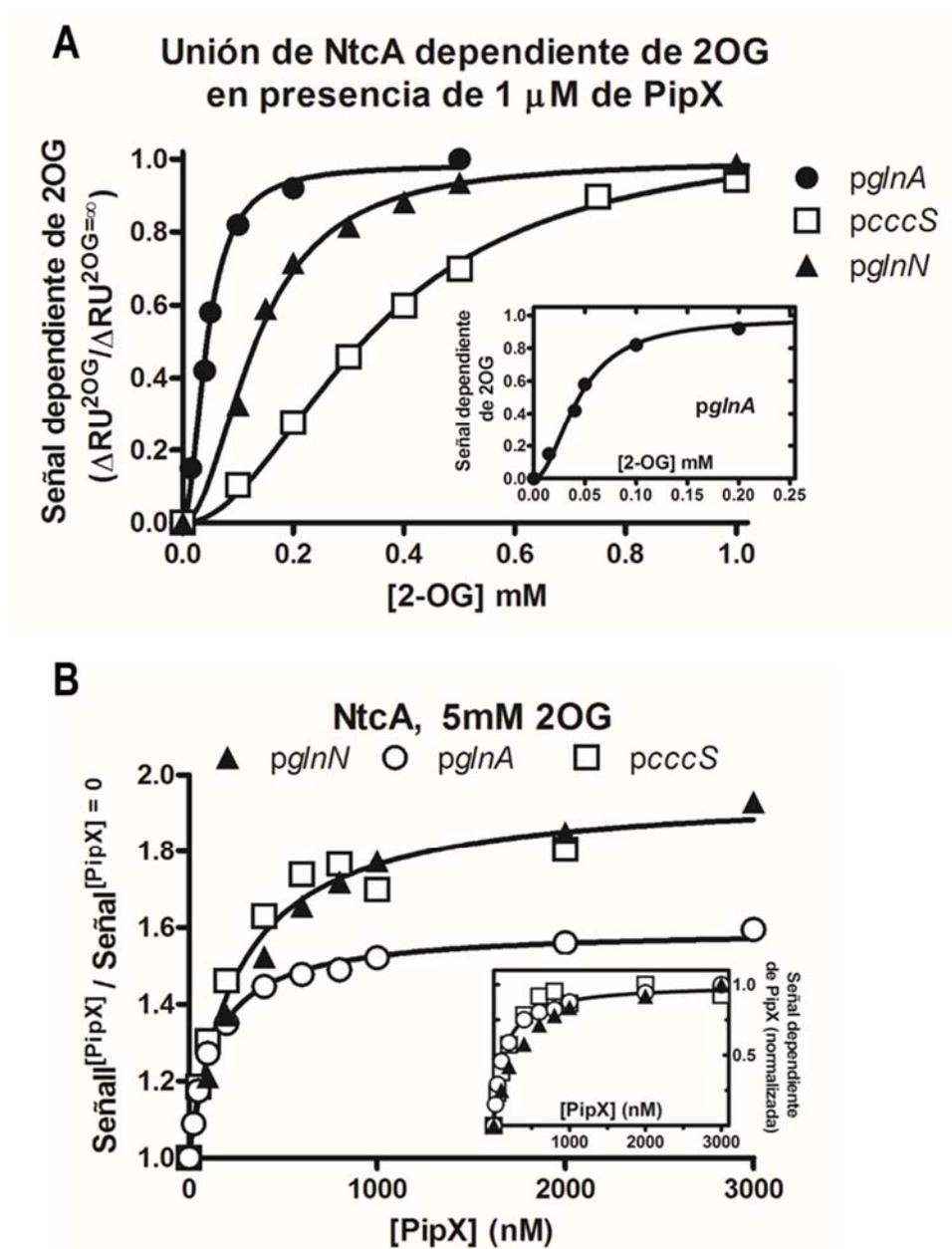


Figura 38. Influencia de PipX sobre (A) la afinidad de NtcA por el 2OG; y (B) sobre la señal generada por NtcA para cada promotor. La concentración de NtcA fue de 400 nM para los chips *pglN N* y *pcccS* y 100nM para el chip *pglN A*. (A) Incremento fraccional de la señal producida por NtcA sobre el chip para el promotor del gen indicado. Excepto para la presencia de 1000 nM PipX, todos los detalles son como para la figura 37A. (B) La concentración de PipX se varió como se indica, representándose el aumento en la señal relativo a la señal sin PipX. En el recuadro se representa el aumento de la señal con PipX como una fracción del aumento máximo en la señal a saturación de PipX.

se analizó dicha unión mediante SPR con los tres promotores utilizados aquí (Fig. 38A y Tabla 7). La dependencia de la señal con respecto a la concentración de 2OG continuó ajustándose a una sigmoide para un número de Hill de 2, y los valores de $S_{0.5}$ para 2OG, aunque disminuidos con respecto a sus valores en ausencia de PipX, continuaron difiriendo para los tres promotores investigados, aumentando en el mismo orden que cuando PipX estaba ausente, de menor a mayor, *pglnA*, *pglnN* y *pcccS*.

El aumento de 40% causado por PipX sobre la señal obtenida por NtcA cuando se une a *pglnA* en presencia de 2OG (Fig. 38B) está próximo a lo esperado para el aumento de masa que resultaría de la unión del complejo NtcA-PipX en vez de NtcA sola (masa del dímero de NtcA ~56 kDa; masa del complejo NtcA-PipX₂, ~82 kDa). Con *pglnN* y *pcccS* los incrementos en la señal fueron algo mayores (Fig. 38B), lo que sugiere que con estos promotores, que no están saturados por NtcA a las concentraciones de NtcA y 2OG utilizadas (Tabla 6), una parte del aumento de la señal dependiente de PipX se debe no solo al aumento de masa debido a la formación del complejo, sino también a una mayor saturación del promotor por NtcA, posiblemente consecuencia de la mayor afinidad de NtcA por el 2OG (ver sección anterior) provocada por PipX, lo que resultaría en un incremento de afinidad por el promotor.

Las representaciones del aumento normalizado en la señal inducida por NtcA frente a la concentración total de PipX virtualmente se superponen para los tres promotores (Fig. 38B recuadro). Si se sustrae de la concentración total de PipX que causa un aumento de señal del 50% del máximo para *pglnA* (135 nM, Fig. 38B) la concentración de PipX unido, (50 nM, correspondiente a hemiocupación de las subunidades de NtcA

presentes, 100 nM) se estima un valor operativo de 85 nM para el K_d para el complejo NtcA-PipX.

A modo de sumario

En conclusión, en estos experimentos de SPR demostramos que NtcA puede unirse al fragmento del promotor dependiente de CRP, pero que CRP no se une a los promotores con cajas NtcA, lo que sugiere que NtcA podría regular genes controlados por CRP. Nuestros datos demuestran de forma concluyente que CRP no se une a su promotor en ausencia de cAMP, mientras que NtcA se une a sus promotores en ausencia de 2OG, aunque su afinidad por dichos promotores aumenta considerablemente en presencia de este ligando (Fig. 35 tabla inferior). También hemos encontrado que la afinidad efectiva del NtcA para el 2OG está fuertemente influenciada por el promotor, y que PipX aumenta esta afinidad de forma importante, habiendo confirmado además que PipX se une a NtcA solo en presencia de 2OG, lo que apoya nuestra visión de que PipX estabiliza la conformación "activa" de NtcA con 2OG unido (Llácer *et al.*, 2010). Por último, la titulación de la unión de PipX a NtcA sugiere que la afinidad de PipX por este regulador transcripcional es mucho mayor que la encontrada por (Llácer *et al.*, 2010) para la unión de PipX a la proteína PII.

Resultados

Capítulo 3

Estructura de PipX en solución, e inmunodeterminación en S. elongatus de proteínas de la red reguladora de PipX

Estructura de PipX en solución

A) El contexto.

La estructura de PipX había sido determinado previamente en nuestro laboratorio en complejo con las proteínas NtcA (Llácer *et al.*, 2010) y PII (Llácer *et al.*, 2010), pero cristalizar PipX sola no había sido posible pese a múltiples intentos. Este es un objetivo de interés, ya que PipX consta de un dominio *tudor-like* rígido y compacto, formado por una hoja beta doblada sobre sí misma en herradura, y de dos hélices alfa C-terminales, de las cuales la segunda exhibe una movilidad muy importante en las tres moléculas de PipX secuestradas por PII, encontrándose en una conformación extendida en dos de ellas y replegada en la tercera (Llácer *et al.*, 2010). Sin embargo, en el complejo de PipX con NtcA ambas hélices están flexionadas (Llácer *et al.*, 2010) y también lo ha estado en los complejos estudiados estructuralmente aquí con NtcA y DNA.

Ya hemos indicado en la Introducción que en un reciente estudio colaborativo de los grupos de la Dra. Contreras y el nuestro (Labella *et al.*, 2016) se ha demostrado que la forma de NtcA con la hélice C-terminal extendida en el complejo con PII participa en un complejo ternario que recluta a un regulador transcripcional homodimérico de la familia de GntR, llamado PlmA. Las evidencias reunidas en dicho trabajo prueban que la unión a PlmA involucra aparentemente de forma exclusiva a PipX, siendo esencial la hélice C-terminal extendida en esa interacción.

La conclusión de todos esos estudios sería que el papel de PII en ese complejo sería de “abridor” de la hélice C-terminal de PipX, propiciando así la interacción de esta con PlmA. Si esta hipótesis es correcta, PipX en ausencia de otras proteínas debería mantener la hélice C-terminal

flexionada, inasequible para interacción con PlmA. De ahí que era clave determinar en qué conformación se encuentra PipX cuando no está formando complejos.

B) Mi contribución

Ante nuestros repetidos fracasos en obtener cristales de PipX (por su baja solubilidad y tendencia a la agregación), establecimos una colaboración con el grupo del Dr. Antonio Pineda (Centro de investigación Príncipe Felipe y Hospital Universitario y Politécnico La Fe), la Dra. Martina Palomino (Centro de Investigación Príncipe Felipe) y el Dr. José Luis Neira (Universidad Miguel Hernández, Elche) para la determinación de la estructura de PipX por resonancia magnética nuclear (RMN), ya que por su tamaño esta proteína sería susceptible de análisis estructural por esta técnica.

Mi papel en ese trabajo fue el de inducir la colaboración, preparar PipX de *S. elongatus* con alta pureza y en las elevadas concentraciones requeridas para los estudios de RMN, poner todos los medios posibles para inducir la solubilidad a la más alta concentración posible de esta muy poco soluble proteína, proporcionarla lo más fresca posible para minimizar problemas de agregación con el paso del tiempo, y sintetizarla marcada con ^{15}N o con y ^{13}C , sin lo cual no hubiera sido posible la resolución de la estructura por la señal exclusiva del protón. Quedaban fuera de nuestro alcance las determinaciones de RMN, realizadas en el instrumento de RMN del Centro de Investigación Príncipe Felipe de Valencia y el análisis de los espectros recogidos, pero sí la interpretación biológica de los resultados obtenidos. La doctora Martina Palomino del grupo del Dr. Pineda fue la persona que se encargó de recoger los

espectros necesarios a las secuencias de pulsos y campos direccionales apropiados, y de la asignación y resolución estructural, aunque el proceso presentó graves dificultades en las que la participación de un gran experto en RMN estructural de macromoléculas, el Dr. José Neira (Universidad de Elche), fue clave para coronar con éxito el proceso y para la realización de análisis particularmente sofisticados y complejos que permitieron alcanzar conclusiones sólidas y robustas sobre dinámica de la proteína.

Como ya he dicho, mi contribución preparativa ha sido clave para propiciar la determinación de la estructura de PipX por RMN, por lo que la resumiré aquí. Expresé PipX de *S. elongatus* PCC7942 (UniProtKB Q7X836) en *E. coli* Rosetta DE3 y la purifiqué esencialmente como ya he descrito en la primera sección de Materiales y Métodos. Un elemento preparativo clave fue la producción de esta proteína marcada con ^{15}N solo o también con ^{13}C . La producción se realizó del mismo modo que para la forma no marcada, excepto por el hecho de que el medio de cultivo no fue LB, sino medio mínimo M9. La proteína, de una pureza $> 97\%$ (pureza estimada por SDS-PAGE en geles de 15% de poliacrilamida, con tinción de azul Coomassie), se colocó mediante diálisis (Slide-A-Lyzer Dialysis Cassette de masa de corte nominal, 3500 MW; de Thermo Scientific) en tampón 50 mM Na fosfato pH 6.0 conteniendo 50 mM NaCl, apropiado para los experimentos de RMN y compatible con una razonable estabilidad de PipX, concentrándose en este medio (a 2 mg/mL mediante ultrafiltración centrífuga, con dispositivos Amicon Ultra de corte nominal 3 kDa; de Millipore). En total realicé 16 preparaciones de 1 L de cultivo primario cada una para generar PipX pura marcada y no marcada. Aunque no los hice yo, haré constar aquí que los experimentos de RMN se

realizaron a una temperatura de adquisición de 290 K en un espectrómetro Bruker Avance-II de 600 MHz, equipado con una criosonda (ver para los detalles de RMN, Forcada-Nadal *et al.* *Biochemistry* 2017; 56:3211-3224) (Forcada-Nadal *et al.*, 2017). Los experimentos se realizaron en 50 mM Na fosfato pH 6.0 y 50 mM NaCl, con una concentración de proteína de 200 μ M, ya que las concentraciones superiores precipitaban.

C) Los resultados del trabajo de RMN.

El detalle de los frutos de esta colaboración puede consultarse en nuestra reciente publicación en la revista *Biochemistry*. No mostraré aquí los datos obtenidos por RMN por no ser yo la persona que ha resuelto esta estructura o ha hecho el estudio dinámico de la proteína. Sin embargo, debido a mi muy fuerte implicación en este proyecto y en dar el impulso clave para la realización del mismo, me gustaría comentar brevemente los resultados (Forcada-Nadal *et al.*, 2017).

Afortunadamente los espectros 2D ^1H - ^{15}N HSQC mostraban buena dispersión, pero la muestra tendía a precipitar durante la adquisición de los espectros, lo que fue una complicación, pues limitó las concentraciones a las que fue posible trabajar (200 μ M), disminuyendo la sensibilidad. Fue posible asignar 1005 NOEs sin ambigüedad aunque fue pequeño el número de contactos a larga distancia que fue posible identificar (solo 119). Experimentos de intercambio de hidrógeno solo identificaron 16 residuos conectados por puentes de hidrógeno, y alternativamente, 37 residuos protegidos del solvente. Estos resultados representaron un grave obstáculo a la determinación de la estructura tridimensional de PipX. Sin embargo, la inspección cuidadosa de los

NOEs observados y el análisis de los desplazamientos químicos hizo posible arrojar luz sobre la estructura de PipX.

Me escapa el procedimiento que hizo posible asignar las resonancias debidas a la cadena principal, sobre la base de espectros HNCA, HNCACB y CBCA(CO)NH. Los NOEs secuenciales continuos de cadena principal facilitaron la asignación de la mayoría de resonancias de dicha cadena, proveyendo así de una confirmación para los espectros 3D. En este proceso se obtuvo evidencia de que las dos prolinas de la molécula adoptan conformaciones de sus enlaces peptídicos en *trans*. Las asignaciones se han depositado en el Biological Magnetic Resonance Data Bank (BMRB; archivo 27041).

Los resultados experimentales junto con el análisis de los desplazamientos químicos estuvieron de acuerdo con el carácter monomérico de PipX libre y refrendaron la arquitectura de la molécula establecida previamente por cristalografía de sus complejos con PII y con NtcA (Llácer *et al.*, 2010; Zhao *et al.*, 2010a) como cinco hebras β antiparalelas en la porción N-terminal, conectadas entre sí por giros, a las que siguen dos hélices α en el extremo C-terminal (Fig. 39). Una inspección meticulosa del espectro NOESY, sugiere que las dos hélices están muy cerca la una de la otra habiendo evidencia de contactos entre los grupos metilo de los residuos L65 y L80, justo los residuos que previamente se habían establecido como clave en la fijación de la forma flexionada de las hélices (Labella *et al.*, 2016). También, la última hélice presenta contactos NOE con la hoja β_2 , indicando que esta hélice está empaquetada contra la hoja β_2 , corroborando así que en solución PipX presenta una estructura flexionada. Así, las estructuras de PipX sola y en

el complejo NtcA-PipX parecen ser idénticas. Este resultado explica muy bien por qué PipX sola o unida a NtcA, en ambos casos con sus hélices replegadas, no es capaz de interactuar con PlmA, mientras sí lo es en complejo con PII, en el que se observa una conformación extendida de la hélice α terminal, hélice cuya delección no impide la formación del complejo PII-PipX pero sí causa la abolición de la interacción con PlmA (Labella et al., 2016). Todo ello da un fuerte espaldarazo a la idea de que PII facilita la adopción de la forma extendida de PipX, propiciando la interacción con PlmA.

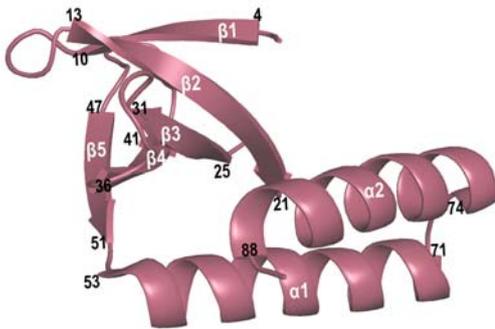


Figura. 39. Estructura de PipX en su conformación flexionada. Se señalan los elementos de estructura secundaria, indicándose en cada uno de ellos los números de residuo inicial y final.

Debido a que la extensión de la hélice C-terminal es un cambio conformacional parecía conveniente estudiar la dinámica de PipX en solución, lo que se realizó en el régimen de tiempo de picosegundos a nanosegundos, de nuevo por procedimientos en que no me he iniciado pero que han incluido la determinación del tiempo de correlación global, τ_c , predicho usando el servidor VADAR (Willard *et al.*, 2003) o HYDROPRO (García De La Torre *et al.*, 2000) para una estructura de PipX con la hélice C-terminal flexionada (Llácer *et al.*, 2010; Zhao *et al.*, 2010a), confrontando luego los valores predichos con los determinados experimentalmente con cuatro aproximaciones experimentales diferentes,

todas ellas basadas en aproximaciones de RMN. El valor medio estimado de τ_c de 9.3 ns a 290 K, utilizado en los cálculos ulteriores de dinámica, fue superior a lo esperado para una proteína de su pequeño tamaño (~10kDa), justificando la discrepancia al intercambio conformacional en la escala de tiempo de microsegundos a milisegundos (Zink *et al.*, 1994; Habazettl *et al.*, 1996; Jin *et al.*, 1998) o a acoplamiento entre la correlación global y los movimientos internos locales (Gümral *et al.*, 2008).

Para caracterizar la dinámica interna de la proteína a nivel de resolución atómica se puso el foco en movimientos muy rápidos (picosegundos a nanosegundos), más rápidos que el tiempo de correlación global, o en movimientos lentos (escala de microsegundos a milisegundos). Aunque la metodología utilizada también me resulta crítica, lo cierto es que el estudio dinámico mostró que PipX es rígida en solución pero presenta varias regiones muy flexibles, como son el extremo C-terminal y los lazos o giros que conectan los elementos de estructura secundaria. También mostró que de entre todos los elementos de estructura secundaria, la hebra β_1 presenta el valor del NOE en promedio más bajo, sugiriendo que posee baja rigidez. También predijo dos zonas de alta movilidad, una correspondiente al final de la hebra β_4 e inicio del giro que le sigue, y la otra correspondiendo a la región que abarca del residuo 70 al 74, porción que conecta las dos hélices (Fig. 39).

A partir de estos estudios se generó un mapa de residuos que presentaron una mayor movilidad en las escalas de tiempo estudiadas, muchos de ellos residuos que las evidencias cristalográficas han probado que intervienen en la formación de los complejos de PipX con NtcA o PII. De

entre ellos, destacan Leu39, Ile52 y Arg70 por presentar el mayor grado de flexibilidad conformacional (Fig. 40). La Leu39 se encuentra al final de β_4 , cerca de una región muy flexible que comprende la secuencia entre la Phe42 y la Thr47, en el lazo entre β_4 y β_5 (Fig. 40). Por otra parte, la Ile52, localizada al final de β_5 , muy móvil, implicado en las interacciones con NtcA y con PII, está en la unión entre β_5 y la primera hélice (Figs. 39 y 40). El residuo con la mayor movilidad es la arginina 70, al final de α_1 (Fig. 40), indicando que la región alrededor de ese residuo podría ser la bisagra que en presencia de PII unido se abre y proyecta la hélice C-terminal lejos del dominio tudor.

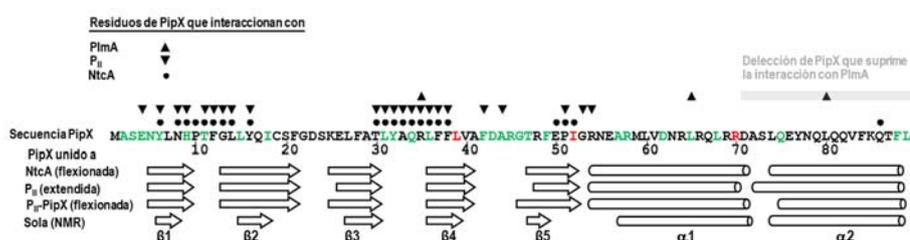


Figura 40. Secuencia de PipX de *S. elongatus* y su estructura secundaria basada en las estructuras cristalinas de los complejos con NtcA (PDB 2XKO) y PII (PDB 2XG8) y en los datos obtenidos por RMN para la proteína aislada. La secuencia se muestra en código de una letra y se resaltan en color los residuos que presentan mayor movilidad siendo los de color rojo los más móviles. Las hélices alfa se muestran como cilindros y las hojas beta como flechas. Los residuos de PipX que interactúan con PII o con NtcA están marcados con círculos o con triángulos invertidos, respectivamente. Los residuos marcados con un triángulo son los que se cree que interactúan con PImA (Labella *et al.*, 2016). La barra gris marca la delección de la hélice α_2 , la cual elimina las interacciones con PImA.

Inmunodeterminación en *S. elongatus* de proteínas de la red reguladora de PipX

A) PII, PipX, PlmA.

En colaboración con el grupo de la Dra. Contreras he determinado por western blot los niveles en *S. elongatus* de las proteínas PII, PipX y PlmA (tal como se ilustra en Materiales y Métodos y en la Fig. 41), Estas determinaciones se realizaron para ayudar a caracterizar “*in vivo*” el sistema PII-PipX-PlmA. Realicé las determinaciones en *S. elongatus* que había crecido en medio con amonio (Labella *et al.*, 2016), condición que favorecería la formación de los complejos PII-PipX. La figura 41 muestra ejemplos representativos de la cuantificación por western blot de esas proteínas (ver Materiales y Métodos). La especificidad de estos inmunoensayos viene avalada por la ausencia de detección de la banda de la proteína ensayada en el mutante sin PII (doble mutante *pipX glnB*; solo se puede eliminar PII si se elimina también PipX (Espinosa *et al.*, 2009) o sin PipX (Fig. 41). No tuvimos un control negativo similar con PlmA puesto que no fue posible preparar una cepa nula para PlmA (Labella *et al.*, 2016).

Los resultados de estos ensayos (Tabla 8), en términos molares de cadena polipeptídica prueban que PII es ~14 veces más abundante que PipX y este a su vez es ~7 veces más abundante que PlmA. Resultados preliminares (no mostrados) indican que el grado de riqueza de nitrógeno no afecta de forma importante a la abundancia relativa de estas tres proteínas.

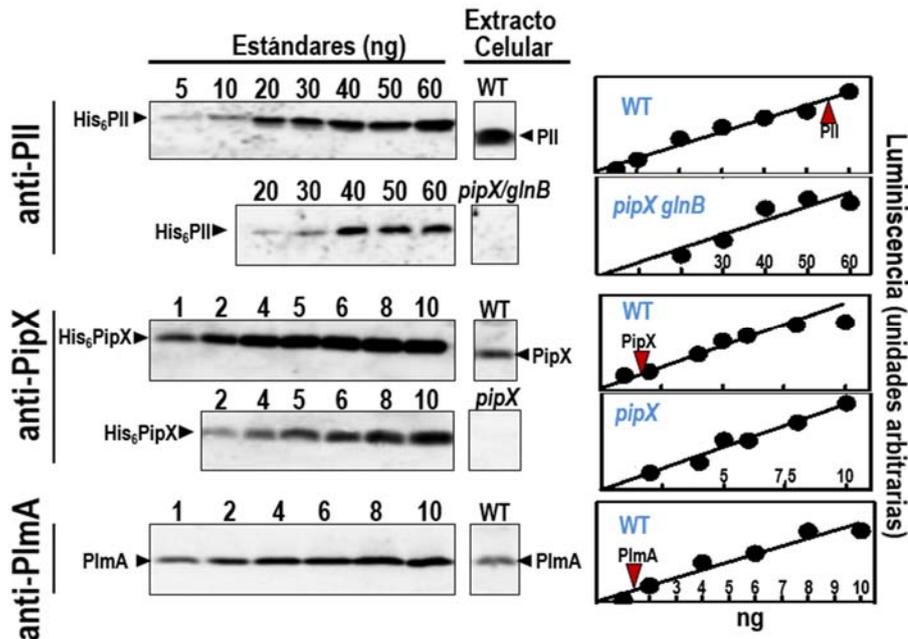


Figura 41. Inmunocuantificación de PII, PipX y PlmA en *S. elongatus*. Se muestran ejemplos ilustrativos de señales de inmunodetección obtenidas con proteínas puras recombinantes (His₆-PII, His₆-PipX y PlmA) (paneles de la izquierda) y extractos totales de cepas mutantes o silvestres (paneles del medio). Las señales cuantificadas con cantidades crecientes de proteínas recombinantes se representan (paneles de la derecha) y la señal obtenida en los extractos fue interpolada en la recta de calibrado, cuya posición se indica con una flecha roja). Los mutantes nulos *pipX* y *pipXglnB* proporcionan controles de la especificidad del inmunoensayo.

Mientras se llevaban a cabo estos ensayos, un estudio de proteómica (Guerreiro *et al.*, 2014) cuantificó las cantidades de todas las proteínas en el proteoma de *S. elongatus*, encontrando para PII, PipX y PlmA relaciones molares similares a las encontradas por nosotros (Tabla 8). Además, ese estudio también proporcionó datos sobre NtcA y NAGK (Tabla 8), revelando que están presentes en cantidades molares muy pequeñas en relación a sus respectivas dianas PipX y PII (las ratios molares entre NtcA/PipX y NAGK/PII, como subunidades, son 0.22 y <0.05, respectivamente) (Labella *et al.*, 2016). En la parte final de la

Discusión se realizan inferencias sobre el sistema de control de PII basados en los resultados cuantitativos resumidos en la Tabla 8.

Tabla 8. Niveles en *S. elongatus* de proteínas implicadas en la red de PipX

	Nivel por mg de proteína total (western blotting) ^a		Abundancia relativa (como subunidades) ^c	
	Como masa	Como subunidades ^b	Western blot	Proteómica (Guerreiro et al., 2014)
	µg/mg	pmol/mg		%
PII	2.2 ± 0.6	177	100	100
PipX	0.13 ± 0.04	12	6.8	7.6
PlmA	0.06 ± 0.02	1.7	1.0	0.7
PipY ^d	0.070 ± 0.004	2.9	1.6	2.0
NtcA	-	-	-	1.7
NAGK	-	-	-	3.6

^aResultados (media ± SD para ≥ 2 determinaciones independientes) obtenidos por interpolación en curvas patrón como se ilustra en la Fig. 41.

^bEl nivel de subunidades por mg de proteína total se ha calculado a partir del valor medio de la columna anterior. Se ha redondeado al entero más cercano en los casos de PII y PipX y al primer decimal más cercano en los casos de PlmA y PipY.

^cLa abundancia relativa de subunidades se da como un porcentaje de la abundancia de las subunidades de PII, a la que se le ha dado un valor arbitrario de 100.

^dEl valor de PipY se ha tomado de los resultados de la última sección de este capítulo.

B) El operón *pipX-pipY*

Usé también la aproximación de inmunodetección mediante western blot (ver Materiales y Métodos) en otra colaboración con el grupo de la Dra. Contreras (Labella *et al.*, 2017) (incluso formando parte de dicho grupo durante un breve periodo) en que se investigaba la funcionalidad del operón de los genes *pipX* y *pipY* de *S. elongatus*, reflejado en una asociación física entre ambos genes en esta especie (y en la mayor parte de los genomas cianobacterianos estudiados), sin casi distancia entre los mismos (Labella *et al.*, 2017).

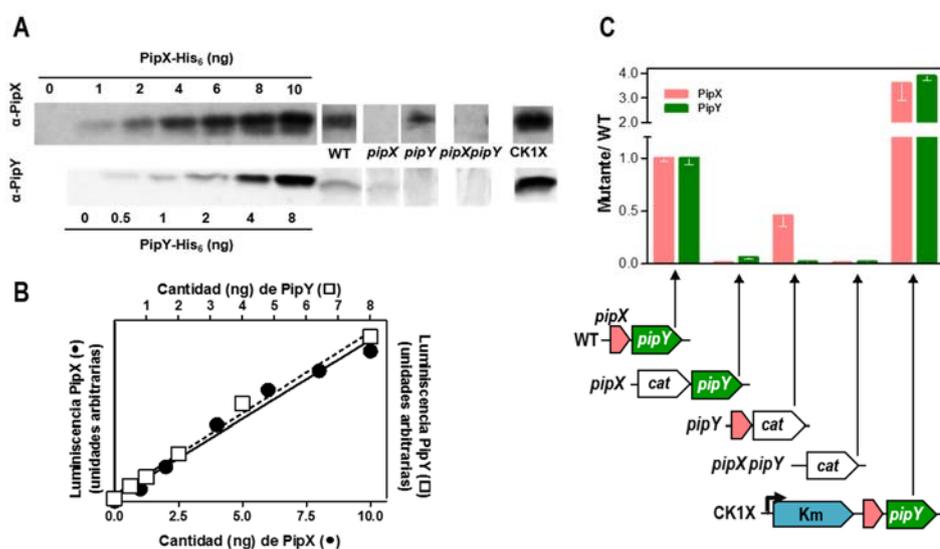


Figura 42. Cuantificación por inmunoensayo de PipX y PipY en *S. elongatus*. (A) Ejemplos ilustrativos de señales de inmunodetección para las cantidades indicadas de proteínas puras recombinantes (His₆-PipX e His₆-PipY) y para 25 μ g (para la detección de PipX) o de 50 μ g (detección de PipY) de proteínas totales del extracto celular indicado. También se representan (B) las señales luminiscentes cuantificadas con cantidades crecientes de proteína recombinante. La señal de los extractos fue interpolada en la recta de calibrado. Los mutantes nulos *pipX* y *pipY* proporcionan respectivos controles específicos del inmunoensayo. (C) Esquema de la composición génica del operón *pipX-pipY* en la cepa silvestre (WT) y en las cepas en las que uno, otro o ambos genes han sido reemplazados por el casete *cat*. También se muestra el esquema de la construcción donde se ha colocado un promotor fuerte inducible por kanamicina precediendo el cluster para sobreexpresar estos genes (CK1X). Las barras en el histograma muestran la relación de los niveles (por mg de proteína total en el extracto) de las proteínas PipX y PipY de las cepas mutantes con respecto al nivel correspondiente en el extracto de las células con genotipo silvestre (WT).

El grupo de la Dra. Contreras había construido cepas de *S. elongatus* en las que se habían inactivado ambos genes o uno u otro de ellos (Labella *et al.*, 2017) (mediante reemplazo por el gen *cat*, ver esquemas del panel C en la Fig. 42). Este grupo corroboró mediante PCR, RFLP y secuenciación, que las construcciones eran las correctas, llevándoles así a probar que *pipY* no es esencial para *S. elongatus*, con curvas de crecimiento en presencia de nitrato o de amonio indiferenciables de las de la cepa WT (Labella *et al.*, 2017).

Tabla 9. Niveles de PipX y pipY en las diferentes cepas estudiadas de *S. elongatus*

Proteína	Nivel en la cepa indicada de <i>S. elongatus</i>				
	WT	<i>pipX</i> ⁻	<i>pipY</i> ⁻	<i>pipX</i> ⁻ <i>pipY</i> ⁻	CK1X
	ng/mg de proteína total				
PipX	126 ± 4	1 ± 1	57 ± 13	1 ± 1	453 ± 88
PipY	70 ± 4	4 ± 1	1 ± 1	1 ± 1	272 ± 13

Resultados (media ± SD para ≥ 2 determinaciones independientes) obtenidos por interpolación en curvas patrón como se ilustra en la Fig. 42.

Mi intervención tuvo por objeto analizar si la inactivación de *pipX* o *pipY* afectaba los niveles de proteína expresada a partir del gen no inactivado. Ya los estudios de RT-PCR del grupo demostraron unos niveles disminuidos del transcrito del gen no inactivado, y un efecto más marcado de la inactivación de *pipX* sobre los niveles de *pipY* que a la inversa (Labella *et al.*; 2017), sin que se afectara un gen testigo ajeno a este operón (*sepF*) usado como control, indicando así la existencia de efectos polares restringidos al operón *pipXY*. Mis ensayos demostraron que la inactivación individual de *pipX* o *pipY* disminuyó los niveles respectivos de las proteínas PipY o PipX a un 7% y un 50%, respectivamente (Fig. 42

y Tabla 9), sugiriendo que los niveles de proteína PipY son fuertemente dependientes del gen *pipX*, como cabría esperar para un efecto polar.

De acuerdo con ello, la expresión inducida a partir de un promotor fuerte (CK1; cepa CK1X) colocado delante de *pipX*, que resulta en un aumento significativo en el nivel de PipX (Espinosa *et al.*, 2010) produjo un aumento de 3.4 y 4.3 veces en los niveles de PipX y PipY, respectivamente (Fig. 42 y Tabla 9). Así, mis resultados obtenidos mediante ensayos de Western blot apoyan la funcionalidad de este operón y las interacciones entre ambos genes.

Discusión

NtcA y la familia de CRP de reguladores transcripcionales

A) Especificidad de reconocimiento.

Dada la considerable similitud entre las cajas NtcA y CRP:

nGTA_nnnnnnnnnTAC_n (caja NtcA)

TGTGA_nnnnnnnTCACA (caja CRP) (las bases en las que se diferencian han sido subrayadas) y la similitud estructural entre ambos reguladores transcripcionales, un aspecto relevante de este trabajo de tesis ha sido establecer cuáles son los rasgos que hacen posible la discriminación entre ambas cajas por uno y otro factor. Se ha abordado esta cuestión desde dos perspectivas, una de ellas estructural, determinando las estructuras de complejos con DNA de NtcA de *S. elongatus* (Fig. 43), estructuras que compararemos en esta discusión con las correspondientes para CRP, y otra funcional, en experimentos de SPR con los promotores y las proteínas de *Synechocystis*.

Desde la perspectiva estructural, una gran parte del reconocimiento de la secuencia del DNA por parte de NtcA se puede atribuir únicamente a cuatro residuos: R186, V187, T190 y R191, ya que estos residuos, que son invariables en NtcA, son los únicos que contactan directamente con bases del DNA. En el caso de CRP el reconocimiento específico de bases del DNA lo llevan a cabo los residuos R180, E181 y R185, de los cuales R180 y R185 están conservados en NtcA (Fig. 44). El papel de R186 (NtcA) y R180 (CRP) en el reconocimiento del DNA es el mismo: el grupo guanidinio de esta arginina forma dos puentes de hidrógeno con los átomos O⁶ y N⁷ de un par G:C invariable localizado a 7 pares de bases del eje de simetría en ambas cajas NtcA y CRP (Fig. 20B Fig. 45A y B). Por otra parte, aunque los residuos R191 de NtcA y R185 de CRP también forman dos puentes de hidrógeno con los átomos O⁶ y N⁷ de una G

correspondiente a la caja de reconocimiento, esta G ocupa diferentes posiciones en las cajas NtcA y CRP.

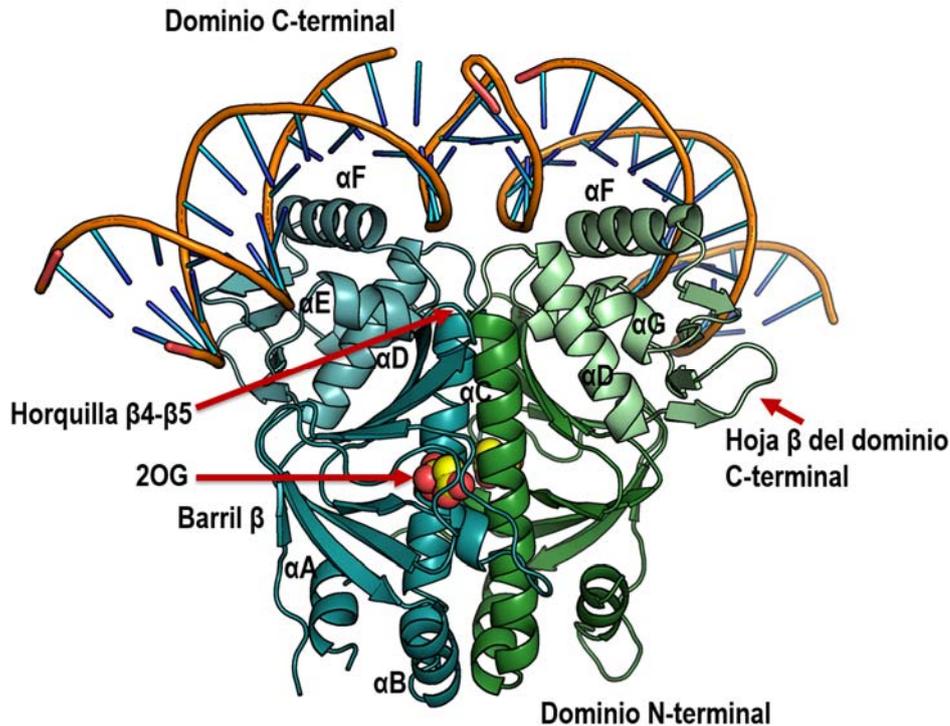


Figura 43. Estructura del complejo NtcA-DNA con 2OG unido (representado en esferas). Una subunidad de NtcA ha sido coloreada en azul y la otra en verde, mostrándose en ambas subunidades el dominio N-terminal de color más oscuro que el C-terminal. Se señalan algunos elementos de estructura secundaria.

Mientras que en la caja CRP esta G pertenece a la secuencia consenso y se encuentra a 5 pares de bases del centro del palíndromo de DNA (Fig. 45B), la caja NtcA no tiene una G en esta posición y la G reconocida pertenece a los ocho nucleótidos que unen las dos partes del palíndromo. Específicamente, R191 de NtcA se une a una G que se encuentra a 3 pares de bases del eje de simetría del DNA (Fig. 20B, Fig. 45A). Un estudio de transcriptómica masiva en *Anabaena sp.* (Mitschke *et al.*, 2011) ha mostrado que esta G se encuentra con una frecuencia del 50% en promotores putativos de NtcA, lo que podría explicar en parte las

diferentes afinidades encontradas para promotores de NtcA que tienen secuencias consenso perfectas, quizá atribuibles a diferencias de secuencia en los nucleótidos centrales no reconocidos como esenciales excepto por su número en la secuencia consenso de la caja NtcA. Por supuesto, posiblemente no es este el único determinante. NtcA curva el DNA, y se ha demostrado que los segmentos ricos en A/T en el DNA exhiben una mayor facilidad para la curvatura (Olson *et al.*, 1998) y que existe una correlación entre el contenido de A/T en diferentes promotores sensibles a NtcA y la afinidad que este factor de transcripción presenta por ellos (Vázquez-Bermúdez *et al.*, 2002b). Así, la secuencia del promotor de *glnA*, por el que NtcA presenta una muy alta afinidad, posee un alto contenido en A/T comparado con otros genes activados por NtcA. De esta forma, la curvatura del DNA y la afinidad de unión de NtcA también podrían estar fuertemente relacionadas.

```

                                1
CRP      -----MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAE 37
NtcA    1 MLANENSLTMTFRELGSGKLP-----LQIEQFERGKTIFFPGDPAE 41

CRP     38 TLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVR 87
NtcA   42 RVYLLVKGAVKLSRVYESGEEITVALLRENSVFGVLSLL-TGQRSDRFYH 90

CRP     88 AK--TACEVAEISYKKFRQLIQVNPDIIMRLSAQMARRLQVTSEKVG NLA 135
NtcA   91 AVAFTPVQLFSVPIEFMQKALIERPELANVMLQGLSSRILQTEMMIETLA 140

CRP   136 FLDVTGRIAQTLLNLAK-----QPDAMTHPDGMQIKITRQEIQIVGCSR 180
NtcA  141 HRDMGSRLVSFLLILCRDFGIPSPDGIT----IDLKLSHQAI AEAIGSTR 186

CRP   181 ETVGRILKMLEDQNLISAHGKTI VVYGTR----- 209
NtcA  187 VTVTRLLGDLRESKLI A IHKKRITVFNPVALSQQFS 222

```

Figura 44. Alineamiento entre las secuencias de aminoácidos de NtcA de *S. elongatus* y CRP de *E. coli*. Los residuos implicados en contactos con bases del DNA y con fosfatos se colorean en rojo y en verde, respectivamente, coloreando en marrón la Q170 de CRP por interactuar tanto con una base como con un fosfato.

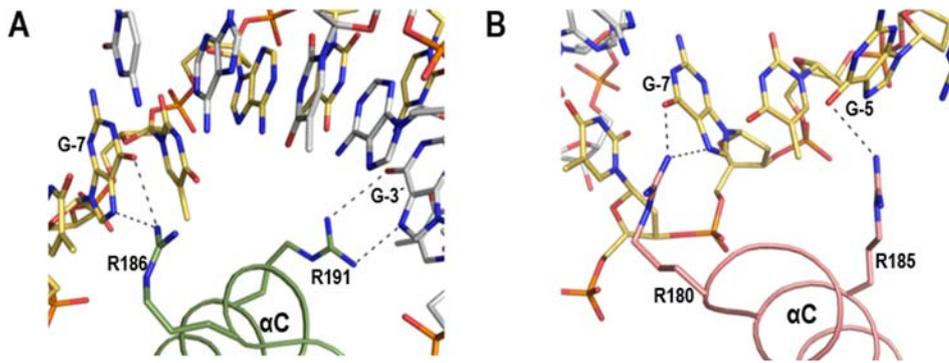
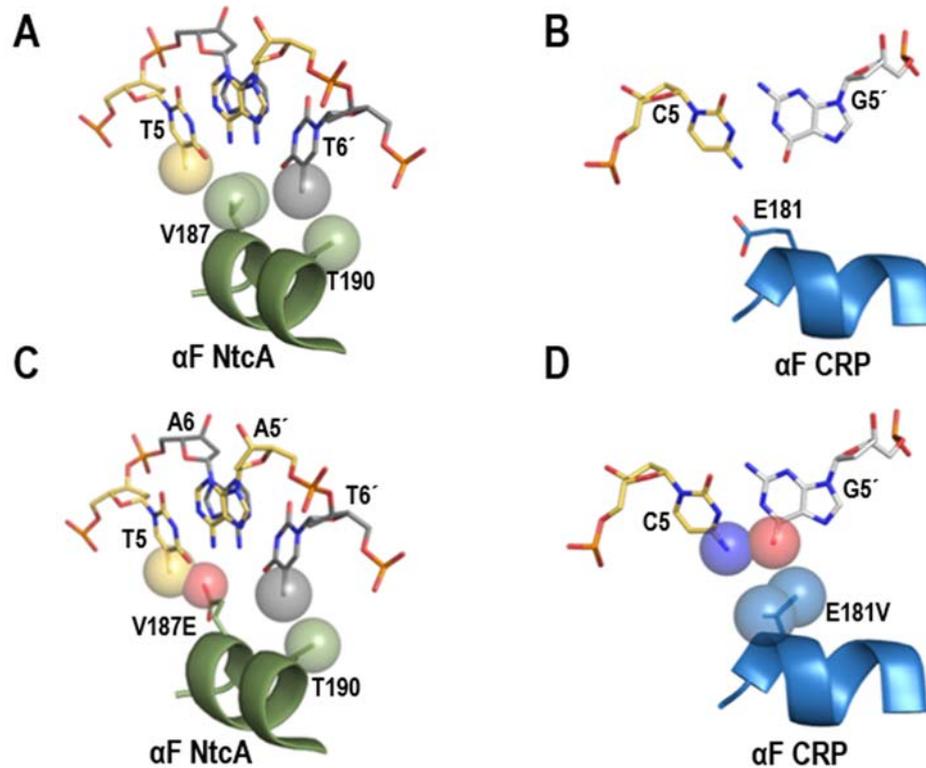


Figura 45. Comparación entre las interacciones formadas por las argininas R186 y R191 de NtcA (A) y R180 y R185 de CRP (B) con el DNA. Los residuos de NtcA se muestran en color verde y los de CRP en rosa. En ambos casos una cadena del DNA se muestra con los carbonos en amarillo y la otra con los carbonos en color gris. La figura correspondiente a CRP ha sido realizada a partir del PDB 1J59.

Los otros dos residuos que interaccionan con bases del DNA en NtcA, V187 y T190, son también invariables y parecen ser importantes para la discriminación entre las cajas NtcA y CRP. En CRP están reemplazados por residuos diferentes, pero también invariables: E181 y G184, de los cuales solo E181 interacciona con bases del DNA. En el caso de NtcA las cadenas laterales de los residuos V187 y T190 atrapan el grupo metilo de la timina perteneciente al par T:A situado a 6 pares de bases del eje de simetría del palindromo (Fig. 46A). En CRP, V187 se encuentra reemplazado por E181, formando este glutámico un puente de hidrógeno con el N4 de la citosina del par G:C situado a -5/+5 pares de bases del eje de simetría (Fig. 46B). Esta interacción es importante, ya que CRP tiene un requerimiento estricto por un par G:C en las posiciones -5/+5 (Ebright et al., 1987).

Nosotros hemos comprobado aquí mediante estudios de retardo en gel (EMSA) que al sustituir en NtcA V187 por un residuo de ácido glutámico, simulando lo que ocurre en CRP, la habilidad de NtcA para

unirse a la caja del promotor del gen *glnA* deja de detectarse por EMSA



(ver Fig. 25).

Figura 46. Representación de las interacciones entre los residuos V187 y E181 de NtcA y CRP respectivamente y sus mutantes con el DNA. Los residuos y las hélices se representan en verde para NtcA y en azul para CRP. (A) Interacciones hidrofóbicas entre el grupo metilo de la timina en la posición 6 del DNA y los residuos V187 y T190 de NtcA. (B) Interacción entre el residuo E181 de CRP y el DNA en la posición 5. (C) Mutante V187E de NtcA, la cadena lateral del glutámico choca con el grupo metilo de la timina en posición 5 (D) Mutante E181V de CRP.

Esta observación se puede explicar observando la estructura de NtcA-DNA, donde a 5 pares de bases del eje de simetría en vez de un par G:C se encuentra un par A:T en la secuencia consenso ideal. La larga cadena lateral del glutámico en el mutante chocaría con el grupo metilo de la timina (Fig. 46C). También se podrían explicar así los resultados de SPR

donde vemos que CRP es incapaz de unirse a los promotores de NtcA estudiados, ya que en ninguno de ellos se cumple el requerimiento de un par G:C en ambas posiciones -5/+5, y el glutámico presente en CRP chocaría con el grupo metilo de la timina. Sin embargo, los estudios de transcriptómica masiva realizados en cianobacterias (Su *et al.*, 2005; Mitschke *et al.*, 2011; Picossi *et al.*, 2014; Giner-Lamia *et al.*, 2017) demuestran que no existe una exigencia absoluta de un par A:T en estas posiciones de la caja de NtcA. En el trabajo de Mitschke (Mitschke *et al.*, 2011) muestran alrededor de 100 cajas de NtcA putativas y solo en dos de ellas se encuentra un par G:C en ambas posiciones -5/+5 de la caja. Menos en esos dos casos, en todos los demás se excluiría así, en aquellas cianobacterias que posean ambos factores de transcripción, como *Synechocystis*, la posibilidad de que CRP se pudiera unir a la caja NtcA, excluyendo así en la práctica la regulación promiscua por CRP mediada a través de cajas NtcA. Por el contrario, la presencia de una valina no debería impedir el reconocimiento de la caja CRP por parte de NtcA, ya que el pequeño tamaño de la cadena lateral de la valina deja espacio para acomodar el par C:G en la posición -5/+5 de la caja CRP (Fig. 46D). Además se ha visto que sustituciones como E181V o E181L en CRP eliminan completamente la necesidad de un par G:C en la posición -5/+5 de la caja y estos mutantes son capaces de unirse a pares A:T aunque con una afinidad reducida en más de 10 veces (Mitschke *et al.*, 2011). Esto concuerda con lo que hemos observado aquí mediante SPR, pues nuestros resultados prueban que NtcA puede unirse a al promotor de *cccS*, aunque con una afinidad unas 10 veces menor que la que presenta CRP por el mismo promotor (ver tabla 6).

El caso de la conservación de E181 en CRP es particularmente interesante si observamos las estructuras de otros miembros de la familia CRP unidos a DNA. Este residuo de CRP se encuentra conservado en FixK₂ (E196) y en TAP (E154) y, como ocurre en CRP, en estas proteínas, este glutámico interactúa con la citosina de un par C:G localizado a 5 pares de bases del eje de simetría. Sin embargo, este residuo es sustituido por una valina en CprK, como ocurre en NtcA y en ambos casos la valina establece contactos hidrofóbicos con el grupo metilo de una timina situada a 5 y 6 pares de bases del eje de simetría del palíndromo en CrpK y NtcA respectivamente. Parece que hay una alta conservación entre estos residuos y la secuencia del DNA que reconocen lo que les confiere gran especificidad.

Además de estas interacciones con las bases, NtcA forma un gran número de contactos con fosfatos del DNA (ver Fig. 20B y Fig. 44). Algunas de estas interacciones son equivalentes a las que se han descrito en CRP (residuos implicados: M144, H175, T185 y T188) pero también incorporando interacciones adicionales no conservadas en CRP (residuos R142, S184 y R197), siendo las mediadas por R142 particularmente interesantes. La cadena lateral de R142 surge del conector entre los dos dominios de NtcA, en el lazo que sigue a la larga hélice interfacial y cruza la línea media del dímero para formar un puente de hidrógeno con un fosfato situado en la parte de la caja NtcA que interactúa con la otra subunidad del factor de transcripción (Fig. 47). Este parece ser un rasgo único de NtcA ya que en los complejos resueltos de otros miembros de la familia CRP unidos a DNA, cada mitad de la caja interactúa exclusivamente con una única subunidad del factor de transcripción. Además de unirse a un fosfato del DNA, R142 forma un puente de

hidrógeno a través de su grupo guanidinio con un glutamato invariable (E61) que surge del giro entre $\beta 4$ y $\beta 5$ en el dominio regulatorio (EBD) de la subunidad adyacente de NtcA (la misma que contiene el fosfato con el que interacciona R142). Esta horquilla $\beta 4$ - $\beta 5$ experimenta movimientos dependiendo de si NtcA se encuentra en su conformación activa o inactiva. Así, R142, al interactuar simultáneamente con E61 por un lado y con un fosfato del DNA por el otro, está conectando el DNA con dos elementos móviles (la hélice C y horquilla $\beta 4$ - $\beta 5$), los cuales están implicados en los cambios desencadenados por la unión del 2OG (y también con su unión a PipX; ver más adelante).

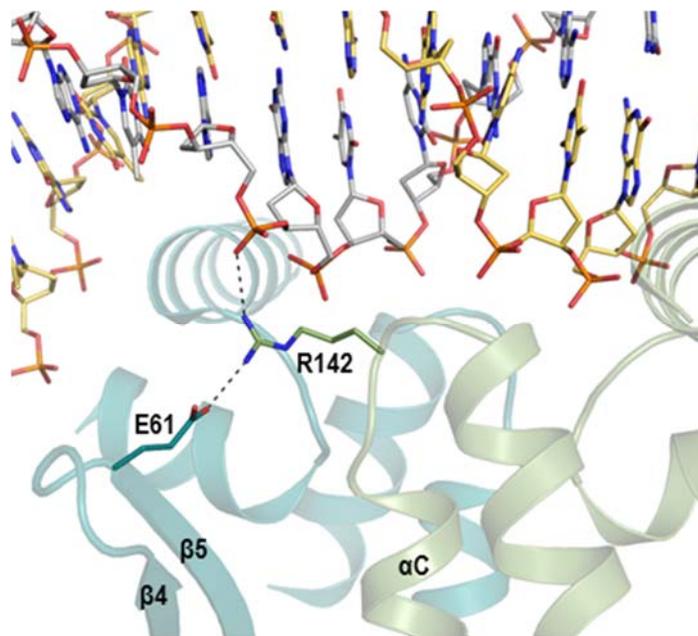


Figura 47. Interacciones formadas por el residuo R142 de NtcA. Una subunidad de NtcA se muestra en azul y la otra en verde. Algunos elementos de estructura secundaria han sido señalados.

B) Comparación estructural de las formas "activa" e "inactiva" de NtcA. Activación por 2OG

Las diferencias observadas entre las estructuras activas e inactivas de NtcA (A y B) se deben a dos tipos de cambios: cambios dentro de la subunidad y cambios en la posición relativa entre las dos subunidades. La estructura de los dominios de unión a DNA (DBDs) es prácticamente la misma en los tres tipos de estructuras (activa, inactiva A e inactiva B) (Fig. 48), aunque sí se dan cambios en las posiciones relativas del DBD de una subunidad con respecto al otro dominio DBD y al resto de la molécula. Sin embargo, al comparar los dominios de unión a ligando (EBDs) se revelan cambios importantes entre ellos (rmsd alrededor de 1.7 Å para 115 C α). Este hecho sugiere que la unión del 2OG al EBD causa reordenamientos estructurales en este dominio, los cuales pueden desencadenar los cambios encontrados en el dímero completo.

La consecuencia más importante de este reordenamiento estructural es el cambio en la posición de las hélices de unión a DNA (α F), las cuales adoptan orientaciones diferentes. Así, mientras en la estructura de NtcA-DNA la distancia entre las dos α Fs es de 34.8 Å, en las formas apo A1, apo A2 y B esta distancia aumenta hasta 37.6, 38.4 y 53.1 Å respectivamente (Fig. 48). De hecho, la distancia entre las dos α Fs en las regiones que interaccionan con los dos surcos mayores consecutivos del DNA (residuos 186 a 191) es solo de 27.7 Å en la estructura NtcA-DNA y sin embargo de 32.8, 33.5 y 45.8 Å en las estructuras apo A1, apo A2, y apo B, respectivamente. Esta mayor distancia entre las α Fs es incompatible con una unión al DNA específica en las dos partes del palíndromo simultáneamente y está causada por cambios en las orientaciones relativas de las hélices C al estar el extremo C-terminal de

estas hélices unido al DBD. En la forma activa las α Cs de ambas subunidades se encuentran en conformación de hélice superenrollada, mientras que en las formas apo hay una disminución en el ángulo entre ellas (relativo al eje de simetría binario del dímero), de forma que adoptan una conformación más paralela, reflejando la falta de contactos del 2OG con residuos de la α C y con la horquilla β 4- β 5 (Fig. 48).

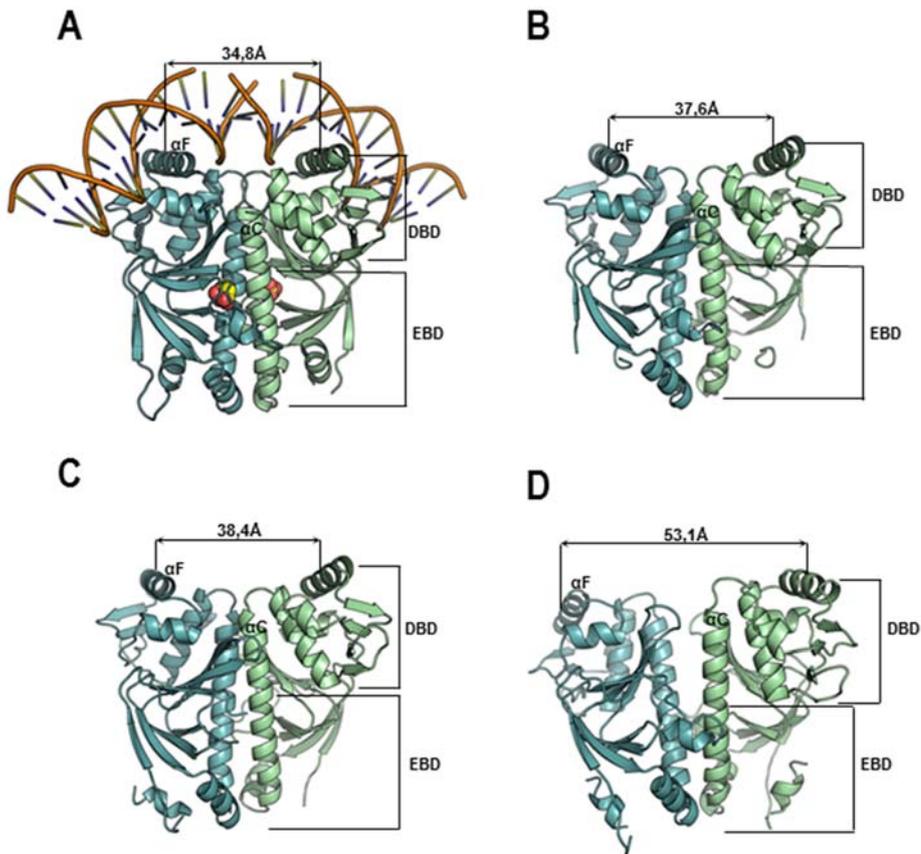


Figura. 48. Distancia entre las α Fs en las estructuras del complejo NtcA-DNA (A), NtcA apo A1 (B), NtcA apo A2 (C) y NtcA apo B (D). Se indican los dominios de unión a DNA (DBD) y a ligando (EBD) en cada estructura, así como las hélices C y F.

C) Transmisión de la señal

Una posible ruta de transmisión de la señal desde el 2OG hasta las hélices F fue descrita en *Anabaena* (Zhao *et al.* 2010b), en la cual parece ser muy importante la interacción del 2OG con el residuo E133* de la hélice C de la subunidad vecina (el asterisco sobre el número del residuo indica su posición en la otra subunidad).

En todas las formas apo de NtcA estudiadas aquí (A1 y A2 y B) o descritas anteriormente, el residuo E133* interacciona con los residuos R55 y Y89; y además, en el caso de apo A, también lo hace con R128 (Fig. 49A,B). Sin embargo, en la forma activa, estas interacciones cambian, siendo el 2OG y R128 los que interaccionan con E133*, mientras que R55 e Y89 pasan a interaccionar con E137* (Fig. 49C). Este residuo, E137*, se encuentra mirando al solvente en la forma apo B (Fig. 49B) pero forma interacciones con E61 y con S150* y H141* en la forma apo A (Fig. 49A). Así, parece que es este cambio en las interacciones formadas por E133* y E137* en las conformaciones activa/ inactiva el responsable de las distintas orientaciones de las hélices C entre NtcAs activa e inactiva.

Además de lo descrito, cabe destacar el papel de la horquilla β 4- β 5. Como se ha descrito previamente, en la conformación activa, E137*, localizado en el último giro de la hélice C, interacciona con el residuo R55 (Fig. 49C) el cual está situado en β 4 de la subunidad vecina, influenciando de esta forma la estructura de la horquilla β 4- β 5. Además, el 2OG forma un puente de hidrógeno con el carboxilato de E133*, desplazándolo de esta manera de su interacción con la horquilla β 4- β 5, la cual queda libre para adoptar la forma competente observada en la conformación activa. Esta

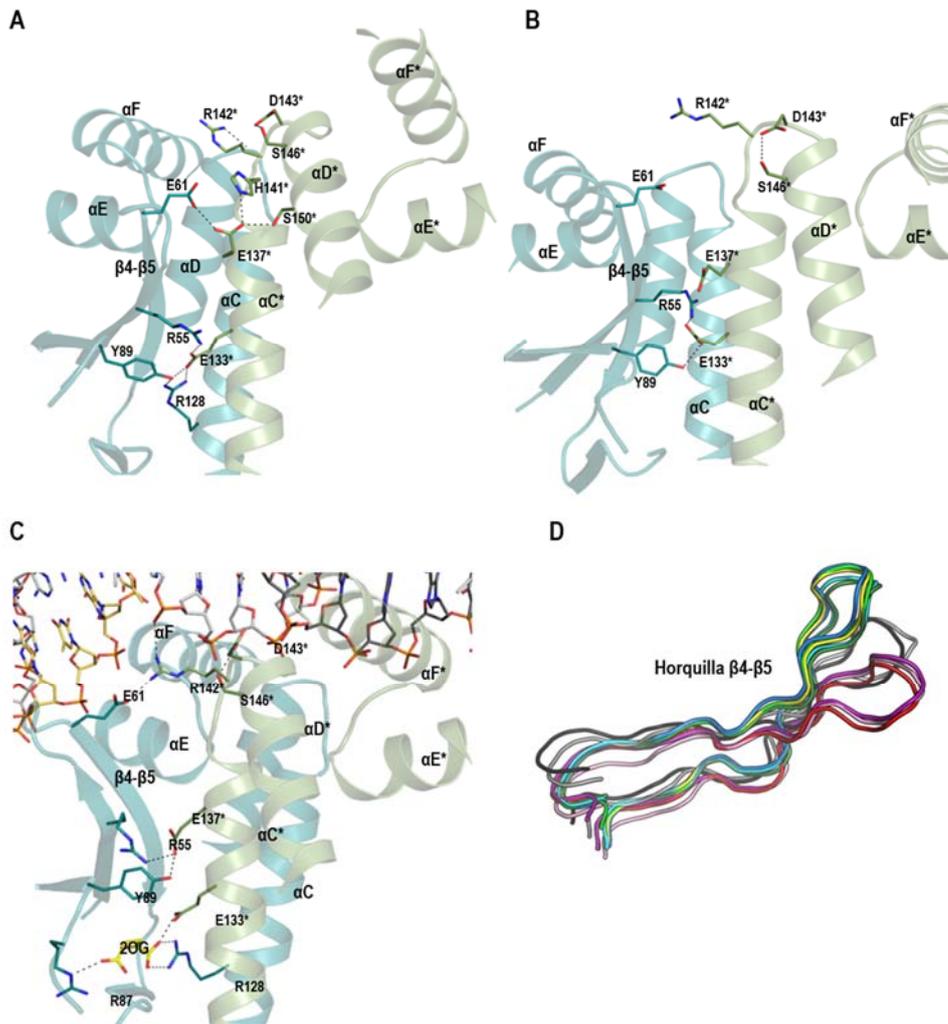


Figura 49. Mecanismo de activación de NtcA basado en las estructuras de NtcA apoA1 (A) y NtcA apoB (B) NtcA-DNA (C) de *S. elongatus*. Representación en cintas mostrando los residuos importantes para la transmisión de la señal desde el sitio de unión del 2OG de una subunidad a las hélices de unión a DNA. Una subunidad de NtcA se ha coloreado en color azul y la otra en verde, indicándose los residuos de esta última subunidad (denominada como vecina en el texto) con un asterisco. Los puentes de hidrógeno se muestran como líneas discontinuas. (D) Comparación de la conformación de la horquilla $\beta 4\text{-}\beta 5$ entre las formas activas e inactivas de NtcA. NtcA-DNA (azul). PipX-NtcA-DNA (verde oscuro), PipX-NtcA (cyan; PDB 2XKO), NtcA-2OG de *S. elongatus* (amarillo, PDB 2XHK), NtcA-2OG de *Anabaena sp.* (verde, PDB 3LA2), NtcA apo A1 (rojo), NtcA apo A2 (rosa), NtcA apo A de *Anabaena sp.* (púrpura, PDB 3LA7), NtcA apoB (negro) y NtcA apo publicada para *S. elongatus* (gris, PDB 2XHP).

horquilla presenta conformaciones muy distintas en NtcA según se encuentre en su forma activa o inactiva, (Fig. 49D) y está conectada directamente con el motivo HTH de la misma subunidad. En ambas formas inactivas es E133* el que interacciona con R55, promoviendo así una conformación distinta de la horquilla β 4- β 5 a la que la que adopta cuando NtcA está unida a 2OG, lo que tiene consecuencias en la estructura del motivo HTH. Además, el hecho de que la conformación de la horquilla β 4- β 5 dependa directamente de la unión de 2OG puede ser una de las claves para entender el segundo mecanismo por el cual el 2OG podría activar la transcripción dependiente de NtcA (además de aumentar la afinidad por los promotores), ya que en otros miembros de la familia, como FNR, CooA y CRP (Bell & Busby, 1994; Leduc *et al.*, 2001; Lawson *et al.*, 2004) la horquilla β 4- β 5 pertenece a la región llamada AR3 (“región activadora 3”), la zona del factor de transcripción que interacciona con la subunidad sigma de la RNA polimerasa. Por otra parte, la conformación de la horquilla β 4- β 5 es importante también para la interacción con PipX, ya que en NtcA "activa" esta horquilla interacciona con el extremo C-terminal de la hélice C-terminal de la subunidad vecina, de forma que la conformación de esta hélice cambia a una apropiada para interaccionar con PipX (Llácer *et al.*, 2010).

En conclusión, parece que tras la unión del 2OG a NtcA se establece un nuevo conjunto de interacciones entre las dos subunidades a través de las hélices C. Particularmente, la interacción del efector con residuos como R128 y E133*, pertenecientes a hélices C contrarias, (Fig. 49C), conecta dichas hélices con el barril β que contiene el 2OG, forzando de esta manera la reorientación de las hélices C. Así, la señal se transmite desde el sitio de unión del 2OG a los extremos C-terminales de las hélices C, los

cuales están pegados al DBD. Además, al final de las hélices C se encuentra localizado el residuo R142, el cual está directamente implicado en contactos con el DNA (Fig. 49C). Por último, el hecho de que la unión de 2OG a una subunidad de NtcA influya en la orientación de ambas hélices C, explicaría por qué existe cooperatividad positiva para la unión de 2OG a la segunda subunidad de NtcA.

D) Activación de la transcripción por otros miembros de la familia de CRP

Las estructuras de varios de los miembros de la familia CRP han sido resueltas tanto en su forma activa unida a ligando como en sus formas apo. En todos los casos, la señal efectora se transmite desde del EBD al DBD a través de la hélice C, ya que el sitio de unión a ligando está conservado topológicamente en todos ellos, aunque los residuos implicados en su unión varían según el miembro de la familia. Una serie de cambios conformacionales inducen entonces a las hélices F a adoptar una distancia y orientación adecuadas para su unión al DNA, la cual activará la transcripción de los genes diana. A pesar de estas características comunes, varias propiedades estructurales distinguen a los distintos miembros de la familia CRP, lo que se traduce en diferentes mecanismos de activación.

A pesar de que CRP es incapaz de unirse al DNA en ausencia de su ligando, el cAMP, se ha descrito una serie de mutaciones en *EcCRP* que aumentan la afinidad de unión al DNA en ausencia de su efector alostérico (Weber & Steitz, 1987; Kim *et al.*, 1992; Eiting *et al.*, 2005) y por lo tanto son responsables de activar la transcripción de una forma independiente de ligando. Algunas de estas mutaciones están localizadas

en αD o en la unión entre EBD y DBD, lo que resulta en interacciones nuevas entre los residuos de estas regiones y aquellos en el final de αC . De esta forma se previene la rotación libre del DBD con la consecuente exposición de las hélices de reconocimiento hacia el solvente como ocurre en la conformación "activa". Una de estas mutaciones activadoras es la sustitución del residuo G141 (localizado en αD) por un aminoácido polar, de forma que al juntarse las hélices C y D, αF queda libre y orientada adecuadamente para interactuar con el DNA (Kim *et al.*, 1992). De la misma forma, el reemplazo de otro residuo situado en αD , A144 por otro con una cadena lateral más voluminosa, independientemente de su naturaleza, resulta también en una unión a DNA de forma independiente de cAMP (Weber *et al.*, 1987; Kim *et al.*, 1992). Semejante a este mutante de CRP es la variante de G145S PrfA, la cual exhibe una mayor afinidad por el DNA a través de la estabilización del dominio HTH. Ambas mutaciones están localizadas en la hélice αD y presentan un mecanismo de activación similar.

Es interesante que estas sustituciones que consiguen una activación independiente de ligando se cumplen de forma natural en la secuencia de NtcA. El residuo correspondiente a G141 de CRP es polar en NtcA (S146) e interactúa con D143, localizado entre los dos dominios, lo que provoca el acercamiento de las hélices αC y αD (Fig. 50). También el residuo correspondiente a A144 CRP es sustituido por un residuo más voluminoso e hidrofóbico en NtcA: V149 (Fig. 50) de forma que empuja y orienta αF para una mejor interacción con el DNA. Además, el equivalente en NtcA de G145S de PrfA es el residuo S150, el cual, al interactuar con T138 al final de αC (Fig. 50), ancla αD con la conexión entre los dos dominios de NtcA. La aparición simultánea de estos

residuos “activadores” en NtcA explican por qué el 2OG solo aumenta moderadamente su afinidad por el DNA ya que la activación de NtcA no depende de una reorientación del DBD con respecto al EBD de su misma subunidad como en CRP u otros reguladores de la familia CRP, sino en cambios en la orientación relativa de α C, que acaba repercutiendo en la distancia entre los DBDs de las dos subunidades.

Algunas de estas mutaciones “activadoras” ocurren también de forma natural en otros miembros de la familia CRP en los cuales se ha visto unión a DNA independiente de ligando. Por ejemplo, el residuo correspondiente a G141 de CRP es un aminoácido polar en los factores de transcripción FixK₂ (E159), SdrP (N121) y CLP (D162) y también el residuo correspondiente a A144 de CRP es sustituido por un aminoácido más voluminoso en el caso de CLP (V165) y de PrfA (L150).

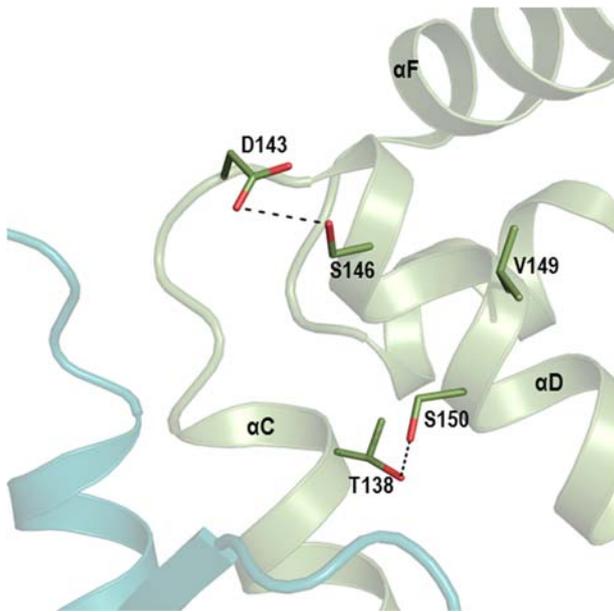


Figura 50. Representación en cintas de NtcA de *S.elongatus* mostrando los residuos que se corresponden con mutaciones activantes descritas en CRP (residuos equivalentes S146, V149 y S150) y las interacciones que forman. Se señalan algunos elementos de estructura secundaria.

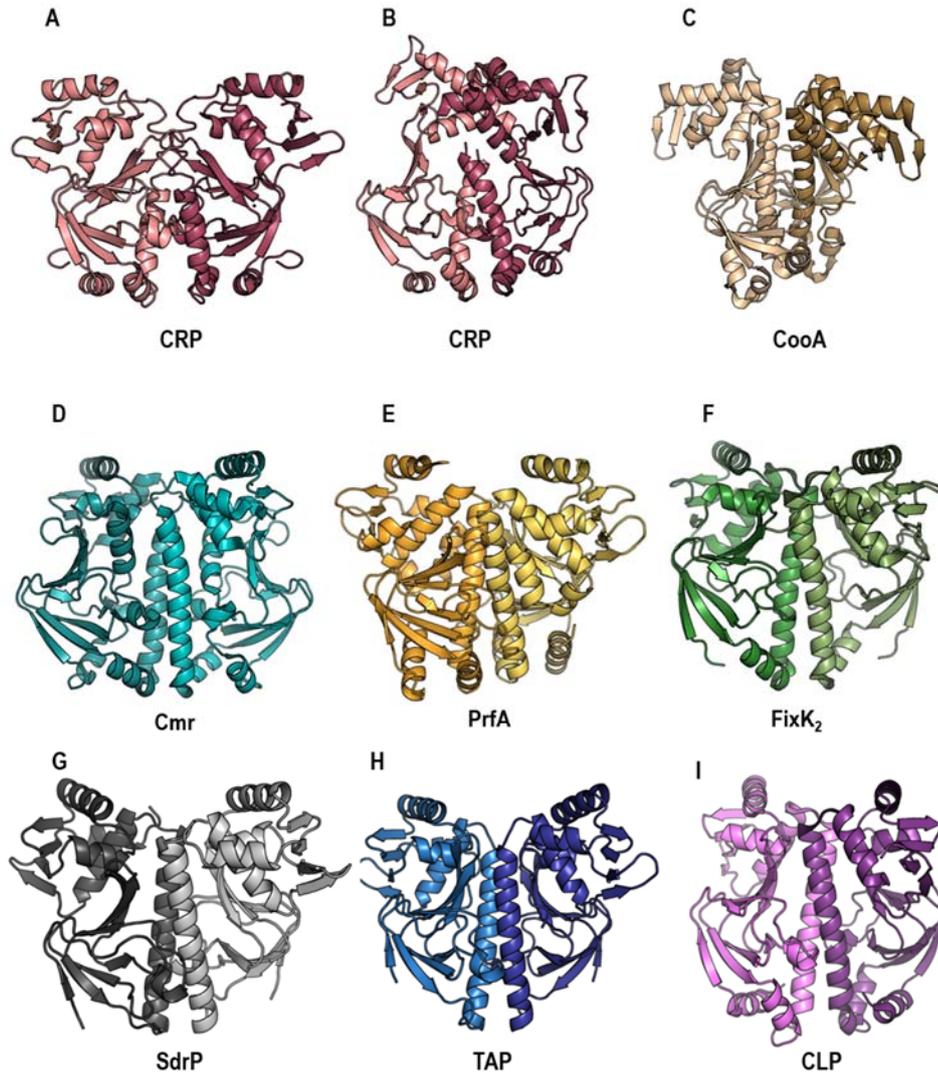


Figura 51. Estructuras de algunos miembros de la familia CRP en ligando. (A) y (B) muestran dos conformaciones distintas de CRP de *E. coli* en ausencia de cAMP (PDBs 2WC2 y 3H1F respectivamente). (C) CooA de *Carboxythermus. hydrogenoformans* (PDB 2FMY) (D) Cmr de *M. tuberculosis* (PDB: 5W5A). (E) PrfA de *L. monocytogenes* (PDB: 2BEO). (F) FixK₂ de *B. japonicum* (PDB: 4I2O), (G) SdpP (PDB: 2ZCW) y (H) TAP (PDB: 5I2D) ambos de *T. thermophilus*, (I) CLP de *Xanthomonas campestris* (PDB: 3IWZ). Nótese la diferencia en la posición de las hélices de reconocimiento de DNA en las estructuras representadas en los paneles A-C, los cuales necesitan ligando para su unión al DNA y el resto, las cuales representan proteínas capaces de unirse al DNA en ausencia de ligando.

Así, la presencia de estas mutaciones activadoras de forma natural en estos factores de transcripción (incluido NtcA) quizá podría explicar su mayor capacidad de unirse al DNA en ausencia de efector que la de otros miembros de la familia que carecen de dichas mutaciones, ya que en los primeros las hélices de reconocimiento estarían siempre expuestas al solvente (Fig. 51).

El complejo PipX-NtcA-DNA

Resolver la estructura del complejo entre PipX-NtcA y el DNA era un objetivo importante de esta tesis para aclarar el papel co-activador de PipX sobre NtcA. Aunque no había evidencias previas (Llácer *et al.*, 2010) una posibilidad era que PipX estuviera contactando directamente con el DNA. Sin embargo, la presente estructura del complejo ternario descarta esta opción al encontrarse PipX muy lejos del DNA (ver Fig. 30). Lo que observamos en la estructura del complejo ternario es que tanto el dímero de NtcA como PipX se encuentran en la misma conformación descrita en el complejo entre ellos en ausencia de DNA, y además NtcA se une al DNA de la misma forma que en el complejo NtcA-DNA. Sin embargo, y aunque la resolución del cristal es baja (3.8Å) (aunque los mapas de densidad electrónica y los valores de $R_{\text{factor}}/R_{\text{free}}$ son razonablemente buenos, ver Tabla 3), podemos observar una serie de interacciones adicionales entre NtcA y el DNA en este complejo ternario (ver Fig. 21) que no se dan en el complejo NtcA-DNA en ausencia de PipX. Esto podría justificar en parte el papel co-activador de PipX al promover puntos de contacto extra entre NtcA y el DNA. En cualquier caso, lo que sí parece claro es que un factor importante en el papel co-activador de PipX es su capacidad de estabilizar la conformación

activa de NtcA (con 2OG unido) (Llácer *et al.*, 2010) de forma que la unión de PipX a NtcA desplaza el equilibrio entre las formas inactivas del factor de transcripción en favor de su forma competente para unir DNA. En concordancia con esta visión, nuestros resultados de SPR muestran cómo la presencia de PipX incrementa la afinidad de NtcA por el 2OG con los tres promotores estudiados (ver Fig. 38A y Tabla 7). Ya que PipX no se une a NtcA en ausencia de 2OG, no debería ser capaz de estimular la unión de la primera molécula de 2OG al dímero de NtcA y por tanto se uniría a la subunidad que ya ha unido este ligando, favoreciendo la unión de la segunda molécula de 2OG a la otra subunidad de NtcA. Sin embargo y dado que hay cooperatividad positiva para la unión de la segunda molécula de 2OG incluso en ausencia de PipX, parece que PipX estaría favoreciendo más la no disociación del 2OG que la unión del mismo a NtcA. Por otro lado, cada molécula de PipX interacciona con el EBD de una subunidad de NtcA y el DBD de la otra (Llácer *et al.*, 2010), lo que ayudaría a orientar la hélice F para que interaccione con el DNA, estabilizando la forma activa. Por tanto, PipX no solo mejora la afinidad por el promotor indirectamente a través de la mejora de la afinidad por el 2OG sino también a través de interacciones con elementos de NtcA implicados finalmente en la unión con el DNA.

Por otra parte, la mayoría de la superficie de PipX (incluyendo las dos hélices) permanece expuesta en el complejo ternario, lo que supone una posibilidad de interacción entre PipX y componentes del complejo transcripcional, como es la RNA polimerasa. De hecho, el dominio Tudor-like (presente en PipX) ha sido identificado en proteínas bacterianas que interaccionan con la RNA polimerasa y con el DNA (Deaconescu *et al.*, 2006; Shaw *et al.*, 2008). Además, en un modelo

realizado por Llácer (Llácer *et al.*, 2010) en el cual se superponen los complejos PipX-NtcA (Llácer *et al.*, 2010) y CRP-DNA- α CTD (Benoff *et al.*, 2002) se puede ver cómo PipX ocuparía una posición próxima a la del α -CTD sin impedimentos estéricos para establecer contactos con él, aunque esto es puramente un modelo no corroborado (ni excluido) experimentalmente.

Análisis funcional de las interacciones en el complejo de activación: estudios de SPR con NtcA y CRP de *Synechocystis*

Reiteraremos aquí que el hecho de que estos dos factores de transcripción homólogos posean secuencias de reconocimiento altamente similares unido a su coexistencia en algunas cianobacterias, como es el caso de *Synechocystis* sp., planteaba la posibilidad de que existiera regulación cruzada entre ambos sistemas regulatorios, de forma que uno de estos factores de transcripción fuera capaz de interactuar con promotores portadores de cajas para el otro regulador. Nuestros resultados muestran cómo CRP es incapaz de unirse a los promotores de NtcA aún en presencia de cAMP, el efector de CRP, de acuerdo con la falta de una G en la posición -5 en los dos promotores con caja NtcA examinados, de acuerdo con la conclusión de (Omagari *et al.*, 2004) de que esta G es esencial para la unión de CRP. La gran selectividad de CRP para su caja frente a cajas NtcA descarta una acción directa de CRP sobre el regulón de NtcA. Por el contrario, NtcA no solo se une a sus propias cajas sino que también es capaz de unirse al promotor CRP de *cccS*, aunque con una afinidad mucho menor que la que presenta CRP para dicho promotor. Asumiendo que los niveles celulares de cAMP son mucho mayores que la K_d de 1.4 μ M estimada aquí por SPR para la unión de este efector a CRP,

parece improbable que NtcA pueda competir *in vivo* con CRP por sus cajas, dada la mayor afinidad del complejo CRP-cAMP que del complejo NtcA-2OG por la caja CRP. Sin embargo, los niveles de cAMP y la expresión de CRP en *Synechocystis* no se han estudiado en detalle, por lo que sería concebible que NtcA pudiera influenciar genes regulados por CRP cuando los niveles de cAMP y/o de CRP sean muy bajos, en particular cuando se den condiciones de bajo nitrógeno, en las que NtcA es abundante y activa.

La completa falta de unión de CRP a su propio promotor observada en nuestros experimentos de SPR en ausencia de cAMP, y la alta afinidad de CRP por cAMP tanto en presencia (nuestros resultados) como en ausencia de promotor (Yoshimura *et al.*, 2000), sugieren que el sitio para cAMP está siempre accesible en CRP y que la unión del cAMP desencadena un cambio conformacional que promueve la unión de CRP a su secuencia de reconocimiento en el DNA. Sin embargo, para la activación de NtcA por 2OG nuestros resultados sugieren un modelo diferente. Puesto que NtcA se une a sus promotores esté 2OG presente o no, y puesto que la afinidad efectiva de NtcA por el 2OG varía dependiendo del promotor, la caja NtcA y el 2OG deben unirse a una misma conformación de NtcA, desplazando el equilibrio hacia ese conformero entre las varias conformaciones de NtcA que puedan coexistir en solución. Desgraciadamente nuestros resultados estructurales no han podido someter a prueba esta posibilidad, ya que solo hemos obtenido estructuras de NtcA unidas a DNA cuando 2OG estuvo también unido, a pesar de que hemos intentado ensayos de cristalización de mezclas de DNA y NtcA en ausencia de 2OG (no reportado aquí por brevedad). Sin embargo, para el regulador transcripcional PrfA perteneciente también a esta familia se han

obtenido estructuras en complejo con DNA sin efector alostérico (GSH) y con GSH sin DNA, y ambas estructuras son iguales (Hall *et al.*, 2016), algo que aplicaría también para NtcA.

Existen promotores de NtcA que presentan cambios en la secuencia consenso perfecta (**GTAN₈TAC**) y a los que denominamos promotores no canónicos. Normalmente los cambios se dan solo en un sitio del consenso (en GTA o en TAC) ya que la unión del factor de transcripción a la secuencia consenso es secuencial, empezando primero en una de las dos subunidades de NtcA y uniéndose después la otra, por lo que al menos una de ellas necesita tener alta afinidad por la secuencia de unión (Jiang *et al.*, 2000). Esto es interesante porque un cierto grado de asimetría es importante para la diversidad funcional ya que genes con motivos de unión que posean diferentes afinidades por el factor de transcripción pueden ser regulados de forma distinta bajo diferentes condiciones fisiológicas (De Crombrughe *et al.*, 1984).

Los cambios más frecuentes en la secuencia consenso de NtcA se dan en las posiciones -5/5 (Mitschke *et al.*, 2011), de hecho, la secuencia **GTN₁₀AC** ha sido considerado esencial (Herrero *et al.*, 2004), aunque con preferencia por una A en la posición -5 y por una T en la posición 5 (Picossi *et al.*, 2014). Se han descrito cambios también en la posición -6/6 de la secuencia consenso, siendo los cambios en la posición -7/7 (el par de bases que interaccionan con R186) extremadamente raros (y solo en un sitio) (Mitschke *et al.*, 2011b; Su *et al.*, 2005). En cualquier caso, todos estos cambios en la secuencia canónica resultan en disminución de la afinidad de NtcA por el promotor.

Nosotros estudiamos por SPR en *Synechocystis* la afinidad que presenta NtcA por uno de estos promotores no canónicos: el promotor del gen *glnN*, cuya secuencia consenso es **GTAN₈GTC** en *Synechocystis* y **GTAN₈AGC** en *S. elongatus*. Al comparar la afinidad de NtcA por *glnN* con la que presenta por el promotor con secuencia consenso perfecta *glnA* (en *Synechocystis*) observamos que efectivamente NtcA tiene mucha mayor afinidad por el promotor *glnA* que por *glnN*, tanto en presencia

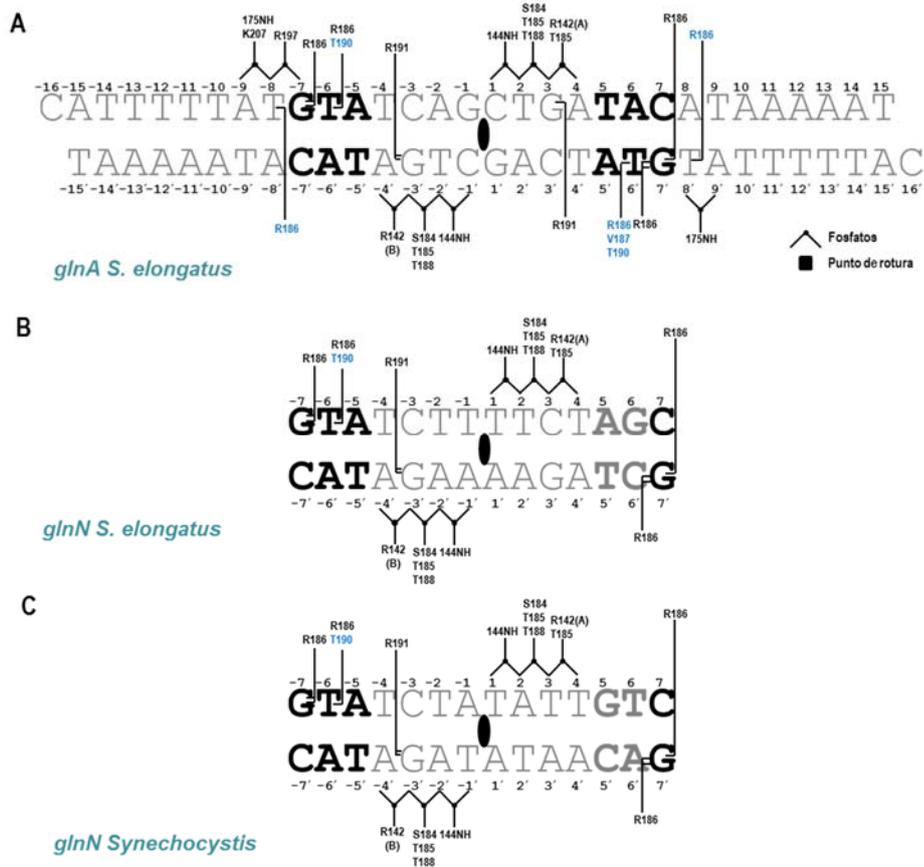


Figura 52. Diferencia entre las interacciones formadas entre NtcA y el promotor del gen *glnA* de *S. elongatus* según nuestra estructura (A) y las interacciones que formaría con el promotor del gen *glnN* en las cianobacterias *S. elongatus* (B) y *Synechocystis* (C). La caja mostrada para NtcA es la secuencia consenso idealizada utilizada para cristalizar el complejo. Sin embargo, en el caso de *glnN* se muestra la secuencia fisiológica. Se resaltan en negro las bases conservadas del consenso.

como en ausencia de 2OG (ver Tabla 6).

Esta menor afinidad por *glnN* se podría explicar observando la estructura de NtcA-DNA, a través de la cual podemos inferir la pérdida de las interacciones entre este promotor (tanto en *S. elongatus* como en *Synechocystis*) con los residuos R191, R186, V187 y T190 de NtcA (Fig. 52). Además de estos contactos, también se perdería el paso pirimidina-purina canónico “TA” de la segunda parte del consenso (posiciones 5 y 6) por un paso purina-purina o purina-pirimidina en *S. elongatus* y *Synechocystis* respectivamente, los cuales tienen menos susceptibilidad energética a la deformación del DNA (Olson *et al.*, 1998).

Los genes *glnA* y *glnN* codifican dos isoformas distintas de la glutamina sintetasa, la cual es un enzima clave para la asimilación de amonio. Las cianobacterias asimilan nitrógeno desde amonio principalmente a través de la ruta glutamina sintetasa-glutamato sintasa (GS-GOGAT) siendo la GS el enzima que cataliza la incorporación de amonio en glutamato para producir glutamina. Se han caracterizado en cianobacterias dos tipos de GS, la clásica y constitutiva (tipo I) codificada por el gen *glnA*, y una segunda GS llamada GS tipo III, codificada por el gen *glnN*, que se encuentra en cianobacterias no diazótrofas (Aldehni & Forchhammer, 2006). La GS tipo III parece jugar un papel importante en la aclimatación a condiciones de carencia de nitrógeno. Por ejemplo, se ha visto en *S. elongatus* que un mutante deficiente en *glnN* se recupera más lentamente de períodos prolongados de carencia de nitrógeno que las células silvestres (Sauer *et al.*, 2000). Además, sólo bajo condiciones de disminución de nitrógeno es posible detectar la expresión de *glnN* (Aldehni & Forchhammer, 2006), habiendo sido incluso imposible

detectar la unión de NtcA a este promotor en ausencia de 2OG (Reyes *et al.*, 1997). Nosotros sí observamos unión de NtcA a *glnN* en ausencia de 2OG, pero con una afinidad muy baja, y, de hecho, NtcA presenta una afinidad más alta por *pglnA* en ausencia de 2OG que por *pglnN* en presencia de este efector (ver Tabla 6). Este dato también podría indicar que la unión de NtcA a sus promotores en ausencia de 2OG ocurriría también *in vivo* (Giner-Lamia *et al.*, 2017) y, por lo tanto, las consecuencias de esta unión independiente de 2OG deberían explorarse, particularmente desde que estudios en otras cianobacterias marinas como *Prochlorococcus* (Domínguez-Martín *et al.*, 2017), muestran una menor sensibilidad al 2OG en estas cianobacterias con respecto a la que poseen las de agua dulce como *Synechocystis* o *S. elongatus*.

Por otra parte, se ha observado que la transcripción del gen *glnN* requiere de la presencia de la proteína PipX para su máxima expresión (Espinosa *et al.*, 2006), apoyando la visión de que PipX estabiliza la conformación activa de NtcA (Llácer *et al.*, 2010). De hecho, *glnN* solo se transcribe cuando hay una severa carencia de nitrógeno (Aldehni & Forchhammer, 2006; Reyes *et al.*, 1997), que es cuando más PipX está disponible para interaccionar con NtcA. Los resultados de SPR también muestran claramente cómo la presencia de PipX aumenta la afinidad efectiva de NtcA por el 2OG dependiendo del promotor, siendo más importante en *pglnN* que en *pglnA* (incrementa 9 y 3 veces su afinidad, respectivamente) (Tabla 7), lo que podría estar relacionado con la diferente sensibilidad a PipX de los distintos genes dependientes de NtcA (Espinosa *et al.*, 2007). De acuerdo con todo ello, se ha visto en un mutante deficiente en *pipX* cómo la inducción transcripcional de *glnN* es más lenta, sugiriendo que PipX permite una adaptación más rápida y eficiente a la carencia de

nitrógeno (Giner-Lamia *et al.*, 2017). En conclusión, nuestros resultados usando SPR, sugieren que los genes que presentan sitios de unión para NtcA menos "perfectos", como es el caso de *glnN*, requieren de una forma más importante las presencias de PipX y de 2OG, ya que gracias a ellos estaría aumentando la concentración de NtcA en su conformación competente para unirse al promotor. Esta modulación de la afinidad de NtcA por 2OG dependiendo no solo de la naturaleza del promotor sino también de la presencia o ausencia de PipX, añade una riqueza extra al sistema regulatorio PipX-2OG-NtcA.

La red de regulación de PII-PipX, y las relaciones con nuevos elementos de dicha red

Los resultado de los experimentos de RMN (Forcada-Nadal *et al.*, 2017) que hemos propiciado y hecho posibles y que hemos resumido en el Capítulo 3 de los Resultados proporcionan por primera vez una visión de PipX en solución, libre de complejos con PII o NtcA. La información estructural sobre PipX sola prueba que su hélice C-terminal está flexionada, explicando por qué PipX sola, sin formar complejo con PII, no es un ligando de PlmA (Labella *et al.*, 2016). Así, este hallazgo fundamenta la idea de que PII, en cuyo complejo con PipX se encontró la hélice C-terminal extendida en dos de las tres moléculas de PipX presentes en el complejo, actúa como "abridor" de la hélice C-terminal de PipX, función que no realiza NtcA al unirse a PipX, pues esta última proteína permanece con la hélice C-terminal "flexionada" en su complejo con NtcA (Llácer *et al.*, 2010). La información dinámica de RMN relativa a la exposición de distintas partes de PipX y sobre la movilidad de ciertos residuos de esta proteína ayuda a proporcionar explicaciones al efecto de

“apertura“ que tiene PII sobre la misma. Este efecto podría estar aumentado por la presencia simultánea de PlmA, ya que esta proteína, por su interacción con la hélice extendida de PipX en el complejo PII-PipX, atraparía la conformación abierta de PipX. Cabría también plantearse si la trimerización impuesta al complejo por PII pudiera ser un factor importante en la formación del complejo con PlmA.

Las evidencias existentes en favor de regulación de la expresión génica por PipX sin que actúe NtcA como mediador (Espinosa *et al.*, 2014) posiblemente sean explicables por la formación del complejo ternario PipX-PII-PlmA (Labella *et al.*, 2016). PlmA es un posible regulador transcripcional (Labella *et al.*, 2016) pero sus funciones y genes diana todavía son desconocidos y por lo tanto no existe aún una respuesta sobre qué genes serían regulados por el complejo ternario de las tres proteínas. En todo caso se ha establecido que el número de genes cuya expresión aumenta en ausencia de PipX es mayor que el número de los genes cuya expresión disminuye (Espinosa *et al.*, 2014), lo que sugiere que predomina la regulación negativa por PipX, algo totalmente diferente a lo que sucede para la regulación dependiente de NtcA, que es fundamentalmente positiva. Si en el complejo ternario PII-PipX-PlmA el dominio de unión a DNA de PlmA está secuestrado (ver más adelante), se explicaría que la interacción de PlmA con PipX tuviera un efecto inhibitorio de la expresión del regulón de PlmA.

Se puede plantear la pregunta de si todas las funciones de regulación por PipX que son NtcA-independientes podrían explicarse por la formación del complejo ternario PipX-PII-PlmA (Labella *et al.*, 2016) o si PipX sola podría tener funciones adicionales. Claramente el estudio de cuáles son

las funciones específicas de PlmA en cianobacterias va a ser un requisito para responder a esta cuestión.

Nuestra cuantificación mediante western blot de los niveles de PlmA, PII, PipX y PipY en *S. elongatus* ha arrojado datos muy similares a los encontrados en un estudio no focalizado de proteómica de *S. elongatus* (Guerreiro *et al.*; 2014) (Tabla 10). Los datos cuantitativos indican que PII es ~14 veces más abundante que PipX y este a su vez es ~7 y ~4 veces más abundante que PlmA y PipY, respectivamente. El estudio de proteómica también proporciona datos sobre NtcA y NAGK (Tabla 10) revelando que están presentes en cantidades molares muy pequeñas en relación a sus respectivas dianas PipX y PII (Tabla 10 y Fig. 53).

Tabla 10. Resumen de las abundancias relativas (como subunidades) de las diferentes proteínas consideradas aquí en la red reguladora de PII-PipX, según nuestros datos (Western blot) y el estudio de Proteómica (Guerreiro *et al.*, 2014)

Proteína	PII	PipX	PlmA	PipY	NtcA	NAGK
Western blot ^a	100	7	1	1.6	-	-
Proteómica ^b	100	7.6	0.7	2	1.7	3.6

^aDatos obtenidos de la tabla 8, a partir del cálculo del nivel de cada una de las proteínas como subunidades en pmol por mg de proteína total.

^bGuerreiro *et al.*; 2014

Dada la estequiometría 1:1 (como subunidades) entre PII y sus proteínas diana en los complejos PII-NAGK y PII-PipX (Llácer *et al.*, 2007, 2010), la reserva de PII parece más que suficiente para secuestrar toda la PipX y toda la NAGK presente en las células. Por otra parte, el exceso molar > 4 de subunidades de PipX sobre subunidades de NtcA sugiere también la posibilidad de que PipX tenga otras dianas además de NtcA, como es el caso de PlmA. Incluso, como el nivel molar de PipX excede por >3 veces

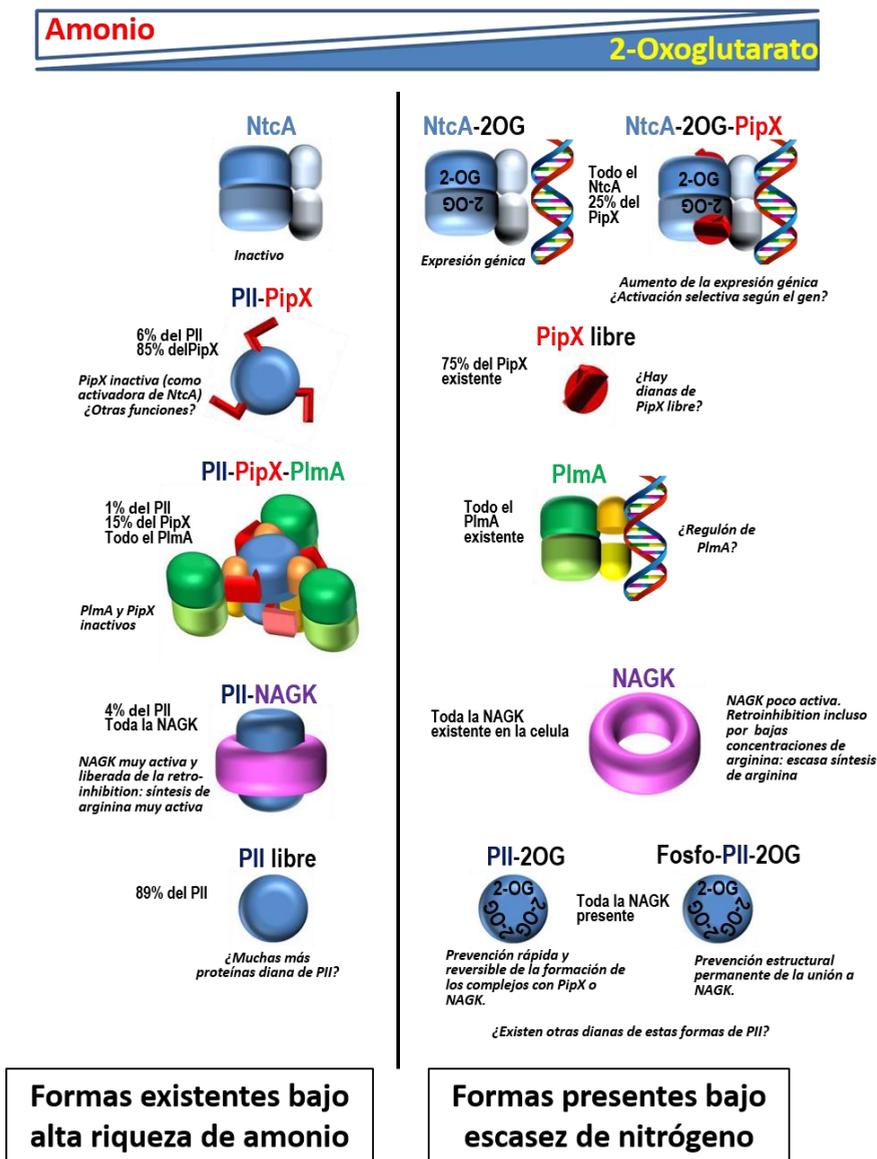


Figura 53. Complejos de proteínas del sistema regulador de PII en *S. elongatus* según la disponibilidad de amonio en la célula. La distribución de cada proteína entre los diferentes complejos se basa en sus niveles en la célula presentados en este trabajo y registrados también en experimentos de proteómica masiva (Guerreiro *et al.*, 2014). El trímero de PipX se ha coloreado en azul. Las hélices C-terminales de PipX (esquemáticas en la conformación extendida) se han coloreado en rojo. Los dominios N-terminal de los dímeros de PlmA se muestra amarillo o naranja mientras que los dominios C-terminal de PlmA se han coloreado en dos tonos de verde. NAGK se muestra en color morado. Los dominios N y C-terminal de NtcA se representa en dos tonos de azul, correspondiendo el más claro de ellos al dominio C-terminal de unión a DNA.

la suma de los niveles molares de subunidades de NtcA y PlmA, parece probable la existencia de dianas adicionales y todavía no identificadas de PipX. Estas inferencias estarían de acuerdo con resultados de transcriptómica masiva (Espinosa *et al.*, 2014) que apoyan el que PipX pueda tener funciones de regulación génica independientes de NtcA y de PII (y por tanto, independientes de PlmA).

En resumen, en condiciones de abundancia de amonio PII actuaría como sumidero tanto para PipX como para PlmA, además de ser un activador permanente de toda la NAGK presente en la célula, aumentando la síntesis

celular de arginina. A su vez, el complejo PII-PipX unirá la mayoría del PlmA existente en la célula (Fig. 53). Como el ADP incrementa de forma importante la afinidad de PII por PipX (Llácer *et al.*, 2010; Zeth *et al.*, 2014), este nucleótido aumentaría el papel secuestrador de PipX por parte de PII (Zeth *et al.*, 2014). El gran exceso molar de PII sobre PipX y NAGK y de PipX sobre PlmA sugiere que en estas condiciones de bajo 2OG, tanto PipX como NAGK y PlmA se encontrarán mayoritariamente formando parte de sus complejos (Fig. 53) (Maheswaran *et al.*, 2004; Labella *et al.*, 2016), mientras que, dada su abundancia, la mayoría de PII seguirá estando libre y disponible para otras interacciones (como se ha descrito con PII y algunos transportadores o la enzima acetyl-CoA carboxilasa (Hisbergues *et al.*, 1999; Lee *et al.*, 2000; Hauf *et al.*, 2016), haciendo posible que PII regule simultáneamente una multiplicidad de dianas distintas.

Cuando las condiciones cambian a fuerte limitación de nitrógeno y los niveles de 2OG se elevan en *S. elongatus*, PII se debe liberar de NAGK y

de PipX y PlmA, lo que debe resultar en una disminución importante en la síntesis de arginina y en una superactivación de NtcA por la acción combinada de 2OG y PipX (Figs. 53 y 54), así como quizá también en la activación del aún sin caracterizar regulón de PlmA, pues la descomposición del complejo PII-PipX debe conllevar obligatoriamente la descomposición del complejo ternario PII-PipX-PlmA. La posibilidad de que el PlmA unido a PII-PipX sea inactivo como regulador transcripcional parece plausible sobre la base de la estructura propuesta para este complejo (Labella *et al.*, 2016). Puesto que residuos de ambas hélices C-terminales de PipX y la Arg35 presente en la parte expuesta del dominio tudor (expuesta en el complejo con PII) están involucrados centralmente en las interacciones con PlmA, PlmA sería atrapado por PipX entre sus hélices C-terminales extendidas y la parte expuesta de su dominio tudor. Dado su pequeño tamaño, el dominio N-terminal de PlmA es el mejor candidato para quedar atrapado por PipX (Fig. 53), dejándolo inaccesible para interactuar con el DNA, inactivando así a PlmA como regulador transcripcional. La deconstrucción de este complejo ternario al unirse a PII el ATP y el 2OG liberaría PlmA, que así quedaría disponible para interactuar con el DNA y para desempeñar sus funciones de regulación génica. Aunque no se ha caracterizado el regulón de PlmA, este podría ser amplio dada la cantidad relativamente alta de moléculas de PlmA (aproximadamente 200 dímeros por célula, (Guerreiro *et al.*, 2014)), similar al de un factor como el NtcA (~500 dímeros/célula (Guerreiro *et al.*, 2014)), que controla un extenso regulón. De este modo la actividad transcripcional podría regularse por PII no solo vía NtcA sino también vía PlmA, subrayando el papel de PII como una pieza clave en el control de la actividad transcripcional en respuesta a las señales

representadas por el 2OG y el par ATP/ADP. Así, en respuesta al 2OG existiría una regulación común por PII/PipX para NtcA y PlmA, aunque con una importante diferencia en cuanto al mecanismo: en el caso de NtcA la unión de PipX lo activa (complejos NtcA-PipX) mientras que a PlmA la unión de PipX (en el complejo con PII) lo inhibe.

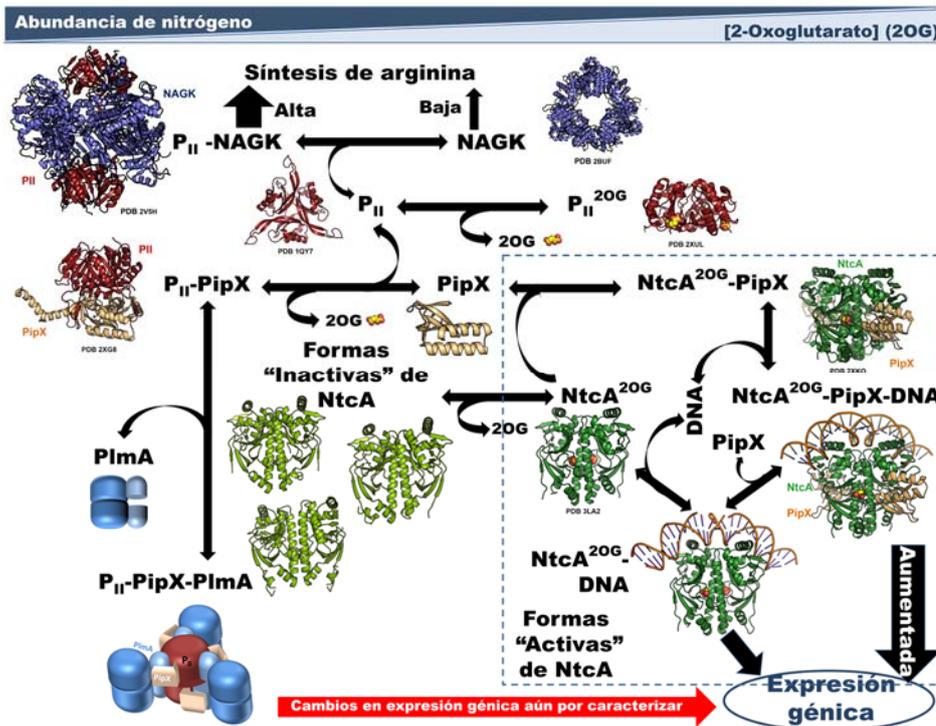


Figura 54. Sistema regulador PII-PipX cianobacteriano con inclusión de las estructuras y los avances realizados en el contexto de este trabajo de tesis. No se ha incluido a PipY, cuyo lugar en la red aún no se ha definido bien. De las interacciones de PII con sus dianas en cianobacterias que no involucran a PipX solo se ha incluido la interacción con NAGK por su alta caracterización estructural y funcional, no reflejándose la interacción recientemente demostrada con la acetil-CoA carboxilasa.

Una cuestión adicional referida a regulación por PII es el hecho de que en *S.elongatus* PII puede fosforilarse *in vivo* en la serina49, aboliendo dicha fosforilación la formación del complejo de PII con NAGK (Llácer *et al.*, 2007). Por el contrario, esta fosforilación no parece tener influencia sobre

la formación del complejo PII-PipX, (Llácer *et al.*, 2010) lo que hacía altamente probable que fuera también irrelevante para la regulación de PlmA, como se probó experimentalmente con mutantes de PII que mimetizan la fosforilación pero que no abolen la formación del complejo ternario, demostrada mediante ensayos de triple híbrido (Labella *et al.*, 2016).

Aunque no se representa en las Figs. 53 o 54 por no existir evidencias de la formación de un complejo físico con PipX (Labella *et al.*, 2017), las evidencias presentadas aquí dejan bien patente la existencia de co-regulación en la expresión de *pipX* y *pipY*. Estas evidencias, el hecho de que ambas proteínas formen un operón, y las evidencias fenotípicas publicadas en (Labella *et al.*, 2017), en particular el hecho de que PipX y PipY contribuyan a la resistencia a antibióticos que son venenos de PLP (por ejemplo D-clicloserina) y la similitud de efectos de ambas sobre la expresión de genes diana de NtcA en experimentos de genómica masiva con *S. elongatus*, son base suficiente para inscribir a PipY dentro de la red relacional de PipX. No siendo PipY un objetivo central de los estudios de esta tesis, dejaré de mencionar investigaciones ulteriores sobre esta proteína en que me he involucrado centralmente, mencionando solo que tanto los detalles moleculares de la interacción funcional entre PipX y PipY, como de la función misma de PipY como máquina molecular distan de estar aclaradas, siendo preciso todavía más trabajo experimental para definir las con precisión. Además, el descubrimiento de PipY como parte de la red de interacción/relaciones/funciones de PipX (Labella *et al.*, 2017) hace concebible que puedan quedar aún otros elementos de dicha red por descubrir, lo que concita a dedicarle aún esfuerzos experimentales adicionales.

Conclusiones

1. Nuestra determinación de la estructura tridimensional del complejo 2-oxoglutarato (2OG)-NtcA-DNA aclara las razones de la especificidad de NtcA por sus secuencias diana, identifica los residuos invariantes R186, V187, T190 y R191 como claves en el reconocimiento por interactuar con las bases del DNA, también identifica similitudes y diferencias con complejos similares de otros miembros de la familia CRP, y racionaliza las peculiaridades regulatorias de NtcA.
2. Nuestra determinación de la estructura del complejo PipX-NtcA-2OG-DNA prueba que PipX es un co-activador de NtcA no por formar nuevos puentes mediados por PipX con el el DNA diana, sino porque la unión de PipX a NtcA induce un aumento de contactos de NtcA con DNA y porque, además, estabiliza el conformero "activo" de NtcA.
3. Nuestra determinación de varias estructuras de NtcA de *S. elongatus* en ausencia de su efector 2OG demuestra la existencia de varias conformaciones apo, "inactivas", de NtcA, lo que se contrapone con la observación de una única conformación "activa" de NtcA presente monótonamente en los complejos de NtcA con su activador 2OG, con este y con el coactivador PipX o con DNA, o en el complejo cuaternario de NtcA con 2OG, PipX y NtcA. Aun así, la exposición de las hélices de unión al DNA en las formas "inactivas," así como ciertos cambios de secuencia en porciones de no interacción con el DNA que mimetizan mutaciones auto-activadoras del paradigma de la familia de estos factores de transcripción, CPR, explican el hecho de que, a diferencia de CRP, NtcA sea capaz de unirse a sus DNAs diana en ausencia de su activador 2OG.

4. Nuestro examen de interacciones cruzadas entre NtcA y CRP con sus DNAs diana mediante experimentos de resonancia de plasmón (SPR) demuestra que, al contrario que CPR para las cajas NtcA, NtcA puede unirse a los promotores diana de CRP, probando también que CRP no se une a su promotor en ausencia de su efector cAMP mientras que NtcA sí lo hace en ausencia de su efector 2OG, aunque con afinidad disminuida.

5. Estos experimentos de SPR prueban también que la afinidad efectiva del NtcA para 2OG está fuertemente influenciada por el promotor al que se une NtcA, y que PipX aumenta esta afinidad de forma importante. Confirman que PipX no se une directamente al DNA, y que solo se une a NtcA en presencia de 2OG.

6. Los resultados de SPR corroboran funcionalmente las inferencias extraídas de la observación de las estructuras de los complejos de NtcA con 2OG, DNA y PipX; apoyan la idea de que PipX estabiliza la conformación "activa" de NtcA con 2OG unido; y aportan evidencias de que la afinidad de PipX por NtcA es mucho mayor que la reportada previamente para la unión de PipX a la proteína PII. Este último dato, junto con la estimación de los niveles de proteínas del sistema de regulación de PipX, podrían ayudar a la modelización cuantitativa de la red de interacciones del regulón PII-PipX-NtcA en *S. elongatus*.

7. Nuestra contribución a los estudios de resonancia magnética nuclear de PipX no unida a sus parejas permite concluir que PipX solo adopta la conformación flexionada que es no competente para unirse a PlmA, dando soporte a la idea de que PII actúa de "abridor" de la hélice C-terminal de PipX, permitiendo interacciones nuevas en ese complejo como las que tienen lugar con la proteína PlmA.

8. Nuestros estudios de inmunocuantificación de PII, PipX y PlmA, junto con resultados cuantitativos de otros con una aproximación de proteómica masiva, permiten proponer qué complejos de estas proteínas existirán en la célula en condiciones de pobreza o de abundancia de amonio. Estos mismos estudios para las proteínas PipX y PipY en varias cepas modificadas de *S. elongatus* han sido clave para identificar una nueva y novedosa función de PipY.

Bibliografía

- Abed RMM, Dobretsov S & Sudesh K (2009)** Applications of cyanobacteria in biotechnology. *J Appl Microbiol* **106**, 1–12.
- Adler SP, Purich D & Stadtman ER (1975)** Cascade control of *Escherichia coli* glutamine synthetase. *J Biol Chem* **250**, 6264–6272.
- Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH & Adams PD (2012)** Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367.
- Agari Y, Kashihara A, Yokoyama S, Kuramitsu S & Shinkai A (2008)** Global gene expression mediated by *Thermus thermophilus* SdrP, a CRP/FNR family transcriptional regulator. *Mol Microbiol* **70**, 60–75.
- Agari Y, Kuramitsu S & Shinkai A (2012)** X-ray crystal structure of TTHB099, a CRP/FNR superfamily transcriptional regulator from *Thermus thermophilus* HB8, reveals a DNA-binding protein with no required allosteric effector molecule. *Proteins* **80**, 1490–1494.
- Aldehni MF & Forchhammer K (2006)** Analysis of a non-canonical NtcA-dependent promoter in *Synechococcus elongatus* and its regulation by NtcA and PII. *Arch Microbiol* **184**, 378–386.
- Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie AGW (2011)** iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.
- Beck C, Knoop H, Axmann IM & Steuer R (2012)** The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics* **13**, 56.
- Bell A & Busby S (1994)** Location and orientation of an activating region in the *Escherichia coli* transcription factor, FNR. *Mol Microbiol* **11**, 383–390.
- Bender DA (2012)** *Amino Acid Metabolism*, 3rd Ed. Chichester, UK. Wiley-Blackwell
- Benoff B, Yang H, Lawson CL, Parkinson G, Liu J, Blatter E, Ebright YW, Berman HM & Ebright RH (2002)** Structural basis of

transcription activation: the CAP- α CTD-DNA complex. *Science* **297**, 1562–1566.

Berg OG & von Hippel PH (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J Mol Biol* **200**, 709–723.

Bonnet M, Kurz M, Mesa S, Briand C, Hennecke H & Grütter MG (2013) The structure of Bradyrhizobium japonicum transcription factor FixK2 unveils sites of DNA binding and oxidation. *J Biol Chem* **288**, 14238–14246.

Borjigin M, Li H, Lanz ND, Kerby RL, Roberts GP & Poulos TL (2007) Structure-based hypothesis on the activation of the CO-sensing transcription factor CooA. *Acta Crystallogr D Biol Crystallogr* **63**, 282–287.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

Broddrick JT, Rubin BE, Welkie DG, Du N, Mih N, Diamond S, Lee JJ, Golden SS & Palsson BO (2016) Unique attributes of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and essential gene analysis. *Proc Natl Acad Sci U S A* **113**, 8344–8353.

Browning DF & Busby SJ (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**, 57–65.

Burillo S, Luque I & Fuentes I (2004) Interactions between the nitrogen signal transduction protein PII and N -acetyl glutamate kinase in organisms that perform oxygenic photosynthesis. *J Bacteriol* **186**, 3346–3354.

Carr PD, Cheah E, Suffolk PM, Vasudevan SG, Dixon NE & Ollis DL (1996) X-ray structure of the signal transduction protein from *Escherichia coli* at 1.9 Å. *Acta Crystallogr D Biol Crystallogr* **52**, 93–104.

Cheah, E., Carr, P., Suffolk, P., and Vasudevan, S (1994) Structure of the *Escherichia coli* signal transducing protein PII. *Structure* **2**, 981-90

Chellamuthu V-R, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD & Forchhammer K (2014) A widespread glutamine-sensing mechanism in the plant kingdom. *Cell* **159**, 1188–1199.

Chen S, Vojtechovsky J, Parkinson GN, Ebright RH & Berman HM (2001a) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: DNA binding specificity based on energetics of DNA kinking. *J Mol Biol* **314**, 63–74.

Chen S, Gunasekera A, Zhang X, Kunkel TA, Ebright RH & Berman HM (2001b) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: alteration of DNA binding specificity through alteration of DNA kinking. *J Mol Biol* **314**, 75–82.

Chin K-H, Lee Y-C, Tu Z-L, Chen C-H, Tseng Y-H, Yang J-M, Ryan RP, McCarthy Y, Dow JM, Wang AH-J & Chou S-H (2010) The cAMP receptor-like protein CLP is a novel c-di-GMP receptor linking cell-cell signaling to virulence gene expression in *Xanthomonas campestris*. *J Mol Biol* **396**, 646–662.

Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.

Commichau FM, Forchhammer K & Stülke J (2006) Regulatory links between carbon and nitrogen metabolism. *Curr Opin Microbiol* **9**, 167–172.

Conroy MJ, Durand A, Lupo D, Li X-D, Bullough PA, Winkler FK & Merrick M (2007) The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. *Proc Natl Acad Sci U S A* **104**, 1213–1218.

Darin N, Reid E, Prunetti L, Samuelsson L, Husain RA, Wilson M, El Yacoubi B, Footitt E, Chong WK, Wilson LC, Prunty H, Pope S, Heales S, Lascelles K, Champion M, Wassmer E, Veggiotti P, de Crécy-Lagard V, Mills PB & Clayton PT (2016) Mutations in PROSC Disrupt Cellular Pyridoxal Phosphate Homeostasis and Cause Vitamin-B₆-Dependent Epilepsy. *Am J Hum Genet* **99**, 1325–1337.

Deaconescu AM, Chambers AL, Smith AJ, Nickels BE, Hochschild A, Savery NJ & Darst SA (2006) Structural basis for bacterial transcription-coupled DNA repair. *Cell* **124**, 507–520.

Decker KB & Hinton DM (2013) Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annu Rev Microbiol* **67**, 113–139.

De Crombrughe B, Busby S & Buc H (1984) Cyclic AMP receptor protein: role in transcription activation. *Science* **224**, 831–838.

Domínguez-Martín MA, López-Lozano A, Clavería-Gimeno R, Velázquez-Campoy A, Seidel G, Burkovski A, Díez J & García-Fernández JM (2017) Differential NtcA Responsiveness to 2-oxoglutarate underlies the diversity of C/N balance regulation in *Prochlorococcus*. *Front Microbiol* **8**, 2641.

Ebright RH, Kolb A, Buc H, Kunkel TA, Krakow JS & Beckwith J (1987) Role of glutamic acid-181 in DNA-sequence recognition by the catabolite gene activator protein (CAP) of *Escherichia coli*: altered DNA-sequence-recognition properties of [Val181]CAP and [Leu181]CAP. *Proc Natl Acad Sci U S A* **84**, 6083–6087

Ebright RH, Ebright YW & Gunasekera A (1989) Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli* lac DNA site. *Nucleic Acids Res* **17**, 10295–10305.

Eiting M, Hagelüken G, Schubert W-D & Heinz DW (2005) The mutation G145S in PrfA, a key virulence regulator of *Listeria monocytogenes*, increases DNA-binding affinity by stabilizing the HTH motif. *Mol Microbiol* **56**, 433–446.

Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.

Espinosa J, Forchhammer K, Burillo S & Contreras A (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Mol Microbiol* **61**, 457–469.

- Espinosa J, Forchhammer K & Contreras A (2007)** Role of the *Synechococcus* PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. *Microbiology* **153**, 711–718.
- Espinosa J, Castells MA, Laichoubi KB & Contreras A (2009)** Mutations at pipX suppress lethality of PII-deficient mutants of *Synechococcus elongatus* PCC 7942. *J Bacteriol* **191**, 4863–4869.
- Espinosa J, Castells MA, Laichoubi KB, Forchhammer K & Contreras A (2010)** Effects of spontaneous mutations in PipX functions and regulatory complexes on the cyanobacterium *Synechococcus elongatus* strain PCC 7942. *Microbiology* **156**, 1517–1526.
- Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F & Contreras A (2014)** PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. *Proc Natl Acad Sci U S A* **111**, E2423–30.
- Fadi Aldehni M, Sauer J, Spielhauer C, Schmid R & Forchhammer K (2003)** Signal transduction protein PII is required for NtcA-regulated gene expression during nitrogen deprivation in the cyanobacterium *Synechococcus elongatus* strain PCC 7942. *J Bacteriol* **185**, 2582–2591.
- Feng Y, Zhang Y & Ebricht RH (2016)** Structural basis of transcription activation. *Science* **352**, 1330–1333.
- Fokina O, Chellamuthu V-R, Forchhammer K & Zeth K (2010)** Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. *Proc Natl Acad Sci U S A* **107**, 19760–19765.
- Forcada-Nadal A, Forchhammer K & Rubio V (2014)** SPR analysis of promoter binding of *Synechocystis* PCC6803 transcription factors NtcA and CRP suggests cross-talk and sheds light on regulation by effector molecules. *FEBS Lett* **588**, 2270–2276.
- Forcada-Nadal A, Palomino-Schätzlein M, Neira JL, Pineda-Lucena A & Rubio V (2017)** The PipX protein, when not bound to its targets, has its signaling C-terminal helix in a flexed conformation. *Biochemistry* **56**, 3211–3224.
- Forchhammer K & Tandeau de Marsac N (1994)** The PII protein in the cyanobacterium *Synechococcus* sp. strain PCC 7942 is modified by

serine phosphorylation and signals the cellular N-status. *J Bacteriol* **176**, 84–91.

Forchhammer K (2004) Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol Rev* **28**, 319–333.

Forchhammer K (2008) PII signal transducers: novel functional and structural insights. *Trends Microbiol* **16**, 65–72.

Fujimori T, Higuchi M, Sato H, Aiba H, Muramatsu M, Hihara Y & Sonoike K (2005) The mutant sll1961, which encodes a putative transcriptional regulator, has a defect in regulation of photosystem stoichiometry in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* **139**, 408–416.

Gallagher DT, Smith N, Kim S-K, Robinson H & Reddy PT (2009) Profound asymmetry in the structure of the cAMP-free cAMP Receptor Protein (CRP) from *Mycobacterium tuberculosis*. *J Biol Chem* **284**, 8228–8232.

García De La Torre J, Huertas ML & Carrasco B (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys J* **78**, 719–730.

García-Domínguez M & Florencio FJ (1997) Nitrogen availability and electron transport control the expression of *glnB* gene (encoding PII protein) in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Mol Biol* **35**, 723–734

García-Domínguez M, Reyes JC & Florencio FJ (2000) NtcA represses transcription of *gifA* and *gifB*, genes that encode inhibitors of glutamine synthetase type I from *Synechocystis* sp. PCC 6803. *Mol Microbiol* **35**, 1192–1201.

Ghosh T, Bose D & Zhang X (2010) Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev* **34**, 611–627.

Giardina G, Rinaldo S, Castiglione N, Caruso M & Cutruzzolà F (2009) A dramatic conformational rearrangement is necessary for the activation of DNR from *Pseudomonas aeruginosa*. Crystal structure of wild-type DNR. *Proteins* **77**, 174–180.

- Giner-Lamia J, Robles-Rengel R, Hernández-Prieto MA, Muro-Pastor MI, Florencio FJ & Futschik ME (2017)** Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp. PCC 6803. *Nucleic Acids Res* **45**, 11800–11820.
- Gray MW, Burger G & Lang BF (2001)** The origin and early evolution of mitochondria. *Genome Biol* **2**, Reviews1018.
- Gruswitz F, O’Connell J & Stroud RM (2007)** Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. *Proc Natl Acad Sci U S A* **104**, 42–47.
- Guerreiro ACL, Benevento M, Lehmann R, van Breukelen B, Post H, Giansanti P, Maarten Altelaar AF, Axmann IM & Heck AJR (2014)** Daily rhythms in the cyanobacterium *synechococcus elongatus* probed by high-resolution mass spectrometry-based proteomics reveals a small defined set of cyclic proteins. *Mol Cell Proteomics* **13**, 2042–2055.
- Gümral D, Nadalin L, Corazza A, Fogolari F, Damante G, Viglino P & Esposito G (2008)** Helix mobility and recognition function of the rat thyroid transcription factor 1 homeodomain - hints from ¹⁵N-NMR relaxation studies. *FEBS J* **275**, 435–448.
- Habazettl J, Myers LC, Yuan F, Verdine GL & Wagner G (1996)** Backbone dynamics, amide hydrogen exchange, and resonance assignments of the DNA methylphosphotriester repair domain of *Escherichia coli* Ada using NMR. *Biochemistry* **35**, 9335–9348.
- Hall M, Grundström C, Begum A, Lindberg MJ, Sauer UH, Almqvist F, Johansson J & Sauer-Eriksson AE (2016)** Structural basis for glutathione-mediated activation of the virulence regulatory protein PrfA in *Listeria*. *Proc Natl Acad Sci U S A* **113**, 14733–14738.
- Hamilton TL, Bryant DA & Macalady JL (2016)** The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans. *Environ Microbiol* **18**, 325–340
- Hart DJ, Speight RE, Cooper MA, Sutherland JD & Blackburn JM (1999)** The salt dependence of DNA recognition by NF-kappaB p50: a detailed kinetic analysis of the effects on affinity and specificity. *Nucleic Acids Res* **27**, 1063–1069.

- Hauf W, Schmid K, Gerhardt ECM, Huergo LF & Forchhammer K (2016)** Interaction of the nitrogen regulatory protein GlnB (PII) with biotin carboxyl carrier protein (BCCP) controls Acetyl-CoA levels in the cyanobacterium *Synechocystis* sp. PCC 6803. *Front Microbiol* **7**, 1700.
- Heinrich A, Maheswaran M, Ruppert U & Forchhammer K (2004)** The *Synechococcus elongatus* PII signal transduction protein controls arginine synthesis by complex formation with N-Acetyl-L-glutamate kinase. *Mol Microbiol* **52**, 1303–1314.
- Herrero A, Muro-Pastor AM & Flores E (2001)** Nitrogen control in cyanobacteria. *J Bacteriol* **183**, 411–425.
- Herrero A, Muro-Pastor AM, Valladares A & Flores E (2004)** Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev* **28**, 469–487.
- Hisbergues M, Jeanjean R, Joset F, Tandeau de Marsac N & Bédu S (1999)** Protein PII regulates both inorganic carbon and nitrate uptake and is modified by a redox signal in *Synechocystis* PCC 6803. *FEBS Lett* **463**, 216–220.
- Hohmann-Marriott MF & Blankenship RE (2011)** Evolution of photosynthesis. *Annu Rev Plant Biol* **62**, 515–548.
- Hoskisson PA & Rigali S (2009)** Chapter 1: Variation in form and function the helix-turn-helix regulators of the GntR superfamily. *Adv Appl Microbiol* **69**, 1–22.
- Huergo LF & Dixon R (2015)** The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol Mol Biol Rev* **79**, 419–435.
- Irmeler A & Forchhammer K (2001)** A PP2C-type phosphatase dephosphorylates the PII signaling protein in the cyanobacterium *Synechocystis* PCC 6803. *Proc Natl Acad Sci U S A* **98**, 12978–12983.
- Ito T, Iimori J, Takayama S, Moriyama A, Yamauchi A, Hemmi H & Yoshimura T (2013)** Conserved pyridoxal protein that regulates Ile and Val metabolism. *J Bacteriol* **195**, 5439–5449.
- Jain D (2015)** Allosteric control of transcription in GntR family of transcription regulators: A structural overview. *IUBMB Life* **67**, 556–563.

Jiang F, Wisén S, Widersten M, Bergman B & Mannervik B (2000) Examination of the transcription factor NtcA-binding motif by *in vitro* selection of DNA sequences from a random library. *J Mol Biol* **301**, 783–793.

Jiang P & Ninfa AJ (2009) Sensation and signaling of alpha-ketoglutarate and adenylylate energy charge by the *Escherichia coli* PII signal transduction protein require cooperation of the three ligand-binding sites within the PII trimer. *Biochemistry* **48**, 11522–11531.

Jin C, Marsden I, Chen X & Liao X (1998) Sequence specific collective motions in a winged helix DNA binding domain detected by ¹⁵N relaxation NMR. *Biochemistry* **37**, 6179–6187. Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD & Wishart DS (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* **31**, 3316–3319.

Joyce MG, Levy C, Gábor K, Pop SM, Biehl BD, Doukov TI, Ryter JM, Mazon H, Smidt H, van den Heuvel RHH, Ragsdale SW, van der Oost J & Leys D (2006) CprK crystal structures reveal mechanism for transcriptional control of halo-respiration. *J Biol Chem* **281**, 28318–28325.

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

Kettenberger H & Cramer P (2006) Fluorescence detection of nucleic acids and proteins in multi-component crystals. *Acta Crystallogr D Biol Crystallogr* **62**, 146–150

Kim J, Adhya S & Garges S (1992) Allosteric changes in the cAMP receptor protein of *Escherichia coli*: hinge reorientation. *Proc Natl Acad Sci U S A* **89**, 9700–9704.

Kolb A, Busby S, Buc H, Garges S & Adhya S (1993) Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem* **62**, 749–795.

Körner H, Sofia HJ & Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27**, 559–592.

- Krissinel E & Henrick K (2004)** Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256–2268.
- Kumar P, Joshi DC, Akif M, Akhter Y, Hasnain SE & Mande SC (2010)** Mapping conformational transitions in cyclic AMP receptor protein: crystal structure and normal-mode analysis of Mycobacterium tuberculosis apo-cAMP receptor protein. *Biophys J* **98**, 305–314.
- Labella JI, Obrebska A, Espinosa J, Salinas P, Forcada-Nadal A, Tremiño L, Rubio V & Contreras A (2016)** Expanding the cyanobacterial nitrogen regulatory network: The GntR-like regulator PlmA interacts with the PII-PipX complex. *Front Microbiol* **7**, 1677.
- Labella JI, Cantos R, Espinosa J, Forcada-Nadal A, Rubio V & Contreras A (2017)** PipY, a member of the conserved COG0325 family of PLP-binding proteins, expands the cyanobacterial nitrogen regulatory network. *Front Microbiol* **8**, 1244.
- Laemmli UK (1970)** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Laichoubi KB, Espinosa J, Castells MA & Contreras A (2012)** Mutational analysis of the cyanobacterial nitrogen regulator PipX. *PLoS ONE* **7**, e35845.
- Lanzilotta WN, Schuller DJ, Thorsteinsson MV, Kerby RL, Roberts GP & Poulos TL (2000)** Structure of the CO sensing transcription activator CooA. *Nat Struct Biol* **7**, 876–880.
- Laurent S, Chen H, Bédu S, Ziarelli F, Peng L & Zhang C-C (2005)** Nonmetabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci U S A* **102**, 9907–9912.
- Lawson CL, Swigon D, Murakami KS, Darst SA, Berman HM & Ebright RH (2004)** Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol* **14**, 10–20.
- Leduc J, Thorsteinsson MV, Gaal T & Roberts GP (2001)** Mapping CooA.RNA polymerase interactions. Identification of activating regions 2 and 3 in CooA, the co-sensing transcriptional activator. *J Biol Chem* **276**, 39968–39973.

- Lee DJ, Minchin SD & Busby SJW (2012)** Activating transcription in bacteria. *Annu Rev Microbiol* **66**, 125–152.
- Lee HM, Flores E, Forchhammer K, Herrero A & Tandeau De Marsac N (2000)** Phosphorylation of the signal transducer PII protein and an additional effector are required for the PII-mediated regulation of nitrate and nitrite uptake in the cyanobacterium *Synechococcus* sp. PCC 7942. *Eur J Biochem* **267**, 591–600.
- Lee MH, Scherer M, Rigali S, James W & Golden JW (2003)** PlmA, a new member of the GntR family, has plasmid maintenance functions in *Anabaena* sp. strain PCC 7120. *J. Bacteriol.* **185**, 4315–4325.
- Leigh JA & Dodsworth JA (2007)** Nitrogen regulation in bacteria and archaea. *Annu Rev Microbiol* **61**, 349–377.
- Levy C, Pike K, Heyes DJ, Joyce MG, Gabor K, Smidt H, van der Oost J & Leys D (2008)** Molecular basis of halorespiration control by CprK, a CRP-FNR type transcriptional regulator. *Mol Microbiol* **70**, 151–167.
- Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I & Rubio V (2007)** The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. *Proc Natl Acad Sci U S A* **104**, 17644–17649.
- Llácer JL, Fita I & Rubio V (2008)** Arginine and nitrogen storage. *Curr Opin Struct Biol* **18**, 673–681.
- Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K & Rubio V (2010)** Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc Natl Acad Sci U S A* **107**, 15397–15402.
- Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS & Richardson DC (2003)** Structure validation by C α geometry: phi, psi and C β deviation. *Proteins* **50**, 437–450.
- Luque I, Flores E & Herrero A (1994)** Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* **13**, 2862–2869.

- Maheswaran M, Urbanke C & Forchhammer K (2004)** Complex formation and catalytic activation by the PII signaling protein of N-acetyl-L-glutamate kinase from *Synechococcus elongatus* strain PCC 7942. *J Biol Chem* **279**, 55202–55210.
- Mandel M & Higa A (1970)** Calcium-dependent bacteriophage DNA infection. *J Mol Biol* **53**, 159–162.
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M & Penny D (2002)** Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A* **99**, 12246–12251.
- Matsui M, Tomita M & Kanai A (2013)** Comprehensive computational analysis of bacterial CRP/FNR superfamily and its target motifs reveals stepwise evolution of transcriptional networks. *Genome Biol Evol* **5**, 267–282.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007)** Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
- McKay DB & Steitz TA (1981)** Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**, 744–749.
- Mérida A, Flores E & Florencio FJ (1992)** Regulation of *Anabaena* sp. strain PCC 7120 glutamine synthetase activity in a *Synechocystis* sp. strain PCC 6803 derivative strain bearing the *Anabaena glnA* gene and a mutated host *glnA* gene. *J Bacteriol* **174**, 650–654.
- Mesa S, Ucurum Z, Hennecke H & Fischer H-M (2005)** Transcription activation in vitro by the *Bradyrhizobium japonicum* regulatory protein FixK2. *J Bacteriol* **187**, 3329–3338.
- Miller JH (1972)** *Experiments in Molecular Genetics*. New York, USA. Cold Spring Harbor Laboratory Press
- Mitschke J, Vioque A, Haas F, Hess WR & Muro-Pastor AM (2011)** Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc Natl Acad Sci U S A* **108**, 20130–20135.

- Mitsui A, Kumazawa S, Takahashi A, Ikemoto H, Cao S & Arai T (1986)** Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. *Nature* **323**, 720–722.
- Muro-Pastor MI, Reyes JC & Florencio FJ (2001)** Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. *J Biol Chem* **276**, 38320–38328.
- Murshudov GN, Vagin AA & Dodson EJ (1997)** Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
- Napoli AA, Lawson CL, Ebright RH & Berman HM (2006)** Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: recognition of pyrimidine-purine and purine-purine steps. *J Mol Biol* **357**, 173–183.
- Ohmori M, Ohmori K & Hasunuma K (1988)** Rapid change in cyclic-3',5'-AMP concentration triggered by a light-off or light-on signal in *Anabaena cylindrica*. *Arch. Microbiol* **150**, 203–204
- Olson WK, Gorin AA, Lu XJ, Hock LM & Zhurkin VB (1998)** DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci U S A* **95**, 11163–11168.
- Omagari K, Yoshimura H, Takano M, Hao D, Ohmori M, Sarai A & Suyama A (2004)** Systematic single base-pair substitution analysis of DNA binding by the cAMP receptor protein in cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS Lett* **563**, 55–58.
- Omagari K, Yoshimura H, Suzuki T, Takano M, Ohmori M & Sarai A (2008)** DeltaG-based prediction and experimental confirmation of SYCRP1-binding sites on the *Synechocystis* genome. *FEBS J* **275**, 4786–4795.
- Painter J & Merritt EA (2006)** Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439–450.
- Palanca C, Pedro-Roig L, Llácer JL, Camacho M, Bonete MJ & Rubio V (2014)** The structure of a PII signaling protein from a halophilic archaeon reveals novel traits and high-salt adaptations. *FEBS J* **281**, 3299–3314.

- Parkinson G, Wilson C, Gunasekera A, Ebright YW, Ebright RH & Berman HM (1996)** Structure of the CAP-DNA complex at 2.5 Å resolution: a complete picture of the protein-DNA interface. *J Mol Biol* **260**, 395–408.
- Passner JM, Schultz SC & Steitz TA (2000)** Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. *J Mol Biol* **304**, 847–859.
- Percudani R & Peracchi A (2009)** The B₆ database: a tool for the description and classification of vitamin B₆-dependent enzymatic activities and of the corresponding protein families. *BMC Bioinformatics* **10**, 273.
- Picossi S, Flores E & Herrero A (2014)** ChIP analysis unravels an exceptionally wide distribution of DNA binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. *BMC Genomics* **15**, 22.
- Plecko B, Zweier M, Begemann A, Mathis D, Schmitt B, Striano P, Baethmann M, Vari MS, Beccaria F, Zara F, Crowther LM, Joset P, Sticht H, Papuc SM & Rauch A (2017)** Confirmation of mutations in PROSC as a novel cause of vitamin B₆-dependent epilepsy. *J Med Genet* **54**, 809-814
- Popovych N, Tzeng S-R, Tonelli M, Ebright RH & Kalodimos CG (2009)** Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc Natl Acad Sci U S A* **106**, 6927–6932.
- Prunetti L, El Yacoubi B, Schiavon CR, Kirkpatrick E, Huang L, Bailly M, ElBadawi-Sidhu M, Harrison K, Gregory JF, Fiehn O, Hanson AD & de Crécy-Lagard V (2016)** Evidence That COG0325 Proteins are involved in PLP Homeostasis. *Microbiology* **162**, 694–706.
- Reddy MCM, Palaninathan SK, Bruning JB, Thurman C, Smith D & Sacchettini JC (2009)** Structural insights into the mechanism of the allosteric transitions of *Mycobacterium tuberculosis* cAMP receptor protein. *J Biol Chem* **284**, 36581–36591.
- Reyes JC & Florencio FJ (1994)** A new type of glutamine synthetase in cyanobacteria: the protein encoded by the *glnN* gene supports nitrogen assimilation in *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **176**, 1260–1267.

- Reyes JC, Muro-Pastor MI & Florencio FJ (1997)** Transcription of glutamine synthetase genes (*glnA* and *glnN*) from the cyanobacterium *Synechocystis* sp. strain PCC 6803 is availability is differently regulated in response to nitrogen availability. *J Bacteriol* **179**, 2678–2689.
- Rubin BE, Wetmore KM, Price MN, Diamond S, Shultzaberger RK, Lowe LC, Curtin G, Arkin AP, Deutschbauer A & Golden SS (2015)** The essential gene set of a photosynthetic organism. *Proc Natl Acad Sci U S A* **112**, 6634–6643.
- Sant’Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC & Schrank IS (2009)** The PII superfamily revised: a novel group and evolutionary insights. *J Mol Evol* **68**, 322–336.
- Sauer J, Dirmeier U & Forchhammer K (2000)** The *Synechococcus* strain PCC 7942 *glnN* product (glutamine synthetase III) helps recovery from prolonged nitrogen chlorosis. *J Bacteriol* **182**, 5615–5619.
- Schultz SC, Shields GC & Steitz TA (1991)** Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **253**, 1001–1007.
- Sharma H, Yu S, Kong J, Wang J & Steitz TA (2009)** Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proc Natl Acad Sci U S A* **106**, 16604–16609.
- Shaw G, Gan J, Zhou YN, Zhi H, Subburaman P, Zhang R, Joachimiak A, Jin DJ & Ji X (2008)** Structure of RapA, a Swi2/Snf2 protein that recycles RNA polymerase during transcription. *Structure* **16**, 1417–1427.
- Stanier RY & Cohen-Bazire G (1977)** Phototrophic prokaryotes: the cyanobacterial. *Annu Rev Microbiol* **31**, 225–74.
- Stanier RY, Deruelles J, Rippka R, Herdman M & Waterbury JB (1979)** Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* **111**, 1–61.
- Steinhauser D, Fernie AR & Araújo WL (2012)** Unusual cyanobacterial TCA cycles: not broken just different. *Trends Plant Sci* **17**, 503–509.

Steinmetz M, Le Coq D, Djemia HB & Gay P (1983) Genetic analysis of *sacB*, the structural gene of a secreted enzyme, levansucrase of *Bacillus subtilis*. *Mol Gen Genet* **191**, 138–144.

Stols L, Gu M, Dieckman L, Raffin R, Collart FR & Donnelly MI (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8–15.

Su Z, Olman V, Mao F & Xu Y (2005) Comparative genomics analysis of NtcA regulons in cyanobacteria: regulation of nitrogen assimilation and its coupling to photosynthesis. *Nucleic Acids Res* **33**, 5156–5171.

Suvorova IA, Korostelev YD & Gelfand MS (2015) GntR family of bacterial transcription factors and their DNA binding motifs: structure, positioning and co-evolution. *PLoS ONE* **10**, e0132618.

Suzuki, T., Yoshimura, H., Ehira, S., Ikeuchi, M., and Ohmori, M. (2007) AnCrpA, a cAMP receptor protein, regulates *nif*-related gene expression in the cyanobacterium *Anabaena sp.* strain PCC 7120 grown with nitrate. *FEBS Lett* **581**, 21–28.

Tandeau de Marsac N & Houmard J (1993) Adaptation of cyanobacteria to environmental stimuli: new steps towards molecular mechanisms. *FEMS Microbiol Lett* **104**, 119–189.

Tanigawa R, Shirokane M, Maeda Si S, Omata T, Tanaka K & Takahashi H (2002) Transcriptional activation of NtcA-dependent promoters of *Synechococcus sp.* PCC 7942 by 2-oxoglutarate in vitro. *Proc Natl Acad Sci U S A* **99**, 4251–4255.

Townsend PD, Jungwirth B, Pojer F, Bußmann M, Money VA, Cole ST, Pühler A, Tauch A, Bott M, Cann MJ & Pohl E (2014) The crystal structures of apo and cAMP-bound GlxR from *Corynebacterium glutamicum* reveal structural and dynamic changes upon cAMP binding in CRP/FNR family transcription factors. *PLoS ONE* **9**, e113265

Tremiño L, Forcada-Nadal A, Contreras A & Rubio V (2017) Studies on cyanobacterial protein PipY shed light on structure, potential functions, and vitamin B₆ -dependent epilepsy. *FEBS Lett* **591**, 3431–3442.

- Tremiño L, Forcada-Nadal A & Rubio V (2018)** Insight into vitamin B₆-dependent epilepsy due to *PLPBP* (previously *PROSC*) missense mutations. *Hum Mutat* **39**, 1002–1013.
- Truan D, Huergo LF, Chubatsu LS, Merrick M, Li X-D & Winkler FK (2010)** A new PII protein structure identifies the 2-oxoglutarate binding site. *J Mol Biol* **400**, 531–539.
- Vagin A & Teplyakov A (2010)** Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.
- Vaguine AA, Richelle J & Wodak SJ (1999)** SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191–205.
- Valladares A, Rodríguez V, Camargo S, Martínez-Noël GMA, Herrero A & Luque I (2011)** Specific role of the cyanobacterial PipX factor in the heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* **193**, 1172–1182.
- Van Heeswijk WC, Westerhoff HV & Boogerd FC (2013)** Nitrogen assimilation in *Escherichia coli*: putting molecular data into a systems perspective. *Microbiol Mol Biol Rev* **77**, 628–695.
- Vázquez-Bermúdez MF, Herrero A & Flores E (2002a)** 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. *FEBS Lett* **512**, 71–74.
- Vázquez-Bermúdez MF, Flores E & Herrero A (2002b)** Analysis of binding sites for the nitrogen-control transcription factor NtcA in the promoters of *Synechococcus* nitrogen-regulated genes. *Biochim Biophys Acta* **1578**, 95–98.
- Vijayan V, Jain IH & O’Shea EK (2011)** A high resolution map of a cyanobacterial transcriptome. *Genome Biol* **12**, R47
- Volbeda A, Darnault C, Renoux O, Nicolet Y & Fontecilla-Camps JC (2015)** The crystal structure of the global anaerobic transcriptional regulator FNR explains its extremely fine-tuned monomer-dimer equilibrium. *Sci. Adv* **1**, e1501086.

- Weber IT & Steitz TA (1987)** Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. *J Mol Biol* **198**, 311–326.
- Weber IT, Gilliland GL, Harman JG & Peterkofsky A (1987)** Crystal structure of a cyclic AMP-independent mutant of catabolite gene activator protein. *J Biol Chem* **262**, 5630–5636.
- Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R.F., Sykes, B.D., and Wishart, D.S. (2003)**. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* **31**, 3316–3319.
- Winn MD, Isupov MN & Murshudov GN (2001)** Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* **57**, 122–133.
- Xiong W, Brune D & Vermaas WFJ (2014)** The γ -aminobutyric acid shunt contributes to closing the tricarboxylic acid cycle in *Synechocystis* sp. PCC 6803. *Mol Microbiol* **93**, 786–796.
- Xu Y, Carr PD, Clancy P, Garcia-Dominguez M, Forchhammer K, Florencio F, Tandeau de Marsac N, Vasudevan SG & Ollis DL (2003)** The structures of the PII proteins from the cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803. *Acta Crystallogr D Biol Crystallogr* **59**, 2183–2190.
- Xu M & Su Z (2009)** Computational prediction of cAMP receptor protein (CRP) binding sites in cyanobacterial genomes. *BMC Genomics* **10**, 23.
- Yoshimura H, Hisabori T, Yanagisawa S & Ohmori M (2000)** Identification and characterization of a novel cAMP receptor protein in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem* **275**, 6241–6245.
- Yoshimura, H., Yoshihara, S., Okamoto, S., Ikeuchi, M., and Ohmori, M. (2002a)**. A cAMP receptor protein, SYCRP1, is responsible for the cell motility of *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* **43**, 460–463.

- Yoshimura H, Yanagisawa S, Kanehisa M & Ohmori M (2002)** Screening for the target gene of cyanobacterial cAMP receptor protein SYCRP1. *Mol Microbiol* **43**, 843–853.
- Yoshimura H, Kaneko Y, Ehira S, Yoshihara S, Ikeuchi M & Ohmori M (2010)** *CccS* and *CccP* are involved in construction of cell surface components in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Cell Physiol* **51**, 1163–1172.
- Zeth K, Fokina O & Forchhammer K (2014)** Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *J Biol Chem* **289**, 8960–8972.
- Zhang S & Bryant DA (2011)** The tricarboxylic acid cycle in cyanobacteria. *Science* **334**, 1551–1553.
- Zhang S & Bryant DA (2015)** Biochemical validation of the glyoxylate cycle in the cyanobacterium *Chlorogloeopsis fritschii* Strain PCC 9212. *J Biol Chem* **290**, 14019–14030.
- Zhao M-X, Jiang Y-L, Xu B-Y, Chen Y, Zhang C-C & Zhou C-Z (2010a)** Crystal structure of the cyanobacterial signal transduction protein PII in complex with PipX. *J Mol Biol* **402**, 552–559.
- Zhao M-X, Jiang Y-L, He Y-X, Chen Y-F, Teng Y-B, Chen Y, Zhang C-C & Zhou C-Z (2010b)** Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proc Natl Acad Sci U S A* **107**, 12487–12492.
- Zink T, Ross A, Lüers K, Cieslar C, Rudolph R & Holak TA (1994)** Structure and dynamics of the human granulocyte colony-stimulating factor determined by NMR spectroscopy. Loop mobility in a four-helix-bundle protein. *Biochemistry* **33**, 8453–8463.

